



UNIVERSITÀ DEGLI STUDI DI FERRARA

DIPARTIMENTO DI ECONOMIA ISTITUZIONI TERRITORIO

Corso Ercole d'Este, 144 - 44100 Ferrara

Quaderno n. 3/2006

March 2006

Evaluating total factor productivity differences
by a mapping structure in growth models

Rosa Bernardini Papalia

Silvia Bertarelli

Quaderni deit

Editor: Giovanni Ponti (ponti@economia.unife.it)

Managing Editor: Marisa Sciutti (sciutti@economia.unife.it)

Editorial Board: Giovanni Masino
Simonetta Renga

<http://newdeit.economia.unife.it/quaderno.phtml>

Evaluating total factor productivity differences by a mapping structure in growth models

Rosa Bernardini Papalia* and Silvia Bertarelli**

Abstract: The paper aims at providing a suitable measure of total factor productivity levels by introducing a mapping structure within the regional conditional convergence framework. Our goal is twofold. First, we introduce an endogenous spatial representation of unobserved fixed effects in the conditional convergence process. Second, we develop a generalized maximum entropy estimation procedure in order to improve the efficiency of the estimates and to account for ill-posed and ill-conditioned inference problems. The proposed approach is applied to assess the existence of catching-up across Italian regions over the period 1960-1996 and to identify the effects of technology and geographic spillovers on the determination of TFP levels.

Keywords: conditional convergence, mapping models, dynamic panel data, maximum entropy estimation, human capital, industrial districts.

JEL classification: C130, C210, C230, O180, O470

* University of Bologna, Dipartimento di Scienze Statistiche, e-mail: bernardini@stat.unibo.it.

** Address for correspondence: University of Ferrara, Faculty of Economics, Via del Gregorio 13, 44100 Ferrara, e-mail: silvia.bertarelli@unife.it

1. Introduction

It is generally established that large differences in income levels across countries and regions are mainly due to differences in total factor productivity (TFP henceforth) levels (Klenow and Rodriguez Clare, 1997; Prescott, 1998; Hall and Jones, 1999). Typically, most of these studies show that about half of cross-country per capita income differences are left unexplained after taking into account for differences in physical or human capital. In addition, in recent TFP literature divergence in the form of twin-peaks phenomena with reference to income distribution emerge, and this evidence is usually attributed to cross country divergence in TFP rather than to factor accumulation rates. Feyrer (2003) finds that long run distributions of output per capita and TFP are bimodal while distributions of capital-output ratio and human capital are unimodal.

The persistence of wide differences in total factor productivity levels among regions is empirically related to heterogeneity in technological characteristics as well as in institutions. Geographical proximity effects are also identified. From a theoretical perspective, the efficiency of an economy may be influenced by productivity improvements driven by firms' innovation decisions. Such activities take the form of new varieties (Romer, 1990) or quality improvement of existing products (Aghion and Howitt, 1992, 1998). While the mechanism driving R&D decisions is analyzed in endogenous growth models, incorporating institutions in growth theories is currently dealt with in the literature, but much work must still be done¹. By 'institutions' we mean various aspects of law enforcement, the functioning of markets, inequality and social conflicts, democracy, political stability, government corruption, the health system, financial institutions, etc. Finally, regional studies have identified geographical spillovers and neighborhood effects that may explain uneven regional development. The theoretical explanation of such phenomena is based on models with convergence clubs, where history and geographical characteristics fundamentally condition the growth path. In all cases, the theoretical implications of technological, institutional, and geographic factors point in favor of multiple equilibria².

Since total factor productivity is not directly observable, three main methodologies have been adopted to calculate it. The first approach is based on the estimation of an aggregate production function, where capital and labour appear as inputs. The second approach, labeled 'growth accounting', computes TFP as a residual once the contribution of factors of production to per capita GDP has been taken into account. Most TFP studies based on this approach, however, ignore the possibility of technology spillovers between physical/human capital accumulation and productivity. The third way to obtain TFP relates to the conditional convergence process, as proposed by Islam (1995) and Caselli *et al.* (1996). In the convergence framework, the investment in human capital is considered to evaluate the importance of such variable on the cross-sectional steady-state level of GDP, as well as its role in determining the effects on TFP cumulated through time. Furthermore, we can also evaluate the presence of geographical spillovers, which have been detected by several regional studies. In this case, the uneven distribution of economic activities may contribute either to the uneven development or to growth convergence of neighbor regions.

¹ For a survey, see Acemoglu *et al.* (2005)

² Evidence in favor of convergence clubs and multiple equilibria can be found in Durlauf and Johnson (1995), Bloom *et al.* (2003), Canova (2004) among others. Multiple steady states and convergence clubs emerge in theoretical models with human capital externalities (Azariadis and Drazen, 1990), liquidity constraints in the accumulation of human capital and physical capital (Galor and Zeira, 1993), innovation and implementation decisions (Howitt and Mayer-Foulkes, 2005), and institutional barriers (Acemoglu *et al.*, 2004).

In the latter framework, two relevant issues emerge. First, what is the role of spatial dependence and spatial heterogeneity across regions? Regional data cannot be regarded as independently generated because of the presence of spatial similarities among neighbouring regions. As a consequence, the standard estimation procedures employed in many studies can produce biased estimates (and/or with large variances) of the convergence rate. Second, which is the best estimator for the associated modelling? These are exactly the questions addressed in this paper. First, the paper aims at providing a suitable measure of the unknown TFP levels by introducing a mapping structure within the conditional convergence framework. We investigate differences in TFP levels by a spatial representation of the regional fixed effects in terms of some characteristics suggested by economic theory. The main (spatial) factors which drive differences in TFP levels across regions are identified without imposing an *a priori* spatial structure in the growth model specification. Second, besides traditional dynamic panel estimators, a generalized maximum entropy estimation procedure is developed with the aim of improving the efficiency of the estimates by dealing with ill-posed and ill-conditioned inference problems. Finally, we apply the proposed methodology to estimate a conditional convergence equation with reference to data collected from CRENOS for Italian regions (sample period 1960-1996), with the purpose to evaluate the convergence process, as well as the role of technological spillovers through human capital accumulation and agglomeration spillovers through the geographic distribution of economic activities (districts), both in the determination of steady-state GDP levels and in the accumulation process of TFP levels. Our procedure is used to detect the presence of multiple equilibria and club convergence, and results are compared to the empirical evaluation of TFP differences proposed by other authors.

The paper is organized as follows. Section 2 introduces the neoclassical growth model deriving the typical conditional convergence equation. Section 3 presents the econometric conditional β -convergence equation, which includes both time and country-specific effects. The approach used to obtain a location map for unobservable variables is also illustrated. Section 4.1 presents estimation issues connected to single cross-section, dynamic panel, and system of SUR equations approaches. A generalized maximum entropy estimation procedure is also developed and discussed in section 4.2. Data description and results are reported in sections 5. The convergence process and mapping analysis are presented in section 5.1 and 5.2. Finally, section 6 concludes and lists some potential advantages and investigations of the proposed approach.

2. The neoclassical growth model

The existence of a negative relationship between the initial GDP per capita/per worker and subsequent growth is a phenomenon, called *b*-convergence³, largely documented in the empirical literature with reference to both cross-country and cross-region analysis. If convergence derives by physical and human capital accumulation (that is capital deepening), initial capital-poor regions have higher marginal productivity of capital, hence faster growth than rich regions. This view is grounded by the Solow (1956) neoclassical model and its extended version by Mankiw, Romer and Weil (1992) or by endogenous growth models that display transitional dynamics, such as the two-sector growth models of Uzawa (1965) and Lucas (1988). However, the fact that poor countries grow faster than rich countries may be also (or only) the effect produced by a process of technological diffusion⁴. Another class of

³ For a complete survey see, for example, Barro and Sala-i-Martin (2004).

⁴ Caselli and Tenreyro (2004) suggest two further factors explaining the existence of per capita income convergence or divergence. The growth process may be linked with structural transformation, where resources move from low-productivity to high-productivity sectors (Lewis, 1954, Imbs and Wacziarg, 2003).

endogenous growth models motivates this second hypothesis, where imitation is less costly than innovation, so that regions initially behind the technology frontier experience faster improvements in technology than the leaders (Abramovitz, 1986; Barro and Sala-i-Martin, 1997; Howitt, 2000).

The theoretical discussion on conditional convergence should be based on a growth equation, which derives from the transitional dynamics of both Solow-type neoclassical growth models and technology catching-up formulations. However, we concentrate only on the former approach, in which we assume TFP levels are different across regions but stationary. The specific formulation we use is the growth model by Cass (1965) and Koopmans (1965).

