45th Conference of the European Regional Science Association
"Land use and Water Management in a Sustainable Network Society"
Vrije Universiteit Amsterdam, 23-27 August 2005

ECONOMIC STRUCTURE, TECHNOLOGY DIFFUSION AND CONVERGENCE: THE CASE OF THE ITALIAN REGIONS

Matteo Lanzafame

Abstract: Taking Italy as a case study, this paper investigates the link between economies' structural similarities and convergence. Specifically, treating technology as sector-specific and modelling technological spillovers as a positive function of the degree of similarity between economies' sectoral features, we propose a modified version of the Solow model and derive an "extended" convergence equation. The latter is then estimated by means of Panel Data procedures and data on the Italian regions over the 1970-1995 period. The results bring empirical support to our approach.

JEL Classification: O40, R11

Keywords: Convergence, technology diffusion, regional growth, Italy

Acknowledgements: I would like to thank, without implicating, Prof. Tony Thirlwall and Dr. Miguel León-Ledesma for helpful comments and suggestions.

Department of Economics, University of Kent, Canterbury, Kent CT2 7NP, United Kingdom

Phone: 0044 1227 827639, Email: ml45@ukc.ac.uk

Dipartimento di Economia, Statistica ed Analisi Geopolitica del Territorio, Università degli Studi di Messina, via Tommaso Cannizzaro n. 278 – 98100, Messina, Italy.

Economic structure, technology diffusion and convergence: The case of the Italian regions

1. Introduction

Italy's long-standing feature of profound regional disparities has always been the object of intense investigation. Although the traditional North-South distinction still retains its relevance, several researchers have in recent years progressively shifted their interest to the regional level and, in particular, to the analysis of "convergence" (or lack thereof) between the Italian regions[1]. In many respects, however, the evidence so far produced is not conclusive. In particular, it does not allow us to draw any unambiguous conclusions as regards the theoretical underpinnings of the convergence process between the Italian regions: is it simply the empirical reflection of the mechanistic operation of neoclassical growth theory principles? Or is it technological transfers that are shaping its evolution, and thus the process could be better explained relying on the technology-gap theories of growth?

Far from being confined to Italy, this is a well-known issue in the convergence literature and distinguishing between the two hypotheses has proven not to be an easy task, not least because the two explanations need not be alternatives [Sala-i-Martin (1996)]. But if the latter is the case, i.e. if both diminishing returns to capital and technological diffusion are playing a role in reducing regional productivity differentials in Italy, another critical question arises, that is whether it may be possible to evaluate the relative importance of the two factors.

In what follows, we formally develop a theoretical framework which is an attempt at bridging both the neoclassical and catch-up theories of growth in dealing with the issues encountered in convergence analysis. As such, its main feature is a different treatment of technological progress, modelled as being partly dependent on inter-regional intra-sector spillovers. As well as assessing its theoretical implications for the study of convergence, we evaluate the empirical relevance of our approach in the case of the Italian regions.

The remainder of the paper is organised as follows: Section 2 dwells upon the theoretical links between our approach and the neoclassical and technology-gap theories of growth; Section 3 introduces a version of the Solow model [Solow (1956)] modified to take account of

[1] See, among others, Mauro and Podrecca (1994), Paci and Pigliaru (1997, 1998), Terrasi (1999), Carmeci and Mauro (2002).

technological diffusion, from which an “extended” convergence equation is formally derived; Section 4 deals with the econometric issues involved in the panel data estimation of the latter, while Section 5 presents and discusses the results of carrying out such an exercise using data for the Italian regions in the 1970-1995 period. Finally, Section 6 concludes.

2. The hypothesis: A “structural channel” for technology diffusion?

The traditional adaptation of the Solow model to the empirical study of convergence embraces the neoclassical notion of technology as an essentially public good. The assumption of an equal (as well as constant) growth rate of technological progress for all the countries or regions in the sample can be justified on the grounds that technological innovations are not only freely available but also introduced and exploited in all production systems within the same unit of time. On the other hand, the literature on the technology-gap approach to economic growth departs from these assumptions and, recognising the undeniable existence of technological differentials between the more- and less-advanced economies, depicts a more complex picture to describe the process of technology diffusion [Gerschenkron (1962), Nelson (1981), Nelson and Wright (1992)]. Among the various hypotheses put forward, the concepts of “social capability” [Ohkawa and Rosovsky (1973), Abramovitz (1986)] and “technological congruence” [Abramovitz (1992, 1994)] emerge as most relevant. Far from being instantaneous, the adoption of external technological innovation is portrayed as a difficult process, held to occur with a (variable-length) temporal lag. The existence of technological differences and gaps, however, does represent an opportunity for laggard countries and regions, and opens up the possibility that, if able to get hold of and exploit the more advanced technology developed elsewhere, the latter may temporarily enjoy a higher growth rate than would otherwise be the case. The result would be that of speeding up the transition of backward economies towards the levels of development of the more advanced ones, thus enhancing the convergence process implied by the neoclassical mechanism.

The two theoretical approaches can conceivably be reconciled considering the stage at which the technological catch-up is complete, for in that instance both can be deemed to treat technology as a public good: if the *steady state* is characterised by the absence of any technology gap, the implication must be that of a frictionless and complete spillover of technological innovations across economies, thus leading to the “neoclassical result” of an equal rate of technological progress among the latter[2].

