The Role of Human Capital and Technological Interdependence in Growth and Convergence Processes: International Evidence

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: January 25, 2005

Abstract

This paper develops a bisectorial growth model with physical and human capital accumulation. Each sector is characterized by a different technology involving different human capital parameters. The model includes human capital externalities together with technological interdependence between economies. It leads to a spatial autoregressive reduced form for the real income per worker at steady state. The structural parameters of the model are recovered and evidence of the insignificance of human capital in explaining per capita growth, that is the human capital puzzle, is reconsidered. In fact, the parameter related to human capital in the consumption good sector is low which is consistent with evidence presented in the growth accounting framework. In contrast it is indeed higher in the education sector. Our model leads to spatial econometric specifications which are estimated on a sample of 89 countries over the period 1960-1995 using maximum likelihood as well as Bayesian estimation methods, which are robust versus outliers and heteroskedasticity. This model yields a spatially augmented convergence equation characterized by parameter heterogeneity. A locally linear spatial autoregressive specification is then estimated providing a different convergence speed estimate for each country of the sample.

KEYWORDS: Conditional convergence, technological interdependence, spatial autocorrelation, parameter heterogeneity, locally linear estimation

JEL: C14, C21, O4

*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

How does human capital or the educational attainment of the labor force affect the output and the growth of an economy? Where has all the education gone? Many recent empirical studies have found that economic growth appears to be unrelated to increases in educational attainment. Using standard growth accounting equation Benhabib and Spiegel's paper (1994) was the first to point out this puzzle: human capital growth has an insignificant, and usually negative effect in explaining per capita income growth. Another influencial paper is that of Pritchett (2001) who supports also the fact that human capital do not enter significantly in empirical growth regressions in cross-section whereas microeconomic evidence would imply the opposite. These results must be viewed in contrast to that of Mankiw, Romer and Weil (1992) who show that human capital explains per capita levels and growth rates if we consider that countries were near their steady-state. These authors elaborate their specification as an augmented Solow growth model.

However, Benhabib and Spiegel (1994) do not reject the important role of human capital in development and growth processes. They only reject the traditional role given to human capital as a separate factor of production. Therefore, treating human capital simply as another factor in growth accounting may misspecify its role, as in the Mankiw, Romer and Weil (1992) model. Following the seminal paper of Nelson and Phelps (1966), Benhabib and Spiegel (1994) present a model which allows human capital levels to affect the speed of technological catch-up and diffusion. The ability of a nation to adopt and implement new technology from abroad is a function of its domestic human capital stock. Their model then includes technological interdependence and a catch-up process to the world leader in technology depending of the absorption ability embodied in the human capital stock of a given country.

This paper proposes to reconsider the puzzle underlined by Benhabib and Spiegel (1994) and the lack of significant evidence in favor of the growth impact of human capital. With that aim, we develop a bisectorial growth model with human capital externalities and technological interdependence in a world with N countries denoted by $i = 1, ..., N$. Let us consider a Cobb-Douglas production function exhibiting constant return to scale and decreasing return in each factor:

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

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 level of technological progress. Taking logarithms of this expression, Benhabib and Spiegel (1994) and Pritchett (1995) show that the estimated coefficient of human capital, β_K, is not significant implying that the growth rate of human capital is not correlated with the growth rate in output per worker across countries. In contrast to the usual growth accounting framework that assigns output or its growth rate to contributions from $K_i(t)$, $H_i(t)$, $L_i(t)$ and $A_i(t)$, Mankiw et al. (1992) rearrange (1) in order to obtain the real per worker income level depending on the physical capital and human capital output ratios:

$$\frac{Y_i(t)}{L_i(t)} = A_i(t)^{\frac{1}{1-\alpha-\beta_K}} \left(\frac{K_i(t)}{Y_i(t)}\right)^{\alpha/(1-\alpha-\beta_K)} \left(\frac{H_i(t)}{Y_i(t)}\right)^{\beta_K/(1-\alpha-\beta_K)} \tag{2}$$

Using the fundamental dynamic equation of Solow (Solow, 1956) describing accumulation of physical and human capital, they find that this specification performs quite well and find structural parameter values close to those expected. Indeed, these contrasting results indicate a puzzle. However, we claim that the Mankiw et al. (1992) model is misspecified so that we modify their model in two manners. First, as underlined by Klenow and Rodriguez-Clare (1997), the physical capital and the human capital sectors can have different production functions, so that we consider a more general bisectorial growth model assuming that the coefficients of human capital are different across sectors. Second, equation (2) can be viewed as $\frac{Y_i(t)}{L_i(t)} = AX$ where A is related to the level of technological progress and X is a composite of the two capital intensities (Klenow and Rodriguez-Clare 1997). As underlined by Romer (1993), ideas gaps are more important than object gaps, in other words, the effect of A are more important in explaining income disparities and growth than X. So, and in accordance with Benhabib and Spiegel (1994), we assume that technological progress is characterized by worldwide global interdependence between countries. In fact, an important characteristic of technology is its capacity to diffuse across borders (Coe and Helpman, 1995 ; Eaton and Kortum, 1996; Keller, 2004). Therefore, we explicitly include this global technological interdependence in our model. Moreover, as the study of transitional dynamics is an important feature in economic growth models, we will also investigate the convergence process.

Ertur and Koch (2005), show that omitting technological interdependence in the traditional Solow growth model, yields a misspecified econometric reduced form and biased structural parameters. In particular, they show that the capital share in income is close to one third without adding human capital as an additional production factor. In this paper, we also obtain a spatial econometric reduced form implying that omitted technological interdependence involves spatial autocorrelation which bias the traditional results about the influence of human capital on growth and development. Our results lead to the rejection of human capital as a simple production factor along with raw labor and physical capital stock. Therefore, when we consider technological interdependence, we reject the Mankiw, Romer and Weil (1992) specification and the results are more in accordance with those of Benhabib and Spiegel (1994) and Nelson and Phelps (1966).

In order to estimate our model we use spatial econometric methods. In fact, when we introduce global technological interdependence in the theoretical model we obtain a spatial autoregressive model as a reduced form, which is estimated using maximum likelihood (Lee, 2004). However, outliers and heteroscedasticity can bias results as underlined by Temple (1998, 1999b) in both the context of the conditional convergence equation and the context of the human capital puzzle. Use of non-parametric and robust estimation methods have thus been advocated. We propose to use an alternative approach relying on Bayesian spatial econometric methods which are robust with respect to heteroskedasticty and outliers (LeSage 1997, 2002). Moreover, our model is characterized by inherent parameter heterogeneity. In fact, as underlined by Durlauf and Johnson (1995) or Durlauf et al. (2001) for instance, parameter heterogeneity in the conditional convergence equation is a fundamental issue that is potentially plaguing empirical results obtained in the growth regression literature. It is well known that technological interdependence leads to clubs (Lucas, 1993; Durlauf and Quah, 1999) because the access to different international levels of technology imply different equilibria. Our bisectorial growth model leads to parameter heterogeneity because each economy has its own specific access to the world technology as in Ertur and Koch (2005). Therefore, we use the Spatial Autoregressive Local Estimation (SALE) method developed by Lesage and Pace (2004) in order to estimate our final model. This method allows taking into account both parameter heterogeneity and spatial autocorrelation, two effects embedded in our theoretical model.

2 The model

2.1 The bisectorial model

Many authors have studied bisectorial growth models including a sector describing the accumulation of a consumption good and another sector describing education and accumulation of human capital (Mulligan and Sala-i-Martin (1993), Rebelo (1991), Mankiw, Romer and Weil (1992), Uzawa (1965), Lucas (1988, 1993). Our model also includes two sectors: a consumption good sector producing an homogeneous good both accumulated as physical capital and consumed at each period and an education sector producing human capital. We suppose that each sector uses a different Cobb-Douglas production function and the two factors are used in the same proportion in each sector. More precisely, the consumption good sector uses a fraction equal to $(1 - s_{i,H})$ of the physical capital stock available in country i and the workers use the same fraction of their time in the productive sector.[1] The education sector uses the remaining fraction $s_{i,H}$ of the physical capital stock and the remaining fraction of the time of workers. In contrast to the endogenous bisectoriel growth models elaborated by Mulligan and Sala-i-Martin (1993), Rebelo (1991) and Lucas (1988, 1993) among others, we assume that both sectors use a nonreproducible input (raw labor).[2]

The consumption good sector uses the production function (1) with the proportion $(1 - s_{i,H})$ of the production factors. The output in this sector $Y_{i,K}(t)$ can be written as follows :

$$Y_{i,K}(t) = A_i(t)((1 - s_{i,H})K_i(t))^{\alpha}((1 - s_{i,H})H_i(t))^{\beta_K}((1 - s_{i,H})L_i(t))^{1-\alpha-\beta_K} \quad (3)$$

[1]This hypothesis is the same as the one in Mankiw, Romer and Weil (1992) and all empirical papers use it for evident algebraic manipulations. Some theoretical papers generalize this hypothesis supposing that economic agents choose their time and physical capital allocation by maximizing an utility function as Lucas (1988), Rebelo (1991) and Mulligan and Sala-i-Martin (1993) for instance.

[2]It is well known that: "All that is required to assure the feasibility of perpetual growth is existence of a "core" of capital goods that is produced with constant returns technologies and without the direct or indirect use of nonreproducible factors" (Rebelo, 1991, p.502). This idea is originally due to Von Neumann (1945).

Physical capital is assumed to be forgone consumption in the following form:

$$\dot{K}_i(t) + \delta K_i(t) + C_i(t) = Y_{i,K}(t) \tag{4}$$

where δ is the constant rate of depreciation of physical capital.[3] Consumption is defined as following: $C_i(t) = (1 - s_{i,K})Y_{i,K}(t)$ where $s_{i,K}$ is the part of the output which is saved in the country i and invested in the physical capital sector. We obtain finally the following equation describing the variation of physical capital per worker:

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

where $k_i(t) = K_i(t)/L_i(t)$ is the per worker physical capital, $y_{i,K}(t) = Y_{i,K}(t)/L_i(t)$ is the per worker real income. Because of hypothesis on decreasing returns, $\alpha + \beta_K < 1$, the physical capital per worker converges to the steady state growth rate denoted by g.[4] At steady state, the capital-output ratio is constant and equal to:

$$\frac{k_i^\star(t)}{y_{i,K}^\star(t)} = \frac{s_{K,i}}{n_i + g + \delta} \tag{6}$$

The per worker physical capital at steady-steady is then defined as follows:

$$k_i^\star(t) = \left(\frac{A_i(t)s_{K,i}}{n_i + g + \delta}\right)^{\frac{1}{1-\alpha}} h_i^\star(t)^{\frac{\beta_K}{1-\alpha}} \tag{7}$$

Human capital is supposed produced with the following specification:

$$\dot{H}_i(t) + \delta H_i(t) = Y_{H,i}(t) \tag{8}$$

where $Y_{H,i}(t)$ is the human capital output. Note that the important asymmetric hypothesis in bisectorial models is that accumulation of physical capital $K_i(t)$ is a perfect substitute for consumption. In other words, the consumption subtracts from $\dot{K}_i(t)$ and not from

[3] We suppose that the constant rate of depreciation of human capital is identical to that of physical capital and furthermore identical for all N countries.

[4] At steady state, each variable in the model grows at the same rate: $g = \frac{g_A}{1-\alpha-\beta_K}$ where g_A is the rate of growth of technical progress.

$\dot{H}_i(t)$ in the educative sector.[5] The human capital output is defined as:

$$Y_{H,i}(t) = A_i(t)(s_{H,i}K_i(t))^{\alpha}(s_{H,i}H_i(t))^{\beta_H}(s_{H,i}L_i(t))^{1-\alpha-\beta_H} \tag{9}$$

with $\alpha+\beta_H < 1$. $s_{i,H}$ is often interpreted as the part of output which is saved and invested in the education sector because of the hypothesis of identical technology in both sectors. We can write the evolution of the human capital stock per worker in the following form:

$$\dot{h}_i(t) = s_{H,i}k_i(t)^{\alpha}h_i(t)^{\beta_H} - (n_i+\delta)h_i(t) \tag{10}$$

As the physical capital per worker growth rate, the human capital per worker growth rate converges to g. At steady state, we have:

$$h_i^{\star}(t) = \left(\frac{A_i(t)s_{H,i}}{n_i+g+\delta}\right)^{\frac{1}{1-\beta_H}} k_i^{\star}(t)^{\frac{\alpha}{1-\beta_H}} \tag{11}$$

After rearranging terms, we obtain finally:

$$\begin{aligned} \ln y_i^{\star}(t) &= \frac{1}{1-\alpha-\beta_K}\ln A_i(t) + \frac{\alpha(1-\beta_H+\beta_K)}{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)}\ln\left(\frac{s_{K,i}}{n_i+g+\delta}\right) \\ &+ \frac{\beta_K}{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)}\ln\left(\frac{s_{H,i}}{n_i+g+\delta}\right) \end{aligned} \tag{12}$$

This equation shows how real income per worker depends on population growth and accumulation of physical and human capital. The qualitative predictions are essentially identical to those of Mankiw et al.(1992). If we consider that both sectors use the same production function, we have $\beta_K = \beta_H$ and we obtain their model. However, the expected values of the coefficients are different. In fact, along the lines of Benhabib and Spiegel (1994) and Pritchett (2001), the coefficient β_K should be close to zero. However, let us now introduce global technological interdependence in the model.

[5]In contrast, the early bisectorial models supposed that capital goods and consumption goods were not perfect substitute and were produced in two different sectors using different production functions (Uzawa, 1963 and Meade, 1961, for instance). Their models are more complicated because they include additional relative prices and are not oriented towards empirical applications.

2.2 The global technological interdependence

One of the well-known characteristics of the Mankiw, Romer and Weil (1992) model is the convergence of all countries to the same growth rate defined by the common exogenous technical progress. Their model implies a global technological interdependence without frictions. In contrast, we assume that technological interdependence is not homogenous between all countries and depends on their connectivity scheme with all other countries. Therefore, we assume that the level of technical progress is defined as following:

$$A_i(t) = \Omega(t)\left(\frac{h_i(t)}{k_i(t)}\right)^{\phi}\prod_{j\neq i}^{N} A_j(t)^{\gamma w_{ij}} \tag{13}$$

This technical progress depends on a part which identically grows at a constant rate μ in all countries: $\Omega(t) = \Omega(0)e^{\mu t}$, where $\Omega(0)$ is initial technology. It depends also on the ratio of accumulated stocks of reproducible factors per worker $k_i(t)$ and $h_i(t)$ with $0 < \phi < 1$ the level of externalities. More precisely, ϕ can be viewed either as externalities generated by human capital accumulation as in Lucas (1988, 1993) or as an essential component for creating new knowledge as in Romer (1990). In fact, human capital may speed the adoption of technology as in Nelson and Phelps (1966), Schultz (1975), Benhabib and Spiegel (1994) or Desdoigts (2004), allowing catching-up with the world technology frontier. It can also be viewed as indexing the fraction of frontier technology which the country can use. We adopt the latter "use" formulation as also adopted by Easterly et al. (1994), Jones (2004), Caselli (1999) and Bils and Klenow (2000), for instance.