We are interested in finding $\{c_t, k_{t+1}\}_{t=0}^{\infty}$ so as to maximize utility:

$$\max U_0 = \sum_{t=0}^{\infty} b^t u(c_t), \quad (1)$$

subject to the resource constraint of the economy, and taking initial k_0 as given:

$$\begin{aligned} c_t + k_{t+1} &\leq (1-d-n)k_t + f(k_t) \\ c_t &\geq 0, \quad k_{t+1} \geq 0, \\ k_0 &> 0 \quad \text{given} \end{aligned} \quad (2)$$

where c_t is consumption, b is the discount factor in the utility function, $0 < b < 1$, $u(\cdot)$ is a well-behaved utility function; $f = Ak^a$ is a well-behaved production function expressed in intensive form ($f = F/AL$), where A states the level of technology, L is labor, and k_t is capital, which is assumed to depreciate at a constant rate d , $0 < d < 1$. Labor supply grows at a constant rate n . By solving the Bellman equation associated with the dynamic programming (DP) problem we find the steady state capital, k^* . The DP problem is the following:

$$\max V(k) = u(c) + bV(k'), \quad (3)$$

subject to:

$$c + k' \leq (1-d)k + f(k). \quad (4)$$

There is a unique V that solves the Bellman equation, which is continuous, differentiable, and strictly concave. Solving the Lagrangian for the DP problem and using the envelope condition we find the Euler equation:

$$\frac{u'(c_t)}{u'(c_{t+1})} = b[1-d-n+f'(k_{t+1})], \quad (5)$$

We may solve equation (5) for c_t by restricting preferences to a specific formulation of the utility function. Otherwise, consumption can be written as a fraction $(1-s)$ of the output:

$$c_t = (1-s(b, d, n))f(k_t), \quad (6)$$

where the propensity to save $s(\cdot)$ depends on the structural parameters of the model. Given (4) it follows that:

$$k_{t+1} - k_t = s(b, d, n)f(k_t) - (d+n)k_t. \quad (7)$$

Equivalently, the growth rate of capital is given by:

$$\frac{k_{t+1} - k_t}{k_t} = g_k = s(b, d, n) \frac{f(k_t)}{k_t} - (d+n) \quad (8)$$

Alternatively, the convergence process is affected by the presence of comparative advantages in integrated trading countries. These two hypotheses are empirically evaluated, in comparison with the capital deepening and the technological catching-up views, by Caselli and Tenreyro (2004) and Bernardini Papalia and Bertarelli (2005), in a different framework.

A steady state equilibrium of the economy exists if and only if $g_k = 0$. We indicate the steady state value of capital with k^* .

By writing the growth equation as a linearization around the steady state k^* we find that:

$$g_k = I(\ln k - \ln k^*), \quad (9)$$

where $I = -(1-e_k)(d+n)$ is the speed of convergence, and e_k is the elasticity of production with respect to capital, evaluated at the steady state k^* ⁵. The same relation is easily obtained for income (in efficiency units):

$$g_y = I(\ln y - \ln y^*), \quad (10)$$

or taking logs:

$$\Delta \ln y = I \ln y - I \ln y^*. \quad (11)$$

A negative value for I reflects local diminishing returns. Such local diminishing returns occur in endogenous-growth models, as well as in the Solow model (Solow, 1956).

3. Heterogeneity and mapping in the convergence equation

The empirical analysis of conditional convergence requires a regression specification, which is formally derived from the neoclassical growth model we have presented in the previous section.

The first step requires the approximation of the steady state output y^* by a set of country-specific controls X , which include s , d , n , etc. That is, let:

$$-I \ln y^* \approx g'X \quad (12)$$

we conclude from equation (10):

$$\Delta \ln y = I \ln y + g'X + \text{error}. \quad (13)$$

Specifically, we may write the following dynamic equation:

$$\ln y_{it} = (1+I) \ln y_{it-1} + g'X_{it-1} + m_i + u_i + u_{it}. \quad (14)$$

The above equation represents a typical conditional convergence regression, where the parameter I captures the negative correlation between the initial income level and the subsequent growth rate (traditionally called b -convergence)⁶. A negative value for I reflects local diminishing returns, which occur both in endogenous and in exogenous growth models. The vector $X_{it-1} = (sk_{it-1}, sh_{it-1}, ndx_{it-1})$ gives the determinants of the steady state output and consists of a set of region-specific explanatory variables suggested by the theory, including sk_{it-1} the saving rate in physical capital, and ndx_{it-1} the sum of the population growth rate, the exogenous technological growth rate, and the depreciation rate. In addition, we also consider sh_{it-1} , the investment rate in human capital, by following the extended version by Mankiw, Romer and Weil (1992). Time effects control for the presence of a time trend component and of a common stochastic trend (the common component of technology). Individual effects capture total factor productivity (TFP) differences.

Traditionally, strong restrictions have been imposed to represent unobserved heterogeneity. In most of the cases, such unobserved components are estimated using either fixed effects methods or random effects methods that do not require data on some specific variables. Another approach has been proposed by Chamberlain (1984), which introduces a

⁵ The model may be extended to consider the accumulation of human capital, as in Mankiw *et al.* (1992). In this framework, the convergence rate depends on the elasticity of output with respect to physical and to human capital.

⁶ In SUR specification complete heterogeneity across regions is assumed and $(1+I)$ and g in eq. (14) become $(1+I)$ and g_n respectively.

regressor as a proxy for unobserved heterogeneity correlated with observable variables. More specifically, the individual-effect term is specified as a function of the explanatory variables, such as the mean of the exogenous variable pertaining to the individual (Mundlak, 1978) or, less restrictively, several lags of the individual exogenous variables (Chamberlain, 1984). Differently, a spatial structure to the variance-covariance matrix may be imposed to model phenomena such as capital flows, labor migration or technological spillovers in the regional growth process.

Our approach aims at investigating regional TFP heterogeneity in cross-country (cross-regions) analysis by means of a mapping representation. In this vein, the idea is to identify the role and the weight of unobserved variables connected to the quality of institutions, and other variables, which are relevant in regional technological and structural characteristics. A mapping model defines the spatial position of a region's TFP level in terms of different unobserved dimensions weighted by the variables' features. In Elrod (1988), a market map has been used as a spatial representation of products being sold by firms, in which the distance between two brands represents a measure of quality differentiation. Holbrook *et al.* (1982) improve these spatial representations by consumer-preference information in the form of points that indicate their relative preferences for products. Specifically, these mapping models are derived from the neoclassical consumption theory and incorporate information about consumers' brand and quality choice decisions.

The variability in both cross-region specific unobserved characteristics and time invariant components is considered and, as in the choice map, the position of the unobserved variables on the M-dimensional map and the country's importance weights for these dimensions are derived. In this framework, the interpretation of the dimensions of maps is aimed at endogenously identifying the determinants of technological and structural differences. The resulting location map can be obtained by using a two-stage process. First, the parameters of the growth model (eq. 14) are estimated and the covariance matrix of unobserved components m_i 's is computed. Then this matrix is used as an input in multidimensional scaling to obtain their locations in a multi-attribute space. As in the choice map representation we assume that the time-invariant effect for region i , m_i , is a linear function of the region's time invariant attributes which lie within a two-dimension map, such as:

$$m_i = w_1 z_{1i} + w_2 z_{2i} + x_i \quad (15)$$

where the parameters w_1 and w_2 are modeled as a function of country's characteristics, (z_{1i} , z_{2i}) are the coordinates representing the location (to be estimated) of the unobserved effect on the map, and x_i is a random error with zero mean.

It should be noted that this approach presents the advantage over the more traditional approaches to simultaneously identify the main spatial factors that provide an indirect measure of unknown invariant TFP (dis)similarities across regions, without imposing a priori spatial structure on the growth model.

4. Methodology and estimation problems

4.1 Parametric and semi-parametric estimation approaches

The empirical analysis on regional b -convergence are almost based on cross-sectional regressions or panel data methods. The single cross-section approach to convergence analysis considers the behavior of the output differences between regional economies over a fixed time interval and uses average data for long periods of 25 or 30 years (Barro, 1991; Levine and Renelt, 1992; Sala-i-Martin, 1997). Most empirical formulations of cross section convergence tests are based on more general models, which also (apart from investment and population growth rates) include a set of socio-economic variables to control for differences

in steady states and asymmetric shocks. These variables are either initial values or average values over the time period. This approach present several problems. First, reducing the time series to a single (average) observation means that not all available information is used; second, an omitted variable bias due to unobserved differences across countries and regions can be produced; third, one or more of the regressors may be endogenous. Finally, it imposes the strong hypothesis of complete regional homogeneity in the parameters of the process that describes the evolution of per-capita GDP.

