

Julie Le Gallo and Bernard Fingleton

## Contents

16.1	Introduction .....	292
16.2	Growth Regressions: From Theory to Empirics .....	292
16.3	Estimating the Rate of Convergence .....	299
16.3.1	Unconditional and Conditional $\beta$ -Convergence .....	300
16.3.2	Space and Growth .....	301
16.3.3	Econometric Issues .....	302
16.3.4	Panel Estimation .....	304
16.3.5	Multiple Regimes and Convergence Clubs .....	307
16.4	Sigma-Convergence and Distribution Approach to Convergence .....	308
16.4.1	$\sigma$ -Convergence .....	308
16.4.2	Studying the Evolution of the Cross-Sectional Distributions .....	309
16.4.3	Distribution Dynamics and Space .....	312
16.5	Conclusions .....	313
	References .....	314

## Abstract

This chapter provides a selective survey of the main developments related to the study of regional convergence. We discuss the methodological issues at stake and show how a number of techniques applied in cross-country studies have been adapted to the study of regional convergence. In doing this, we focus on the two main strands of growth econometrics: the regression approach where

---

J. Le Gallo (✉)

CRESE, Université de Franche-Comté, Besançon, France

e-mail: [jlegallo@univ-fcomte.fr](mailto:jlegallo@univ-fcomte.fr)

B. Fingleton

Department of Economics, University of Strathclyde, Glasgow, Scotland, UK

e-mail: [bf100@cam.ac.uk](mailto:bf100@cam.ac.uk)

predictions from formal neoclassical and other growth theories have been tested using cross-sectional and panel data and the distribution approach, which typically examines the entire distribution of output per capita across regions. In each case, we show how the analysis of regions rather than countries emphasizes the need to take proper account of spatial interaction effects.

---

## 16.1 Introduction

Given the persistent disparities in aggregate growth rates between countries and even within countries, the question whether incomes are converging across regions has received a lot of attention in the last two decades. From a theoretical point of view, regional growth modeling has been largely motivated by work done at the cross-country level, notably by Barro (1991), Barro and Sala-i-Martin (1995), and Mankiw et al. (1992), who developed empirical models based on the Solow-Swan economic growth model. These neoclassical models have as a major prediction the convergence of countries or regions to an equilibrium at which growth settles down to a constant rate, referred to as the steady state. Set against this are numerous variants on the basic theory and more radical departures from neoclassical principles, which allow non-convergent outcomes.

This chapter provides an overview of the main developments related to the study of regional convergence. We discuss the methodological issues at stake and show how a number of techniques applied in cross-country studies have been adapted to the study of regional convergence. In doing this, we focus on the two main strands of growth econometrics: the regression approach where predictions from formal neoclassical and other growth theories have been tested using cross-sectional and panel data and the distribution approach, which examines the entire distribution of regions. In each case, we show how the analysis of regions rather than countries emphasizes the need to take proper account of spatial interaction effects.

The chapter is organized as follows. In Sect. 16.2, we present a simple theoretical framework for two regions describing the neoclassical growth model. Section 16.3 provides a survey on the regression approach based on the concept of  $\beta$ -convergence and its spatial extensions. Section 16.4 examines the distribution dynamics approach together with exploratory spatial data analysis techniques. Section 16.5 concludes.

---

## 16.2 Growth Regressions: From Theory to Empirics

Consider two regions, each of which is governed by the same production technology, although there are differences between the regions, which lead them to separate parallel growth paths. The production technology can be described as

$$Y_{jt} = K_{jt}^{\alpha} (A_t H_{jt})^{\beta} \quad (16.1)$$

in which  $Y_{jt}$  is the level of output (GDP) in region  $j$  at time  $t$ ,  $K_{jt}$  denotes the level of capital in region  $j$  at time  $t$ ,  $A_t$  is labor augmenting technology (total factor productivity), and  $H_{jt}$  is the level of skilled labor. Dividing variables on both sides by  $A_t H_{jt}$ , we have output and capital per unit of effective labor:

$$\tilde{y}_{jt} = \tilde{k}_{jt}^\alpha \quad (16.2)$$

where  $\tilde{y}_{jt} = Y_{jt}/A_t H_{jt}$  and  $\tilde{k}_{jt} = K_{jt}/A_t H_{jt}$ . In writing this, we assume that  $\alpha + \beta = 1$ , that is, constant returns to scale, with capital's share of income equal to  $\alpha$  and augmented labor's equal to  $1 - \alpha$ , with diminishing returns to capital and augmented labor.

Consider now the dynamics entailed by this model. First, we assume that technology  $A$  grows at the constant rate  $g$  and raw labor  $L$  grows at the rate  $n_1$  in region 1 and  $n_2$  in region 2. For the moment, this is the only difference assumed between the regions. Second, assume that skilled labor  $H$  is determined by the years of schooling ( $c$ ) and the rate of return per year of schooling ( $\phi$ ) that raw labor experiences. The product  $c\phi$  determines the rate at which raw labor turns into skilled labor. Finally, the level of capital  $K$  is determined by the investment rate  $I$  and the depreciation rate  $d$  of existing capital, with investment equal to a share  $s$  of output  $Y$ . We capture the dynamics with the following system:

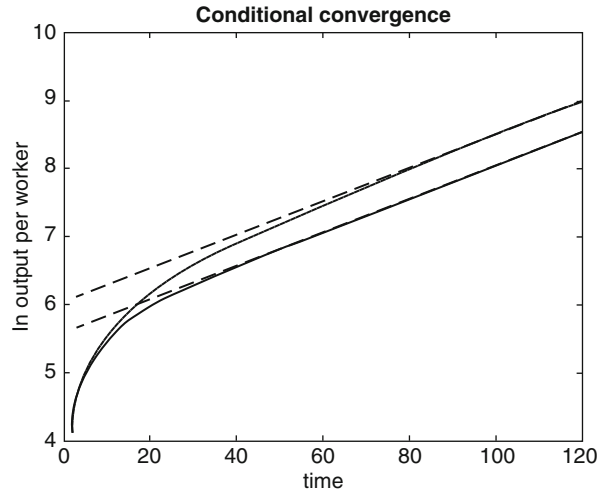
$$\begin{aligned} A_t &= A_{t-1} (1 + g) \\ L_{jt} &= L_{jt-1} (1 + n_j) \\ H_{jt} &= L_{jt} (1 + \phi c) \\ K_{jt} &= I_{jt-1} (1 - d) K_{jt-1} \\ I_{jt} &= s Y_{jt} \end{aligned} \quad (16.3)$$

Figure 16.1 shows the evolution of the system based on some assumptions (for visual effect rather than realism) about initial values and parameters  $\alpha$ ,  $g$ ,  $n_1$ ,  $n_2$ ,  $c\phi$ ,  $d$ , and  $s$ . We assume that  $g = 0.025$ ,  $s = 0.5$ ,  $A = 110$ ,  $K = 88.875$ ,  $L = 20$ ,  $Y = 90$ ,  $c = 9$ ,  $\phi = 0.1$ ,  $\alpha = 0.333$ ,  $d = 0.025$ ,  $n_1 = 0.01$ , and  $n_2 = 0.1$ . While both regions start from the same position, they move onto different steady-state paths of growth in output per worker as a result of their differing labor force growth rates.