Therefore, as recently proposed by Ertur and Koch (2005), the technological interdependence is modeled by the last term: the level of technology of a country i depends on the level of technology in other countries denoted by $j = 1, ..., N$, and $j \neq i$. γ describes the level of technological interdependence between economies. However, each country has a different access to this international technology because of the connectivity parameters denoted by w_{ij}. We suppose that these terms are non negative, non stochastic and finite; we have also $0 \leq w_{ij} \leq 1$ and $w_{ij} = 0$ if $i = j$. We also assume that $\sum_{j\neq i}^{N} w_{ij} = 1$ for $i = 1, ..., N$. These terms are specific for each country and describe the access of a given country i to the international technology. Note that that if $w_{ij} = 1$ if j is the leading edge technology and 0 elsewhere we obtain a model close to those of Benhabib and Spiegel

(1994) or Howitt (2000), for instance. However, our model implies a richer technological structure.

Taking logarithms of expression (13), we obtain:

$$\ln A_i(t) = \ln \Omega(t) + \phi \ln\left(\frac{h_i(t)}{k_i(t)}\right) + \gamma \sum_{j \neq i}^{N} w_{ij} \ln A_j(t) \tag{14}$$

and after some straightforward manipulations:

$$\ln A_i(t) = \frac{1}{1-\gamma} \ln \Omega(t) + \frac{\phi}{1-\gamma} \ln\left(\frac{h_i(t)}{k_i(t)}\right) + \frac{\gamma}{1-\gamma} \sum_{j \neq i}^{N} w_{ij} (\ln A_j(t) - \ln A_i(t)) \tag{15}$$

This equation shows that the level of technology of a country i is high if its of ratio human capital to physical capital is high. Moreover, the last term shows that the level of technology is high if the country i is far away its technology frontier defined by others countries' technology level. Each country has its own "world" technology frontier which is specific because of the parameters w_{ij} defining the connectivity structure between all countries. When all countries are at their steady states, $A_i(t)$ for $i = 1, ..., N$, grows at rate $g_A = \frac{\mu}{1-\gamma}$ so that each variables in all countries, $k_i(t)$, $h_i(t)$, $y_{K,i}(t)$ and $y_{H,i}(t)$ grows at the same rate defined by: $g = \frac{\mu}{(1-\alpha-\beta_K)(1-\gamma)}$.
This global interdependence structure implies that countries must be analyzed as an interdependent system. Taking equation (14) in matrix form, we have:

$$A = \Omega + \phi(h - k) + \gamma W A \tag{16}$$

with A the $(N \times 1)$ vector of the logarithms of the level of technology, k the $(N \times 1)$ vector of the logarithms of the aggregate level of physical capital per worker, h the $(N \times 1)$ vector of the logarithms of the aggregate level of human capital per worker and W the $(N \times N)$ Markov-matrix with connectivity terms w_{ij}. We can resolve (16) for A, if $\gamma \neq 0$ and if $1/\gamma$ is not an eigenvalue of W:[6]

$$A = (I - \gamma W)^{-1}\Omega + \phi(I - \gamma W)^{-1}(h - k) \tag{17}$$

[6]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 $|W - (1/\gamma)I| \neq 0$.

we can develop (17) if $|\gamma| < 1$ and regroup terms to obtain:

$$A = \frac{1}{1-\gamma}\Omega + \phi(h-k) + \phi \sum_{r=1}^{\infty} \gamma^r W^{(r)}(h-k) \tag{18}$$

where $W^{(r)}$ is the matrix W to the power of r. For a country i, we have:

$$A_i(t) = \Omega(t)^{\frac{1}{1-\gamma}} \left(\frac{h_i(t)}{k_i(t)}\right)^{\phi} \prod_{j=1}^{N} \left(\frac{h_j(t)}{k_j(t)}\right)^{\phi \sum_{r=1}^{\infty} \gamma^r w_{ij}^{(r)}} \tag{19}$$

The level of technology in a country i depends on its own human capital to physical capital ratio and on the same ratio in other countries j in its neighborhood defined by the connectivity parameters w_{ij}. Replacing (19) in the production function (3) written per worker, we obtain:

$$y_{i,K}(t) = \Omega(t)^{\frac{1}{1-\gamma}} k_i(t)^{u_{K,ii}} h_i(t)^{v_{K,ii}} \prod_{j \neq i}^{N} k_j(t)^{u_{K,ij}} \prod_{j \neq i}^{N} h_j(t)^{v_{K,ij}} \tag{20}$$

with: $u_{K,ii} = \alpha - \phi\left(1 + \sum_{r=1}^{\infty} \gamma^r w_{ii}^{(r)}\right)$, $v_{K,ii} = \beta_K + \phi\left(1 + \sum_{r=1}^{\infty} \gamma^r w_{ii}^{(r)}\right)$ and $v_{K,ij} = -u_{K,ij} = \phi \sum_{r=1}^{\infty} \gamma^r w_{ij}^{(r)}$. The terms $w_{ij}^{(r)}$ are the elements of row i and column j of the matrix W to the power of r, and $y_{i,K}(t) = Y_{i,K}(t)/(1 - s_{i,H})L_i(t)$ the level of output per worker. We can note that the social elasticity of income per worker in a country i with respect to all physical and human capital is identical to the private elasticity, that is: $\alpha + \beta_K < 1$. However, this model implies heterogeneity in the parameters of the production function since they are specific to each country because of the connectivity parameters w_{ij} defining the global technological interdependence. In fact, each country accesses differently to the world technology so that each country converges locally to its own technology frontier defined by the last term of our technology function (13).

2.3 Steady-state

Using the production function (20), we can rewrite the fundamental dynamic equation of Solow as follows:

$$\dot{k}_i(t) = s_{K,i}\Omega(t)^{\frac{1}{1-\gamma}} k_i(t)^{u_{K,ii}} h_i(t)^{v_{K,ii}} \prod_{j \neq i}^{N} k_j(t)^{u_{K,ij}} \prod_{j \neq i}^{N} h_j(t)^{v_{K,ij}} - (n_i + \delta)k_i(t) \tag{21}$$

Since the production function is characterized by decreasing returns, equation (21) implies that the physical capital growth rate of country i, for $i = 1, ..., N$, converges to a balanced growth rate defined by g. Therefore:

$$k_i^\star(t) = \Omega(t)^{\frac{1}{(1-\gamma)(1-u_{K,ii})}} \left(\frac{s_{K,i}}{n_i + g + \delta}\right)^{\frac{1}{1-u_{K,ii}}} h_i^\star(t)^{\frac{v_{K,ii}}{1-u_{K,ii}}} \prod_{j \neq i}^{N} k_j^\star(t)^{\frac{u_{K,ij}}{1-u_{K,ii}}} \prod_{j \neq i}^{N} h_j^\star(t)^{\frac{v_{K,ij}}{1-u_{K,ii}}} \tag{22}$$

The per worker physical capital at steady state in a country i depends positively on its own human capital and on the human capital to physical capital ratio in neighboring countries because of technological interdependence. In the same way, the fundamental dynamic equation for the human capital can be rewritten as follows:

$$\dot{h}_i(t) = s_{H,i}\Omega(t)^{\frac{1}{1-\gamma}} k_i(t)^{u_{H,ii}} h_i(t)^{v_{H,ii}} \prod_{j \neq i}^{N} k_j(t)^{u_{H,ij}} \prod_{j \neq i}^{N} h_j(t)^{v_{H,ij}} - (n_i + \delta) h_i(t) \tag{23}$$

The human capital growth rate in country i, for $i = 1, ..., N$, converges to a balanced growth rate defined by g. Therefore:

$$h_i^\star(t) = \Omega(t)^{\frac{1}{(1-\gamma)(1-v_{H,ii})}} \left(\frac{s_{H,i}}{n_i + g + \delta}\right)^{\frac{1}{1-v_{H,ii}}} k_i^\star(t)^{\frac{u_{H,ii}}{1-v_{H,ii}}} \prod_{j \neq i}^{N} k_j^\star(t)^{\frac{u_{H,ij}}{1-v_{H,ii}}} \prod_{j \neq i}^{N} h_j^\star(t)^{\frac{v_{H,ij}}{1-v_{H,ii}}} \tag{24}$$

As the per worker physical capital at steady state, the per worker human capital at steady state also depends on the human capital to physical capital ratio in neighboring countries because of technological interdependence. We can replace these two expressions of physical and human capital at steady state for $i = 1, ..., N$ in the production function in order to obtain the real income per worker at steady state:

$$
\begin{aligned}
\ln y^{\star}_{i,K}(t) &= \frac{1-\beta_H+\beta_K}{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)+\phi(\beta_K-\beta_H)}\ln\Omega(t) \\
&+ \frac{(\alpha-\phi)(1-\beta_H+\beta_K)}{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)+\phi(\beta_K-\beta_H)}\ln\left(\frac{s_{i,K}}{n_i+g+\delta}\right) \\
&+ \frac{\beta_K+\phi}{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)+\phi(\beta_K-\beta_H)}\ln\left(\frac{s_{i,H}}{n_i+g+\delta}\right) \\
&- \frac{\alpha(1-\beta_H+\beta_K)}{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)+\phi(\beta_K-\beta_H)}\gamma\sum_{j\neq i}^{N} w_{ij}\ln\left(\frac{s_{j,K}}{n_j+g+\delta}\right) \\
&- \frac{\beta_K}{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)+\phi(\beta_K-\beta_H)}\gamma\sum_{j\neq i}^{N} w_{ij}\ln\left(\frac{s_{j,H}}{n_j+g+\delta}\right) \\
&+ \frac{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)}{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)+\phi(\beta_K-\beta_H)}\gamma\sum_{j\neq i}^{N} w_{ij}\ln y^{\star}_{j,K}(t) \qquad (25)
\end{aligned}
$$

This bisectorial growth model has the same qualitative predictions as the Mankiw et al. model about the influence of the domestic saving rate, human capital investment rate and population growth rate on the real income per worker of a country i at steady state. However, this equation is written in implicit form since the real income per worker at steady state of a country i depends positively on the real income per worker at steady state in countries $j = 1, ..., N$, $j \neq i$. In order to determine the qualitative and quantitative influences of each variables on the real income per worker at steady state, we need to write this equation in explicit form and compute the elasticities describing these influences.[7] Because of global technological interdependence, the real income per worker at steady state of a country i depends positively on the saving rate in physical capital and human capital of all other countries and negatively on their population rate of growth.

2.4 Convergence

Considering the relations describing the growth of physical and human capital, the transitional dynamics can be quantified by using a log linearization of equations (21) and (23) around the steady state, for $i = 1, ..., N$:

$$\frac{d\ln k_i(t)}{dt} = (n_i+g+\delta)[(\ln y_i(t)-\ln y^{\star}_i)-(\ln k_i(t)-\ln k^{\star}_i)] \qquad (26)$$

[7] See Appendix for the expressions of these elasticities.

and:

$$\frac{d\ln h_i(t)}{dt} = (n_i + g + \delta)[(\ln y_{H,i}(t) - \ln y^{\star}_{H,i}) - (\ln h_i(t) - \ln h^{\star}_i)] \tag{27}$$

Replacing these expression in the production function written in logarithms :

$$\begin{aligned}\frac{d\ln y_i(t)}{dt} &= \frac{\mu}{1-\gamma} + \sum_{j=1}^{N} u_{ij}\frac{d\ln k_i(t)}{dt} + \sum_{j=1}^{N} v_{ij}\frac{d\ln h_i(t)}{dt} \\ &= \frac{\mu}{1-\gamma} + \left[\sum_{j=1}^{N}\frac{1}{\Theta_j}(u_{ij} + \frac{1}{\Theta_j^H}v_{ij})(n_j + g + \delta) - \Lambda_i\right](\ln y_i(t) - \ln y^{\star}_i) \quad (28)\end{aligned}$$

The expression between brackets represents the convergence speed, denoted by λ_i. This expression depends on three scale parameters. First, it depends on the parameter Λ_i representing the fact that each country has a different population growth rate n_i. If all countries have the same population growth rate, in the context of unconditional convergence for example, the scale parameter reduces to 1. The second parameter, Θ_j^H is the parameter representing the differences between technologies. In fact, this parameter represents a proportionality coefficient specific to each country as follows:

$$\ln y_i(t) - \ln y^{\star}_i = \Theta_j^H(\ln y_{H,i}(t) - \ln y^{\star}_{H,i}) \quad \text{for} \quad i = 1, ...N \tag{29}$$

If all countries use the same technology in the physical capital and the human capital sectors, as in the MRW model, for instance, this scale parameter drops. Finally, the parameter Θ_j represents the effects of global technological interdependence between all countries. It represents the relations between the gaps of countries in respect to their own steady states:

$$\ln y_i(t) - \ln y^{\star}_i = \Theta_j(\ln y_j(t) - \ln y^{\star}_j) \quad \text{for} \quad j = 1, ...N \tag{30}$$

If this parameter is equal to 1, the countries i and j are in the same distance in respect to their own steady states. If this parameter is higher than 1 (respectively lower than 1) the country i is farther (respectively closer) to its own steady than the country j.

The solution for $\ln y_i(t)$, subtracting $\ln y_i(0)$, the real income per worker at some initial date, from both sides, is:

$$\begin{aligned}\ln y_i(t)-\ln y_i(0) &= (1-e^{-\lambda_i t})\frac{\mu}{1-\gamma}\frac{1}{\lambda_i}-(1-e^{-\lambda_i t})\ln y_i(0)\\ &+ (1-e^{-\lambda_i t})\ln y_i^* \qquad (31)\end{aligned}$$

The model predicts convergence since the growth of real income per worker is a negative function of the initial level of income per worker, but only after controlling for the determinants of the steady state.

We rewrite equation (31) in matrix form: $G = DC - Dy(0) + Dy^*$ where G is the $(N \times 1)$ vector of growth rates of real income per worker, $y(0)$ is the $(N\times 1)$ vector of the logarithms of the initial level of real income per worker, y^* is the $(N \times 1)$ vector of the logarithms of real income per worker at steady state, C is the $(N \times 1)$ vector of constants and D is the $(N \times N)$ diagonal-matrix with $(1-e^{-\lambda_i t})$ terms on the main diagonal. Introducing equation (25) in matrix reduced form, premultiplying both sides by the inverse of $D(I-\rho W)^{-1}$ and rearranging terms we obtain:

$$\begin{aligned}G &= \Delta - Dy(0) + \frac{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)}{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)+\phi(\beta_K-\beta_H)}\gamma DWy(0)\\ &+ \frac{(\alpha-\phi)(1-\beta_H+\beta_K)}{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)+\phi(\beta_K-\beta_H)}DS_K\\ &+ \frac{\beta_K+\phi}{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)+\phi(\beta_K-\beta_H)}DS_H\\ &- \frac{\alpha(1-\beta_H+\beta_K)}{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)+\phi(\beta_K-\beta_H)}\gamma DWS_K\\ &- \frac{\beta_K}{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)+\phi(\beta_K-\beta_H)}\gamma DWS_H\\ &+ \frac{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)}{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)+\phi(\beta_K-\beta_H)}\gamma DWD^{-1}G \qquad (32)\end{aligned}$$

where Δ is an inessential constant.[8] Finally, the growth rate of real income per worker is a negative function of the initial level of income per worker described in the diagonal matrix D, but only after controlling for the determinants of the steady state. More specifically, the growth rate of real income per worker depends positively on its own saving rate in physical capital, on its own investment rate in human capital and negatively on its own population growth rate. Moreover, it depends also, in the same direction, on the same

[8]The expression of this constant is: $\Delta = D\left(C + \frac{1-\beta_H+\beta_K}{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)+\phi(\beta_K-\beta_H)}\Omega\right)$

variables in the neighboring countries because of technological interdependence. We can observe that the growth rate is higher the larger the initial level of income per worker and the higher the growth rate in neighboring countries. Finally, the last term of the equation (32) shows that the growth rate of a country i depends on the growth rate of neighboring countries weighted by their convergence speed and by the connectivity terms. In section 5, we test the predictions of this model. We then show how technological interdependence may influence growth and conditional convergence.