A dynamic panel data analysis has been introduced to address omitted variable and/or endogeneity issues and alternative estimators, which address the bias problems of the single cross-section regressions, are proposed. An important advantage of panel data analysis is that it controls for unobservable or unspecified differences between regions. The assumptions of correlation between the errors and the explanatory variables of the model determine the appropriate estimator for panel regressions. The ways of controlling for unobservable omitted variables depend upon whether a fixed or random effects model is used. A fixed effects estimator is typically chosen when unobservable region-specific and/or time-specific effects are assumed to be correlated with the observed explanatory variables. The within-region transformation of the fixed effects estimator, applied by including region-specific and/or time-specific dummy variables, eliminate such effects thereby lessening the possibility of simultaneity bias in the coefficients on the observed variables. A random-effects estimator is typically chosen when unobservable region-specific effects are assumed to be uncorrelated with the observed explanatory variables. However, there could remain residual specification problems, arising from simultaneity, mis-specified dynamics and/or measurement errors. This effect is produced when there exists correlation of the error term over time. Techniques are being developed to correct for this source of bias. Within this framework, both the Arellano-Bond (1991) first differenced generalized method of moments (GMM) estimator, and the Blundell and Bond (1998) system generalized method of moments (SYS-GMM) estimator are able to account for unobserved specific effects and to allow for the endogeneity of the regressors. Arellano and Bond (1991) proposed a way of obtaining more efficient estimators once the model is differentiated, by using all the orthogonality conditions that exist between lagged values of the endogenous variables and the disturbances in the model. They derived a Generalized Method of Moments estimator using as instruments lagged levels of the dependent variable, the predetermined variables, and also differences of the strictly exogenous variables. The GMM estimator provides consistent estimates but, when the series are close to a random walk, the instruments are only weakly correlated with the endogenous variables and the GMM estimator suffers from serious finite sample bias. The Blundell-Bond estimator (1998, SYS-GMM) is derived by adding the original equations in levels to the system. With these additional moment conditions it is possible to increase efficiency. The SYS-GMM estimator provides consistent coefficients estimates when the measured region characteristics are all uncorrelated with the unobserved region effects. When it is possible to utilize prior information, a Local Generalized Method of Moments, LGMM, (Bernardini Papalia, 1999) can represent a suitable alternative estimation procedure. The main advantage of the LGMM method is that it is possible to introduce prior information in the form of a parametric model without assuming a parametric form for the unknown regression function. Another approach to control for the Least Squares with Dummy Variable (LSDV) finite sample bias, introduced by Kiviet (1995) proceeds by estimating a small sample correction to the LSDV estimator. Monte Carlo studies show that for small T KIVIET estimates seem to be more attractive than GMM.

In a dynamic panel data context where the number of time series observations (T) is relatively large or of the same order of magnitude as the number of regions (N), there are two procedure commonly used: the Mean Group (MG) and the Pooled Mean Group (PMG)

estimators (Pesaran *et al.*, 1998, Pesaran and Smith, 1995). The MG approach consists of estimating separate regressions for each region and calculating the coefficients means. In order to obtain a gain in efficiency, the PMG estimator imposes homogeneity on long-run coefficients but allows the short run coefficients and error variances to differ across groups. The PMG approach is inefficient in samples with a small number of individuals, since any individual outlier may severely influence average coefficients.

To take into account the possibility of cross-sectional correlations, the regional relationships can be treated as a system of seemingly uncorrelated regression equations (SUR) (Andrés *et al.*, 2004). Nevertheless, in both SUR and panel estimation, it is possible to allow for some restricted serial correlation in the error process (Liang and Zeger, 1986, Keane and Runkle, 1992). If one relaxes the assumption that the error variances are equal, one can weight observations accordingly to obtain weighted least squares estimators as in Fisher (1993). For variance-covariance matrix of errors non-diagonal, one possibility is to collapse some of the cross-sectional units into group averages to make $N < T$ or a further possibility is to replace the required inverse of variance-covariance matrix by its generalized inverse. Recently, some empirical studies have used the spatial econometric framework for testing regional convergence in presence of spatial dependence as for the presence of similarities among neighboring regions (Meliciani and Peracchi, 2004, Brasili, 2005). A semi-parametric spatial-covariance model of regional growth behavior is proposed by Basile and Gress (2005) with the aim of simultaneously taking account the problems of non-linearity (presence of multiple regimes) and of spatial dependence. The concept of multiple regimes is based on endogenous growth models characterized by the possibility of multiple, locally stable, steady state equilibria. The basic idea underlying these models is that the level of per-capita GDP on which each economy converges depends on some initial conditions and that, according to these characteristics, some economies converge to one level and others converge to another (Liu and Stengos, 1999).

Convergence study may also be implemented within a time series approach. Bernard and Durlauf (1995) define time series convergence in output in two countries (regions) to be the equality of their long run output forecasts. Convergence is tested by looking for unit roots or deterministic trends in the deviation of per capita output from a reference economy or from the sample average. See, among others, Evans and Karras (1996 a, b), Bernard and Durlauf (1995, 1996).

4.2. A Maximum Entropy-based approach

In this section we propose a generalized maximum entropy (GME) estimation approach that retains the flexibility of the SUR approach in allowing for correlated shocks across regions and that can be implemented when the T is not sufficiently large, and in particular if $N < T$ and/or in presence of small samples.

The entropy-based estimation procedure shares some of the characteristics of Stein Rule estimators and Bayesian approaches to estimation (see Judge *et al.*, 1988 and Zellner, 1996, 1997, 1999). There is now a considerable body of work, which applies the entropy criterion to a wide class of models (Marsh *et al.* 1988, Golan *et al.* 1996, 1997). As regards traditional estimation techniques, the formulation of the constrained maximization problem in the maximum entropy view does not require: (i) the use of restrictive parametric assumptions on the model; (ii) the formulation of hypotheses regarding the form of the distribution of the objective variables. Restrictions expressed in terms of inequality can be introduced and it is possible to calibrate the precision in the estimation. Good results are produced in the case of small-sized samples, in the presence of high numbers of explanatory parameters and variables (highly correlated). By means of a relative measure of uncertainty (Normalized Entropy) different cases can be compared, verifying the informative contribution of every restriction

introduced. The extra-sample information is used to define both the range of variation of the parameter values and the restrictions to be introduced into the optimization phase of the estimation procedure.

We start by considering the maximum entropy formulation relative to a seemingly unrelated system of N equations which allows for covariance between the disturbances across different regions where the i -th model (equation) is given by:

$$Y_i = X_i b_i + e_i \quad (16)$$

for $i=1, \dots, N$, where y_i and ε_i are of dimension $(T \times 1)$, X_i is $(T \times K_i)$ and β_i is $(K_i \times 1)$. Here $Y_i = \ln Y_i$ and X_i includes sk_i, sh_i, ndx_i and $\ln y_{i,t-1}$.

Stacking all the equations, the system approach considers the N equations, in the form (1), for each region as:

$$Y = Xb + e \quad (17)$$

$$\text{where } Y \equiv \begin{bmatrix} y_1 \\ \dots \\ y_N \end{bmatrix} \quad X \equiv \begin{bmatrix} X_1 & 0 & \dots & 0 \\ 0 & X_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & X_N \end{bmatrix}$$

$$b \equiv \begin{bmatrix} b_1 \\ \dots \\ b_N \end{bmatrix} \quad e \equiv \begin{bmatrix} e_1 \\ \dots \\ e_N \end{bmatrix} \quad \text{with } i = 1, \dots, N \quad k = 1, \dots, K$$

where Y and ε are each of dimension $(NT \times 1)$, X is a block diagonal matrix of dimension $(NT \times K)$ with $K = \sum K_i$ and $\beta = (\beta_1, \beta_2, \dots, \beta_N)'$ is an unknown vector of dimension $(NK \times 1)$.

Under the GME framework we recover simultaneously the unknown parameters β , the unknown errors by defining an inverse problem, which is based only on indirect, partial or incomplete information. We also assume that the equation error are contemporaneously correlated but uncorrelated over time. Consequently, the covariance matrix for ε may be written as:

$$\Phi = \Sigma \otimes I_T \quad (18)$$

where Σ is an $(N \times N)$ positive definite symmetric matrix, \otimes is the Kronecker product operator and I_T is an identity matrix of dimension T .

The GME estimator is a more general version of the ME in which no weight is placed on the noisy component. It is obtained by maximizing the sum of the entropy corresponding to the probabilities of parameter, and the entropy from the noisy probabilities, subject to data consistency conditions and adding up constraints on the probabilities. It should be noted that, unlike ML estimators, the Generalized Maximum Entropy (GME) approach does not require explicit error distributional assumptions.