To a certain extent, the subject can be related to the various studies, mainly inspired by the advent of the so called “New Economic Geography”, which have focused on the relevance of cross-regional (or cross-country) spillovers taking a *spatial perspective* [Fingleton (2001), Maurseth (2001), Rey and Montuory (1999)]. In the context of this literature, the interaction between regional economies is held to be dependent on geographical location, so that it is stronger the closer the spatial proximity. One simple way of assessing the merits of this hypothesis in relation to the convergence phenomenon would involve the introduction of a “spatially-weighted” measure of the productivity level (or growth rate) in surrounding regions in an “informal growth equation” [Temple (1999)], alongside a set of control variables and the logarithm of the initial level of productivity. The significance of the “spatial variable”, then, would give a measure of the importance of regional spillovers.

However, though valuable in other respects, the implementation of such a procedure would not bring us far in answering the questions posed above. On the one hand, the concept of regional interaction being used in this framework is a fairly broad one. Our interest, however, rests solely in its technological aspects, so that the need arises of devising a narrower definition, with the aim of isolating as much as possible the regional growth effects of technological transfers from those of other factors[3]. On the other hand, relying on the geographical approach, this technique neglects the possibility of a different diffusion channel for regional spillovers.

To be more precise, the empirical evidence on the existence of sectoral productivity growth rate differentials seems to suggest that the scope for technological innovation varies significantly across productive activities [Salter (1960), Fagerberg (2000)]. The implication is that technological advances may be primarily “sector-specific” and, thus, subject to an “intra-sector” transmission process, as opposed to the “across-sector” type. Indeed, this conjecture is consistent with the (broader) concept of “technological congruence” put forward by Abramovitz (1992, 1994) as one of the determinants of technology transfers.

---

[2]The technology-gap approach has been mainly concerned with cross-country convergence, to the extent that some authors have suggested the idea of “national innovation systems” to qualify the technological differences which hinder technology diffusion among countries [see Lundvall (1992) and Nelson (1993)]. Taking this view, the conjecture of complete technology diffusion in the long-run is clearly more plausible in a regional context.

[3]Among such factors are, for instance, the inter-regional migration of labour or regional input-output production linkages.

Hence, if the aforementioned process of intra-sector productivity convergence is, at least partly, determined by technology diffusion, the extent to which a less-advanced economy can benefit from the technological improvements developed externally and, indeed, the convergence process as a whole, may depend on the characteristics of the regions' production structures. That is, technological spillovers may be more related to *structural distances* than to geographical proximities, their size and significance dependent on the degree of similarity between economies' sectoral compositions.

Pursuing this conjecture and drawing on the procedures used within the "spatial econometric perspective" approach[4], we start our investigation by designing a measure of *structural distances* defined as

$$D_{ij}(t) = \frac{2}{K_{ij}(t)} \tag{1}$$

where, for each time *t*, $K_{ij}(t)$ is the Krugman Specialisation index (or K-index) between the regions i and j (for $i \neq j$) developed by Midelfart-Knarvirk et al. (2000). This is defined as

$$\mathrm{K_i(t)} = \sum_{\mathrm{k}} \mathrm{abs}\left[\mathrm{v_i^k(t) - v_j^k(t)}\right]$$

where, for any region *i*,

$$\mathrm{v_i^k(t)} \equiv \frac{\mathrm{x_i^k(t)}}{\sum_{\mathrm{k}} \mathrm{x_i^k(t)}}$$

and $\mathrm{x_i^k(t)}$ denotes region *i*'s value added[5] in sector k at time t. For each point in time, $K_{ij}(t)$ is thus constructed as the sum over the k sectors of the absolute differences between the sector shares of value added in regions *i* ($\mathrm{v_i^k}$) and *j* ($\mathrm{v_j^k}$). Its value ranges between zero and two and

[4] Specifically, our treatment is here akin to Fingleton and McCombie (1998).

[5] In preference to value added, Midelfart-Knarvirk et al. (2000) employ the gross value of output as a measure of activity level, on the grounds that this makes the results of the analysis less likely to be biased by the effects of structural shifts in outsourcing to other sectors. This option was precluded by data unavailability in our case.

increases with the degree of specialisation, i.e. it is higher the more a region's production structure differs from that of the other. Since $K_{ij}(t)$ *increases* with the degree of structural dissimilarities and $0 \le K_{ij}(t) \le 2$, $D_{ij}(t)$ *falls* when specialization rises and $0 \le D_{ij}(t) \le \infty$. Subsequently, we normalise $D_{ij}(t)$ as follows

$$W_{ij}(t) = \frac{D_{ij}(t)}{\sum_j D_{ij}(t)} \tag{2}$$

to obtain $0 \le W_{ij}(t) \le 1$ and $\sum_j W_{ij}(t) = 1$,excluding the case in which $D_{ij}(t) = \infty$ for at least one j, which would occur only if two regions had a perfectly identical production structure. Implementing this transformation across all regions results in a matrix of "structural weights", which can then be used to construct, for each region, a "structurally-weighted" measure of the growth rate of external technological progress. Specifically, for each region i, one suitable variable for such a role, which we name $\dot{X}_i(t)$, is

$$\dot{X}_i(t) = \sum_j W_{ij} \dot{y}_j(t) \qquad \text{(for } i \ne j \text{ and } \dot{y}_j(t) \ge 0\text{)} \qquad (3).$$

The dot-notation is adopted to indicate the exponential growth rate of a variable and $\dot{y}_j(t)$ refers to the rate of growth of labour productivity in region j at time t.