3 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 Mankiw et al. (1992) non-oil sample, for which PWT 6.1 provide data.

We mesure n as the average growth of the working-age population (ages 15 to 64). For this, we have computed the number of workers like Caselli (2004): $RGDPCH \times POP/RGDPW$, 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 measured as the average share of gross investment in GDP as in Mankiw et al. (1992). Finally, the variable s_H, labeled SCHOOL in Mankiw et al. (1992), is the average percentage of a country's working-age population in secondary school. More specifically, Mankiw et al. (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 the data provides by Bernanke and Gurkaynak (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 Markov-matrix W defined in equation (16) corresponds to the so-called spatial weight matrix commonly used in spatial econometrics to model spatial interdependence between regions or countries (Anselin 1988; Anselin and Bera 1998; Anselin 2001). 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.

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 and 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 (e.g. inverse distance, inverse squared distance, negative exponential etc.). The critical cut-off may be the same for all countries or may be defined to be specific to each country leading in this 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 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.[9] 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 $[w1_{ij}]$:

$$w1_{ij} = w1^*_{ij} / \sum_j w1^*_{ij} \quad \text{with} \quad w1^*_{ij} = \begin{cases} 0 \ \text{ if } & i = j \\ d_{ij}^{-2} & \text{otherwise} \end{cases} \tag{33}$$

[9]Klenow and Rodriguez-Clare (2005, p. 28-29) suggest that use of pure geographical distance could capture trade and FDI related spillovers.

where d_{ij} is the great-circle distance between country capitals.[10]

4 Impact of saving, human capital, population growth and technological interdependence on real income

4.1 Econometric specification

In this section, we evaluate the impact of saving, human capital, population growth and global technological interdependence on real income. Taking equation (25), we find that the real income per worker along the balanced growth path at a given time ($t = 0$ for simplicity) is:

$$\begin{aligned} \ln\left[\frac{Y_i}{L_i}\right] &= \beta_0 + \beta_1 \ln s_{K,i} + \beta_2 \ln s_{H,i} + \beta_3 \ln(n_i + 0.05) \\ &+ \theta_1 \sum_{j \neq i}^{N} w_{ij} \ln s_{K,j} + \theta_2 \sum_{j \neq i}^{N} w_{ij} \ln s_{H,j} \\ &+ \theta_3 \sum_{j \neq i}^{N} w_{ij} \ln(n_j + 0.05) + \rho \sum_{j \neq i}^{N} w_{ij} \ln\left[\frac{Y_j}{L_j}\right] + \varepsilon_i \end{aligned} \tag{34}$$

where $\frac{1-\beta_H+\beta_K}{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)+\phi(\beta_K-\beta_H)} \ln \Omega(0) = \beta_0 + \varepsilon_i$, for $i = 1, ..., N$, with β_0 a constant and ε_i a country-specific shock since the term $\Omega(0)$ reflects not just technology but also resource endowments, climate, etc. and so it may differ across countries. We suppose also that $g + \delta = 0.05$ as is common in the literature since Mankiw et al. (1992) and Romer (1989). Finally, we have the following theoretical constraints between coefficients: $\beta_1 + \beta_2 = -\beta_3 = \frac{(\alpha-\phi)(1-\beta_H+\beta_K)+(\beta_K+\phi)}{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)+\phi(\beta_K-\beta_H)}$ and $\theta_3 = -(\theta_1 + \theta_2) = \frac{\alpha(1-\beta_H+\beta_K)+\beta_K}{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)+\phi(\beta_K-\beta_H)}\gamma$. Thus equation (34) is our basic econometric specification in this section.

Rewriting this equation in matrix form, we have:

$$y = X\beta + WX\theta + \rho W y + \varepsilon \tag{35}$$

[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.

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 physical capital investment rate, the vector of the logarithms of the human capital investment rate and the vector of the logarithms of the physical capital effective rate of 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[11] and Wy the endogenous spatial lag variable. $\beta' = [\beta_0, \beta_1, \beta_2, \beta_3]$, $\theta' = [\theta_1, \theta_2, \theta_3]$ and $\rho = \frac{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)}{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)+\phi(\beta_K-\beta_H)}\gamma$ is the spatial autoregressive parameter. ε is the $(N \times 1)$ vector of independently and identically distributed errors with mean zero and variance σ^2. In the spatial econometrics literature, 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). The model with the endogenous spatial lag variable and the explanatory variables only is referred to as the mixed regressive, spatial autoregressive model (SAR).[12]

For ease of exposition, equation (26) may also be written as follows:

$$y = \widetilde{X}b + \rho Wy + \varepsilon \tag{36}$$

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:

$$y = (I - \rho W)^{-1}\widetilde{X}b + (I - \rho W)^{-1}\varepsilon \tag{37}$$

Since for row-standardized spatial weight matrices $|\rho| < 1$ and $w_{ij} < 1$, the inverse matrix in equation (32) can be expanded into an infinite series as:

$$(I - \rho W)^{-1} = (I + \rho W + \rho^2 W^2 + ...) \tag{38}$$

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 the investment rate and the physical capital effective rate of depreciation in i, but also by those in all other locations

[11] 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 constant itself.

[12] see Anselin 1988; Anselin and Bera 1998; Anselin, 2001.

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:

$$Var(y) = \sigma^2 (I - \rho W)^{-1} (I - \rho W')^{-1} \tag{39}$$

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 (34), the variance at each location, are related to the neighborhood structure and therefore are not constant, leading to heteroskedasticity even though the initial process (30) is not heteroskedastic.

It also follows from the reduced form (32) that the spatially lagged variable Wy is correlated with the error term since:

$$E(Wy\varepsilon') = \sigma^2 W (I - \rho W)^{-1} \neq 0 \tag{40}$$

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). 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 (31) is given by:

$$\begin{aligned} \ln L(\beta', \rho, \sigma^2) \;\; = \;\; & -\frac{N}{2}\ln(2\pi) - \frac{N}{2}\ln(\sigma^2) + \ln|I - \rho W| \\ & -\frac{1}{2\sigma^2}\left[(I - \rho W)y - \widetilde{X}b\right]'\left[(I - \rho W)y - \widetilde{X}b\right] \end{aligned} \tag{41}$$

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:

$$|I-\rho W| = \prod_{i=1}^{N}(1-\rho\omega_i) \quad \Longrightarrow \quad \ln|I-\rho W| = \sum_{i=1}^{N}\ln(1-\rho\omega_i) \tag{42}$$

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:

$$\hat{\beta}_{ML}(\rho) = (\widetilde{X}'\widetilde{X})^{-1}\widetilde{X}'(I-\rho W)y \tag{43}$$

$$\hat{\sigma}^2_{ML}(\rho) = \frac{1}{N}\left[(I-\rho W)y - \widetilde{X}\hat{\beta}_{ML}(\rho)\right]'\left[(I-\rho W)y - \widetilde{X}\hat{\beta}_{ML}(\rho)\right] \tag{44}$$

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 $\hat{\sigma}^2_{ML}(\rho) = \left[\frac{(\hat{e}_O-\rho\hat{e}_L)'(\hat{e}_O-\rho\hat{e}_L)}{N}\right]$.

Substitution of (38) and (39) in the log-likelihood function (36) yields a concentrated log-likelihood as a non-linear function of a single parameter ρ:

$$\begin{aligned}\ln L(\rho) &= -\frac{N}{2}\left[\ln(2\pi)+1\right] + \sum_{i=1}^{N}\ln(I-\rho\omega_i) \\ &\quad -\frac{N}{2}\ln\left[\frac{(\hat{e}_O-\rho\hat{e}_L)'(\hat{e}_O-\rho\hat{e}_L)}{N}\right]\end{aligned} \tag{45}$$

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 (40). 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.[13] 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:

[13]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).

AsyVar$[b', \rho, \sigma^2] =$

$$\begin{bmatrix} \frac{1}{\sigma^2}\widetilde{X}'\widetilde{X} & \frac{1}{\sigma^2}(\widetilde{X}'W_A\tilde{X}b)' & 0 \\ \frac{1}{\sigma^2}\widetilde{X}'W_A\tilde{X}b & tr\left[(W_A + W_A')W_A\right] + \frac{1}{\sigma^2}(W_A\widetilde{X}b)'(W_A\widetilde{X}b) & \frac{1}{\sigma^2}trW_A \\ 0 & \frac{1}{\sigma^2}trW_A & \frac{N}{2\sigma^4} \end{bmatrix}^{-1} \tag{46}$$

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 as well as the Bayesian method (LeSage, 1997, 2002).

Furthermore, as outlined in Temple (1998, 1999a and b), heteroskedasticity and outliers are well known problems in the empirical growth regression literature. Use of non-parametric and robust estimation methods have thus been advocated (for instance least trimmed squares and reweighed least squares in Temple, 1998, 1999b). 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 SAR 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). Details with regard to the implementation of this extended estimation method for the linear and SAR models are presented in Appendix 2.

4.2 Results

We begin by estimating the traditional augmented Solow model with $\beta_K = \beta_H$ as a benchmark. OLS estimation results using a heteroskedasticity consistent covariance matrix estimator (White, 1980) are presented in the first column of Table 1. Our results with regard to its qualitative predictions are essentially identical to those of Mankiw et al. (1992) or Bernanke and Gürkaynak (2003), since the coefficients on saving, human capital investment and population growth have the predicted signs and are significant. However, in contrast to the result obtained by Bernanke and Gürkaynak (2003) using version 6.0 of PWT, the overidentifying restriction is not rejected with a $p-$value of 0.065.[14] The estimated capital share (α) and human capital share (β) are close to one fourth and one third respectively as expected. The estimated value of the capital share, α, is lower than in Mankiw et al. (1992) but one fourth is the lower bound commonly admitted for this parameter in the literature.

[14]Ertur and Koch (2005) note however that the overidentifying restriction is rejected in the textbook Solow model using version 6.1 of PWT as also underlined by Bernanke et al. (2003) using versions 5.6 or 6.0.

Bayesian heteroskedastic estimation results relying on Markov Chain Monter Carlo (MCMC) methodology using the procedure outlined in Geweke (1993) are reported in the second column of Table 1. We use 25000 replications with 5000 burn-in replications discarded. The value for the hyperparameter r is set to 4 reflecting our prior belief in the presence of heteroskedasticity and potentiel outliers. Diffuse priors are used for other model parameters as our main concern is here on robustness analysis versus those issues. In the third column 95% Highest Posterior Density Intervals (HPDI) for Bayesian estimates are reported in brackets. It is readily seen that again all the coefficients of interest have the predicted signs and are significant. Furthermore, it seems here that Bayesian heteroskedastic estimates and standard deviations are very close to those obtained using OLS-White for the unrestricted as well as the restricted models. This implies that heteroskedasticity or outliers, if present, do not clearly impact on OLS results. We can further explore this issue by examining a plot of the mean of v_i draws. Provided we use a small value for r, the presence of heterogeneity or outliers will be indicated by large v_i estimates that deviate substantially from unity. The upper part of Figure 1 confirms the presence in the unrestricted model of some large v_i values exceeding 4, for Congo, Ghana, Jamaica and Zaire. For the restricted model Zimbabwe adds to those. The role of these large v_i values is to accommodate outliers or observations containing large variances by down-weighting these observations in the estimation procedure. However heteroskedasticity or outliers does not seem here to be influential in both the unrestricted and restricted models.

Last but not least, we note that Moran's I test (Cliff and Ord, 1981) against spatial autocorrelation in the error terms rejects the null hypothesis of absence of spatial autocorrelation implying that OLS estimates are at best inefficient and at worst biased and non-consistent. Bayesian heteroskedastic estimates obtained using the linear model are also plagued by the presence of spatial autocorrelation.

[Table 1 around here]

One of our main points in this paper is that the augmented Solow model is misspecified since it omits variables due to technological interdependence and physical capital externalities. In fact, the econometric specification of our theoretical model is, in matrix form:

$$
\begin{aligned}
y^{\star} &= \frac{1}{1-\alpha-\beta_K}\Omega + \frac{\alpha(1-\beta_H+\beta_K)}{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)}S_K \\
&+ \frac{\beta_K}{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)}S_H + \frac{\phi}{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)}(I-\gamma W)^{-1}(h^{\star}-k^{\star}) \\
&+ \frac{1}{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)}(I-\gamma W)^{-1}\varepsilon \qquad (47)
\end{aligned}
$$

Therefore, the error term in the Mankiw et al. (1992) model contains omitted information since we can rewrite it:

$$
\begin{aligned}
\varepsilon_{MRW} &= \frac{\phi}{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)}(I-\gamma W)^{-1}(h^{\star}-k^{\star}) \\
&+ \frac{1}{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)}(I-\gamma W)^{-1}\varepsilon \qquad (48)
\end{aligned}
$$

We also note the presence of spatial autocorrelation in the error term even if there are no human capital externalities, that is $\phi = 0$, and then the presence of technological interdependence between all countries through the inverse spatial transformation $(I-\gamma W)^{-1}$.

[Figure 1]

Maximum likelihood estimation results of our SDM model for levels are presented in the first column and Bayesian heteroskedastic estimation relying on the MCMC procedure suggested by LeSage (1997) are reported in the second column of Table 2. In the last column 95% Highest Posterior Density Intervals (HPDI) for Bayesian estimates are reported in brackets. As for Table 1, we use 25000 replications with 5000 burn-in replications discarded. The value for the hyperparameter r is set to 4 reflecting our prior belief in the presence of heteroskedasticity and potentiel outliers. We use an Uniform prior $U[-1,+1]$ for the spatial autocorrelation parameter ρ as suggested by LeSage (1997, 2002). Diffuse priors are used for other model parameters.

First, all the coefficients of interest, except the spatial lag of human capital, have the predicted signs and all are significant except the coefficients of the spatial lags of both human capital and population rate of growth. We can note also the highly significant value of the spatial autocorrelation coefficient, ρ, showing the important role played by spatial autocorrelation in empirical growth specifications.

[Table 2 around here]

Second, the joint theoretical restrictions are not rejected since the $p-$value of the LR-test is 0.355. The implied value of the capital share in the income, α, is close to one half which is higher than the traditional result generally found in the Mankiw et al. (1992) framework. This suggest that the introduction of human capital in a framework with a different production function in each of the two sectors and with technological interdependence between countries does not allow to get closer to the commonly predicted value for this parameter.

The implied γ parameter is about 0.55. This parameter shows that technological interdependence between economies must be considered both in theoretical and empirical work since it implies spatial autocorrelation flawing the traditional estimation methods.

Moreover, we can note that the parameters describing the influence of human capital in each of the two sectors are different. In fact, the parameter for the consumption good sector is very low ($\beta_K = 0.010$). This result is consistent with those of Benhabib and Spiegel (1994) and Pritchett (2001) who show that human capital does not play a role as a simple production factor. However, it seems to play a natural role in the education sector since its value is higher ($\beta_H = 0.221$). Moreover, human capital plays an important role as facilitating the use of technology. In fact, ϕ is equal to 0.219.