In the GME estimation the objective is to recover probability distributions for unknown parameters and errors. Each parameter is treated as a discrete random variable with a compact support and M possible outcomes, $2 \leq M \leq \infty$. The uncertainty about the outcome of the error

process is represented by treating each error as a finite and discrete random variable with J possible outcomes, $2 \leq J \leq \infty$. To this end, we start by choosing a set of discrete points, the support space $v = [v_1, v_2, \dots, v_M]'$ of dimension $M \geq 2$, that are at uniform intervals and symmetric around zero. Each error term has corresponding unknown weights $w_j = [w_{j1}, w_{j2}, \dots, w_{jM}]'$ that have the properties of probabilities $0 \leq w_{jm} \leq 1$ and $\sum_m w_{jm} = 1$ ⁷. The choice of the support spaces represents a way to impose prior restrictions on the parameter estimates.

Re-parameterizing the set of equations (17), so that $\beta = Zp$ and $\varepsilon = Vw$, yields:

$$\begin{bmatrix} y_1 \\ \dots \\ y_N \end{bmatrix} = \begin{bmatrix} X_1 & & \\ & X_2 & \\ & & \dots \\ & & & X_N \end{bmatrix} \begin{bmatrix} Z_1 p_1 \\ \dots \\ Z_N p_N \end{bmatrix} + \begin{bmatrix} V_1 w_1 \\ \dots \\ V_N w_N \end{bmatrix} \quad (19)$$

where $p = (p_1, p_2, \dots, p_K)'$ and $w = (w_1, w_2, \dots, w_N)'$ are the unknown signal and noise probabilities we wish to recover, and Z, V_1, V_2, \dots, V_N are the corresponding parameter supports for β and ε as previously defined.

Given the data consistency (19) and the covariance's relationship (18) the GME objective function relative to our formulation problem may be formulated as:

$$\max_{p_i, w_i} H(p, w) = -p' \ln p - w' \ln w \quad (20)$$

for $i = 1, 2, \dots, N$.

subject to:

i) data consistency conditions:

$$\begin{bmatrix} y_1 \\ \dots \\ y_N \end{bmatrix} = \begin{bmatrix} X_1 & & \\ & X_2 & \\ & & \dots \\ & & & X_N \end{bmatrix} \begin{bmatrix} Z_1 p_1 \\ \dots \\ Z_N p_N \end{bmatrix} + \begin{bmatrix} V_1 w_1 \\ \dots \\ V_N w_N \end{bmatrix} \quad (21)$$

(ii) adding-up constraints:

$$1_{K_i} (1_{K_i} \otimes 1'_M) p_i, \quad \text{for } i = 1, 2, \dots, N; K = \sum_i K_i; \quad (22)$$

$$1_{T_i} (1_{T_i} \otimes 1'_J) w_i, \quad \text{for } i = 1, 2, \dots, N; T = \sum_i T_i.$$

where $p = (p_1, p_2, \dots, p_K)'$ and $w = (w_1, w_2, \dots, w_N)'$.

The solution to the system of equations related to the first-order condition produce the following point estimates:

⁷ By increasing the weight on the ε component of entropy, we improve the accuracy of the estimation (decrease the mean square error of the estimates of p), while by increasing the weight on the p component of entropy we improve the predictive power.

$$b_k = \sum_{m=1}^M p_{km} z_m \quad k = 1, \dots, K \quad \text{and} \quad \varepsilon_i = \sum_{j=1}^J w_{ij} v_j \quad i = 1, \dots, N. \quad (23)$$

If there exists additional non-sample information from theory and/or empirical evidence, over that contained in the consistency and adding-up constraints, for the p probabilities, it may be introduced in the form of known q probabilities, by means of the cross-entropy formalism (Kullback, 1959).

Letting q and u be the prior probabilities for p and w respectively, the generalized cross entropy, GCE, formulation is given by:

$$\min_{p_i, w_i} H(p, w) = p' \ln \frac{p}{q} + w' \ln \frac{w}{u} \quad (24)$$

for $i = 1, 2, \dots, N$.

subject to:

- i) data consistency conditions (21),
- ii) adding-up constraints (22),

where $p = (p_1, p_2, \dots, p_K)'$ and $w = (w_1, w_2, \dots, w_N)'$.

The solution p in (24) is a function of prior information, the data and a normalization factor. If the q s are specified such that each of the choices is equally likely to be selected (uniform distributions), the GCE result reduces to the GME result. To allow the possibility of non-zero covariances for errors it is possible to specify, within the GME and GCE formulations, an additional set of restrictions which is based on a particular error covariance structure and incorporates the known *a priori* information of contemporaneous correlations among the disturbance terms in the equations of the system (Bernardini Papalia, 2002).

For the empirical analysis on regional convergence two points of interest should be noted here. First, the system estimation approach facilitates testing of hypotheses involving cross equation restrictions such as testing the equality of total factor productivities in two neighboring regions. Second, by using the GME/GCE procedure it is possible to derive an estimator even if the number of regions involved, N , is more than the number of time periods, T , and the corresponding variance-covariance matrix for errors is singular.

5. Data and results

The analysis of conditional convergence across Italian regions is based on CRENOS data set covering the period 1960-1996, and TFP levels are calculated using the methodology proposed by Islam (1995) and Caselli *et al.* (1996). This approach allows for differences in individual technologies, but assumes that such differences are stationary, so that technology catching-up is ruled out by assumption⁸. This hypothesis is supported by Bianchi and Menegatti (2005), who assess the relative importance of capital deepening rather than technology diffusion processes. Indeed, the authors find evidence of a strong impact of the former component in explaining the convergence process across Italian regions.

Since we are interested in evaluating the effects of technology and geographical spillovers, we introduce two variables to take account of such effects. The first class of spillovers is connected to the possibility of imitation processes from some countries/regions, which perform R&D in favor of other countries/regions, which take advantage of such technology

⁸ The same methodology can be extended to analyse cases in which TFP differences are not stationary, by estimating the convergence equation over some sub-periods, as proposed by Di Liberto *et al.* (2004). Alternatively, we could compute TFP levels without imposing stationarity within the 'growth accounting' approach by using the methodology proposed by Dowrick and Nguyen (1989) and Benhabib and Spiegel (1994).

improvements. The diffusion of technology is conditioned by some characteristics of the follower, such as the degree of trade liberalization, social, legal and geographic factors, and most importantly the stock of human capital. Specifically, the higher the accumulation of human capital, the higher are TFP levels. In this vein, the investment in human capital is considered to evaluate its role in determining the (dynamic) effects on the accumulation of TFP, as well as the importance of such variable on the cross-sectional (static) steady-state level of GDP⁹.

Geographical spillovers refer to positive knowledge external effects produced by some located firms and affecting the production process of firms located elsewhere. In this case, the uneven distribution of economic activities may contribute to the uneven development only if the geographical spillovers are local. If instead they are global such uneven distribution may contribute to growth convergence, because the knowledge accumulation in rich regions improves productivity of all the firms wherever they are located. In the case of Italy, production activities are mainly concentrated in districts. As a consequence, we try to capture the effects of the distribution of economic activities on the steady state value of GDP and on the dynamic behavior of TFP both in terms of convergence or divergence. As Baumont *et al.* (2000) observe, if the geographic distribution of rich and poor regions is rather stable through time, it is plausible to justify this evidence on the ground of the cumulative nature of both growth and agglomeration processes and by the fact that history matters through initial conditions.

In what follows, we use per capita GDP and other economic aggregates at constant 1990 prices. Real per capita GDP is calculated as a ratio of real GDP and population, the saving rate in physical capital is given by the ratio of investment and GDP, and the investment rate in human capital is the ratio of enrollment in secondary school and population of age 14-19. Moreover, we add to the population growth rate, a constant value of 0.05 to take account of the exogenous technological growth rate and the depreciation rate. The district variable is a measure of the local degree of industrial district diffusion over the Italian regions; specifically, we consider the relative number of district identified by ISTAT in a region over the total number of Italian districts in 1991. All final data are expressed in logs and are calculated as 5-year averages to eliminate the business cycle component¹⁰.

In our basic specification we assume (i) complete regional homogeneity in the parameters of the conditional convergence model by using pooled OLS; (ii) heterogeneous intercepts to accommodate level effects across regions, with LSDV and system GMM estimators; (iii) we then proceed by assuming complete heterogeneity in the coefficients of the production function and in the rate of technological progress, by treating the relationship which describes the evolution of regional per-capita GDP as a system of seemingly unrelated regression equations (SUR-GME). In the latter case, cross-sectional correlation is taken into account. Problems connected to collinearity of regressors and to the singularity of the error covariance matrix emerged when the system of SUR equations specification has been adopted. Therefore we have employed the generalized maximum entropy approach, developed in section 4.2, and have taken averages of regional coefficients (MG-GME).