Two things need be noted as regards the expression we devise for $\dot{X}_i(t)$. Firstly, as the values of the K-index used in their calculation, the structural weights $(W_{ij}\text{'}s)$ are themselves time-dependent. However, in order to avoid an "excessive" volatility in their values, we consider their rate of change as being discrete and not continuous, so that they are constant *within* each time period *t*. To emphasise this assumption, in (3) the structural weights are denoted by W_{ij} and not $W_{ij}(t)$. Secondly, in this context, the restriction on the value of $\dot{y}_j(t)$ to be positive is dictated by the role of this variable as a proxy for the growth rate of technological progress and, to some extent, allows us to isolate the effects of the latter from those of other factors influencing regional interaction.

Within the framework of an informal growth regression, the impact of structural similarities on convergence trends could then be tested according to the significance of $\dot{X}_i(t)$ in an equation of the following form

$$\dot{y}_i(t) = \alpha + \beta \ln y_i(t_1) + \sum_{j=1}^{N} \theta_j Z_i^j(t) + \pi \dot{X}_i(t) \qquad (4),$$

where $y_i(t_1)$ is the sample period's initial-year level of labour productivity, the $Z_i^j(t)'s$ are a set of N control variables and the expectation is to find a positive value for π, the elasticity of labour productivity growth in region i with respect to $\dot{X}_i(t)$, the "structurally-weighted" growth rate of external technological progress.

However, as first shown by Mankiw, Romer and Weil (1992) (henceforth "MRW") and Islam (1995) (henceforth "Islam") in the context of, respectively, cross-section and panel data estimations, a formal derivation of the convergence equation from the Solow model presents various advantages. In particular, it provides the "correct" specification of the equation, in the sense of being consistent with the model's assumptions, as well as explicit formulations for the estimated coefficients in terms of the model's structural parameters, which can then be retrieved. In our case, this option offers the additional benefit of lending itself to a more thorough explanation and understanding of the theoretical implications of our approach. In the next section, therefore, we proceed to the derivation of a convergence equation from a version of the neoclassical growth model modified to take account of "structurally-weighted" technological spillovers. In view of the empirical testing we will subsequently undertake, and taking into account the characteristics of our dataset, in doing so we follow closely Islam's procedure and notation.

3. The modified version of the Solow model and the extended convergence equation

The starting point for the derivation of the convergence equation from the Solow model is the specification of its production function, which is characterised by constant returns to scale and labour-augmenting technological progress. In its familiar Cobb-Douglas form, it can be expressed as

$$Y(t) = K(t)^{\alpha} \left(A(t)L(t) \right)^{1-\alpha} \qquad 0 < \alpha < 1 \tag{5}$$

where *Y* is output, *K* is capital, *L* is labour, *A* is a shift-parameter accounting for the level of technology or the "effectiveness of labour", α is the elasticity of output with respect to capital and *t* refers to time. The growth rates of both *L* and *A* are assumed to be exogenous and constant, so that

$$L(t) = L(0)e^{nt} \tag{6}$$

and

$$A(t) = A(0)e^{gt} \tag{7}$$

where *n* and *g* are, respectively, the growth rates of labour and technology. In keeping with the original formulation of the model, we let the evolution of the labour input into production be described by equation (6). As regards the *A* term, however, because of the role played by technological spillovers, its rate of growth needs to be modelled differently.

Taking as a basis (7), we consider the growth rate of $A(t)$ as being made up of four different components. More specifically, a constant term *p* is here let to reflect the effects of "country-wide" technological progress and of all the factors that influence the effectiveness of labour in all regions contemporaneously (national institutions, aggregate policy changes, etc…). However, the existence of technological spillovers between regions implies a certain degree of diversity between regional technological systems and innovations. Thus, we introduce a region-specific component of technological progress $\left(q_i \right)$, again assumed to be constant over time. In each point in time, some technological innovations are developed and introduced in any region *i*, which directly increase $A_i(t)$ and are potentially exploitable by other regions as well. More precisely, according to the respective degree of structural similarity, each region benefits from the rise in labour efficiency in the rest of the country, which is here proxied by the growth rate of labour of productivity, in excess of *p*. If in the short-run many factors can hinder it, in the long-term such a technological transfer is assumed to be frictionless and complete. Lastly, to allow for a variable rate of technological progress in the short-run and across regions, the process of technological innovation is assumed to be subject to a random shock, $\upsilon_i(t)$, which is region-specific and serially uncorrelated, so that $E\left(\upsilon_j(t)\upsilon_i(t)\right)=0$ for $i \neq j$, $E\left(\upsilon_i(t)\upsilon_i(s)\right)=0$ for $t \neq s$ and $E\left(\upsilon_i(t)\right)=0$.

Thus, assuming a linear relation, our modification to the traditional Solow model leads to the following specification of the growth rate of *A(t)*:

$$\frac{d\ln A_i(t)}{dt} \equiv \dot{A}_i(t) = p + q_i + \left[\pi\left(\sum_j W_{ij}\frac{d\ln y_j(t)}{dt}\right) - \sum_j W_{ij}p\right] + \upsilon_i(t),$$

where the term in squared brackets reflects the inter-regional diffusion of technological progress and $\sum_j W_{ij}p$ is subtracted from $\pi\left(\sum_j W_{ij}\frac{d\ln y_j(t)}{dt}\right)$, the structurally-weighted *external* productivity growth, to avoid double counting the effect of *p* on $\dot{A}_i(t)$. Rearranging and using (3), we obtain

$$\dot{A}_i(t) = q_i + \pi\dot{X}_i(t) + \upsilon_i(t) \quad (8).$$

The expression in (8) describes the evolution of *A(t)* at each point in time. However, because of the assumptions of frictionless technological diffusion $(\pi = 1)$ and $\upsilon_i(t)=0$, the steady-state equivalent of equation (8) simplifies to

$$\dot{A}_i(t) = q_i + \dot{X}_i \quad (9),$$

which denotes a constant rate of growth of *A(t)* in the long-run.