Bayesian heteroskedastic estimates and standard deviations are again close to the maximum likelihood values indicating their robustness with regard to heteroskedasticity and outliers. Posterior model probabilities (PMP) also show evidence in favor of the restricted model. Potential outliers for both the unrestricted and the restricted models are reported in the lower part of Figure 1. Only a few posterior mean v_i values exceed 4: Hong Kong, Jamaica and Tanzania are therefore potential outliers. Heteroskedasticity or outliers do not seem to influence much the maximum likelihood estimation results as also readily seen on Figures 2 and 3 representing gaussian kernel density estimates of posterior distribution compared with simulated maximum likelihood distributions for the parameters of interest both in the unrestricted and restricted models.

[Figures 2 and 3 around here]

More specifically, we can test the absence of human capital externalities represented by ϕ. In fact, if $\phi = 0$ in specification (34), we obtain:

$$\begin{aligned}\ln\left[\frac{Y_i}{L_i}\right] &= \beta'_0 + \beta'_1 \ln s_{K,i} + \beta'_2 \ln s_{H,i} + \beta'_3 \ln(n_i + 0.05)\\ &+ \theta'_1 \sum_{j\neq i}^{N} w_{ij} \ln s_{K,j} + \theta'_2 \sum_{j\neq i}^{N} w_{ij} \ln s_{H,j}\\ &+ \theta'_3 \sum_{j\neq i}^{N} w_{ij} \ln(n_j + 0.05) + \gamma \sum_{j\neq i}^{N} w_{ij} \ln\left[\frac{Y_j}{L_j}\right] + \varepsilon_i \end{aligned} \quad (49)$$

with: $\beta'_1 + \beta'_2 = -\beta'_3 = \frac{\alpha(1-\beta_H+\beta_K)+\beta_K}{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)}$ and $\theta'_3 = -(\theta'_1 + \theta'_2) = \frac{\alpha(1-\beta_H+\beta_K)+\beta_K}{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)}\gamma$. Therefore, we have the following non-linear constraints: $\theta'_1 + \beta'_1\gamma = 0$, $\theta'_2 + \beta'_2\gamma = 0$ and $\theta'_3 + \beta'_3\gamma = 0$. Specification (49) is the so-called constrained spatial Durbin model, which is formally equivalent to a spatial autoregressive error model written in matrix form:

$$y = X\beta' + \varepsilon_{MRW} \qquad \text{and} \qquad \varepsilon_{MRW} = \gamma W \varepsilon_{MRW} + \varepsilon \quad (50)$$

where $\beta' = [\beta'_0\ \beta'_1\ \beta'_2\ \beta'_3]$ and ε_{MRW} is the same as before with $\phi = 0$. Therefore our model reduces in that case to the Mankiw et al. (1992) model with spatial autocorrelation in the error term. Maximum likelihood and Bayesian heteroscedatic MCMC estimation results are presented in Table 3. We reject these non-linear constraints and therefore the null hypothesis $\phi = 0$ and we conclude that there are indeed significant human capital externalities in our model. The $p-$value of the common factor tests in the constrained and unconstrained model are 0.01 and 0.02 respectively. Computation of posterior model probabilities show evidence in favor of the unconstrained SDM model since their values are respectively 0.998 against 0.002 and 0.987 against 0.013.

[Table 3 around here]

5 Impact of saving, human capital, population growth and technological interdependence on growth

In this section, we assess the predictions for conditional convergence of our model in two polar cases. First, like Mankiw et al. (1992) and Barro and Sala-i-Martin (1992), we suppose that the speed of convergence is identical for all countries and we refer to this case as the homogenous model, which is our benchmark model. Second, we estimate a model with complete parameter heterogeneity and we refer to this case as the heterogenous model, which is the full econometric specification of our theoretical model.

5.1 Homogenous model

In this section estimate equation (32): we first assume that the speed of convergence is homogenous and so identical for all countries: $\lambda_i = \lambda$ for $i = 1, ..., N$. Rewrite equation (32), dividing both sides by T, in the following form:

$$\begin{aligned} \frac{[\ln y_i(t) - \ln y_i(0)]}{T} &= \beta_0 + \beta_1 \ln y_i(0) + \beta_2 \ln s_{K,i} + \beta_3 \ln s_{H,i} + \beta_4 \ln(n_i + 0.05) \\ &+ \theta_1 \sum_{j \neq i}^{N} w_{ij} \ln y_j(0) + \theta_2 \sum_{j \neq i}^{N} w_{ij} \ln s_{K,j} + \theta_3 \sum_{j \neq i}^{N} w_{ij} \ln s_{H,j} \\ &+ \theta_4 \sum_{j \neq i}^{N} w_{ij} \ln(n_j + 0.05) + \rho \sum_{j \neq i}^{N} w_{ij} \frac{[\ln y_j(t) - \ln y_j(0)]}{T} + \varepsilon_i \quad (51) \end{aligned}$$

where β_0 is a constant,
$\beta_1 = -\frac{(1-e^{-\lambda T})}{T}$,
$\beta_2 + \beta_3 = -\beta_4 = \frac{(1-e^{-\lambda T})}{T} \frac{(\alpha-\phi)(1-\beta_H+\beta_K)+(\beta_K+\phi)}{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)+\phi(\beta_K-\beta_H)}$,
$\theta_1 = \frac{(1-e^{-\lambda T})}{T} \frac{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)}{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)+\phi(\beta_K-\beta_H)} \gamma$,
$\theta_4 = -(\theta_2 + \theta_3) = \frac{(1-e^{-\lambda T})}{T} \frac{\alpha(1-\beta_H+\beta_K)+\beta_K}{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)+\phi(\beta_K-\beta_H)} \gamma$,
$\rho = \frac{(1-e^{-\lambda T})}{\Gamma} \frac{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)}{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)+\phi(\beta_K-\beta_H)} \gamma$, where Γ is a scale parameter which should be equal to one when convergence speeds are identical for all countries.

In matrix form, we have a non-constrained spatial Durbin model which is estimated in the same way as the model in section 4. We test the convergence predictions of the Mankiw at al. (1992) model in Table 4 using OLS-White in the first column and heteroscedastic bayesian MCMC in the second column. Our results are essentially identical to those of Mankiw et al. (1992). 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 for the growth rate of working-age population. The results also support the predicted signs of each variables.[15]

[15] We note also that each variable is highly significant in contrast to Mankiw et al. (1992) where the growth rates of working-age population were not significant at 5%.

[Table 4 around here]

Estimation results for the equation imposing the restriction that the coefficients sum to zero are presented in the bottom part of Table 4. We find that this restriction is not rejected ($p-$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 Mankiw et al. (1992), these regressions give a somewhat larger weight to physical capital and a somewhat smaller weight to human capital. The implied values 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 the Mankiw et al. (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 4: Botswana, Hong-Kong, Jamaica, Zaire and Zambia have posterior mean v_i values exceeding 4 in both cases.

Once again, we claim that the Mankiw et al. (1992) model is misspecified since it omits variables due to technological interdependence and human capital externalities. Therefore, as in Section 4, the error terms of Mankiw et al. (1992) contain omitted information and are spatially autocorrelated as also indicated by Moran's I tests in Table 4.

[Figure 4 around here]

In table 5, 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 importance 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 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 4 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 maximum likelihood estimate of the coefficient of the spatial lag of population growth rate is represented in Figure 5.

[Figure 5 around here]

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 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 4) and down-weighted in the MCMC estimation procedure. However, the Bayesian heteroskedastic estimates remain close to those obtained by maximum likelihood as it can also be readily seen on Figure 6 representing gaussian kernel density estimates of posterior distributions compared with simulated maximum likelihood distributions for the parameters of interest in the restricted models. Significance test results are not affected.

[Figure 6 around here]

The implied value of the convergence speed λ is higher than that found by Mankiw et al. (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 lower than the value obtained in section 4. Its value is close to 0.4, the upper bound generally admitted in the literature. The value of the coefficient for global technological interdependence is close to 0.7 showing once again its importance on growth and convergence processes.

The implied value of human capital parameters are more consistent with those expected. On the one hand, the human capital share in income β_K is very low in the consumption good sector showing the weak role played by human capital where it is considered as a simple production factor along the lines of Benhabib and Spiegel (1994) and Pritchett (2001). On the other hand, the human capital share in the education sector, β_H, is close to 0.6, a higher value than the one third obtained by Mankiw et al. (1992). It seems more consistent with the value we would expect since some authors as Kleenow and Rodriguez-Clare (1997) and Kendrick (1976) present evidence that the technology for producing human capital is more intensive in labor than is the technology for producing other goods. Finally, the value of human capital externalities is lower than in section 4: its value is close to 0.06 estimated by ML and 0.08 estimated by bayesian heteroskedastic MCMC. We can test, for the same reason as in section 4, the absence of human capital externalities with the common factor test.[16] The common factor tests allow to reject the absence of physical capital externalities in the model. 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, that is in favor of rejection of the absence of human capital externalities.

[Table 6 around here]

5.2 Heterogenous model

In recent papers, Durlauf (2000, 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), Lee, Pesaran and Smith (1997) or Evans (1998) 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.

[16]More precisely, the null hypothesis of this test is jointly $\phi = 0$ and $\Gamma = 1$, that is there is no physical capital externalities and the scale parameter is equal to 1. However, if there is no human capital externalities, the scale parameter disappears and is equal to 1.

The empirical methodology we propose takes into account the heterogeneity embodied in our bisectoriel growth model. Reconsider equation (32), dividing both sides by T:

$$\begin{aligned}
\frac{[\ln y_i(t) - \ln y_i(0)]}{T} &= \beta_{0i} + \beta_{1i}\ln y_i(0) + \beta_{2i}\ln s_{K,i} + \beta_{3i}\ln s_{H,i} + \beta_{4i}\ln(n_i+g+\delta) \\
&+ \theta_{1i}\sum_{j\neq i}^{N} w_{ij}\ln y_j(0) + \theta_{2i}\sum_{j\neq i}^{N} w_{ij}\ln s_{K,j} + \theta_{3i}\sum_{j\neq i}^{N} w_{ij}\ln s_{H,j} \\
&+ \theta_{4i}\sum_{j\neq i}^{N} w_{ij}\ln(n_j+g+\delta) + \rho_i\sum_{j\neq i}^{N} w_{ij}\frac{[\ln y_j(t) - \ln y_j(0)]}{T} + \varepsilon_i
\end{aligned} \tag{52}$$

with $\beta_{0i} = \frac{(1-e^{-\lambda_i T})}{T}(\frac{g}{1-\gamma}\frac{1}{\lambda_i} + \frac{1-\beta_H+\beta_K}{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)+\phi(\beta_K-\beta_H)}\Omega)$,
$\beta_{1i} = -\frac{(1-e^{-\lambda_i T})}{T}$,
$\beta_{2i}+\beta_{3i} = -\beta_{4i} = \frac{(1-e^{-\lambda_i T})}{T}\frac{(\alpha-\phi)(1-\beta_H+\beta_K)+(\beta_K+\phi)}{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)+\phi(\beta_K-\beta_H)}$,
$\theta_{1i} = \frac{(1-e^{-\lambda_i T})}{T}\frac{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)}{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)+\phi(\beta_K-\beta_H)}\gamma$,
$\theta_{4i} = -(\theta_{2i}+\theta_{3i}) = \frac{(1-e^{-\lambda_i T})}{T}\frac{\alpha(1-\beta_H+\beta_K)+\beta_K}{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)+\phi(\beta_K-\beta_H)}\gamma$
$\rho_i = \frac{(1-e^{-\lambda_i T})}{\Gamma_i}\frac{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)}{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)+\phi(\beta_K-\beta_H)}\gamma$.
The term Γ_i is a scale parameter reflecting the effects of the speeds of convergence in the neighboring countries. 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 model in (52). The original specification was proposed by LeSage and Pace (2004) and labeled spatial autoregressive local estimation (SALE). This specification is used in Ertur et al. (2003) and Ertur and Koch (2005). We also implement here this specification and label it the local SDM model:

$$U(i)y = U(i)X\beta_i + U(i)WX\theta_i + \rho_i U(i)Wy + U(i)\varepsilon \tag{53}$$

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 and estimation results are presented in Figures 7 to 9 footnotecomplete results are displayed in Table 7 in Appendix 3. Countries are ordered by continent (Africa, America, Asia and Europe) and increasing latitude in each continent.

[Figures 7 to 9 around here]

We note strong evidence for parameter heterogenity like Durlauf et al. (2001) or Ertur and Koch (2005). This heterogeneity is furthermore linked to the geographical location of the observations. For sub-sample sizes of $m = 30, 45$ and 60, in Figures 7 to 9 respectively, we report the local impact of each variable after considering the explicit form of the SDM heterogenous convergence model with the same analytical methodology to evaluate elasticities as in section 2.

First consider $m = 30$, 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. Note that all these results are quite sensitive to the choice of the sub-sample size.

5.3 Counterfactual income kernel density estimates

The effect of the different theoretical frameworks on the world income dynamics are estimated by applying kernel density methods (Silverman, 1986). 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 kernelk that is also chosen by Di Nardo, Fortin and Lemieux (1996) and Desdoigts (2004). The results are presented in Figure 10.

[Figure 10 around here]

In the upper left box of Figure 10, univariate density estimates of the world real output per worker is displayed. Blue line represents the density estimates in 1960 and the red line represents the density estimates in 1995. We can note that the red line shows more structure than the blue 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 (in red line) with the distributions implied by the Mankiw et al. (1992) model, our homogenous and heterogenous models. First, in the upper right box of figure 10, we can note that the Mankiw et al. (1992) model does not capture the structure of the observed income distribution. In fact, there is only one mode whereas there are tree modes in the observed distribution. Our bisectorial growth model captures two of tree modes in the observed income distribution. However, when we consider the local estimation of our heterogenous model, we capture the tree modes of the observed distribution. The latter model seems to perform the best.

6 Conclusion

In this paper we develop a bisectorial growth model with consumption good sector and human capital sector. This model is more general than the Mankiw et al. (1992) model since the two sectors use different production functions and furthermore technological interdependence between countries is introduced. Implied econometric specifications are estimated by spatial econometric technics: standard 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 against the Mankiw et al. (1992) model for different. First, as also underlined by Benhabib and Spiegel (1994) or Pritchett (2001), we find that human capital does not influence growth as a simple production factor. Nevertheless, it plays an important role in the education sector as expected and as a technological use parameter in the technological progress function. Moreover, the estimated value of the human capital parameter in the education sector is higher than the estimated value of the physical capital parameter as expected. Evidence with regard to the different impacts of human capital parameters shows the importance of assuming different production functions between the two sectors of a bisectorial growth model. 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 Mankiw et al. (1992) model.

References

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

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

[3] 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.

[4] Barro R.J., Sala-i-Martin X. (1992), "Convergence", *Journal of Political Economy*, 100, 223-251.

[5] Bernanke B.S., Gurkaynak R.S. (2003), "Is Growth Exogeneous? Taking Mankiw, Romer and Weil Seriously", In: *NBER Macroeconomics Annual*, Bernanke B.S., Rogoff K. (eds.), MIT Press, Cambridge, 11-57.

[6] Benhabib J. and Spiegel M.M. (1994), "The role of human capital in economic development: evidence from aggregate cross-country data", *Journal of Monetary Economics*, 34, 2, 143-74.

[7] Bils M. and Klenow P.J. (2000), "Does schooling cause growth?", *American Economic Review*, 90, 5, 1160-1183.

[8] Brock W.A., Durlauf S.N. (2001), "Growth Empirics and Reality", *The World Bank Economic Review*, 15, 229-272.

[9] Canova (2004), "Testing for Convergence Clubs in Income per capita: A Predictive Density Approach", *International Economic Review*, 45, 49-77.