As a first step of the analysis, we consider the convergence equation without the district variable and compare results with and without human capital obtained by pooled OLS, LSDV, and system GMM estimators (see Table 1). Results obtained with the generalized maximum entropy approach are summarized for all regions in table 2, while average values of

⁹ The theoretical model considering the accumulation of human capital in the Solovian framework is developed by Mankiw *et al.* (1992). Aiyar and Feyrer (2002) find that human capital plays a substantial role in determining the dynamic path of TFP, for a sample of 86 heterogeneous countries over the period 1960-1990.

¹⁰ Other studies have taken averages over 5-year periods, like Islam (1995) and Caselli *et al.* (1996) among others.

regional coefficients (MG-GME) are reported with other estimators in table 1. Among the regressors, only the coefficients on the lagged dependent variable and on the population growth rate are significant, while both the coefficient on the physical capital and human capital investment shares are not always significant¹¹. More specifically, the system GMM coefficient estimates of y_{t-1} are not comprised between the LSDV and the pooled OLS estimates, showing a possible bias according to the procedure suggested by Bond *et al.* (2001)¹². As for the population growth rate ndx we find a negative sign consistently with theoretical expectations. Even if it is not always significant, we find negative signs for the estimated saving propensity variable sk . This result is consistent with other evidence on Italian regions (Carmeci and Mauro, 2004) while it is not supported by theory. As to human capital, our results display the most common feature that can be found in this literature. Enrollment rates lose their significance when the individual effects are incorporated into the convergence equation¹³. However, the presence of the human capital variable improves the explanation of the model, by reducing TFP heterogeneity across regions, as documented by mapping graphs in section 5.1. Furthermore, we observe that neglecting the role of human capital differentials, one produces estimates of the convergence rates for some northern Italian regions (Val D'Aosta, Lombardia, Trentino Alto-Adige, Veneto, Liguria), which are not consistent with theoretical expectations. The inclusion of the human capital variable produces for these northern regions a speed of convergence lower than the mean rate of convergence, which is consistent with the theoretical expectation. In this case, we can conclude that human capital has not static effects on GDP levels.¹⁴

At the second step of the analysis the contribution of agglomeration economies within industrial districts is evaluated by means of the district variable. Results are summarized/reported in table 4. In this case, with the introduction of the district variable we find no (or negligible) effects on the steady-state level of GDP (compare tables 1 and 4), so we can state that industrial districts has not static effects on GDP levels.

5.1 Evaluation of the convergence process for Italian regions

We confirm conditional convergence across Italian regions with heterogeneity. By considering different estimators, two main results emerge. First, we find conditional convergence for all estimators considered. The estimated speed of convergence is biased downwards assuming complete homogeneity across actually different regions, that is the convergence rate computed in the pooled OLS framework is too slow (0.014 without human capital, 0.017 with human capital). The LSDV estimate is biased upward, since the endogeneity issue is not dealt with. Such problems may be overcome by using system-GMM and GME estimators. However, while the former approach display very low values of the convergence rate (0.005-0.009), the GME approach give an average value of 0.12 without human capital (0.11 with human capital). This is in line with other papers, which compare several techniques differing in the degree of homogeneity of parameters across countries/regions. When parameter heterogeneity increases, higher values of the speed of

¹¹ The lagged value of per capita GDP may be a proxy of either past capital stocks or past technology levels. As noted by Bianchi and Menegatti (2005), for Italian regions the prevalence of the capital deepening effect over the technology catching up hypothesis excludes the latter possibility.

¹² It is well known that OLS is biased upwards and LSDV is biased downwards in dynamic panels as reported in section 4. These authors suggest that a consistent estimate should lie between the two.

¹³ In Caselli *et al.* (1996) the human capital variable is even negative and significant. The robustness of human capital in growth equations has been extensively discussed in recent years (see for example De la Fuente and Doménech, 2001).

¹⁴ Static and dynamic effects of human capital have been defined by Aiyar and Feyrer (2002).

convergence are obtained. The highest level of the speed of convergence is located in Southern Italy (as expected): Puglia (11.9%) in the regression equation without district, Basilicata (14.6%) with district variable. The lowest level is calculated for a Northern region: Emilia-Romagna (8.9%) without district, Trentino Alto Adige (7.6%) with district variable.

Evidence of persistent gaps between regional GDP levels feed the debate on the existence of an economic dualism, the so-called 'Mezzogiorno issue'. Indeed, only a partial convergence process has been detected by previous studies, given that the level of per capita/per worker GDP of the poorest region is still below that of the richest one. Barro and Sala-i-Martin (1991) find, during the period 1950-85, a slow (around 2%) absolute convergence process for per worker GDP. Subsequent studies show the absolute convergence process only until the mid seventies (Galli and Onado, 1990; Faini, Galli and Giannini, 1992; Mauro and Podrecca, 1994; Paci and Pigliaru, 1995; Di Liberto, 1994; Bianchi and Menegatti, 1997; Paci and Saba, 1998) and the persistence of economic dualism. In addition, Paci and Pigliaru (1995) find evidence of conditional divergence for the Italian regional data (sample period 1970-1989) within a cross section approach. However, the presence of a conditional convergence process is quite accepted when heterogeneity among regions is assumed (Di Liberto, 1994; Bianchi and Menegatti, 1997).

Our evidence may be also compared to the empirical evaluation of TFP differences proposed by Leonida *et al.* (2004), in a growth accounting approach. Other studies deal with the calculation of TFP levels, within an aggregate production approach, such as Ascari and Di Cosmo (2005), among others. The authors identify several determinants of TFP, in terms of economic and institutional factors. Finally, our mapping analysis is linked to regional studies where spatial correlation among Italian regions is modelled exogenously (Arbia *et al.*, 2005).

Our findings are confirmed by Andrés *et al.* (2004), who directly compare the estimated rates of convergence for OECD countries in cross-section, fixed effects, and mean group regressions. The authors interpret such puzzle in the following way. If the heterogeneity is well modelled by differences in the intercept (system GMM) all regions approach a 'unique' steady state at low rates. If the real world is much more heterogeneous, so that a MG-GME approach should be preferred, data points in favour of multiple steady state equilibria, with very high convergence rates and therefore a limited transitional dynamics. This observation points to our second result: Italian regions seem to be located in a multiple equilibrium regime. We will show it in next section.

5.2 Unique vs. multiple steady states

In the previous section, we have concluded that the heterogeneity of the results obtained with different estimators points to two alternative models. The first one is characterized by a unique steady state and regions converge at low rates. The second one comprises multiple steady states with higher rates of convergence. Mapping analysis points in favour of a multiple regime hypothesis, so we give empirical support to growth models with multiple equilibria and club convergence, in accordance with other studies focusing on Italian regions (Di Liberto and Symons, 1998; Mauro and Podrecca, 1994; Cellini and Scorcu, 1995).

In our analysis, regional effects are an indirect measure of TFP levels. With reference to the theoretical model of section 2, TFP is given by the level of technology A (see Eq. 2) and the regional rate of convergence depends on inputs parameters and on population growth rate. The level of technology A does not influence it. However, the dimensions of the map indicate groups of regions with 'homogeneous' speed of convergence after taking account of differences in human capital and of the uneven distribution of economic activities. Specifically, the richest regions of Northern Italy attain the highest growth convergence rates, while Southern regions experienced the lowest speeds of convergence (see table 5). This

finding is obtained without imposing any *a priori* restriction on data. In addition, we find geographic localisation to have a prevailing role. Maps show two separated groups of regions, located in the North-Center and South of Italy with a bi-modal distribution of real per capita GDP (sigma convergence). Besides the result of two different regimes, we also observe that convergence clubs are spatially concentrated.

With the introduction of the human capital variable, even though human capital has not static effects on GDP levels, it plays a substantial role in determining the cumulated path of TFP. Moreover, with the introduction of the district variable, the (dynamic) effects on TFP can be detected by comparing maps with and without districts (figures 1 and 2). In general, mapping analysis suggest that spillovers are spatially concentrated and confirm the presence of a spatial dependence related to geographical and non-geographical spillovers effects, as argued by Paci and Pigliaru (1995). Non-observable heterogeneity identified by mapping analysis may also be imputed to differences in institutional factors and qualitative elements, such as the efficiency of human capital and infrastructure, for which an appropriate measure is difficult to obtain.

The need for a third dimension is not emerged by the values of the Stress indicator and the map of figure 2 seems to better represent the group-structure in two dimensions. The first dimension, the vertical axe in figure 2, can be interpreted as a separation between the Northern-Central and Southern regions, apart for Lazio. The second dimension (see below the horizontal axe in figure 2) contributes to isolate the regions that present some anomalies (Sicilia, Basilicata, Val D'Aosta, and Lazio), showing the relevance of the geographical location of a region as a source of (dis)advances in a context where economic activity is not homogeneously distributed in space but it is concentrated in some areas. The second dimension shows a clear separation of some Italian regions that present remarkable differences and/or anomalies that can be interpreted in terms of: (i) growth rates estimates, as for the Lazio case (C12) which shows some strong differences in the growth rates relative to the model specification with and without the human capital variable (see table 3); (ii) a mix of differences in the local conditions that may include public and social institutions, transport infrastructure, local expenditure in research and development etc., as for the Val D'Aosta (C2), Basilicata (C17) and Sicilia (C19) regions.