Convergence to equilibrium is determined by the fundamental assumption of the neoclassical growth model that there are diminishing returns. To show this, consider the derivative of output per unit of effective labor with respect to capital per unit of effective labor:

$$\begin{aligned} \frac{\partial \tilde{y}_{jt}}{\partial \tilde{k}_{jt}} &= \alpha \tilde{k}_{jt}^{\alpha-1} > 0, & \lim_{\tilde{k} \rightarrow 0} \left[ \frac{\partial \tilde{y}_{jt}}{\partial \tilde{k}_{jt}} \right] &= \infty, & \lim_{\tilde{k} \rightarrow \infty} \left[ \frac{\partial \tilde{y}_{jt}}{\partial \tilde{k}_{jt}} \right] &= 0 \\ \frac{\partial^2 \tilde{y}_{jt}}{\partial \tilde{k}_{jt}^2} &= (\alpha - 1) \alpha \tilde{k}_{jt}^{\alpha-2} < 0 \end{aligned} \quad (16.4)$$

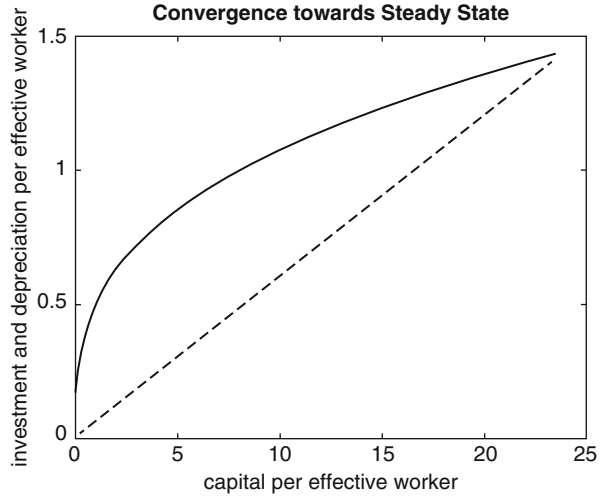
**Fig. 16.1** Conditional convergence for two regions



The first derivative is positive but goes to 0 as  $\tilde{k}_{jt} \rightarrow \infty$ , indicating that although the marginal product of capital is positive, capital deepening in the form of additional amounts of capital produces a diminishing rate of return (these are the Inada conditions).

The steady state to which the economy evolves is determined by the fact that although increasing income produces increasing investment, as shown by  $I_{jt} = sY_{jt}$ , there is a simultaneously occurring increase over time in aggregate depreciation, the most obvious component of which is due to capital depreciation, but which also depends on the growth in the effective number of workers. Moreover, while depreciation per effective worker is linear in capital per effective worker, investment is nonlinear, reflecting the diminishing marginal product of capital. This is shown in Fig. 16.2, which is the outcomes if we run our model and plot investment  $sY_{jt}/(A_tH_{jt})$  (solid line) and depreciation per effective worker  $(n_j + g + d)K_{jt}/(A_tH_{jt})$  (dotted line) against capital per effective worker  $K_{jt}/(A_tH_{jt})$  using the data for region  $j = 1$ . Figure 16.2 shows that at low levels of capital per effective worker, investment is at a higher level than “depreciation.” However, with diminishing returns, the gap between the investment and depreciation schedule narrows progressively to the point where all savings are absorbed offsetting the effects of depreciation and effective labor force growth. Beyond this point, although additional income would generate additional savings and investment, the curvilinear savings schedule is now below the linear depreciation schedule, and the change in capital per effective worker becomes negative, and so the system moves back in the direction of falling income toward the equilibrium point. Thus, we have a stable equilibrium at which investment is just sufficient to balance the effects of depreciation and effective labor force growth and maintain the level of capital per effective worker.

**Fig. 16.2** Investment vs capital per effective worker; First region



**Fig. 16.3** Investment vs income per effective worker; First region

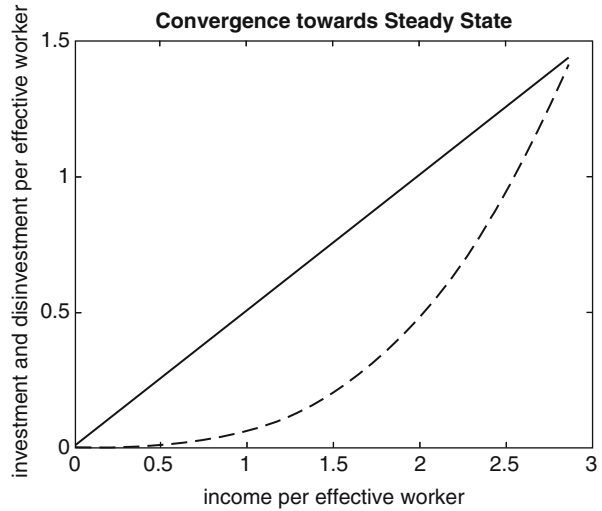
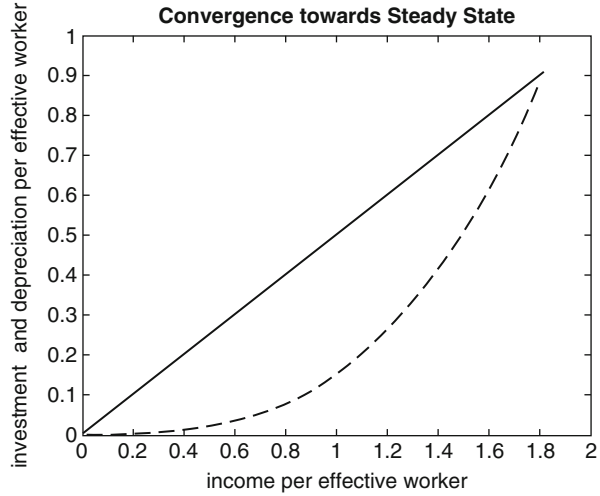


Figure 16.3 plots the same data but with income per effective worker  $Y_{jt}/(A_t H_{jt})$  as the horizontal axis. Thus, using the data for region 1, this identifies the stable equilibrium point for income per effective worker as 2.86. Figure 16.4 is the equivalent data for the second region. Here, we see the effect of faster labor force growth, which produces a lower equilibrium point at about 1.81.

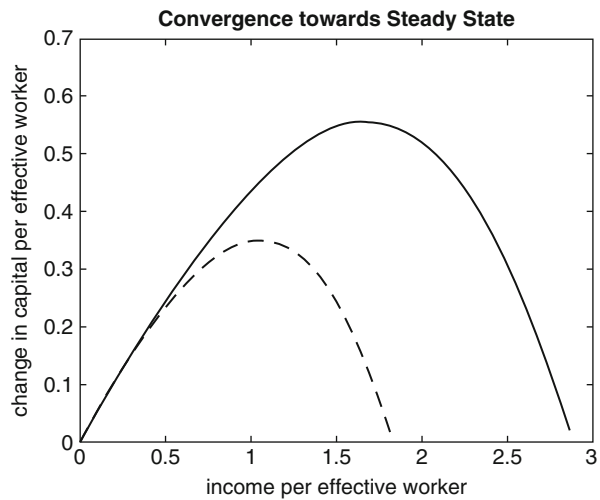
Figure 16.5 plots the two components of the right-hand side of the equation showing how capital per effective worker evolves, which is equal to

$$\dot{\tilde{k}}_t = s\tilde{y}_t - (n + d + g)\tilde{k}_t \tag{16.5}$$

**Fig. 16.4** Investment vs income per effective worker; First region



**Fig. 16.5** Change in capital per effective worker



where  $\dot{\tilde{k}}_t$  is the derivative of  $\tilde{k}_t$  with respect to time. From this, it is possible to obtain the equilibrium point equal to  $\dot{\tilde{k}}_t = 0$ , so that as we have shown graphically  $s\tilde{y}_t = (n + d + g)\tilde{k}_t$ . Figure 16.5 shows the evolution of  $\dot{\tilde{k}}_t$  identifying our two equilibrium income per effective worker points at which  $\dot{\tilde{k}}_t = 0$  for our two regions. It follows that at equilibrium

$$s\tilde{y}_j^* = (n_j + d + g)\tilde{k}_j^* \tag{16.6}$$

Hence,

$$\begin{aligned} s\tilde{k}_j^{*\alpha} &= (n_j + d + g)\tilde{k}_j^* \\ \tilde{k}_j^* &= \left[ \frac{s}{n_j + d + g} \right]^{\frac{1}{1-\alpha}} \end{aligned} \quad (16.7)$$

and the equilibrium output per effective worker is

$$\tilde{y}_j^* = \tilde{k}_j^{*\alpha} = \left( \frac{s}{n_j + d + g} \right)^{\frac{\alpha}{1-\alpha}} \quad (16.8)$$