Now, going back to the building blocks of the model, our treatment of technological progress leads to the substitution of (7), the Solovian formulation of *A(t)*, with

$$A(t) = A(0)e^{qt+\pi\ln X(t)+\varepsilon(t)} \quad (10).$$

Dropping the *i* subscript, (10) is obtained by direct integration of (8) and in it, for each region *i*, $\ln X_i(t) = \sum_j W_{ij} \ln y_j(t)$. Defining output and capital in efficiency units as, respectively, $\hat{y} = Y/AL$ and $\hat{k} = K/AL$ and assuming both the rate of capital depreciation (δ) and the share of output that is saved and invested (s) are constant, the evolution of the capital stock over time is given by

$$\frac{d\hat{k}(t)}{dt} = s\hat{y}(t) - \left(n + q + \pi\dot{X}(t) + \delta\right)\hat{k}(t) \qquad (11).$$

Taking account of the steady state value of $\dot{A}_i(t)$ given in (9), the expression for the steady state level of $\hat{k}$ is

$$\hat{k}^* = \left[\frac{s}{\left(n + q + \dot{X} + \delta\right)}\right]^{\frac{1}{1-\alpha}} \qquad (12).$$

Output per unit of labour is $\frac{Y(t)}{L(t)} = y(t) = \hat{k}(t)^{\alpha} A(t)$, which in logarithmic form becomes $\ln y(t) = \alpha \ln \hat{k}(t) + \ln A(t)$. Integrating (9) in order to obtain the steady-state value of $\ln A(t)$ and using the latter together with (12), for the steady-state level of labour productivity we have

$$\ln y(t) = \frac{\alpha}{1-\alpha}\ln s - \frac{\alpha}{1-\alpha}\ln\left(n + q + \dot{X} + \delta\right) + \ln A(0) + qt + \ln X(t) \qquad (13).$$

Equation (13) reflects closely the specification for the steady state value of $\ln y(t)$ derived by MRW. Indeed, apart from the inclusion of $\dot{X}$ as an additional variable in the sum within parenthesis and, critically, the presence of $\ln X(t)$ on the RHS, the similarity would turn into equivalence when (13) is considered at a given point in time and it is postulated that $\ln A(0) = a + \xi$, so that

$$\ln y(t) = a_i + \frac{\alpha}{1-\alpha}\ln s - \frac{\alpha}{1-\alpha}\ln\left(n+q+\dot{X}+\delta\right)+\xi \qquad (13')$$

The error term ξ, which MRW assume to be unit (country or region) specific and uncorrelated with the explanatory variables, allows for across-unit random differences in the $\ln A(0)$ term and makes OLS estimation of (13') feasible. Yet, there is a key difference between (13') and MRW's correspondent specification, namely the unit-specific intercept a_i. Modelling the growth rate of $A(0)$ as being just a constant, MRW can proceed assuming any across-unit deviation in $\ln A(0)$ to be random, which leads to a common intercept for all units in the sample. However, if the growth rate of $A(0)$ is treated as being partly dependent on a unit-specific feature, such as the country's or region's production structure, then the implication is that of an individual intercept for each unit in the sample, in our case $a_i = \ln A(0) + q_i + \ln X_i(t)$. Although cross-section estimation of (13') would still be possible with the introduction of unit-specific dummy variables, the existence of individual effects can be better accommodated within a panel data framework. Further, other considerations point to the choice of the latter as a more appropriate estimation procedure. Specifically, MRW's assumption of no correlation between the level of technological efficiency and the other regressors is generally seen as not easily justifiable [Temple (1999)]. Allowing to control for individual heterogeneity, a panel data approach provides a way around this problem and, hence, a better setting for the analysis of the issues at hand.

Following Islam, therefore, in order to substantiate formally the latter statement, we now turn to the analysis of the out-of-steady-state behaviour of the model. This can be studied by taking a first-order Taylor approximation around the steady-state, which gives

$$\frac{d\ln \hat{y}(t)}{dt} = \lambda\left[\ln \hat{y}^* - \ln \hat{y}(t)\right] \qquad (14)$$

where, $\hat{y}^*$ is the steady-state level of output per effective unit of labour, $\hat{y}(t)$ is, as usual, its actual value at any time *t* and $\lambda = (1-\alpha)\left(n+q+\dot{X}+\delta\right)$ is the rate of convergence.

It can be noted that $\dot{X}$, the variable accounting for the across-region diffusion of technology, is one of the determinants of the convergence rate and that, just like *n*, *q* and δ, its effects on λ are "filtered" by the $(1-\alpha)$ term, the labour elasticity of output under the assumption of constant returns to scale. Thus, the modified version of the Solow model that we put forward formally shows that, if the presence of technological diffusion between regions (or countries) is allowed for, the convergence process cannot be solely ascribed to neoclassical principles. At the same time, the impossibility of fully disentangling the convergence effects of diminishing returns to capital from those of technological diffusion within a Solovian framework remains. Indeed, as long as technology enters the production function in a "factor-augmenting" fashion, this feature of the model is unavoidable. In what follows, however, it will be shown that our approach allows making some progress in the exploration of this issue.