[10] Caselli F. (1999), "Technological revolutions", *American Economic Review*, 89, 1, 78-102.

[11] Caselli F. (2004), "Accounting for cross-country income differences", *National Bureau of Economic Research*, Working Paper 10828, October 2004, forthcoming In: Aghion P. and Durlauf S. (eds.), *The Handbook of Economic Growth.*

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

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

[14] Desdoigts, A. (1999), "Patterns of economic development and the formation of clubs", *Journal of Economic Growth*, 4, 3, 305-330.

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

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

[17] Durlauf S.N. (2000), "Econometric analysis and the study of economic growth: A skeptical perspective", In: Backhouse R., Salanti A. (eds.), *Macroeconomic and the Real World*, Oxford University Press, Oxford.

[18] Durlauf S.N. (2001), "Manifesto for a Growth Econometrics", *Journal of Econometrics*, 100, 65-69.

[19] Durlauf S.N., Johnson P. (1995), "Multiple regimes and cross-country growth behavior", *Journal of Applied Econometrics*, 10, 365-384.

[20] Durlauf S.N., Kourtellos A., Minkin A. (2001), "The local Solow growth model", *European Economic Review*, 45, 928-940.

[21] Durlauf S.N., Quah D. (1999), "The New Empirics of Economic Growth", In: Taylor J. and Woodford M. (eds), *Handbook of Macroeconomics*, Elsevier Science, North-Holland.

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

[23] Ertur C. and 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.

[24] Ertur C., Le Gallo J., LeSage J. (2003) "Local versus Global Convergence in Europe: A Bayesian Spatial Econometrics Approach", Regional Economics Applications Laboratory-Technical Series, 03-T-28, University of Illinois at Urbana-Champaign.

[25] Easterly W., King R., Levine R. and Rebelo S. (1994), "Policy, technology adoption and growth", In: Pasinetti L. and Solow R.M. (eds), *Economic growth and the structure of long-term development*, New-York: St. Martin's Press, 75-89.

[26] Evans P. (1998), "Using Panel Data to Evaluate Growth Theories", *International Economic Review*, 39, 2, 295-306.

[27] Gelfand, Alan E., and A.F.M Smith. (1990), "Sampling-Based Approaches to Calculating Marginal Densities", *Journal of the American Statistical Association*, 85, 398-409.

[28] Geweke J. (1993), "Bayesian Treatment of the Independent Student t Linear Model," *Journal of Applied Econometrics*, 8:19-40.

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

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

[31] Islam N. (1995), "Growth Empirics: A panel Data Approach", *Quarterly Journal of Economics*, 110, 1127-1170.

[32] Jones C.I. (2002),"Sources of U.S. economic growth in a world of ideas",*American Economic Review*, 92, 220-239.

[33] 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.

[34] Keller W. (2004), "Geographic Localization of International Technology Diffusion", *American Economic Review*, 92, 120-142.

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

[36] Klenow P.J., Rodriguez-Clare A. (1997), "The neoclassical revival in growth economics: has it gone to far?", In: *NBER Macroeconomics Annual*, Bernanke B.S., Rotemberg J. (eds.), MIT Press, Cambridge, 73-103.

[37] Klenow P.J., Rodriguez-Clare A. (2004), "Externalities and Growth", *National Bureau of Economic Research*, Working Paper 11009, December 2004, forthcoming In: Aghion P. and Durlauf S. (eds.), *The Handbook of Economic Growth*.

[38] Lange K.L., Little R.J.A. , Taylor J.M.G. (1989), "Robust Statistical Modeling Using the t Distribution," *Journal of the American Statistical Association*, 84:881-896.

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

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

[41] Lee K., Pesaran M.H. and Smith R. (1997), "Growth and Convergence in a Multi-Country Empirical Stochastic Solow Model", *Journal of Applied Econometrics*, 12, 357-392.

[42] LeSage, J.P. (1997) "Bayesian Estimation of Spatial Autoregressive Models", *International Regional Science Review*, 20(1&2), 113-129.

[43] LeSage, J.P. (2002) "Application of Bayesian methods to spatial econometrics", *mimeo*, University of Toledo, Department of Economics, 51p.

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

[45] Lucas R.E. (1988), "On the Mechanics of Economic Development", *Journal of Monetary Economics*, 22, 3-42.

[46] Lucas R.E. (1993), "Making a miracle", *Econometrica*, 61, 2, 251-271.

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

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

[49] Mulligan C.B. and Sala-i-Martin X. (1993), "Transitional dynamics in two-sector models of endogenous growth", *Quarterly Journal of Economics*, 108, 3, 739-773.

[50] Nelson R.R. and Phelps E.S. (1966), "Investment in humans, technological diffusion, and economic growth", *American Economic Review*, 56, 2, 69-75.

[51] Ord J.K. (1975), "Estimation Methods for Models of Spatial Interaction", *Journal of American Statistical Association*, 70, 120-126.

[52] Pritchett L. (2001), "Where has all education gone?", *World Bank Economic Review*, 15, 3, 367-391.

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

[54] Romer P.M. (1989), "Capital Accumulation in the Theory of Long Run Growth", In: Barro R.J. (ed.), *Modern Business Cycle Theory*, Cambridge, M.A., Harvard University Press, 51-127.

[55] Romer P.M. (1990), "Endogenous technological change", *Journal of Political Economy*, 98, S71-S102.

[56] Romer P.M. (1993), "Idea Gaps and Object Gaps in Economic Development", *Journal of Monetary Economics*, 32, 543-573.

[57] Rebelo S. (1991), "Long-Run Policy Analysis and Long-Run Growth", *Journal of Political Economy*, 99, 3, 500-521.

[58] Schultz T.W. (1975), "The value of the ability to deal with disequilibria", *Journal of Economic Literature*, 13, 3, 827-46.

[59] Solow R.M. (1956), "A contribution to the theory of economic growth", *Quarterly Journal of Economics*, 70, 65-94.

[60] Sheather S.J. and Jones, M.C. (1991), "A reliable data-based bandwith selection method for kernel density estimation", *Journal of the Royal SStatistical Society, Series B*, 53, 683-690.

[61] Temple J. (1998), "Robusness Tests of the Augmented Solow Model," *Journal of Applied Econometrics*, 13: 361-375.

[62] Temple J. (1999a), "The New Growth Evidence" *Journal of Applied Econometrics*, 37: 112-156.

[63] Temple J. (1999b), "A positive effect of human capital on growth" *Economics Letters*,65, 131-134.

[64] Uzawa H. (1965), "Optimum technical change in an agragative model of economic growth", *International Economic Review*, 6, 18-31.

[65] Von Neumann J. (1945), "A model of General Equilibrium", *Review of Economic Studies*, 13, 1-9.

[66] White H. (1980), "A heteroscedasticity consistent covariance matrix estimator and a direct test for heteroscedasticity", *Econometrica*, 48, 817-838.

[67] Young A. (1995), "The Tyranny of Numbers: Confronting the Statistical Realities of the East Asian Growth Experience", *Quarterly Journal of Economics*, 110, 641-680.

Appendix 1: Steady state of real income per worker in the bisectorial model

Steady-state in the bisectorial model

It is possible to explain a very useful relation between physical capital output and human capital output as following:

$$\frac{y_{K,i}(t)}{y_{H,i}(t)} = h_i(t)^{\beta_K - \beta_H} \tag{54}$$

When both sectors use the same technology ($\beta_K = \beta_H$), this ratio is equal to 1, as in the Mankiw, Romer and Weil (1992) model. At steady state, we can write:

$$\frac{h_i^\star(t)}{y_{K,i}^\star(t)} = \frac{h_i^\star(t)}{y_{H,i}^\star(t)} \frac{y_{H,i}^\star(t)}{y_{K,i}^\star(t)} = \left(\frac{s_{H,i}}{n_i + g + \delta} \right) h_i^\star(t)^{\beta_H - \beta_K}$$

We can write finally the ratio human capital-output at steady-state in the following form:

$$\frac{h_i^\star(t)}{y_{K,i}^\star(t)} = \left(\frac{s_{H,i}}{n_i + g + \delta} \right)^{\frac{1}{1+\beta_K - \beta_H}} y_{K,i}^\star(t)^{\frac{\beta_H - \beta_K}{1+\beta_K - \beta_H}}$$

Replacing both physical capital and human capital ratio at steady-state in the equation (), we obtain the per worker income at steady-state in logarithms form:

$$\ln y_{i,K}^\star(t) = \frac{1}{1-\alpha-\beta_K} \ln A_i(t) + \frac{\alpha}{1-\alpha-\beta_K} \ln \left(\frac{k_i^\star(t)}{y_{i,K}^\star(t)} \right) + \frac{\beta_K}{1-\alpha-\beta_K} \ln \left(\frac{h_i^\star(t)}{y_{i,K}^\star(t)} \right)$$

so that:

$$\begin{aligned} \ln y_{K,i}^\star(t) &= \frac{1}{1-\alpha-\beta_K} \ln A_i(t) + \frac{\alpha}{1-\alpha-\beta_K} \ln \left(\frac{s_{K,i}}{n_i + g + \delta} \right) \\ &+ \frac{\beta_K}{1-\alpha-\beta_K} \frac{1}{1+\beta_K - \beta_H} \ln \left(\frac{s_{H,i}}{n_i + g + \delta} \right) + \frac{\beta_H - \beta_K}{1+\beta_K - \beta_H} \ln y_{K,i}^\star(t) \end{aligned}$$

After rearranging terms, we obtain finally the equation in the text:

$$\begin{aligned} \ln y_{i,K}^\star(t) &= \frac{1}{1-\alpha-\beta_K} \ln A_i(t) + \frac{\alpha(1-\beta_H+\beta_K)}{(1-\beta_H) - \alpha(1-\beta_H+\beta_K)} \ln \left(\frac{s_{K,i}}{n_i + g + \delta} \right) \\ &+ \frac{\beta_K}{(1-\beta_H) - \alpha(1-\beta_H+\beta_K)} \ln \left(\frac{s_{H,i}}{n_i + g + \delta} \right) \end{aligned}$$

Steady-state in the bisectorial model with technological interdependence

In order to determine the equation describing the real income per worker of country i at steady state, we rewrite the production function () in matrix form: $y = A + \alpha k + \beta_K h$ and replacing A by its expression in matrix form, we obtain after some algebraic manipulation:

$$y = \Omega + (\alpha - \phi)k + (\beta_K + \phi)h - \alpha\gamma W k - \beta_K \gamma W h + \gamma W y \tag{55}$$

Replacing k and h by their expressions at steady state and in matrix form:

$$k = S_K + y \quad \text{and} \quad h = \frac{1}{1 - \beta_H + \beta_K}(S_H + y) \tag{56}$$

so that:

$$\left(1 - \alpha + \phi - \frac{\beta_K + \phi}{1 - \beta_H + \beta_K}\right) y = \Omega + (\alpha - \phi)S_K + \frac{\beta_K + \phi}{1 - \beta_H + \beta_K} S_H - \alpha\gamma W S_K$$
$$- \frac{\beta_K}{1 - \beta_H + \beta_K}\gamma W S_H + \left(1 - \alpha - \frac{\beta_K}{1 - \beta_H + \beta_K}\right)\gamma W y \tag{57}$$

After rearranging terms, we obtain:

$$\begin{aligned}
y \;=\;& \frac{1 - \beta_H + \beta_K}{(1 - \beta_H) - \alpha(1 - \beta_H + \beta_K) + \phi(\beta_K - \beta_H)}\Omega \\
+\;& \frac{(\alpha - \phi)(1 - \beta_H + \beta_K)}{(1 - \beta_H) - \alpha(1 - \beta_H + \beta_K) + \phi(\beta_K - \beta_H)} S_K \\
+\;& \frac{\beta_K + \phi}{(1 - \beta_H) - \alpha(1 - \beta_H + \beta_K) + \phi(\beta_K - \beta_H)} S_H \\
-\;& \frac{\alpha(1 - \beta_H + \beta_K)}{(1 - \beta_H) - \alpha(1 - \beta_H + \beta_K) + \phi(\beta_K - \beta_H)}\gamma W S_K \\
-\;& \frac{\beta_K}{(1 - \beta_H) - \alpha(1 - \beta_H + \beta_K) + \phi(\beta_K - \beta_H)}\gamma W S_H \\
+\;& \frac{(1 - \beta_H) - \alpha(1 - \beta_H + \beta_K)}{(1 - \beta_H) - \alpha(1 - \beta_H + \beta_K) + \phi(\beta_K - \beta_H)}\gamma W y
\end{aligned} \tag{58}$$

Elasticities

Take equation () and subtracting $\frac{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)}{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)+\phi(\beta_K-\beta_H)}\gamma W y$ from both sides, and premultiplying both sides by $(I - \rho W)^{-1}$, with $\rho = \frac{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)}{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)+\phi(\beta_K-\beta_H)}\gamma$ we obtain:

$$
\begin{aligned}
y &= \frac{1-\beta_H+\beta_K}{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)+\phi(\beta_K-\beta_H)}\,(I-\rho W)^{-1}\,\Omega \\
&+ \frac{1-\beta_H+\beta_K}{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)+\phi(\beta_K-\beta_H)}\,(I-\rho W)^{-1}\,((\alpha-\phi)I-\alpha\gamma W)\,S_K \\
&+ \frac{1}{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)+\phi(\beta_K-\beta_H)}\,(I-\rho W)^{-1}\,((\beta_K+\phi)I-\beta_K\gamma W)\,S_H
\end{aligned}
$$

Deriving this expression respectively in respect to the vector S_K and the vector S_H, we obtain the expression of elasticities in matrix form:

$$
\Xi^K = \frac{1-\beta_H+\beta_K}{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)+\phi(\beta_K-\beta_H)}\,(I-\rho W)^{-1}\,((\alpha-\phi)I-\alpha\gamma W) \quad (59)
$$

and:

$$
\Xi^H = \frac{1}{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)+\phi(\beta_K-\beta_H)}\,(I-\rho W)^{-1}\,((\beta_K+\phi)I-\beta_K\gamma W) \quad (60)
$$

Developing the expression $(I-\rho W)^{-1}$ as following: $I+\rho W+\rho^2W^2+\rho^3W^3+...$ or: $I+\sum_{r=1}^{\infty}W^r\rho^r$, we obtain:

$$
\begin{aligned}
\Xi^K &= \frac{(\alpha-\phi)(1-\beta_H+\beta_K)}{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)+\phi(\beta_K-\beta_H)}I \\
&+ \frac{(1-\beta_H+\beta_K)((\alpha-\phi)\rho-\alpha\gamma)}{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)+\phi(\beta_K-\beta_H)}\left(\frac{1}{\rho}\sum_{r=1}^{\infty}W^r\rho^r\right) \quad (61)
\end{aligned}
$$

and:

$$
\begin{aligned}
\Xi^H &= \frac{\beta_K+\phi}{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)+\phi(\beta_K-\beta_H)}I \\
&+ \frac{(\beta_K+\phi)\rho-\beta_K\gamma}{(1-\beta_H)-\alpha(1-\beta_H+\beta_K)+\phi(\beta_K-\beta_H)}\left(\frac{1}{\rho}\sum_{r=1}^{\infty}W^r\rho^r\right) \quad (62)
\end{aligned}
$$

Appendix 2: Bayesian MCMC estimation

MCMC sampler for the heteroscedastic linear model

MCMC sampler for the heteroscedastic SAR model

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

Model	MRW OLS-White	MRW Bayesian het	
Dependent variable	$\ln y_i(1995)$	$\ln y_i(1995)$	95% HPDI
Obs.	89	89	
constant	6.844	7.863	$[5.943, 9.741]$
	(1.118) (0.000)	(1.164)	
$\ln s_i$	0.519	0.555	$[0.334, 0.785]$
	(0.137) (0.000)	(0.136)	
$\ln h_i$	0.901	0.899	$[0.735, 1.063]$
	(0.097) (0.000)	(0.010)	
$\ln(n_i + 0.05)$	-2.329	-1.994	$[-2.668, -1.342]$
	(0.377) (0.000)	(0.406)	
Moran's I test $(W1)$	0.353		
	(0.000)		
Restricted regression			
constant	9.216	9.260	$[9.061, 9.458]$
	(0.129) (0.000)	(0.121)	
$\ln s_i - \ln(n_i + 0.05)$	0.575	0.619	$[0.408, 0.828]$
	(0.135) (0.000)	(0.128)	
$\ln h_i - \ln(n_i + 0.05)$	0.920	0.899	$[0.730, 1.066]$
	(0.105) (0.000)	(0.102)	
Moran's I test $(W1)$	0.370		
	(0.000)		
Test of restriction	3.502 (Wald)		
	(0.065)		
Implied α	0.230	0.246	
Implied β	0.369	0.356	

Notes: coefficient estimates and standard deviations as well as p-values (in parentheses) are reported for OLS using White heteroskedasticity consistent covariance matrix estimator. 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.