This means that equilibrium output is

$$Y_{jt}^* = \left( \frac{s}{n_j + d + g} \right)^{\frac{\alpha}{1-\alpha}} A_t H_{jt} \quad (16.9)$$

and equilibrium output per worker is

$$\frac{Y_{jt}^*}{L_{jt}} = \left( \frac{s}{n_j + d + g} \right)^{\frac{\alpha}{1-\alpha}} \frac{A_t H_{jt}}{L_{jt}} \quad (16.10)$$

Hence we have

$$\ln y_{jt}^* = \ln A_t + \frac{\alpha}{1-\alpha} \ln s - \frac{\alpha}{1-\alpha} \ln(n_j + d + g) + \ln \left( \frac{H_{jt}}{L_{jt}} \right) \quad (16.11)$$

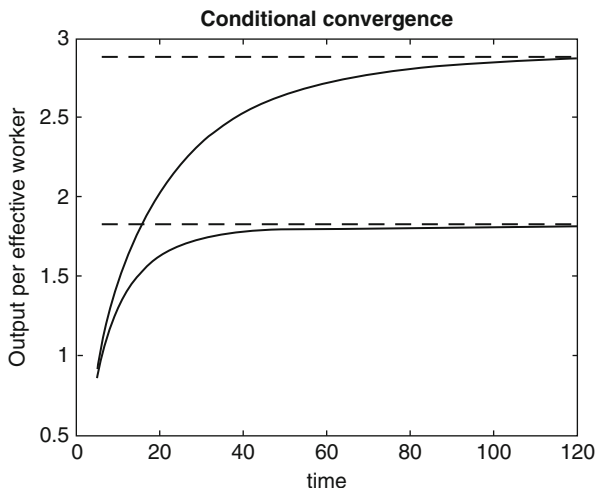
Equation (16.11) provides the equilibrium level of output per worker as traced by the broken lines of Fig. 16.1 for our two regions. It shows a steady growth, at rate  $g$ , but with different levels at any one point in time on account of the different labor force growth rates. In terms of output per unit of effective labor, we have seen from Fig. 16.5 and earlier that this converges to a constant 2.86 for region 1 and 1.81 for region 2. Following Eq. (16.10), the evolution toward this steady state is given by the constant

$$\ln \left( \frac{Y_{jt}^*}{A_t H_{jt}} \right) = \frac{\alpha}{1-\alpha} \ln s - \frac{\alpha}{1-\alpha} \ln(n_j + d + g) \quad (16.12)$$

This is illustrated by Fig. 16.6.

We have given a highly stylized account of the determinants of regional growth, with regional differences existing purely as a consequence of differences in the rate of growth of labor. Thus, we have assumed that depreciation, returns to scale, the rate of technical progress, initial levels of technology, skilled labor, capital, and the savings rate are equal across our regions. Nevertheless, we see that this simple

**Fig. 16.6** Evolution of output per effective worker



difference has consequences for the equilibria to which each region converges and the rate of convergence.

There is much interest in estimating convergence rates. As a result of linearizing the steady-state dynamics using a Taylor series expansion, we find that, approximately, the growth of output per effective worker is given by the gap between log level of output per effective worker and the log equilibrium level, thus

$$\frac{\partial \ln(\tilde{y}_{jt})}{\partial t} = -(1 - \alpha)(n_j + d + g)(\ln(\tilde{y}_{jt}) - \ln(\tilde{y}_{jt}^*)) \tag{16.13}$$

where the rate of convergence is  $\beta_j = (1 - \alpha)(n_j + d + g)$ . Note that for the parameters values in our example,  $\beta_1 = 0.04$  and  $\beta_2 = 0.1$ , which compares with  $\beta = 0.02$  (the so-called 2 percent rule) suggested by Barro and Sala-i-Martin (1995) which has in fact been observed in many growth studies. Integrating and writing in per worker terms, we obtain

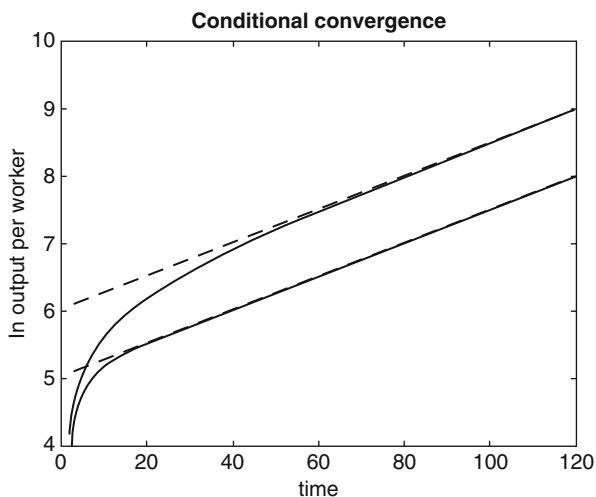
$$\frac{1}{T} \ln\left(\frac{Y_{jt}}{L_{jt}}\right) = k - \frac{1}{T} e^{-\beta_j T} \ln\left(\frac{Y_{jt-T}}{L_{jt-T}}\right) \tag{16.14}$$

With large  $\beta_j T$  the left-hand side is equal to  $k$ , which is proportional to the equilibrium level of output per worker.

One interesting prediction from the neoclassical growth model is the phenomenon of “catching up.” Consider two regions starting from different levels of output per worker. If we keep the equilibrium growth path for each the same for simplicity, then we find that there is faster growth in the initially poorer region. However, the prediction is more complex when both starting level and equilibrium path are different, as is more likely in the real world, as a result, for example, of a lower level of capital endowment and faster labor force growth rate. In our simulation,



**Fig. 16.7** Convergence for two regions



the initially poorer country experiences a short-lived spurt of possibly faster growth at the outset, but over the longer term, we see growth moving sooner onto the equilibrium growth path entailing a lower equilibrium level of output per worker (obtained by setting  $K = 1$  and  $n_2 = 0.4$ ). Figure 16.7 illustrates this outcome.

It should be noted that in the neoclassical growth model described above, there is no explicitly modeled spatial interaction between the regions, although the common growth rate of technology  $g$  may be interpreted as implying perfect diffusion of technological change across the regions. The model also assumes that the investment share of output is fixed over time. Once regional differences in investment behavior, innovation diffusion, and interregional migration are taken into account, the extent of catching up will strongly depend on the strength of these types of spatial interaction (see Nijkamp and Poot 1998).

### 16.3 Estimating the Rate of Convergence

In this section, we review the main econometric issues associated with the estimation of the rate of convergence.

The debate on convergence has given rise to numerous empirical studies with often contradictory results due, partially, because various conceptions of convergence were tested and because various methodological approaches and procedures of tests have been used (cross section, panel data, temporal series, etc.). The first developments concern the idea of convergence-catching up, which is associated with the concept of  $\beta$ -convergence. This is based on the relationship between initial output and subsequent growth.

There are however two main approaches allowing the test of this hypothesis: absolute  $\beta$ -convergence and conditional  $\beta$ -convergence. Take again Eq. (16.11), which gives the equilibrium output per worker. This level depends upon several

parameters  $\theta = (g, n, c, \phi, d, s, \alpha)$ . If all elements of  $\theta$  are similar for all regions, which then only differ by their initial effective per worker capital, then there is absolute  $\beta$ -convergence. If some elements of  $\theta$  differ between regions, as was the case in our simulations, then there is conditional  $\beta$ -convergence.

### 16.3.1 Unconditional and Conditional $\beta$ -Convergence

Consider first the simplifying case where all regions are structurally identical and have access to the same technology. They differ only by their initial conditions. In this case, they converge toward the same steady state and have the same growth rate at steady state. It is only in this case that poor regions grow faster than rich ones and eventually catch them up in the long run.