Going back to the derivation of the convergence equation, the process of adjustment described by (14), implies that

$$\ln \hat{y}(t_2) = \left(1 - e^{-\lambda\tau}\right)\ln \hat{y}^* - e^{-\lambda\tau}\ln \hat{y}(t_1) \tag{15}$$

where $\tau = \left(t_2 - t_1\right)$ and $\hat{y}(t_1)$ is output per effective unit of labour at some initial point in time (t_1). Subtracting $\hat{y}(t_1)$ from both sides and rearranging gives

$$\ln \hat{y}(t_2) - \ln \hat{y}(t_1) = \left(1 - e^{-\lambda\tau}\right)\left(\ln \hat{y}^* - \ln \hat{y}(t_1)\right) \qquad (16).$$

The steady state value of labour productivity is

$$\hat{y}^* = \left(\hat{k}^*\right)^{\alpha} = \left[\frac{s}{\left(n + q + \dot{X} + \delta\right)}\right]^{\frac{\alpha}{1-\alpha}} \tag{17}$$

and using this expression in (16) gives

$$\begin{aligned} \ln \hat{y}(t_2) - \ln \hat{y}(t_1) = {} & \left(1 - e^{-\lambda\tau}\right)\frac{\alpha}{1-\alpha}\ln s \\ & - \left(1 - e^{-\lambda\tau}\right)\frac{\alpha}{1-\alpha}\ln\left(n + q + \dot{X} + \delta\right) - \left(1 - e^{-\lambda\tau}\right)\ln \hat{y}(t_1) \end{aligned} \qquad (18).$$

Equation (18) formalises the temporal evolution of output in efficiency units, $\hat{y}(t)$. For estimation purposes, however, we need to turn from the latter to the evolution of output per unit of labour, $y(t)$. The relation between the two variables can be disclosed as follows:

$$\hat{y}(t) = \frac{Y(t)}{L(t)A(t)} = \frac{Y(t)}{L(t)} \cdot \frac{1}{A(0)e^{qt+\pi \ln X(t)+\varepsilon_i(t)}} ,$$

so that, taking the logarithms of both sides we obtain

$$\begin{aligned} \ln \hat{y}(t) &= \ln\left(\frac{Y(t)}{L(t)}\right) - \ln A(0) - qt - \pi \ln X(t) + \upsilon_i(t) \\ &= \ln y(t) - \ln A(0) - qt - \pi \ln X(t) + \upsilon_i(t) \end{aligned} .$$

Substituting this expression in (18) gives

$$\begin{aligned} \ln y(t_2) - \ln y(t_1) = &\left(1-e^{-\lambda\tau}\right)\frac{\alpha}{1-\alpha}\ln s - \left(1-e^{-\lambda\tau}\right)\frac{\alpha}{1-\alpha}\ln\left(n+q+\dot{X}+\delta\right) \\ &-\left(1-e^{-\lambda\tau}\right)\ln y(t_1) + \left(1-e^{-\lambda\tau}\right)\ln A(0) + q\left(t_2 - e^{-\lambda\tau}t_1\right) \\ &+\left(1-e^{-\lambda\tau}\right)\pi \ln X(t_1) + \pi \ln X(t_2) - \pi \ln X(t_1) + \upsilon_i(t_2) \end{aligned} \qquad (19).$$

Focusing for a moment on the third line of the above equation and neglecting the error term, we notice that $\left\{\left(1-e^{-\lambda\tau}\right)\pi \ln X(t_1) + \pi \ln X(t_2) - \pi \ln X(t_1)\right\}$ can be expressed as $\left\{\pi\left(\ln X(t_2) - e^{-\lambda\tau}\ln X(t_1)\right)\right\}$. This rearrangement, however, would not allow for the imposition of our identifying condition for $\dot{X}(t)$ as a proxy for external technological progress, namely $\dot{y}_j(t) \geq 0$. We thus opt for a different formalisation and simplify the above expression as $\left\{\left(1-e^{-\lambda\tau}\right)\pi \ln X(t_1) + \pi\dot{X}(t_2)\right\}$, where the values of $\dot{y}_j(t)$ used in the construction of $\dot{X}(t_2)$ are strictly non-negative[6]. It may be noted that this expression also

[6] In assuming the growth rate of labour productivity as a proxy for technological progress we follow a fairly well-established practice in the literature. Moreover, notice that in the context of the neoclassical growth model the growth rates of productivity and technological progress are equal when the economy is in steady state, so

lends itself to a clearer economic interpretation: ceteris paribus, in each region *i* and each point in time *t*, the growth rate of labour productivity will be faster the higher the previous period's technological level in the rest of the country, indicating the size of the potential technological transfer and here proxied by $\ln X(t_1)$, and the faster the rate at which the latter is growing at time *t* $\left(\dot{X}(t)\right)$. Notice that the $\ln X(t_1)$ term will be the more significant the wider the technological gap between economies. On the contrary, when the latter is small, backward economies have nearly exhausted the advantages related to the reduction of their *technological backlog* and, just like those already on the technological frontier, will benefit solely from *external technological progress*, denoted by the $\pi\dot{X}(t)$ term.