6. Remarks and conclusions

This paper evaluates unobservable total factor productivity levels admitting heterogeneity across countries and regions, within the framework of conditional convergence processes. Two elements of novelty have been introduced. The first one refers to the strategy of estimating unobserved TFP differentials across regions. The second one refers to the GME estimation procedure that has been proposed to deal with ill-posed and ill-conditioned inference problems in analyzing conditional convergence processes across regions via growth models.

With reference to the first element, the proposed approach draws from the choice of a mapping methodology to model the existing interregional inequalities in total factor productivity levels. The method of producing a map for locations of unobserved components is explicitly based on growth theory and infers the dimensions of the map in addition to their relative positions. More specifically, (i) we introduce spatial heterogeneity at the level of the unobserved variables, and (ii) we explicitly produce a location map for unobserved components from the growth theory. The idea of modeling unobserved state variables by directly introducing a mapping model in the dynamic framework is appealing in several perspectives. First, we can introduce relative weights to calibrate the model. Second, we exploit available information to characterize the relevant conditions influencing decisions on

unobserved variables without imposing restrictive assumptions on the model. Third, it is possible to interpret the dimensions of the map in terms of additional unobservable qualities. Finally, it is possible to analyze the evolution over sub-periods of the existing interregional inequalities in productivity.

With reference to the second element, from a methodological point of view, several advantages of the GME/GCE estimation procedures can be pointed out. First, the maximum entropy-based estimators are most efficient relative to traditional estimators in particular when data constraints for each observation are included in the maximum entropy-based problem formulations. Second, they are able to produce estimates in models where the number of parameters exceeds the number of data points and in models characterized by a non-scalar identity covariance matrix. Third, prior information can be introduced by adding suitable constraints in the formulation without imposing strong distributional assumptions.

The proposed approach has been employed to assess the existence of conditional convergence across Italian regions between 1960-1996. Remarkable differences in TFP levels have been detected also taking account of differences in human capital and in the distribution of economic activities. In synthesis, our results strongly support the presence of TFP heterogeneity across Italian regions and also the idea that Italian regions tend to converge to different steady state levels of per capita GDP. The key role of both technology spillovers through human capital accumulation and agglomeration economies within industrial districts as the relevant determinants of TFP differences has been confirmed by our analysis. Finally, human capital has not static effects on GDP levels.

Further investigation should be done in order to:

- I. analyze the evolution over time of cross-region TFP distribution by considering different sub-periods (70s, 80s, 90s). The idea is to analyze convergence by testing the hypothesis that the current degree of TFP heterogeneity across Italian regions is not stationary;
- II. analyze the evolution of regional productivity disparities in Italian manufacturing sectors by introducing a decomposition of TFP gaps in terms of industry-mix, differential and allocative components;
- III. employ information on clustering schemes identified by the mapping analysis in estimating a club convergence model;
- IV. examine the results produced by imposing *a priori* a mapping structure on the covariance matrix in the maximum entropy problem formulation and by directly estimating the location of the unobserved variables. This approach should reproduce the distribution of unobserved variables locations across countries (regions), while accounting for the effects of heterogeneity structures on unobserved TFP' behavior;
- V. introduce other technological and institutional factors, besides human capital and districts, which are relevant in the Italian experience to account for such large TFP differences;

References

Abramovitz M. (1986) Catching-up, forging ahead and falling behind, *Journal of Economic History*, 46, p. 385-406.

Acemoglu D. Aghion P. Zilibotti F. (2004) Distance to frontier, selection, and economic growth, forthcoming *Journal of the European Economic Association*.

Acemoglu D. Johnson S. Robinson J. (2005) Institutions as the fundamental cause of long-run growth, Aghion P. Durlauf S. (eds.) *Handbook of Economic Growth*, forthcoming.

Aghion P. Howitt P. (1992) A model of growth through creative destruction, *Econometrica*, 60(2), p. 323-351.

Aghion P. Howitt P. (1998) Capital accumulation and innovation as complementary factors in long-run growth, *Journal of Economic Growth*, 3(2), p. 111-30.

Aiyar S. Feyrer J. (2002) A contribution to the empirics of total factor productivity, mimeo, Dartmouth College.

Andrés J. Bosca J.E. Doménech R. (2004) Sigma-convergence in the OECD: transitional dynamics or narrowing steady state differences, *Economic Inquiry*, 42, p. 141-149.

Arbia G. Basile R. Piras G. (2005) Using spatial panel data in modelling regional growth and convergence, ISAE working paper n. 55.

Arellano M. Bond S. (1991), Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations, *Review of Economic Studies*, 58, p. 277-97.

Ascari G. Di Cosmo V. (2005) Determinants of total factor productivity in the Italian regions, University of Pavia, mimeo.

Azariadis C. Drazen A. (1990) Threshold externalities in economic development, *Quarterly Journal of Economics*, 105(2), p. 501-526.

Barro R.J. (1991) Economic Growth in a cross section of countries, *Quarterly Journal of Economics*, 106, pp. 407-444.

Barro R.J. Sala-i-Martin X. (1991) Convergence Across States and Regions, *Brooking Papers on Economic Activity*, 1, p. 107-158.

Barro R.J. Sala-i-Martin X. (1997) Technological Diffusion, Convergence, and Growth, *Journal of Economic Growth*, 2, p. 1-26.

Barro R.J. and Sala-i-Martin X. (2004) *Economic Growth*, MIT Press, London.

Basile R. Gress B. (2005) Semi-parametric spatial auto-covariance models of regional growth in Europe, *Région et Développement*, 21, 93-117.

Baumont C. Ertur C. Le Gallo J. (2000) Geographic spillover and growth: a spatial econometric analysis for European regions, LATEC working paper n. 7, University of Burgundy, Dijon.

Benhabib J. Spiegel M. (1994) The role of human capital in economic development: evidence from aggregate cross-country data, *Journal of Monetary Economics*, 34, p. 143-73.

Bernard A. Durlauf S.N. (1995) Convergence in international output, *Journal of Applied Econometrics*, 10, p. 97-108.

Bernard A. Durlauf S.N. (1996) Interpreting tests of the convergence hypothesis, *Journal of Econometrics*, 71(1-2), p. 161-173.

Bernardini Papalia R. (1999) Local generalized method of moments estimation based on kernel weights: an application to panel data, *Journal of Applied Statistics*, 26, 8, 1005-1015.

Bernardini Papalia, R., (2002). Analyzing partial-incomplete economic data generation processes using maximum entropy, in: *Proceedings of the International Conference on Improving Surveys*, Copenhagen.

Bernardini Papalia, R., Bertarelli S. (2005) Estimating sources of the convergence process within a regional framework, mimeo.

Bianchi C. Menegatti M. (1997) Differenziali regionali di produttività e crescita economica: un riesame della convergenza in Italia nel periodo 1970-1994, *Studi Economici*, 52(63), p. 15-42.

Bianchi C. Menegatti M. (2005) Neoclassical versus technological convergence: an empirical analysis applied to the Italian regions, working paper n. 172, Dipartimento di economia politica e metodi quantitativi, Università di Pavia.

Bloom D. Canning D. Sevilla J. (2003) Geography and poverty traps, *Journal of Economic Growth*, 8, p. 355-378.

Blundell R. Bond S. (1998) Initial conditions and moment restrictions in dynamic panel data models, *Journal of Econometrics*, 87, p. 115-43.

Bond S. Hoeffler A. Temple J. (2001) GMM estimation of empirical growth models, CEPR Discussion Paper n. 3048.

Brasili C. (2005) *Cambiamenti strutturali e convergenza economica nelle regioni dell'Unione europea*. CLUEB Bologna.

Canova F. (2004) Testing for convergence clubs in income per capita: a predictive density approach, *International Economic Review*, 45(1), p. 49-77.

Carmeci G. Mauro L. (2004) A positive effect of investment on Italian regional growth, *Rivista Internazionale di Scienze Economiche e Sociali*, 51(3), p. 423-445.

Caselli F. Esquivel G. Lefort F. (1996). Reopening the convergence debate: a new look at cross-country growth empirics, *Journal of Economic Growth*, September, 1(3), 363-90.

Caselli F. Tenreyro S. (2004) Is Poland the Next Spain?, Public Policy Discussion Papers n. 04-8, Federal Reserve Bank of Boston.