When cross-sectional data are available for two periods, initial period 0 and final period  $T$ , then Barro and Sala-i-Martin (1995) show that the hypothesis of this assumption of unconditional  $\beta$ -convergence is usually tested using the following model:

$$\frac{1}{T} \ln \left( \frac{y_{iT}}{y_{i0}} \right) = \alpha + \beta \ln(y_{i0}) + u_i \quad u_i \rightarrow iid(0, \sigma_u^2) \quad (16.15)$$

where  $i = 1, \dots, N$ ,  $N$  is the number of regions in the sample;  $Y_{iT}$  is the per capita output (measured, for instance, by income or per capita GDP) for region  $i$  at time  $t$ ,  $t = 0$  or  $T$ ;  $(1/T) \cdot \log(y_{iT}/y_{i0})$  is the average growth rate of per capita output between the two dates; and  $\alpha$  and  $\beta$  are the unknown parameters to be estimated. There is unconditional  $\beta$ -convergence if  $\beta$  is negative and significant. The rate of convergence between regions can then be estimated as  $\gamma = -\ln(1 + T\beta)/T$ .

Consider now the case of regions with different steady states. Then, as we showed before, the growth rate of a region is positively related to the distance that separates it from its own steady state. This is the concept of conditional  $\beta$ -convergence. In order to test for this assumption, it is necessary to hold constant the steady states specific to each region. This may be done by adding in Eq. (16.15) explanatory variables that control the heterogeneity of the long-term path:

$$\frac{1}{T} \ln \left( \frac{y_{iT}}{y_{i0}} \right) = \alpha + \beta \ln(y_{i0}) + \gamma X_i + u_i \quad u_i \rightarrow iid(0, \sigma_u^2) \quad (16.16)$$

Where  $X_i$  is the vector of variables adjusting for the steady state of region  $i$ . As before, there is conditional  $\beta$ -convergence if  $\beta$  is negative and significant. The additional variables can be divided in two groups. On the one hand, state variables in accordance with the Solow-Swan model or some version of it must be introduced. As in Eq. (16.11), these are physical capital, human capital, and population growth rate. On the other hand, empirical studies often include numerous control variables, the expected effects of which correspond to their influence on the position of the steady state. For instance, Durlauf et al. (2005) identify 145 potential

growth determinants. This concept of convergence is compatible with a high degree of inequality if the regional steady states are very different. The question then is why the steady states of some regions remain so low.

### 16.3.2 Space and Growth

While lots of papers analyzing convergence at subnational scales initially employed techniques used in cross-country analysis, there is recognition that countries and regions are not interchangeable. Indeed, regions usually display a greater deal of openness, and various forms of regional interdependencies exist. Consequently, a vast strand of the regional science literature has made use of spatial econometric techniques and specifications to analyze regional convergence. We briefly review here some of the main issues at stake.

One major issue associated with the spatial dimension of the data is spatial autocorrelation in the error terms. Indeed, in the cross-sectional context, units are spatially organized and the *iid* assumption usually imposed in convergence specifications is overly restrictive. Various specifications are appropriate to control for spatial dependence; we present here the most commonly used. Consider Eq. (16.16) in matrix form such as  $y = X\gamma + u$  where  $y$  is a vector containing the observations of average regional growth rates and  $X$  is the matrix containing the observations on all explanatory variables: constant term, initial income, and all the other control variables – and  $u$  is the vector of error terms.

In the spatial lag model, a spatially lagged variable  $Wy$  is added as an additional explanatory variable:

$$y = \rho Wy + X\gamma + u \quad (16.17)$$

where  $W$  is the spatial weight matrix and  $\rho$  is the spatial autoregressive parameter. The error term  $u$  is *iid*. The spatial lag  $Wy$  is always endogenous so that this specification should be estimated using maximum likelihood or instrumental variables. Particular attention should be given to the interpretation of the coefficients in this model as they only include the direct marginal effects of an increase in the associated explanatory variables, excluding all indirect induced effects (LeSage and Pace 2009 and ► Chap. 77, “Interpreting Spatial Econometric Models” in this handbook).

The spatial error model is a special case of a nonspherical error covariance matrix in which the spatial error process is based on a parametric relation between a location and its neighbors. In the spatial autoregressive specification, the error vector  $u$  takes the form

$$u = \lambda Wu + \varepsilon \quad (16.18)$$

where  $\varepsilon$  is *iid* and  $\lambda$  is the spatial autoregressive parameter. Conversely, the moving average specification can be expressed as

$$u = \gamma W\varepsilon + \varepsilon \quad (16.19)$$

Both models can be estimated using maximum likelihood or generalized method of moments. The two specifications differ in the terms of the range of spatial dependence in the variance-covariance matrix and of the diffusion process they imply. In particular, in the first case, the spillovers are global: a random shock in one observation impacts upon the income of all the regions in the sample. In the second case, the spillovers remain local: a shock in location  $i$  only affects the regions directly interacting with  $i$ , that is, the regions  $j$  for which  $w_{ij} \neq 0$ .

In the convergence context, both models have been extensively used to capture regional interdependence (Rey and Le Gallo 2009). Interestingly, some cross-country studies also acknowledge the need of taking spatial dependence into account and hence use spatial econometric techniques. Models incorporating spatial lags of the dependent and independent variables (spatial Durbin model) or higher-order spatial models have also been suggested (for a recent review, see Fischer and Wang 2011). As the spatial Durbin model encompasses the spatial lag and the spatial error model, it can be used as a basis for specification search (see ► Chap. 27, “Classical Contributions: Von Thünen, Weber, Christaller, Lösch” in this handbook for more details on specification search in cross-sectional spatial models).

Finally, note that a recent trend of the literature consists in providing sound theoretical foundations for the inclusion of spatial dependence in  $\beta$ -convergence models. For instance, Ertur and Koch (2007) show how a spatial Durbin model version of the  $\beta$ -convergence model can be obtained from a theoretical growth model with Arrow-Romer externalities and spatial externalities that imply inter-economy technology interdependence. Likewise Fingleton and Lopez-Bazo (2006) introduce substantive spatial externalities in the neoclassical convergence equation and show how this leads to a different steady-state level of output per unit of effective labor than would otherwise occur.

### 16.3.3 Econometric Issues

Although the conditional  $\beta$ -convergence approach has given rise to hundreds of studies, it has also been widely criticized.

#### 16.3.3.1 Endogeneity of Explanatory Variables

In a regression setup, error terms are often correlated with the explanatory variables, leading to endogeneity and inconsistent estimates. In  $\beta$ -convergence models, there are numerous sources of endogeneity.

The first source of correlation between errors and some explanatory variables is simultaneity: some explanatory variables are not exogenous, they are determined simultaneously with growth rates, and thus they may affect growth but also depend on growth. For instance, given the Solow-Swan framework, state variables such as investment, initial per capita GDP, or human capital are equilibrium outcomes, as are regional growth rates. More generally, the causality versus the correlation issue is a prevalent one in growth econometrics. On the one hand, this implies biased estimation. On the other hand, this calls into question the interpretation of regression results and the extent to which these variables affect the steady-state levels.

Finding appropriate instruments, that is, variables that are correlated with the endogenous explanatory variables but uncorrelated with, or orthogonal to, the error terms, is a difficult task. Indeed, appropriate instrumental variables are rarely available. Since growth can be explained by numerous determinants, it is difficult to identify instruments that are correlated with the endogenous variables and yet can legitimately be eliminated from the regression. Moreover, as the effect of some variables on growth may be delayed, using lagged explanatory variables as their exogenous instruments is not optimal either.

The second source of correlation between errors and explanatory variables is measurement errors or errors in variables. This is of particular concern in growth regressions. Indeed, many countries build databases in which the accuracy of the variables is undoubtedly measured with error, and also in many cases, pragmatic decisions have to be made to use a variable that is only a proxy of a true variable. When the initial per capita GDP is mismeasured, the attenuation bias tends to bias the estimates of  $\beta$  in favor of the  $\beta$ -convergence hypothesis. For instance, Temple (1998) argues that the famous result of conditional convergence of economies at a rate of 2 % per year could be entirely due to measurement error.