Reintroducing the *i* subscript for each region, adding $\ln y_i(t_1)$ to both sides and rearranging, (19) is thus expressed as

$$\begin{aligned}\ln y(t_2) &= \left(1-e^{-\lambda\tau}\right)\frac{\alpha}{1-\alpha}\ln s - \left(1-e^{-\lambda\tau}\right)\frac{\alpha}{1-\alpha}\ln\left(n+q_i+\dot{X}_i+\delta\right)+e^{-\lambda\tau}\ln y_i(t_1)\\ &+\left(1-e^{-\lambda\tau}\right)\ln A(0)+q_i\left(t_2-e^{-\lambda\tau}t_1\right)+\left(1-e^{-\lambda\tau}\right)\pi\ln X_i(t_1)+\pi\dot{X}_i(t_1)+\upsilon_i(t_2)\end{aligned} \tag{20}$$

or, using the conventional panel data notation

$$y_{it} = \beta\ln y_{i,t-1}+\sum_{j=1}^{4}\theta_j Z_{it}^j+\mu_i+\eta_t+\upsilon_{it} \tag{21}$$

where $y_{it}=\ln y(t_2)$ and, on the right-hand-side, $\mu_i=\left(1-e^{-\lambda\tau}\right)\ln A(0)$ is the individual (region-specific) effect, $\eta_t=q_i\left(t_2-e^{-\lambda\tau}t_1\right)$ is the time-effect and the remaining variables and parameters are

$$\begin{aligned}&y_{i,t-1}=\ln y(t_1),\ Z_{it}^1=\ln s,\ Z_{it}^2=\ln\left(n+q_i+\dot{X}_i+\delta\right),\ Z_{it}^3=\ln X_{i,t-1},\ Z_{it}^4=\dot{X}_{it}\\ &\beta=e^{-\lambda\tau},\ \theta_1=\left(1-e^{-\lambda\tau}\right)\frac{\alpha}{1-\alpha},\ \theta_2=-\left(1-e^{-\lambda\tau}\right)\frac{\alpha}{1-\alpha},\ \theta_3=\left(1-e^{-\lambda\tau}\right)\pi,\ \theta_4=\pi.\end{aligned} \tag{22}$$

---

that, since we are analysing the model behaviour in the neighbourhood of the steady state, such an assumption can be brought into play fairly confidently.

The *extended* convergence equation (20) typifies the implications of our approach. These are, perhaps, better revealed and discussed when (20) is compared to the *classical* formalisation which, as derived by Islam, is expressed as

$$\begin{aligned}\ln y(t_2) = &\left(1-e^{-\lambda\tau}\right)\frac{\alpha}{1-\alpha}\ln s - \left(1-e^{-\lambda\tau}\right)\frac{\alpha}{1-\alpha}\ln(n+g+\delta) + e^{-\lambda\tau}\ln y(t_1) \\ &+ \left(1-e^{-\lambda\tau}\right)\ln A(0) + g\left(t_2 - e^{-\lambda\tau}\right) + \upsilon(t_2)\end{aligned} \tag{23}$$

where, as in (7), g is the constant growth rate of technological progress, common across regions, and all the other variables retain their usual meaning.

Evidently, the main difference between the two equations is the presence, in log-level and growth rate form, of $X(t)$ on the RHS of (20). Thus, the first implication of our model is that, if technological diffusion is of an intra-sector type, economies' structural differences need to be taken account of as an additional factor on which convergence is conditional. In other words, a number of economies, identical in all other respects (investment rates included), will still not be subject to *β-convergence*, unless they are as structurally similar as to share the same technology. This is because, just like a higher saving rate in the original Solow model, higher values of $\ln X(t-1)$ and/or $X(t)$ will shift the production function upwards, so that those economies characterised by a more "favourable" economic structure, i.e. one allowing them to exploit the technological innovations developed elsewhere relatively more easily, will enjoy a higher steady state productivity *level*.

As regards the steady state productivity *growth rate*, however, the implications of our model are more problematic. The neoclassical equation result of a common convergence rate $[\lambda = (1-\alpha)(n+g+\delta)]$ across economies relies on the assumptions of equal growth rates of technological progress (g) and employment (or population) (n), as well as an equal rate of capital depreciation (δ). The public-good hypothesis for technology justifies the assumption of a common *g* but both *n*, which in actual estimation is treated as a variable (n_{it}), and δ may, of course, be different across units. When the latter occurs, the speed of convergence to the steady state will be different too. Until recently, these were usually regarded as minor problems, since it was argued that the estimated λ would provide an *average* value of the convergence rate [Islam (1996)]. In the case of our model, however, more fundamental problems arise on the theoretical side. In fact, the treatment of technological progress as sector-specific implies that, as long as economies are structurally different, the long-run productivity growth rate they convergence to may be different as well. Thus, the panel data study of convergence will yield an *average* estimate of the speed at which each economy converges to *its own* steady state growth rate, a state of affairs which Islam (2003) terms "Weak Conditional Convergence"[7]. Because of technology diffusion, while "converging" economies will also grow more structurally similar: as the best technologies are gradually adopted in each sector across economies, within-sector productivity levels and growth rates will converge as well, raising the degree of structural similarity and thus reinforcing the technological exchange. In the approach we put forward, the two processes are intertwined and reinforce each other in the transitional dynamics to the steady state. The long-run limit to how similar economies can grow, both in terms of productivity levels and structural features, will be set by the degree at which technology is non-transferable, in turn dependent on such factors as the similarity of resource endowments, technological congruence, etc… The more similar these "fundamentals", the closer the unit-specific steady state growth rates and the better the approximation that the panel data estimate of λ will provide in each case[8].