Table 2: ML and Bayesian heteroscedastic estimation results of the SDM model for levels.

Model Dependent variable Obs. / Weight matrix	SDM-level ML $\ln y_i(1995)$ 89 / $(W1)$	SDM-level Bayesian het $\ln y_i(1995)$ 89 / $(W1)$	95% HPDI
constant	0.804	1.213	$[-1.596, 4.074]$
	(1.576) (0.614)	(1.721)	
$\ln s_i$	0.486	0.537	$[0.353, 0.720]$
	(0.094) (0.000)	(0.112)	
$\ln h_i$	0.624	0.581	$[0.397, 0.774]$
	(0.093) (0.000)	(0.114)	
$\ln(n_i + 0.05)$	-1.504	-1.410	$[-2.128, -0.694]$
	(0.455) (0.001)	(0.435)	
$W \ln s_j$	-0.746	-0.805	$[-1.216, -0.387]$
	(0.216) (0.000)	(0.254)	
$W \ln h_j$	0.081	0.136	$[-0.233, 0.505]$
	(0.210) (0.705)	(0.223)	
$W \ln(n_j + 0.05)$	-0.176	-0.144	$[-1.298, 0.933]$
	(0.714) (0.807)	(0.699)	
$W \ln y_j$	0.607	0.602	$[0.409, 0.769]$
	(0.092) (0.000)	(0.110)	
Restricted regression			
constant	3.714	3.717	$[2.167, 5.468]$
	(0.887) (0.000)	(1.005)	
$\ln s_i - \ln(n_i + 0.05)$	0.526	0.591	$[0.408, 0.771]$
	(0.093) (0.000)	(0.110)	
$\ln h_i - \ln(n_i + 0.05)$	0.624	0.556	$[0.368, 0.747]$
	(0.095) (0.000)	(0.115)	
$W[\ln s_j - \ln(n_j + 0.05)]$	-0.553	-0.606	$[-0.962, -0.237]$
	(0.197) (0.005)	(0.221)	
$W[\ln h_j - \ln(n_j + 0.05)]$	-0.016	0.051	$[-0.304, 0.410]$
	(0.209) (0.940)	(0.217)	
$W \ln y_j$	0.624	0.626	$[0.438, 0.791]$
	(0.091) (0.000)	(0.108)	
Test of restriction	4.396 (LR)	0.199 PMP unrest.	
	(0.355)	0.801 PMP rest.	
Implied α	0.463	0.513	
Implied β_k	0.010	-0.034	
Implied β_h	0.221	0.180	
Implied ϕ	0.219	0.237	
Implied γ	0.555	0.550	

Notes: coefficient estimates and standard deviations as well as p-values (in parentheses) are reported for ML estimation.
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.
LR means Likelihood Ratio and PMP stands for Posterior Model Probability.

Table 3: The spatial autoregressive error model and common factor tests.

Model	SEM-level ML		Sem-level Bayesian het	
Dependent variable	$\ln y_i(1995)$		$\ln y_i(1995)$	95% HPDI
Obs. / Weight matrix	91 / $(W1)$		91 / $(W1)$	
constant	7.834		8.065	$[5.743, 10.341]$
	(1.254)	(0.000)	(1.398)	
$\ln s_i$	0.553		0.593	$[0.377, 0.812]$
	(0.095)	(0.000)	(0.132)	
$\ln h_i$	0.696		0.744	$[0.551, 0.936]$
	(0.091)	(0.000)	(0.118)	
$\ln(n_i + 0.05)$	-1.758		-1.786	$[-2.594, -0.968]$
	(0.457)	(0.000)	(0.496)	
ρ	0.678		0.602	$[0.384, 0.776]$
	(0.078)	(0.000)	(0.118)	
Common factor test	11.001 (LR)		0.998 PMP unrest.	
SDM vs. SEM	(0.012)		0.002 PMP rest.	
Restricted regression				
constant	9.184		9.236	$[8.967, 9.504]$
	(0.165)	(0.000)	(0.163)	
$\ln s_i - \ln(n_i + 0.05)$	0.562		0.613	$[0.401, 0.826]$
	(0.095)	(0.000)	(0.129)	
$\ln h_i - \ln(n_i + 0.05)$	0.704		0.752	$[0.561, 0.940]$
	(0.090)	(0.000)	(0.116)	
ρ	0.692		0.607	$[0.406, 0.783]$
	(0.076)	(0.000)	(0.115)	
Test of restriction	1.128		0.420 PMP unrest.	
	(0.569)		0.580 PMP rest.	
Common factor test	7.734 (LR)		0.987 PMP unrest.	
SDM vs. SEM	(0.021)		0.013 PMP rest.	

Notes: coefficient estimates and standard deviations as well as p-values (in parentheses) are reported for ML estimation.
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.
LR means Likelihood Ratio and PMP stands for Posterior Model Probability.

Table 4: 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)	
$\ln s_i$	1.423	1.340	$[0.864, 1.838]$
	(0.321) (0.000)	(0.296)	
$\ln h_i$	1.334	1.338	$[0.889, 1.782]$
	(0.266) (0.000)	(0.0.271)	
$\ln(n_i + 0.05)$	-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)	
$\ln s_i - \ln(n_i + 0.05)$	1.496	1.446	$[1.002, 1.907]$
	(0.312) (0.000)	(0.276)	
$\ln h_i - \ln(n_i + 0.05)$	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	0.017	0.017	
Half-life	53	53	
Implied α	0.364	0.354	
Implied β	0.323	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. 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.

Table 5: ML and Bayesian heteroscedastic estimation results of the SDM homogenous convergence model.

Model	SDM-convergence ML	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	$[-3.438, 10.775]$
	(3.928) (0.798)	(4.338)	
$\ln y_i(1960)$	-1.718	-1.667	$[-1.999, -1.331]$
	(0.214) (0.000)	(0.203)	
$\ln s_i$	1.431	1.544	$[1.124, 1.955]$
	(0.232) (0.000)	(0.253)	
$\ln h_i$	1.196	1.034	$[0.626, 1.456]$
	(0.253) (0.000)	(0.252)	
$\ln(n_i + 0.05)$	-3.790	-3.707	$[-5.433, -2.017]$
	(1.128) (0.000)	(1.037)	
$W \ln y_i(1960)$	1.346	1.426	$[0.708, 2.087]$
	(0.357) (0.000)	(0.420)	
$W \ln s_j$	-1.461	-1.277	$[-2.326, -0.225]$
	(0.578) (0.012)	(0.638)	
$W \ln h_j$	-0.256	-0.314	$[-1.260, 0.662]$
	(0.600) (0.670)	(0.585)	
$W \ln(n_j + 0.05)$	1.590	2.966	$[0.031, 5.828]$
	(1.878) (0.397)	(1.764)	
$W\left(\frac{\ln y_j(1995)-\ln y_j(1960)}{35}\right)$	0.504	0.517	$[0.320, 0.824]$
	(0.109) (0.000)	(0.114)	
Restricted regression			
constant	3.110	3.505	$[-1.608, 8.869]$
	(2.743) (0.257)	(3.195)	
$\ln y_i(1960)$	-1.713	-1.695	$[-2.020, -1.369]$
	(0.214) (0.000)	(0.199)	
$\ln s_i - \ln(n_i + 0.05)$	1.481	1.547	$[1.146, 1.940]$
	(0.228) (0.000)	(0.242)	
$\ln h_i - \ln(n_i + 0.05)$	1.200	1.067	$[0.669, 1.481]$
	(0.252) (0.000)	(0.248)	
$W \ln y_i(1960)$	1.461	1.401	$[0.700, 2.059]$
	(0.340) (0.000)	(0.415)	
$W[\ln s_j - \ln(n_j + 0.05)]$	-1.208	-1.323	$[-2.190, -0.414]$
	(0.497) (0.010)	(0.541)	
$W[\ln h_j - \ln(n_j + 0.05)]$	-0.472	-0.280	$[-1.162, 0.618]$
	(0.556) (0.396)	(0.546)	
$W\left(\frac{\ln y_j(1995)-\ln y_j(1960)}{35}\right)$	0.528	0.528	$[0.331, 0.817]$
	(0.106) (0.000)	(0.114)	
Test of restriction	1.544 (LR)	0.029 PMP unrest.	
	(0.819)	0.971 PMP rest.	
Implied λ	0.026	0.026	
Half-life	40.5	40.5	
Implied Γ	0.969	0.929	
Implied α	0.398	0.440	
Implied β_k	0.116	0.087	
Implied β_h	0.586	0.522	
Implied ϕ	0.061	0.081	
Implied γ	0.731	0.697	

Notes: coefficient estimates and standard deviations as well as p-values (in parentheses) are reported for ML estimation.
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.

Table 6: The spatial autoregressive error convergence model and common factor tests.

Model	SEM-convergence ML	SEM-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	10.852	10.645	$[4.554, 16.992]$
	(3.380) (0.001)	(3.796)	
$\ln y_i(1960)$	-1.525	-1.442	$[-1.844, -1.053]$
	(0.205) (0.000)	(0.240)	
$\ln s_i$	1.533	1.503	$[1.003, 2.018]$
	(0.236) (0.000)	(0.308)	
$\ln h_i$	1.300	1.244	$[0.742, 1.737]$
	(0.252) (0.000)	(0.301)	
$\ln(n_i + 0.05)$	-4.177	-3.907	$[-5.7401, -2.098]$
	(1.094) (0.000)	(1.117)	
ρ	0.503	0.488	$[0.264, 0.695]$
	(0.108) (0.000)	(0.133)	
Implied λ	0.021	0.019	
Half-life	46	49	
Common factor test	7.987 (LR)	1.000 PMP unrest.	
SDM vs. SEM	(0.092)	0.000 PMP rest.	
Restricted regression			
constant	14.288	13.692	$[10.107, 17.382]$
	(1.858) (0.000)	(2.225)	
$\ln y_i(1960)$	-1.512	-1.440	$[-1.845, -1.046]$
	(0.206) (0.000)	(0.244)	
$\ln s_i - \ln(n_i + 0.05)$	1.570	1.556	$[1.067, 2.047]$
	(0.236) (0.000)	(0.299)	
$\ln h_i - \ln(n_i + 0.05)$	1.328	1.266	$[0.768, 1.762]$
	(0.252) (0.000)	(0.303)	
ρ	0.509	0.494	$[0.261, 0.704]$
	(0.107) (0.000)	(0.133)	
Implied λ	0.021	0.019	
Half-life	46	49	
Test of restriction	1.434	0.827 PMP unrest.	
	0.488	0.173 PMP rest.	
Common factor test	7.877 (LR)	1.000 PMP unrest.	
SDM vs. SEM	(0.049)	0.000 PMP rest.	

Notes: coefficient estimates and standard deviations as well as p-values (in parentheses) are reported for ML estimation.
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.

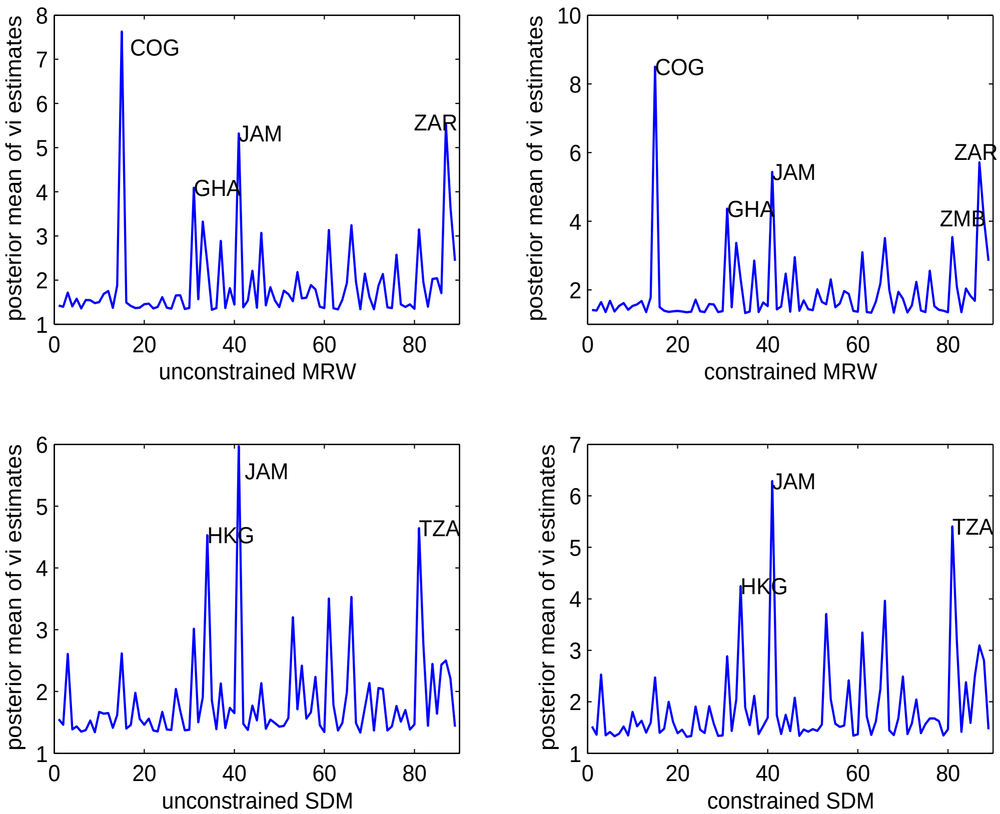

Figure 1: Posterior mean of v_i estimates for the MRW and SDM models in levels.