Correcting for this is not an easy task and is further complicated in the presence of spatial error autocorrelation. Indeed, Le Gallo and Fingleton (2012), using Monte Carlo simulations, show that OLS and instrumental variable estimation, which do not take into account spatial error autocorrelation, outperforms GMM-based and ML estimation. These results would indicate that measurement error plus a disturbance process involving spatial dependence is best accommodated by an estimation method that ignores spatial dependence. Clearly, the interaction between spatial autocorrelation and measurement errors, which are both easy to find in  $\beta$ -convergence models, should be further investigated.

The third source of correlation between errors and explanatory variables is omitted variables. In practice, it is unlikely that researchers are able to find all the variables controlling for the differences in steady states between regions. Hence, the error term in conditional  $\beta$ -convergence models will probably contain a number of omitted variables correlated with the included regressors, though if in the unlikely event they are orthogonal to the included regressors, then there is no problem. Trying to solve this by increasing the number of explanatory variables typically runs into the problem of simultaneity and possibly multicollinearity. Note that LeSage and Fischer (2008) have shown that the existence of omitted explanatory variables exhibiting nonzero covariance with variables included in the model yields a data-generating process for a growth regression that includes both an endogenous spatial lag and exogenous spatial lags (spatial Durbin growth model).

### 16.3.3.2 Robustness of Explanatory Variables

This critique relates to the choice of control variables and is linked to the lack of robustness of conditional  $\beta$ -convergence regression models. Indeed, the finding of conditional  $\beta$ -convergence and the subsequent estimation of the convergence rate is dependent upon a specific choice for the set of control variables. The lack of consensus about the most important growth determinants amplifies this problem:

if most regressors included in the empirical analysis are found to be statistically significant in some specification, it means that there are as many growth theories as the number of significant regressors and that it is impossible to distinguish between them. This is referred to as the problem of observational equivalence of competing theories, which is common in macroeconomic analysis generally.

Confronted by the variety of explanatory variables available for use in these regressions, Levine and Revelt (1992) employ extreme bound analysis, which consists of estimating the upper and the lower extreme bounds of a coefficient of a variable of interest across a range of different model specifications. The variable is considered to be robust if the coefficients at these extreme bounds are significant and if they maintain their signs and statistical significance across a diverse range of other included variables. Using this approach, they show that most variables tested turn out to be insignificant given additional control variables.

This approach has been criticized as being excessively conservative. More recently, the use of model averaging and Bayesian model averaging has been advocated in order to guide in the choice of control variables (Fernandez et al. 2001; Sala-i-Martin et al. 2004). In a spatial context, an additional source of uncertainty pertains to the choice of the spatial weights matrix. A Bayesian model averaging approach for selecting appropriate explanatory variables together with an appropriate spatial weights matrix has been suggested by LeSage and Fischer (2008). An alternative is to explain the variation in results by means of meta-analysis (Abreu et al. 2005).

### 16.3.4 Panel Estimation

If unmodeled region-specific unobserved effects on output levels are present, this implies a link between the error terms and initial output per capita. In order to correct for this, a number of researchers advocate convergence analysis via the use of panel data (Islam 1995). We have a choice as to how we model the individual effects: fixed effects, essentially dummy variables, one per region, or random effects, in which the individual-specific (region) effect is captured as a random variable. The setup of fixed effects models follows on naturally from the pure cross-sectional growth models considered thus far, typically having the form

$$\begin{aligned} \ln y_{it} &= \gamma_t + \alpha \ln(y_{it-\tau}) + X'_{it}\beta + a_i + u_{it} \quad t = 2, \dots, T \\ u_{it} &\rightarrow iid(0, \sigma_u^2) \end{aligned} \quad (16.20)$$

which can be written as a growth equation as follows:

$$\begin{aligned} \Delta \ln y_{it} &= \gamma_t + (\alpha - 1) \ln(y_{it-\tau}) + X'_{it}\beta + a_i + u_{it} \quad t = 2, \dots, T \\ \Delta \ln y_{it} &= \ln y_{it} - \ln y_{it-\tau} \\ u_{it} &\rightarrow iid(0, \sigma_u^2) \end{aligned} \quad (16.21)$$

where growth  $\Delta \ln y_{it}$  is measured between period  $t$  and some previous period  $t - \tau$  (usually  $\tau \geq 5$  years to avoid business cycle effects). In this approach, all the unobserved time-invariant regional heterogeneity is captured by individual-specific effects, denoted by  $a_i$ . Following Eq. (16.11), the matrix  $X$  includes other possibly time-varying factors affecting growth. In addition, growth depends on the start-of-period level  $\ln(y_{it-\tau})$ , so the estimate of the coefficient  $\alpha$  gives the rate of convergence. The term  $\gamma_t$  represents time (dummy variable) effects that are constant across locations.

The presence of the lagged dependent variable together with the time-invariant effect  $a_i$  in Eq. (16.20) renders OLS inconsistent even when the transient disturbances  $u_{it}$  are not serially correlated. The most obvious way to fix this is to first difference the data, so that the individual-specific (fixed or random) effects are eliminated. Thus, our differenced specification is

$$\Delta \ln y_{it} = \Delta \gamma_t + \alpha \Delta \ln(y_{it-\tau}) + \Delta X'_{it} \beta + \Delta u_{it} \quad t = 3, \dots, T \quad (16.22)$$

While the convergence parameter  $\alpha$  is identified in Eq. (16.20), eliminating the time-invariant individual-specific effects does not solve the problem of inconsistent and biased parameter estimation via OLS because the lagged dependent variable is correlated with  $u_{it}$ , and there is also potential endogeneity of other regressors (including measurement error), omitted variables and spatial dependence. Rather, the reason to first difference is to create instruments that are not correlated with the individual effects.

With regard to spatial dependence, this can exist as a result of direct autoregressive interaction across space of the dependent variable, as a consequence of a spatial error process, or both. A good, comprehensive summary for static spatial panel models is provided in Chap. 12 of Pirotte (2011). If we add a spatially lagged dependent variable to the difference equation we obtain:

$$\Delta \ln y_{it} = \Delta \gamma_t + \alpha \Delta \ln(y_{it-\tau}) + \rho \Delta W_N \ln y_{it} + \Delta X'_{it} \beta + \Delta u_{it} \quad (16.23)$$

The variable  $\Delta W_N \ln y_{it}$  is also endogenous, as in the pure cross-section case. While difference-GMM estimation may appear to be appropriate, by using lagged levels of variables as instruments, it does typically create a weak instrument problem. One estimator that can potentially deal with these problems is the system GMM estimator (Arellano and Bond 1991; Bond et al. 2001); this estimates Eq. (16.23) combining both the difference equation and the corresponding levels equation, with lagged first differences as instruments for the levels equation, and lagged levels for the equation in first differences. One should however use this cautiously because, using all available lags of variables as instruments, this estimator in particular presents significant practical problems relating to overfitting and thus failure to purge endogeneity. The solution seems to restrict the number of lags employed as internally generated instruments so as to clearly satisfy the relevant diagnostics, but one may still have use external instruments in order to obtain the

necessary instrument orthogonality for consistent estimation. For the additional moments conditions associated with the levels equation to be orthogonal, it is sufficient for the variables to be mean stationary, having controlled for common time effects  $\gamma_t$ .

The other form of spatial dependence in panel models involves the disturbances. Pirotte (2011) classifies static spatial panel models according to whether the spatial disturbance process is autoregressive (SAR), or a moving average process (SMA), and whether the individual effects are considered to be fixed (deterministic or FE), or random effects (RE). If the random individual effects are not spatially autocorrelated, but the transient component of the compound error is, then he refers to the model as RE-SAR or RE-SMA. If the spatial error process applies in the same way to both transient and individual error components, so that the spatial process is at the level of the compound errors and not its individual components, then this is referred to as SAR-RE or SMA-RE, according to whether we are considering an autoregressive or moving average specification. If however the individual effects are fixed, and spatial effects are restricted to the transient errors, then the model is referred to as FE-SAR or FE-SMA according to whether we have an autoregressive or moving average process.