All of these are issues brought into play by the questioning of the neoclassical public good assumption for technology and are, thus, strictly related to the effects of the potential heterogeneity of steady state growth rates. Leaving aside the already mentioned theoretical consequences, Lee, Pesaran and Smith (1999) (henceforth LPS) note that the econometric problems may be significant as well. As shown by Pesaran and Smith (1995) (henceforth PS), as long as the regressors are serially correlated, the Panel Data estimation of the traditional convergence equation (23) under the wrong assumption of a homogenous *g*, inducing serial correlation in the disturbance, will lead to biased and inconsistent estimates, the problem being more serious the higher the variance of the actual g_i's across the units. Specifically,

[7] This notion is clearly problematic for the traditional meaning of convergence. Indeed, in a previous work, the same author had argued that when "heterogeneity in growth rates is allowed, convergence becomes in essence, an empty construct" [Islam (1996, p. 326)].

[8] These reflections may be conveniently linked to the concept of and the literature on "club convergence" [see, among others, Baumol and Wolff (1988), Durlauf and Johnson (1995)], for, according to our approach, the degree of structural similarity may well qualify as one suitable criterion to select the members of a convergence club. If a group of economies are very structurally dissimilar, during their transitional dynamics the amount of technology diffusion will be generally small and its pace slow. Further, if the structural dissimilarity reflects primarily wide differences in the aforementioned fundamentals, the steady state growth rate differences are likely to be wide as well. In such circumstances, a "no convergence" outcome may be possible and, indeed, likely, so that the absence of "global convergence", which many studies found empirical evidence of, becomes less surprising.

the probability limit of the estimated β, the lagged dependent variable parameter, tends to unity and that of the $\theta's$ tends to zero. To tackle these issues, LPS propose a stochastic version of the Solow model in which steady state growth rates are explicitly allowed to vary across units. Using the Summers and Heston (1991) data set and taking in consideration 102 non-oil-producing countries over the period 1965-1989, they estimate their model using time series methods and find considerably higher convergence rates than those usually obtained in the literature, i.e. an average of about 30 per cent against the traditional 2-3 per cent. However, the mean group estimator employed by LSP suffers from a small sample bias which, as the authors themselves note, "can be important even for *T* as large as 30." (LPS, p. 368). Moreover, the use of annual data raises some concerns, since the estimated coefficients may well be capturing the average frequency of the business cycle, a problem that, as is to be made clear later on, may be serious for Panel Data estimations as well, but is certainly even more severe for time series regressions.

Our approach follows a different route, modelling the heterogeneity of technology growth rates as partly dependent on a deterministic component, $\dot{X}(t)$. This reduces the impact of the econometric problems ascertained by PS in the estimation of (20). Specifically, the parameter-heterogeneity-induced bias will be the less significant the smaller the variances of $\upsilon_i(t)$ and the q_i[9]. Nonetheless, these factors do represent a concern for the Panel Data estimation of (20) in the case of the Italian regions, so that, ideally, the mean-group estimator used by LPS should be implemented. However, given the features of our dataset, the aforementioned problems with the mean group estimator are likely to be very serious in our case[10], so that the cure proposed by PS may well be worse than the disease. Thus, we opt for the use of Panel Data procedures, which itself involves a series of problems, as discussed in the next section.

---

[9] The values of the K-index (not reported, available upon request) show a high and increasing degree of structural similarity between the Italian regions, suggesting they can fairly confidently be defined as a "club", sharing the same (or a not significantly different) long-run growth rate. Thus, the heterogeneity in q_i may not represent too serious a problem in our case.

[10] Preliminary application of the mean group procedure seems to suggest these concerns are justified. Apart from the short time dimension of our panel, which for the estimation of (20) reduces to $T = 25$, the presence of significant "time effects", ascertained by LSDV regressions, represents a serious problem.

4. Data and Panel Data estimation issues

Our dataset is a balanced panel of twenty regions and twenty-six years of annual observations over the 1970-1995 period which, from a purely econometric point of view, poses a series of important questions for the estimation of extended convergence equation (20).

The possible presence of a unit root in the level of labour productivity is a first concern, which we address using a battery tests. In addition to the ADF, the Perron (1997) test, which allows for the presence of an endogenously determined structural break, and the Kwiatkowski et al. (1992) (henceforth KPSS) test, are relied upon. As is well known, however, univariate tests have very low power when applied to a relatively short time-series and/ or variables characterised by a high degree of persistence. Exploiting cross-sectional as well as time-series variation in the data, panel unit root tests have been shown to be more powerful than their univariate counterparts, so that their application is particularly useful in these circumstances. We, thus, also employ one such test, developed by Im et al. (2003) (henceforth IPS), to further investigate the issue[11]. Furthermore, as is to be made clear later on, one of the techniques relied upon (i.e. the Anderson-Hsiao estimator) requires the first-differencing of the variables, thus removing any worry related to the possible presence of a unit root.