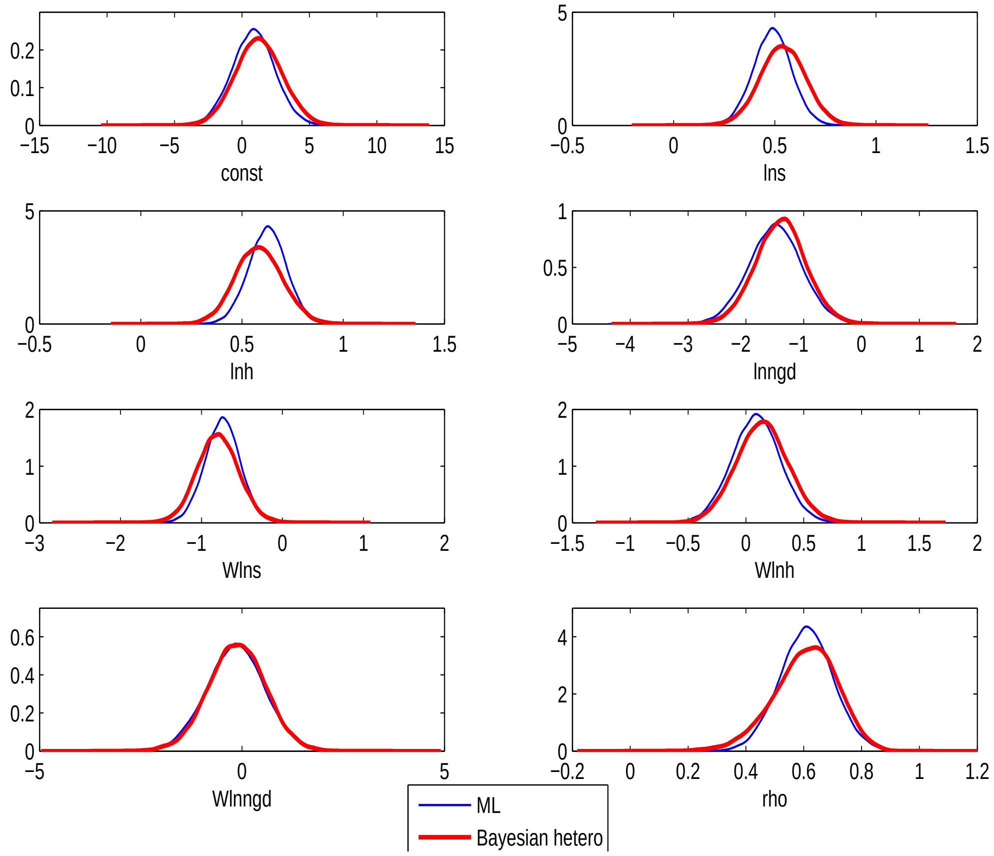

Figure 2: Kernel density estimates of Bayesian posterior and simulated maximum likelihood distributions of the parameters for the unconstrained SDM model in levels.

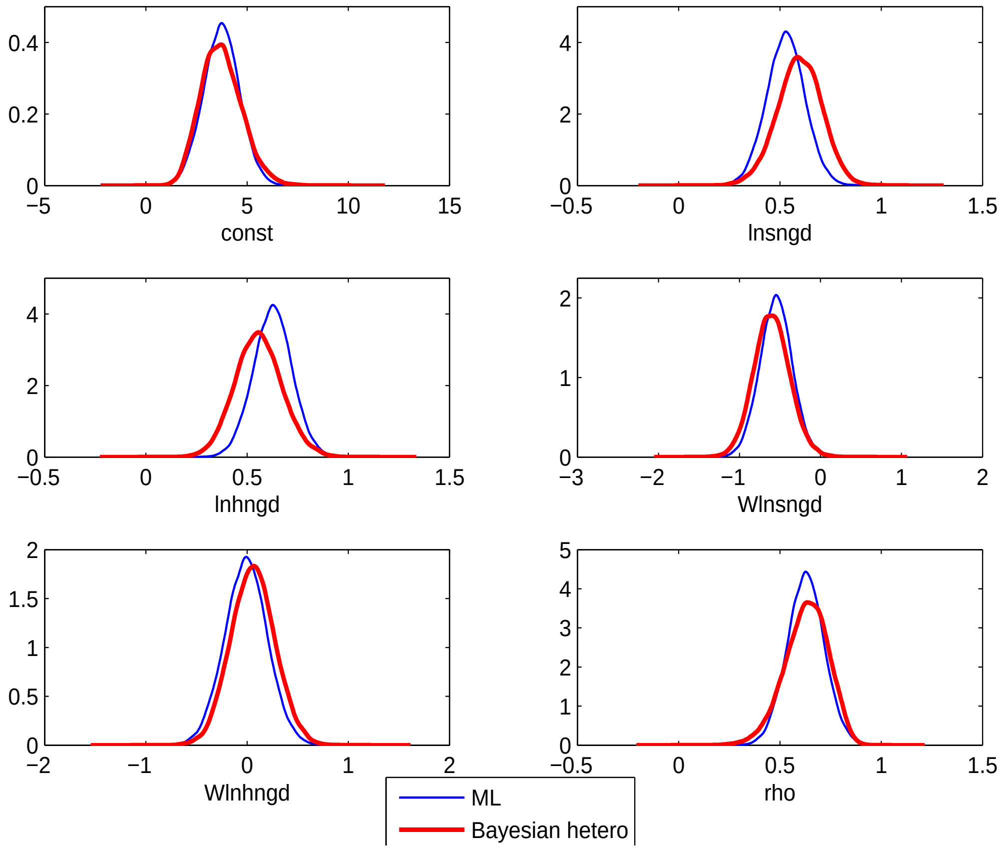

Figure 3: Kernel density estimates of Bayesian posterior and simulated maximum likelihood distributions of the parameters for the constrained SDM model in levels.

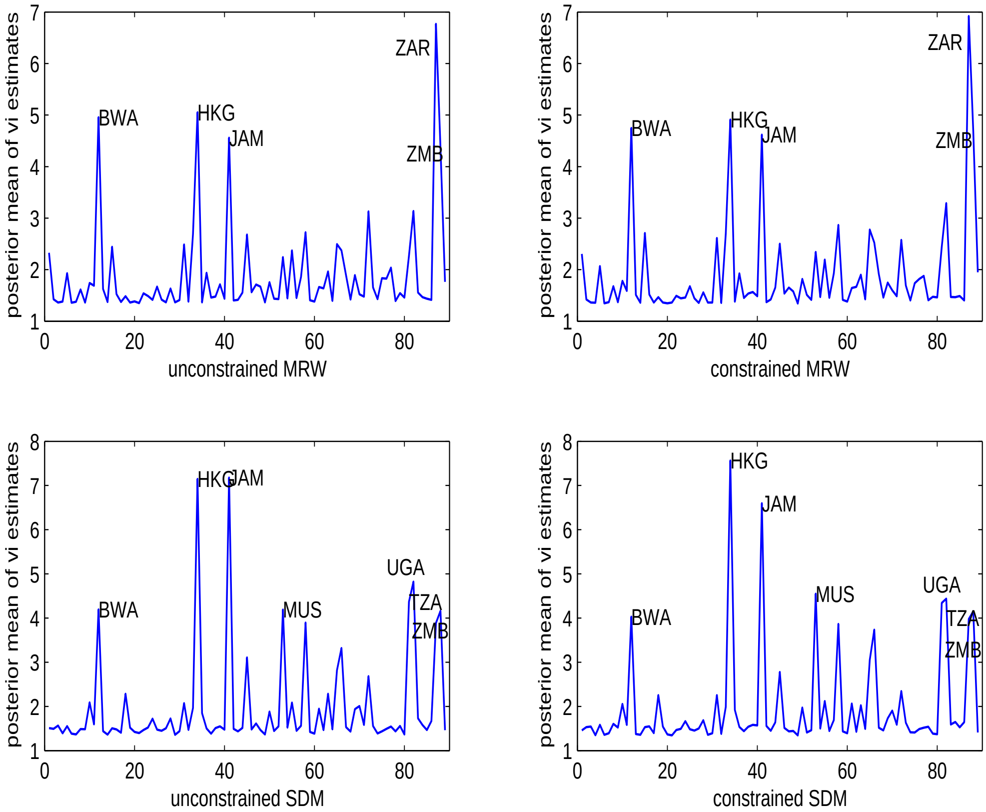

Figure 4: Posterior mean of v_i estimates for the MRW and SDM homogenous convergence models.

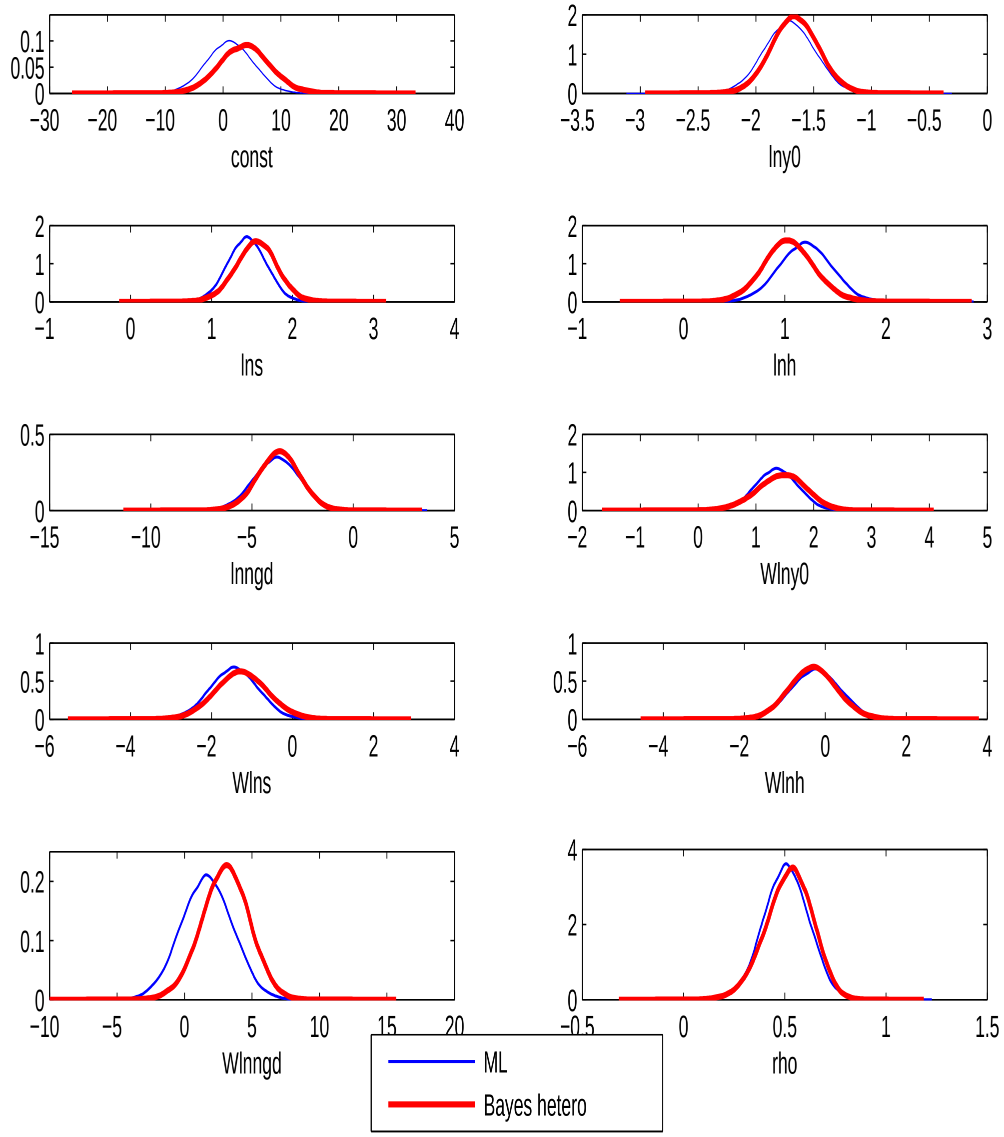

Figure 5: Kernel density estimates of Bayesian posterior and simulated maximum likelihood distributions of the parameters for the unconstrained SDM homogenous convergence model

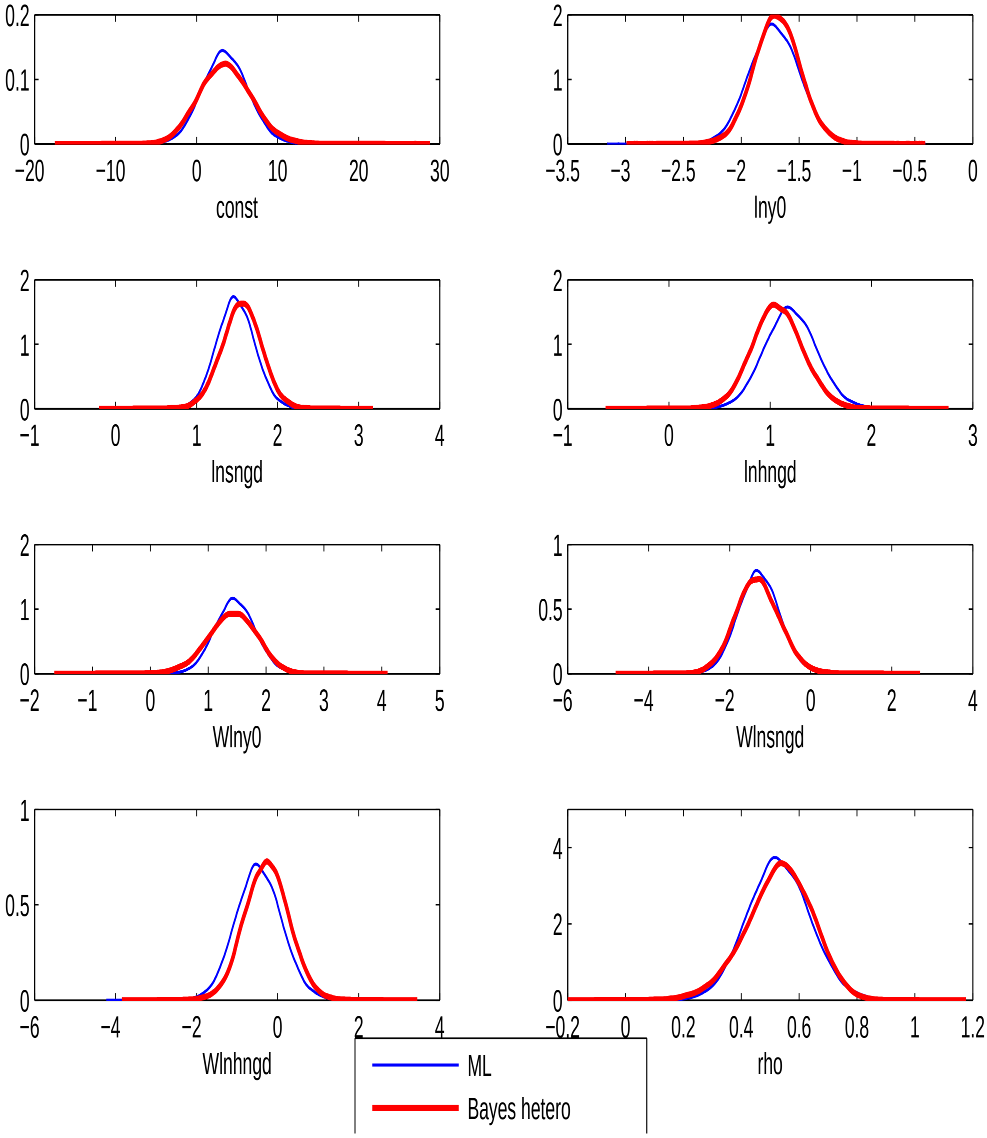

Figure 6: Kernel density estimates of Bayesian posterior and simulated maximum likelihood distributions of the parameters for the constrained SDM homogenous convergence model