Accordingly, introducing the RE-SAR (or the FE-SAR) specification to our levels model gives

$$\begin{aligned} \ln y_{it} &= \gamma_t + \alpha \ln(y_{it-\tau}) + X'_{it}\beta + a_i + u_{it} \\ u_{it} &= \lambda M_N u_{it} + \xi_{it} \end{aligned} \quad (16.24)$$

where  $M_N$  is an  $(N \times N)$  matrix specific to time  $t$ , where  $N$  is the number of regions (and therefore  $M_N$  has similar properties to  $W_N$ ) and  $a_i$  are random (or fixed) effects. The two forms of interactions can also be combined and one might even extend the spatial dependence in the error to include both the transient errors and the individual effects to give the spatial autoregressive equivalent of SAR-RE:

$$\begin{aligned} \ln y_{it} &= \gamma_t + \alpha \ln(y_{it-\tau}) + \rho W_N \ln y_{it} + X'_{it}\beta + \psi_{it} \\ \psi_{it} &= a_i + u_{it} \\ \psi_{it} &= \lambda M_N \psi_{it} + \xi_{it} \\ \xi_{it} &\sim iid(0, \sigma_\xi^2) \end{aligned} \quad (16.25)$$

Alternatively, the equivalent of SMA-RE entails the moving average error process involving both individual and transient errors (Fingleton 2008) with  $\psi_{it} = \lambda M_N \xi_{it} + \xi_{it}$ .

With spatially dependent (moving average or autoregressive) errors combined with an endogenous spatially autoregressive spatial lag, the GMM approach typically has several stages, first one uses instrumental variables, assuming no spatial error process, to obtain consistent estimates of the residuals. These then become the basis for GMM estimates of the error process parameters. Finally, the data are



purged of the error dependence and consistent estimates obtained via instrumental variables in the final stage. Overall, with these more complex models, it is evident that methods based on GMM are the most versatile because they can handle multiple endogeneity and are robust to alternative error distributions, issues that are problematic under maximum likelihood.

### 16.3.5 Multiple Regimes and Convergence Clubs

As we have shown above,  $\beta < 0$  is consistent with the assumptions of the neoclassical growth model. However, this condition is also potentially consistent with economic alternatives, such as endogenous growth models or models with poverty traps. For instance, Azariadis and Drazen (1990) develop an endogenous growth model characterized by the possibility of multiple, locally stable steady-state equilibria. Which of these different equilibria a region will be converging to depends on the range to which its initial conditions belong? In other words there are convergence clubs, that is, groups of economies whose initial conditions are near enough to make group members converge toward the same long-term equilibrium. From an empirical point of view, the existence of convergence clubs can be inferred from the fact that while absolute  $\beta$ -convergence is frequently rejected for large samples of countries and regions, it is usually accepted for more restricted samples of economies belonging to the same geographical area.

While the Arariadis-Drazen model does not exhibit convergence since different initial conditions lead to different steady states, Bernard and Durlauf (1996) show that the data generated by this model will not necessarily lead to the finding that  $\beta \geq 0$ . Therefore, tests for  $\beta$ -convergence have low power against the alternative hypothesis of multiple steady states. The problem is then to distinguish evidence of club convergence from that of conditional convergence.

From an econometric point of view, the existence of multiple equilibria is characterized by parameter heterogeneity in convergence regressions. A vast range of techniques has been used in order to detect convergence clubs. Some use a priori criteria to define club members, such as belonging to the same geographical zone or having similar initial incomes. Durlauf and Johnson (1995) use regression trees (CART algorithm) where initial income and literacy rates are used to detect the convergence clubs. In the context of regional data, a number of authors have made use of exploratory spatial data analysis (ESDA) to detect spatial regimes in the data. In particular, Moran scatter plots and Getis-Ord statistics facilitate the detection of spatial clusters of high values of regional incomes and spatial clusters of low values of regional incomes. The hypothesis of  $\beta$ -convergence is then tested on each group (see, for instance, Ertur et al. 2006).

At the extreme, rather than partitioning the sample into regimes based on some structural characteristics, parameter heterogeneity might also be region specific. For instance, in Eq. (16.15), region-specific parameters  $\alpha_i$  and  $\beta_i$  must be estimated. While varying coefficient models might be used for that purpose (see ► Chap. 73, “Geographically Weighted Regression” in this handbook for

a presentation of these models), we note that for regional samples, similarities in legal and social institutions, as well as culture and language, might create spatially local uniformity in economic structures. This leads to situations where convergence rates are similar for regions located nearby in space. In order to capture this combination of parameter heterogeneity and local similarity, spatial autoregressive local estimation (SALE) model has been suggested by Pace and LeSage (2004).

## 16.4 Sigma-Convergence and Distribution Approach to Convergence

We now turn to alternative concepts of convergence that have been used in the literature on regional growth.

### 16.4.1 $\sigma$ -Convergence

In this approach, convergence is linked to the study of the dynamic evolution of some indicator of dispersion of output per capita between regions. The focus is then on whether this indicator increases or decreases over time. Two indicators of cross-sectional dispersion are commonly used: the standard deviation of log income or the coefficient of variation coefficient of this distribution.

Specifically, the test of  $\sigma$ -convergence consists of comparing an indicator of dispersion, computed at the end of the period, to the value of this indicator computed at the beginning of the period. There is  $\sigma$ -convergence if this indicator decreases over time. Formal tests using regression specifications have also been suggested by Carre and Klomp (1997) and Egger and Pfaffermayr (2009).

It is possible to show that  $\beta$ -convergence is a necessary but not a sufficient condition to  $\sigma$ -convergence. The point of departure is the absolute  $\beta$ -convergence equation where the dependent variable is the cumulated growth rate:

$$\ln(y_{iT}/y_{i0}) = a + \beta \ln(y_{i0}) + u_i \quad (16.26)$$

This equation is rewritten as

$$\ln(y_{iT}) = a + (1 + \beta) \ln(y_{i0}) + u_i \quad (16.27)$$

By taking the variance of each term in this equation, we have  $V[\ln(y_{iT})] = (1 + \beta)^2 V[\ln(y_{i0})] + V(u_i)$ , from which it is easy to show that

$$VR = \frac{V[\ln(y_{iT})]}{V[\ln(y_{i0})]} = \frac{(1 + \beta)^2}{R^2} \quad (16.28)$$

where  $R^2$  is the multiple correlation coefficient associated with Eq. (16.27).

From this, it is evident that  $\beta$ -convergence ( $\beta < 0$ ) is a necessary but not a sufficient condition for  $\sigma$ -convergence ( $VR < 1$ ). In fact, the final result depends upon two opposite effects. The first is the existence of  $\beta$ -convergence implying mean reversion. The second is linked to the existence of specific shocks to which the regions are submitted and that permanently generate per capita output dispersion.  $\sigma$ -convergence is the result of these two mechanisms and exists if the beneficial effects of mean reversion dominate the negative effects of perturbations affecting the regions.

This concept has been subject to a number of criticisms, the first of which obviously concerns the dependence of  $\sigma$ -convergence on the initial date. Second, it only focuses on the second moment of the distribution and is not informative about other moments that may be of interest, such as skewness or kurtosis. Third, interpreting measures of dispersion is not straightforward when distributions are not unimodal, and it is often the case that we encounter multimodality and twin-peakedness in practice. Fourth, it is subject to a spatial identification problem. Indeed, given a map of  $N$  incomes with a sample variance  $\sigma^2$  then there are  $N!$  spatial permutations on the map that would have the same sample variance.