Cass D. (1965) Optimum Growth in an Aggregative Model of Capital Accumulation, *Review of Economic Studies*, 32, 233-240.

Cellini R. Scorcu A. (1995) How many Italies?, Working paper n. 215, Dipartimento di Scienze Economiche, Università degli Studi di Bologna.

Chamberlain G. (1984) Panel Data, in S. Griliches and M.D. Intriligator (eds.), *Handbook of Econometrics*, North-Holland, Amsterdam, 1247-1318.

De la Fuente A. Doménech R. (2001) Schooling data, technological diffusion, and the neoclassical model, *American Economic Review*, Papers and Proceedings, 91(2), p. 323-327.

Di Liberto A. (1994) Convergence Across Italian Regions, Nota di Lavoro N. 68, Fondazione ENI Enrico Mattei, Milano.

Di Liberto A. Mura R. Pigliaru F. (2004) How to measure the unobservable: a panel technique for the analysis of TFP convergence, Crenos working paper n. 5.

Di Liberto A. Symons J. (1998) Human capital stocks and the development of Italian regions: a panel approach, Crenos working paper n. 4.

Dowrick S. Nguyen D. (1989) OECD comparative economic growth 1950-85: catch-up and convergence, *American Economic Review*, 79, p. 1010-30.

Durlauf S. Johnson P. (1995) Multiple regimes and cross country growth behaviour, *Journal of Applied Econometrics*, 10(4), p. 365-84.

Elrod, T. (1988) Choice Map: Inferring a Product-Market Map from Panel Data, *Marketing Science*, 7(1), 21-40.

Evans P. Karras G. (1996a) Convergence revisited, *Journal of Monetary Economics*, 37, p. 249-265.

Evans P. Karras G. (1996b) Do economies converge? Evidence from a panel of US States, *Review of Economics and Statistics*, 78, p. 384-388.

Faini R. Galli G. Giannini C. (1992) Finance and Development, the case of Southern Italy, Banca D'Italia, Temi di discussione n. 170, giugno.

Feyrer J. (2003) Convergence by parts, manuscript (mimeo). Dartmouth College.

Fisher S. (1993) The role of macroeconomic factors in growth, *Journal of Monetary Economics*, 485-512.

Galli G. Onado M. (1990) Dualismo territoriale e sistema finanziario, in *Il sistema finanziario nel Mezzogiorno*, Banca d'Italia, numero speciale del Contributo all'analisi Economica, p. 1-47.

Galor O. Zeira J. (1993) Income distribution and macroeconomics, *Review of Economic Studies*, 60(1), p. 35-53.

Golan, A., Judge, G., Miller, D. (1996). *Maximum entropy econometrics: robust estimation with limited data*. Wiley, New York.

Golan, A., Judge, G., Miller, D. (1997). The Maximum Entropy Approach to Estimation and Inference: An Overview, in T. B. Fomby and R. C. Hill (eds.), *Applying Maximum Entropy to Econometric Problems*, Jai Press Ltd. England, London, 3-24.

Hall R.E. Jones C.I. (1999) Why do some countries produce so much more output per worker than others, *Quarterly Journal of Economics*, 114, p. 83-116.

Holbrook M.B., Moore W.L., Winer R.S. (1982) Constructing Joint Spaces from Pick-Any Data: A New Tool for Consumer Analysis, *Journal of Consumer Research*, 9, 99-105.

Howitt P. (2000) Endogenous Growth and Cross-Country Income Differences, *American Economic Review*, 90(4), p. 829-846.

Howitt P. Mayer-Foulkes D. (2005) R&D, Implementation and stagnation: a Schumpeterian theory of convergence clubs, *Journal of Money, Credit and Banking*, 37, p. 147-77.

Imbs J. Wacziarg R. (2003) Stages of Diversification, *American Economic Review*, 93(1), p. 63-86.

Islam N. (1995) Growth empirics : a panel data approach, *Quarterly Journal of Economics*, 110, p. 1127-1170.

Judge G. Giffiths W.E. Ltukephol H Lee T.S. (1988) *Introduction to the theory and practice of econometrics*. Wiley, New York.

Keane M., Runkle D. (1992) On the Estimation of Panel Data with Serial Correlation when Instruments are not Strictly Exogenous, *Journal of Business and Economic Statistics*, 10(1), 1-9.

Kiviet J. (1995) On bias, inconsistency, and efficiency of various estimators in dynamic panel data models, *Journal of Econometrics*, 68, p. 53-78.

Klenow P.J. Rodriguez Clare A. (1997) Economic growth: a review essay, *Journal of Monetary Economics*, 40, p. 597-617.

Koopmans T.C. (1965) On the Concept of Optimal Growth, in *The Econometric Approach to Development Planning*, North Holland, Amsterdam.

Kullback (1959) *Information theory and statistics*. Wiley, New York.

Lee K. Pesaran M.H. Smith R. (1995) Growth and convergence in a multi-country empirical stochastic Solow model, *Journal of Applied Econometrics*, 12(4), p. 357-392.

Le Gallo J. Ertur C. (2003) Exploratory spatial data analysis of the distribution of regional per capita GDP in Europe, 1980-1995, *Papers in Regional Science*, 82(2), p. 175-201.

Leonida L. Petraglia C. Murillo-Zamorano L.R. (2004) Total factor productivity and the convergence hypothesis in the Italian regions, *Applied Economics*, 36, p. 2187-2193.

Levine R. Renelt D. (1992) A sensitivity analysis of cross-country growth regressions, *American Economic Review*, 82(5), p. 942-963.

Lewis W.A. (1954) Economic development with unlimited supplies of labor, *Manchester School of Economics and Social Studies*, 22, p. 139-191.

Liang K.Y. Zeger S.L. (1986) Longitudinal Data Analysis using Generalised Linear Models, *Biometrika*, 73, 1, 13-22.

Liu Z. Stengos T. (1999) Non-linearities in cross country growth regressions: a semiparametric approach, *Journal of Applied Econometrics*, 14(5), p. 527-538.

Lucas R.E. (1988), On the mechanics of economic development, *Journal of Monetary Economics*, 22, p. 3-42.

Mankiw N.G. Romer D. Weil D.N. (1992) A contribution to the empirics of growth, *Quarterly Journal of Economics*, 107, p. 407-437.

Marsh L. Mittlehammer R.C. Cardell S. (1988) *A Structural-Equation GME Estimator*. Selected Paper for the AAEA Annual Meeting, 1998, Salt Lake City.

Mauro L. Podrecca E. (1994) The Case of Italian Regions: Convergence or Dualism?, *Economic Notes*, 24(3), p. 447-472.

- Meliciani V. Peracchi F. (2004) Convergence in per capita GDP across European regions: a reappraisal, CEIS Tor Vergata Research Paper Series working paper n. 58.
- Mundlak Y. (1978) On the Pooling of Time Series and Cross-Section Data, *Econometrica*, 36, 69-85.
- Paci R. Pigliaru F. (1995) Differenziali di crescita tra le regioni italiane: un'analisi cross-section, *Rivista di Politica Economica*, 85(X), p. 3-34.
- Paci R. Saba F. (1998) The empirics of regional economic growth in Italy 1951-1993, *Rivista Internazionale di Scienze Economiche e Sociali*, 45, p. 515-542.
- Pesaran M.H. Smith R. (1995) Estimating long-run relationships from dynamic heterogeneous panels, *Journal of Econometrics*, 68, p. 79-113.
- Prescott E.C. (1998) Needed: A Theory of Total Factor Productivity, *International Economic Review*, 39(3), p. 525-51.
- Romer P.M. (1990) Endogenous technological change, *Journal of Political Economy*, 98(5), p. S71-S102.
- Sala-i-Martin X. (1997) I just ran two million regressions, *American Economic Review*, 87(2), p. 178-83.
- Solow R. (1956) A contribution to the theory of economic growth, *Quarterly Journal of Economics*, 70, p. 65-94.
- Uzawa H. (1965) Optimal technical change in an aggregative model of economic growth, *International Economic Review*, 6, p. 18-31.
- Zellner A. (1996) Models, prior information, and Bayesian Analysis, *Journal of Econometrics*, 75, 51-68.
- Zellner A. (1997) A Bayesian Method of Moments (BMOM): Theory and Applications. *Advances in Econometrics*, Vol. 12. *Applying Maximum Entropy to Econometric Problems*, T.B. Formby and R.C. Hill, (Eds.) 85-105. JAI Press, Greenwich.
- Zellner A. (1999) New Information Based Econometric Methods in Agricultural Economics: Discussion, *American Journal of Agricultural Economics*, 81, 3, 742-746.