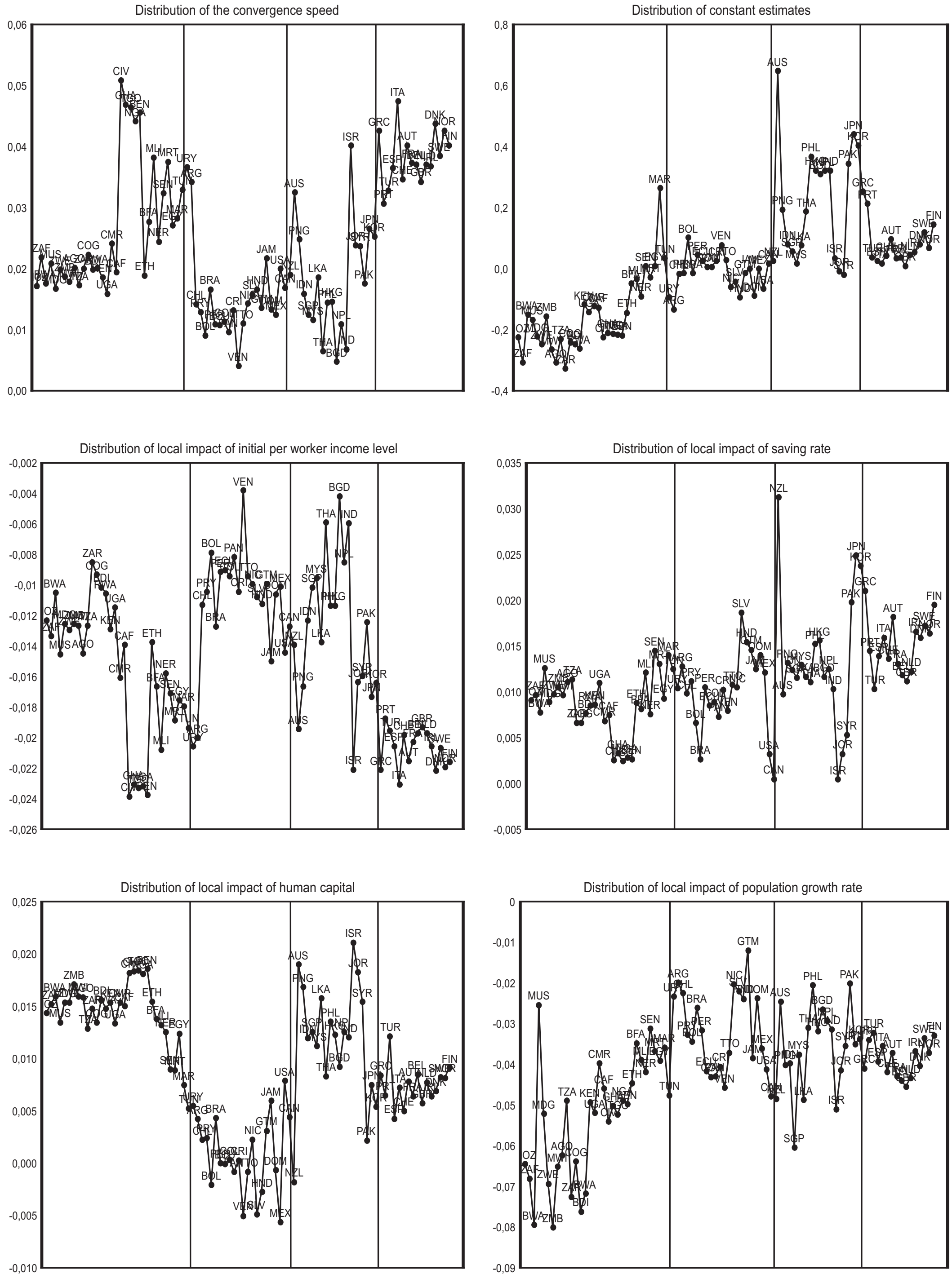

Figure 7: Distributions of local impacts of variables of interest in the heterogenous SDM convergence model with $m = 30$.

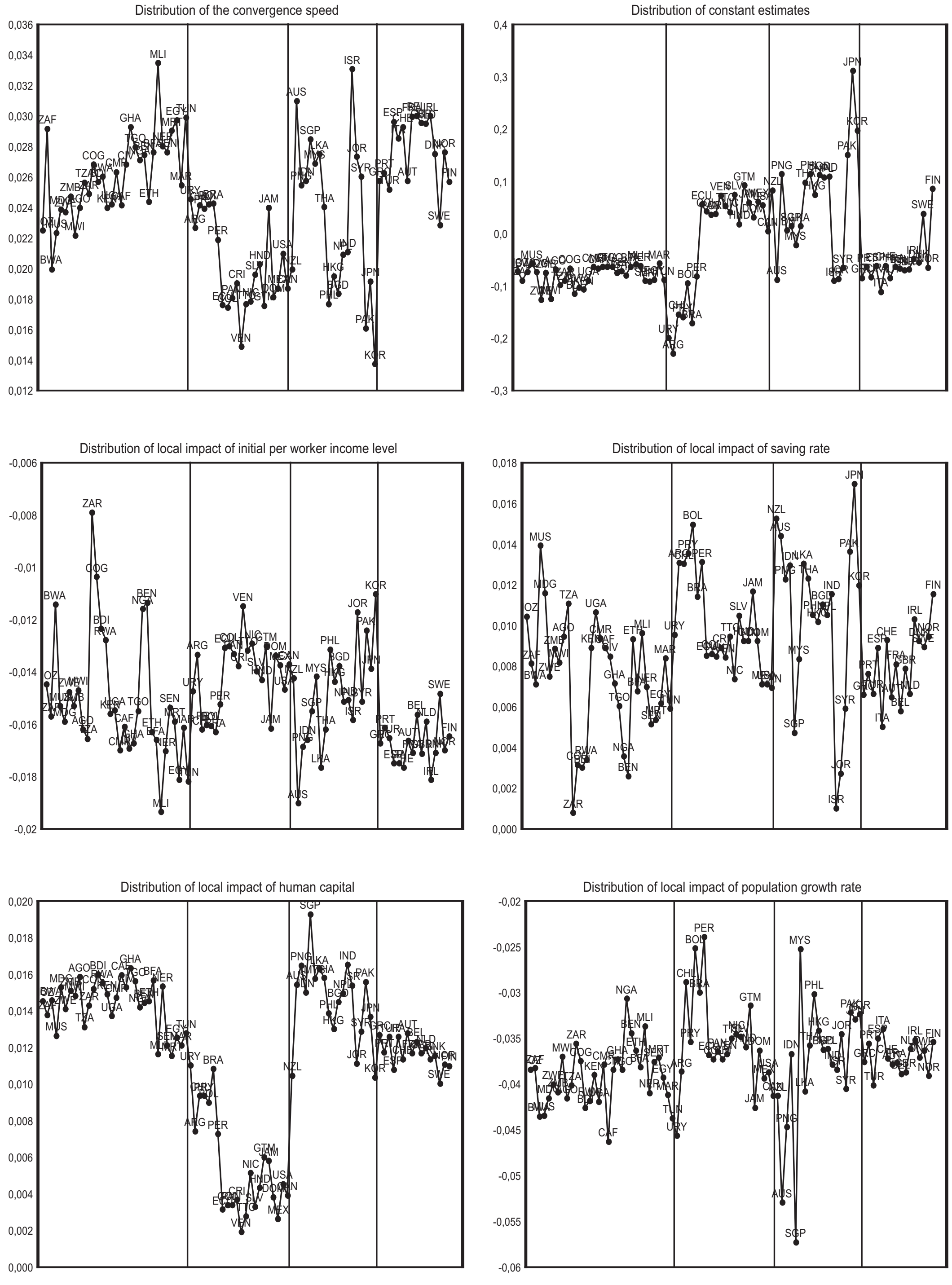

Figure 8: Distributions of local impacts of variables of interest in the heterogenous SDM convergence model with $m = 45$.

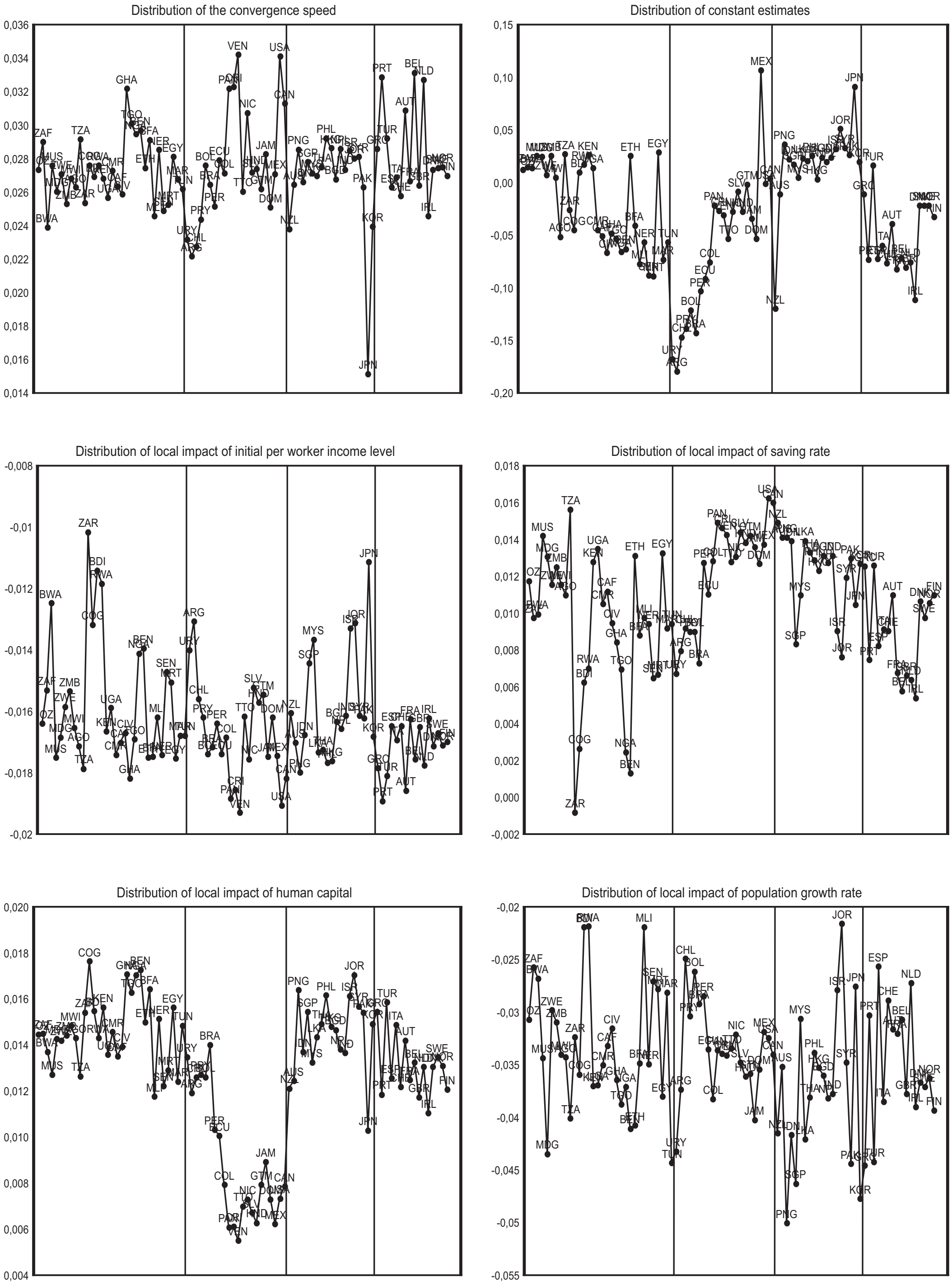

Figure 9: Distributions of local impacts of variables of interest in the heterogenous SDM convergence model with $m = 60$.

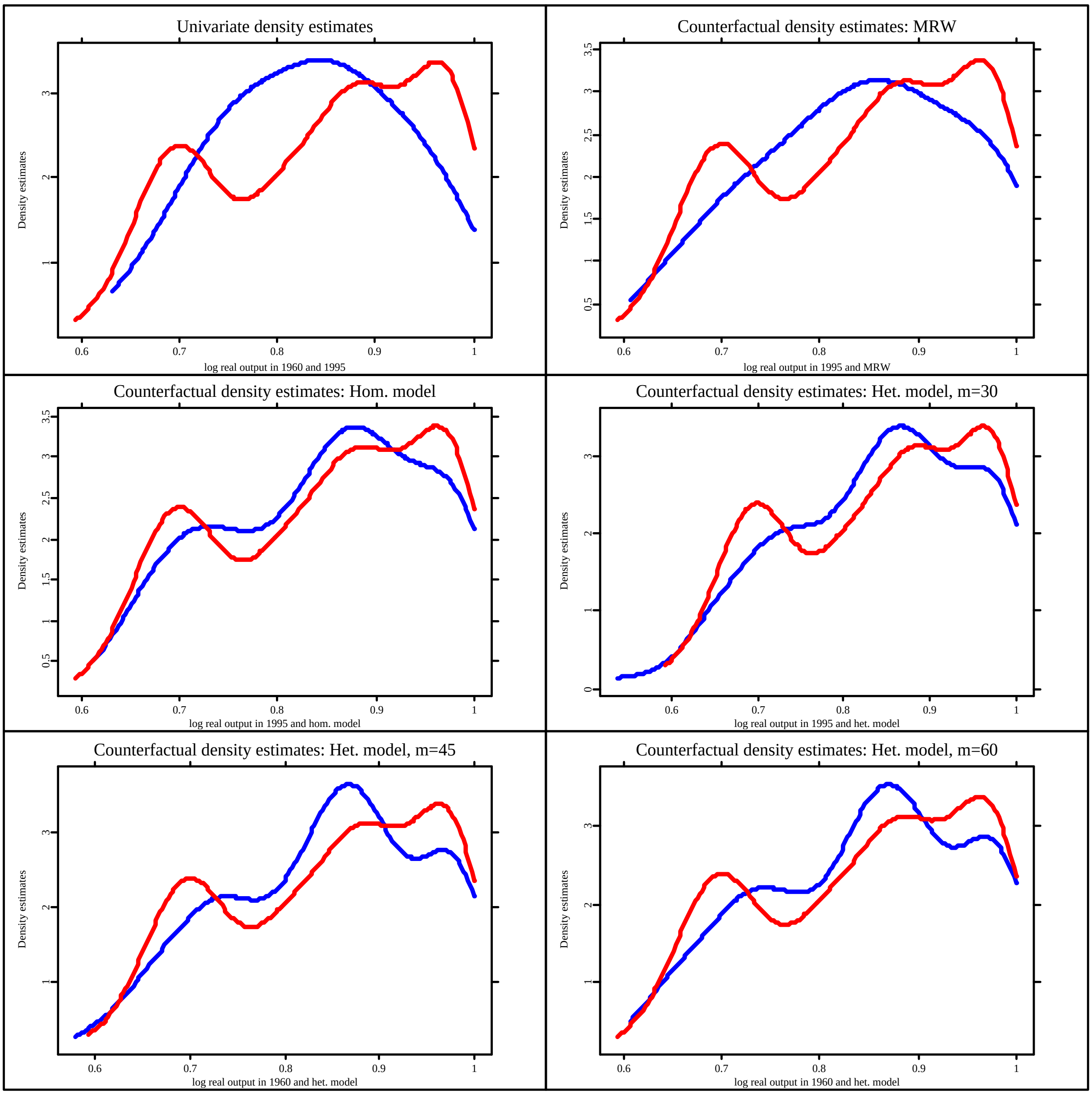

Figure 10: Counterfactual kernel density estimates