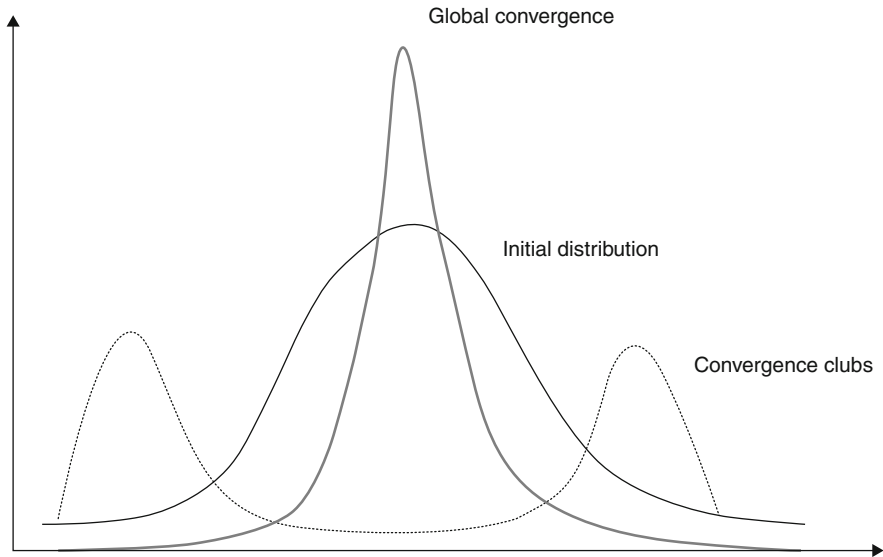
Finally, Quah (1993) forcefully argues that it does not provide meaningful information about income dynamics nor about the mobility of regions within a distribution. For instance, if two regions exchange their relative position between the initial and final date while the gap between the two remains unchanged, then the standard error of this distribution is constant over the period even if the situation of the two regions has changed radically.

## 16.4.2 Studying the Evolution of the Cross-Sectional Distributions

In the light of these criticisms, Quah (1993, 1996) argues that the cross-sectional distributions of income should be considered in their entirety rather than just computing one synthetic indicator such as dispersion. Indeed, with regard to  $\sigma$ -convergence, it tells us nothing about distribution dynamics. Rather, the evaluation of distribution dynamics can be accomplished on the basis of two criteria: the study of the evolution of income level distributions and analysis of the position of the regions or groups of regions within distributions.

Concerning the first point, the method consists of comparing the cross-sectional distributions of regional income at different points in time and evaluating the degree to which the location and shape of these distributions changes.

One possibility is to estimate, using nonparametric smoothing methods, such as kernel estimates, the density function of income for the sample, and examining the changes in the form of this density. For instance, Fig. 16.8 represents two possible ways in which the distribution might evolve over time, each representing two types of convergence. If, given the initial distribution, the regions in the sample evolve toward a tighter distribution, then there is global convergence of all regions toward the same level of income. On the contrary, if the distribution becomes bimodal or multimodal, then the regions converge toward different levels, which is symptomatic of different



**Fig. 16.8** Density functions for three different convergence issues

convergence clubs. In order to go beyond simple visual impression, tests for multimodality can also be undertaken (Henderson et al. 2008).

It is also possible to estimate cross-sectional distribution densities using mixture models, which are weighted sums of component distributions. In this case, one can say that convergence occurs when the distributions are better approximated over time by a small number of components, while multiple components are an indication of multiple regional steady states. The number of components can be evaluated using a bootstrap LR test.

Concerning the second point on the position of the regions or groups of regions within the distributions, we observe that shape dynamics does not directly address this issue. Nonetheless, it may be of interest to study whether, for a given time period, the regions have changed their relative position in the income distribution, that is, which regions move up and down in this distribution.

One method that allows detection of the movements of the regions from one period to another consists of estimating transition matrices or Markov chains. These are constructed using a discretization of the distribution of income into several classes (using for instance quartiles or quintiles of the distribution). Transition matrices allow one to estimate the probabilities of passage from one income class to another, or of remaining in the same income class, over time. If the probabilities of passage from one class to another are high, then mobility is high. If the probability of staying in the same class is high, then mobility is low. By extension, it is possible to detect whether the level of income is tending toward homogeneity or, on the contrary, if distinct groupings of regions with different incomes are emerging and being maintained over time. Formal mobility indices may also be computed, while the ergodic distribution,

that is the long-term distribution, allows one to see the type of convergence mechanism that is at work. Concentration of the frequencies in the median class would imply convergence to the mean, while concentration of the frequencies in several of the classes, that is, a multimodal limit distribution, may be interpreted as a tendency toward stratification into different convergence clubs.

In order to operationalize this, some strong assumptions are usually made, such as stationarity of transition probabilities and a first-order process. Formally, denote  $F_t$  as the cross-sectional distribution of income at time  $t$  relative to the sample average. A set of  $K$  different GDP classes is defined. If the frequency of the distribution follows a first-order stationary Markov process, then the  $(K \times 1)$  vector  $F_t$ , indicating the frequency of the regions in each class at time  $t$ , is described by the following equation:

$$F_{t+1} = MF_t \quad (16.29)$$

where  $M$  is the  $(K \times K)$  transition probability matrix representing the transition between the two distributions. If the transition probabilities are stationary, that is, if the probabilities between two classes are time invariant, then

$$F_{t+s} = M^s F_t \quad (16.30)$$

The ergodic distribution of  $F_t$  is approached as  $s$  tends toward infinity in Eq. (16.30). Such a distribution exists if the Markov chain is regular, *that is*, if and only if for some  $N$ ,  $M^N$  has no zero entries. In this case, the transition probability matrix converges to a limiting matrix  $M^*$  of rank 1. The existence of an ergodic distribution,  $F^*$  is then characterized by

$$F^*M = F^* \quad (16.31)$$

Each row of  $M^t$  tends to the limit distribution as  $t \rightarrow \infty$ . According to Eq. (16.31), this limit distribution is therefore given by the eigenvector associated with the unit eigenvalue of  $M$ . The estimation of the transition matrix is based on maximum likelihood estimation.

As indicated, strong assumptions must usually be made to estimate such transition matrices. Moreover, the results are sensitive to the number and size of the groups of observations used to discretize the data. In fact, discretization of the state space may significantly alter the probabilistic properties of the data.

To overcome this problem of sensitivity of the results to the discretization, stochastic kernels have been suggested. They are the continuous counterpart of transition probability matrices. Formally, if  $f_{X(t)}$  is the regional income density for  $n$  regions in period  $t$ , then the evolution of the cross-sectional distribution is modeled as

$$f_{X(t+s)} = \int_{-\infty}^{\infty} M_{t,s} f_{X(t)} dx \quad (16.32)$$

where  $M_{t,s}$  is the stochastic kernel representing where points in  $f_{X(t)}$  move to in  $f_{X(t+s)}$ . The estimation of this kernel may be based on an estimate of the conditional distribution. In order to explore the transitional dynamics provided by this approach, three-dimensional representations and two-dimensional contour plots are used. For example, polarization or convergence clubs in the per capita GDP distribution are reflected in peaks in the 3D kernel or by concentrated values in the contour plot. Fischer and Stumpner (2008) introduce three-dimensional stacked conditional density plots and highest density regions plot for the visualization of the transition function.

### 16.4.3 Distribution Dynamics and Space

As in the confirmatory econometric analysis of growth and convergence, the spatial dimension of the data invalidates some of the restrictive assumptions regarding random sampling on which  $\sigma$ -convergence and distribution dynamics rest. We briefly consider in this section how this impacts on the measures of convergence and distribution dynamics. First, we note that the concepts of convergence that have been developed in the preceding sections must be adjusted to take into account spatial autocorrelation in the data. Secondly, we observe that much work has been done in exploratory spatial data analysis (ESDA) and Exploratory Space-Time Data Analysis (ESTDA), and their application to convergence and growth analysis has led to interesting new insights.

Regarding the first point, consider the  $\sigma$ -convergence measure presented earlier. We have already pointed out that it is uninformative with regard to the morphology of the distribution and the degree of intradistributional mobility. Moreover, in a spatial context, the presence of spatial dependence complicates the interpretation of, and inference based on, this concept. For instance, Rey and Dev (2006) show that the sample variance also reflects the level and structure of spatial dependence in the data. This should be purged in order to correctly interpret this concept of convergence.