Appendix

ITALIAN REGIONS	
PIEMONTE	C1
VAL D'AOSTA	C2
LOMBARDIA	C3
TRENTINO ALTO ADIGE	C4
VENETO	C5
FRIULI VENEZIA GIULIA	C6
LIGURIA	C7
EMILIA ROMAGNA	C8
TOSCANA	C9
UMBRIA	C10
MARZO	C11
LAZIO	C12
ABRUZZO	C13
MOLISE	C14
CAMPANIA	C15
PUGLIA	C16
BASILICATA	C17
CALABRIA	C18
SICILIA	C19
SARDEGNA	C20

Table 1: Results for different estimators (without district variable)

	no human capital				with human capital			
	SGMM	LSDV	Pooled OLS	MG-GME	SGMM	LSDV	Pooled OLS	MG-GME
ly	0.977*	0.527*	0.933*	0.535	0.956*	0.524*	0.918*	0.581
lsk	-0.065*	0.027	-0.035	-0.013	-0.072	0.041	-0.042**	-0.018
lsh					0.018	0.040	0.040	0.003
lndx	-0.152*	-0.028	-0.164*	-0.001	-0.220*	-0.116**	-0.231*	-0.001
Rate of conv.	0.005	0.128	0.014	0.125	0.009	0.129	0.017	0.109

* 5% significant level; ** 10% significant level

Table 2: Max Entropy approach (without human capital and district variables)

Region	ly	lsk	lndx	Rate of conv.
PIE	0.582	-0.013	-0.003	0.108
VDA	0.390	-0.005	-0.002	0.188
LOM	0.418	-0.028	-0.003	0.175
TAA	0.432	-0.004	-0.002	0.168
VEN	0.597	-0.012	-0.001	0.103
FVG	0.547	-0.010	-0.001	0.121
LIG	0.410	-0.024	-0.003	0.178
EMR	0.577	-0.013	-0.001	0.110
TOS	0.583	-0.017	-0.001	0.108
UMB	0.576	-0.015	0.000	0.110
MAR	0.574	-0.007	0.000	0.111
LAZ	0.592	-0.018	-0.003	0.105
ABR	0.550	-0.005	0.000	0.119
MOL	0.589	0.010	0.000	0.106
CAM	0.576	-0.015	-0.001	0.110
PUG	0.584	-0.011	-0.001	0.107
BAS	0.532	-0.034	0.000	0.126
CAL	0.524	-0.011	0.000	0.129
SIC	0.535	-0.009	-0.001	0.125
SAR	0.531	-0.015	-0.001	0.127
Average	0.535	-0.013	-0.001	0.125

Table 3: Max Entropy approach (with human capital and without district variables)

Region	ly	lsk	lsh	lndx	Rate of conv.
PIE	0.621	-0.009	0.004	-0.002	0.095
VDA	0.576	-0.002	0.001	-0.001	0.110
LOM	0.626	-0.018	0.005	-0.002	0.094
TAA	0.635	-0.003	0.004	-0.001	0.091
VEN	0.622	-0.012	0.005	-0.001	0.095
FVG	0.629	-0.011	0.004	-0.001	0.093
LIG	0.596	-0.157	0.002	-0.002	0.104
EMR	0.642	-0.011	0.003	-0.001	0.089
TOS	0.607	-0.012	0.002	-0.001	0.100
UMB	0.629	-0.017	0.003	-0.001	0.093
MAR	0.632	-0.010	0.003	-0.001	0.092
LAZ	0.124	-0.023	0.007	-0.005	0.418
ABR	0.660	-0.012	0.003	0.000	0.083
MOL	0.622	0.001	0.003	0.000	0.095
CAM	0.591	-0.009	0.002	-0.001	0.105
PUG	0.552	-0.011	0.002	-0.001	0.119
BAS	0.571	-0.009	0.003	-0.001	0.112
CAL	0.559	-0.006	-0.006	-0.001	0.116
SIC	0.564	-0.009	0.002	-0.001	0.115
SAR	0.563	-0.015	0.003	-0.001	0.115
Average	0.581	-0.018	0.003	-0.001	0.109

Table 4: Results for different estimators (with human capital and district variables)

	SGMM	LSDV	Pooled OLS	MG-GME
ly	0.95021*	0.52377*	0.90612*	0.58242
lsk	-0.06467	0.04113	-0.01320	-0.00053
lsh	0.01940	0.04043	0.04613**	0.00343
lndx	-0.22035*	-0.11611**	-0.23020*	-0.00132
district	0.02270	dropped	0.05392*	-0.00017
Rate of conv.	0.01022	0.12934	0.01972	0.10811

* 5% significant level; ** 10% significant level

Table 5: Max Entropy approach (with human capital and district variables)

Region	ly	lsk	lsh	lndx	district	Rate of conv.
PIE	0.62153	-0.00036	0.00404	-0.00227	-0.00027	0.09511
VDA	0.53141	-0.00033	0.00777	-0.00153	-0.00052	0.12644
LOM	0.62982	-0.00072	0.00486	-0.00220	-0.00018	0.09247
TAA	0.68391	-0.00018	0.00602	-0.00140	-0.00035	0.07599
VEN	0.62363	-0.00049	0.00476	-0.00147	-0.00020	0.09444
FVG	0.63108	-0.00045	0.00410	-0.00139	-0.00036	0.09206
LIG	0.62038	-0.00089	0.00267	-0.00211	-0.00037	0.09549
EMR	0.64344	-0.00045	0.00343	-0.00146	-0.00022	0.08819
TOS	0.60926	-0.00050	0.00017	-0.00134	-0.00024	0.09910
UMB	0.63161	-0.00069	0.00305	-0.00052	0.00313	0.09190
MAR	0.63255	-0.00039	0.00314	-0.00088	-0.00026	0.09160
LAZ	0.12286	-0.00096	0.00721	-0.00539	-0.00097	0.41935
ABR	0.66054	-0.00046	0.00337	-0.00027	-0.00016	0.08294
MOL	0.62229	0.00005	0.00282	0.00013	-0.00033	0.09487
CAM	0.58066	-0.00060	0.00413	-0.00084	-0.00020	0.10872
PUG	0.62006	-0.00053	0.00323	-0.00116	-0.00052	0.09559
BAS	0.48209	-0.00138	0.00920	-0.00018	-0.00048	0.14592
CAL	0.56248	-0.00024	-0.01368	-0.00071	-0.00024	0.11508
SIC	0.52374	-0.00054	0.00455	-0.00058	-0.00037	0.12935
SAR	0.61515	-0.00057	0.00380	-0.00089	-0.00037	0.09718
Average	0.58242	-0.00053	0.00343	-0.00132	-0.00017	0.10811

Figure 1. Two-dimensional MDS solution (GME estimates, without District variable)

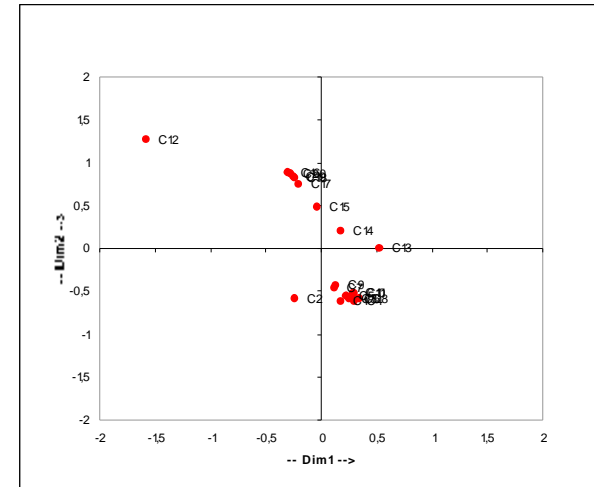
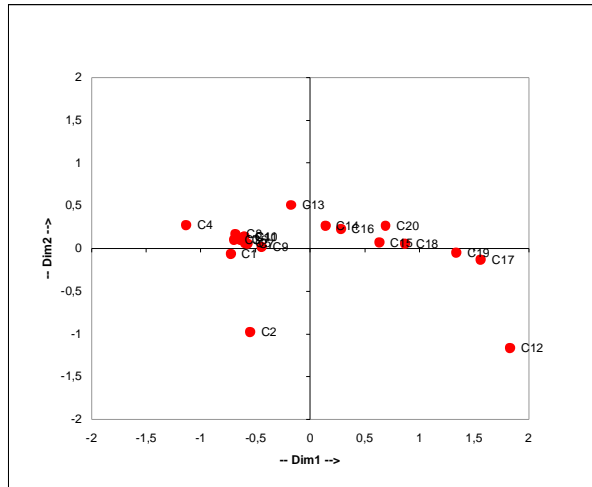


Figure 2. Two-dimensional MDS solution (GME estimates, with District variable)



Notes