Similarly, spatial autocorrelation has been incorporated into measures of intradistributional dynamics. In the case of discrete Markov chains, Rey (2001) extends the approach by estimating transition matrices subject to the spatial lag of the income values for each region. This allows one to analyze how the spatial environment affects the transition probabilities of a region through the income distribution. It is usually found that the probabilities of a given region staying in the same class or of moving up one class are ameliorated when the region is surrounded by other wealthy regions.

Spatial autocorrelation must also be considered when analyzing the shapes of per capita GDP distributions and when estimating stochastic kernels. This is done using regional conditioning, that is, basing density function and kernel estimation on a region's income expressed relative to its geographical neighbors. A formal inferential framework to test hypotheses about distribution dynamics in the presence of spatial effects still needs to be developed however.

Regarding the second point, traditional convergence measures can be usefully augmented by the different ESDA and ESTDA measures. First, the classical Moran's  $I$  statistics is naturally used to assess the level of spatial dependence in the income series and its evolution over time.

Second, local measures of spatial autocorrelation can also be used. In particular, local spatial instability is studied by means of the Moran scatterplot, which plots the spatial lag of standardized income against the original values. The four different quadrants of the scatterplot correspond to the four types of local spatial association between a region and its neighbors: HH denotes a region with a high value surrounded by other regions with high values and LH indicates a low value region that is surrounded by regions with high values, etc. Quadrants HH and LL (resp. LH and HL) refer to positive (resp. negative) spatial autocorrelation indicating local spatial clustering of similar (resp. dissimilar) values. This approach has been used extensively to analyze the evolution of the spatial distribution of income in several regional samples. Whenever these Moran scatterplots are constructed for several years, Rey (2001) has suggested using the discrete Markov methodology: in any period, there are four possible states: HH, LL, LH, and HL so that between any two periods, 16 different spatial transitions are possible, which can be summarized by a transition probability matrix.

---

## 16.5 Conclusions

We have reviewed alternative approaches to regional growth and convergence empirics, focusing on the various methodological problems and solutions that have been offered in the literature. Clearly, there is no obvious consensus regarding the most appropriate approach, or modeling strategy, or even whether convergence is actually a real phenomenon, or simply a feature of the theoretical model that has been the dominant one in the literature, the neoclassical growth model.

There is a point at which some of these approaches do give comparable conclusions, however. Indeed, we can obtain estimates of the time it will take for economies to converge both from the neoclassical growth model and from the Markov chain approach. According to Fingleton (1999), for the regions of the EU, the time needed to achieve neoclassical (conditional) convergence will be of the order of 200–300 years. This is simply due to diminishing returns to capital setting in very slowly, that is, effectively  $\alpha$  is close to 1 in the model of Sect. 16.2. Under the Markov model, convergence to the ergodic distribution should, it is estimated, take a similar amount of time, at least 300 years. Of course, the latter is stochastic convergence, implying constant probabilities of different income states but allowing movement of regions across income states. It is evident that, for the EU at least, convergence of some sort, if it occurs at all, will not be a rapid phenomenon and be characterized by distributed income levels rather than the homogeneity associated with unconditional neoclassical convergence.

## References

- Abreu M, De Groot HLF, Florax RJGM (2005) A meta-analysis of  $\beta$ -convergence: the legendary 2 %. *J Econ Surv* 19:389–420
- Arellano M, Bond S (1991) Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Rev Econ Stud* 58:277–297
- Azariadis C, Drazen A (1990) Threshold externalities in economic development. *Q J Econ* 105:501–526
- Barro RJ (1991) Economic growth in a cross section of countries. *Q J Econ* 106:407–443
- Barro RJ, Sala-i-Martin X (1995) *Economic growth theory*. McGraw-Hill, Boston
- Bernard A, Durlauf S (1996) Interpreting tests of the convergence hypothesis. *J Econ* 71:161–173
- Bond S, Hoeffler A, Temple J (2001) GMM estimation of empirical growth models. Centre for Economic Policy Research Discussion Paper No. 3048
- Carree M, Klomp L (1997) Testing the convergence hypothesis: a comment. *Rev Econ Stat* 79:683–686
- Durlauf S, Johnson P (1995) Multiple regimes and cross-country growth behavior. *J Appl Econ* 10:365–384
- Durlauf S, Johnson P, Temple J (2005) Growth econometrics. In: Aghion P, Durlauf S (eds) *Handbook of economic growth*. North-Holland, Amsterdam
- Egger P, Pfaffermayr M (2009) On testing conditional sigma-convergence. *Oxf Bull Econ Stat* 71:453–473
- Ertur C, Koch W (2007) Growth, technological interdependence and spatial externalities: theory and evidence. *J Appl Econom* 22:1023–1062
- Ertur C, Le Gallo J, Baumont C (2006) The European regional convergence process, 1980–1995: do spatial dependence and spatial heterogeneity matter? *Int Reg Sci Rev* 29:2–34
- Fernandez CE, Ley E, Steel M (2001) Model uncertainty in cross-country growth regressions. *J Appl Econ* 16:563–576
- Fingleton B (1999) Estimates of time to economic convergence: an analysis of regions of the European Union. *Int Reg Sci Rev* 22:5–35
- Fingleton B (2008) A generalized method of moments estimator for a spatial panel model with an endogenous spatial lag and spatial moving average errors. *Spat Econ Anal* 3:27–44
- Fingleton B, Lopez-Bazo E (2006) Empirical growth models with spatial effects. *Pap Reg Sci* 85:177–198
- Fischer MM, Stumpner P (2008) Income distribution dynamics and cross-region convergence in Europe. Spatial filtering and novel stochastic kernel representations. *J Geogr Syst* 10:109–139
- Fischer MM, Wang J (2011) *Spatial data analysis: models, methods and techniques*. Springer, Heidelberg/New York
- Henderson D, Parmeter C, Russell R (2008) Modes, weighted modes, and calibrated modes: evidence of clustering using modality tests. *J Appl Econ* 5:607–638
- Islam N (1995) Growth empirics: a panel data approach. *Q J Econ* 110:1127–1170
- Le Gallo J, Fingleton B (2012) Measurement errors in a spatial context. *Reg Sci Urban Econ* 42:114–125
- LeSage JP, Fischer MM (2008) Spatial growth regressions: model specification, estimation and interpretation. *Spat Econ Anal* 3:275–304
- LeSage J, Pace K (2009) *Introduction to spatial econometrics*. CRC Press, Boca Raton
- Levine R, Revelt D (1992) A sensitivity analysis of cross-country growth regressions. *Am Econ Rev* 82:942–963
- Mankiw NG, Romer D, Weil DN (1992) A contribution to the empirics of economic growth. *Q J Econ* 107:407–437
- Nijkamp P, Poot J (1998) Spatial perspectives on new theories of economic growth. *Ann Reg Sci* 32:7–37
- Pace R, LeSage J (2004) Spatial autoregressive local estimation. In: Getis A, Mur J, Zoller H (eds) *Spatial econometrics and spatial statistics*. Palgrave Macmillan, Basingstoke



- Pirotte A (2011) *Econometrics of panel data*. Economica, Paris
- Quah D (1993) Empirical cross-section dynamics in economic growth. *Eur Econ Rev* 37:426–434
- Quah D (1996) Empirics for economic growth and convergence. *Eur Econ Rev* 40:1353–1375
- Rey SJ (2001) Spatial empirics for economic growth and convergence. *Geogr Anal* 33:195–214
- Rey SJ, Dev B (2006)  $\sigma$ -convergence in the presence of spatial effects. *Pap Reg Sci* 85:217–234
- Rey SJ, Le Gallo J (2009) Spatial analysis of economic convergence. In: Mills TC, Patterson K (eds) *Palgrave handbook of econometrics*, vol 2, *Applied econometrics*. Palgrave MacMillan, Basingstoke
- Sala-i-Martin X, Doppelhofer G, Miller R (2004) Determinants of long-term growth: a Bayesian averaging of classical estimates (BACE) approach. *Am Econ Rev* 94:813–835
- Temple J (1998) Robustness tests of the augmented Solow model. *J Appl Econ* 13:361–375