A second problem regards the length of the time-intervals to be used in breaking-up the entire sample period in the process of passing from Cross-Section to Panel Data estimation. Islam notes that yearly data regressions may be significantly affected by short-run variations and, as a solution to the problem, he chooses to average the variables over 5-year intervals, in order to smooth out business cycle volatility. This procedure, however, is not devoid of drawbacks. On the one hand, the choice of averaging the variable over whatever n-year interval is, at least to a certain degree, inevitably arbitrary and, although conventionally applied to purge the data from short-run influences, it may well result in the imposition of a different (unpredictable) bias. Furthermore, from a theoretical viewpoint, the assumption of constant s, n and X is less defensible the more the chosen time-span is longer than one year. Finally, this option entails a reduction of the number of observations and, hence, degrees of freedom, which could have serious consequences. In our case, averaging the variables over 5-year intervals, the total number of observations drops from 520 to 100, so that the advantages from averaging may be outweighed by the negative consequences on the precision of our estimates. Although the concerns raised by Islam and others [see Temple (1999)] should not be overlooked, all of

[11] To save space, we do not report the details of the various tests. The reader is referred to the quoted studies.

these considerations point to the choice of annual data as the most preferable. Our way of dealing with these issues will be to carry out the estimations using both annual data and 5-year averages, in order to identify any significant differences[12].

Further problems are related to the choice of the appropriate estimation technique. As already exemplified in (21), the formalisation of the extended convergence equation in (20) results in the conventional "Error Component" (EC) model. The most common estimation procedure for this model is Least Squares Dummy Variables (LSDV) and, as Islam points out, in this case the fixed effects specification should be preferred to the random effects. Indeed, the latter is inconsistent when the explanatory variables are correlated to the individual effects and, as recalled when discussing the advantages of the Panel Data approach to the empirical study of convergence, it is the existence of this correlation, again postulated on theoretical grounds, that forms the basis for the main critique of cross-section estimations.

However, the presence of a lagged dependent variable on the RHS of the equation makes equation (20) a "Dynamic Panel Data Model" (DPDM) and the presence of individual effects makes LSDV estimates biased in such instances, with the coefficient on the lagged dependent variable (β) being more severely affected. Nickell (1981) has formally derived an expression for such a bias, showing that it is inversely related to the time dimension of the panel (i.e. it goes to zero as $T \to \infty$) and a number of techniques have been proposed for estimation of DPDMs [Anderson and Hsiao (1981), Arellano and Bond (1991, 1995), Blundell and Bond (1998)], so that the question arises of which one is to be chosen. Several studies, relying on Monte Carlo simulations, have tried to shed some light on this issue and the general conclusion which can be drawn from their results is that the most appropriate technique changes with the size of the panel.

Concerning themselves with the estimation of DPDMs in the context of macroeconomic panel datasets, characterised by relatively large time dimensions for a comparatively small number of units (regions, countries, etc…), Judson and Owen (1999) compare OLS and LSDV estimates to the performances of the Anderson and Hsiao (1981) (hereafter AH) estimator, of two GMM procedures [proposed by Arellano and Bond (1991)] and a corrected LSDV estimator developed by Kiviet (1995). Among other things, their findings show that the bias of the coefficients on the independent variables $(Z_{it}^{j}\text{'}s)$ is "relatively small and cannot be used

[12] As will be seen later on, no major difference is found between the two sets of results. Although the LSDV regressions suggest the presence of significant "time effects", the use of time dummies or the between-group transformation of the data seem to have been very effective in correcting for any potential bias.

to distinguish between estimators" and suggest that "…when $T = 20$, GMM or AH may be chosen…….Because the efficiency of the AH estimator increases substantially as T gets larger, the computationally simpler AH may be justified when T is large enough." [Judson and Owen (1999, p. 13)]. According to the results of their simulations, in our case OLS should be favoured when dealing with 5-year averages and the AH estimator should be relied upon when using yearly data. The empirical testing of our model is carried out following these indications.

For completeness purposes, LSDV estimates are also provided and another estimation technique, based on assumptions regarding the distribution of the residuals different from those of the EC method, is also implemented. This procedure has been proposed by Beck and Katz (1995) as an alternative to the Parks' method [Parks (1967)] for Time-Series Cross-Section (TSCS) data and allows for the presence in the disturbances of both heteroscedasticity and contemporaneous correlation. The latter may represent a problem for the reliability of EC estimates, so that we provide TSCS regressions of (20) as a further robustness check[13].

The technical details of the empirical investigation of the model are provided in the next subsection.

Structure of the model

In its general (matrix) form, the error component model can be described as:

$$y_{it} = \alpha + Z_{it}\beta + v_{it} \tag{24}$$

where $i = 1.........N$, $t = 1.........T$ and

$$v_{it} = \mu_i + \eta_t + \upsilon_{it} \tag{25}$$

where *i* refers to units, *t* denotes time periods, Z_{it} is a vector of *k* exogenous variables, μ_i is an unobservable individual specific effect, η_t is an unobservable time specific effect and υ_{it} is

[13] Our modeling of technology diffusion between regions implies a certain degree of simultaneity in the determination of $\ln y_{it}$ across i's. As long as the "small region" assumption is a valid one, i.e. as long as no one single region is driving productivity growth in the rest of the country, the EC estimations will remain reliable. Correcting for potential contemporaneous correlation, the TSCS regressions provide a test of this assumption.

an idiosyncratic effect. If μ_i and η_t are parameters and $\upsilon_{it} \sim (0, \sigma_\upsilon^2)$, then (24) is referred to as a “fixed effects error component model” and can be estimated by least squares with dummy variables (LSDV). If $\mu_i \sim (0, \sigma_\mu^2)$, $\eta_t \sim (0, \sigma_\eta^2)$ and $\upsilon_{it} \sim (0, \sigma_\upsilon^2)$ are random disturbances, independent of each other and among themselves, as well as uncorrelated with the Z_{it} 's, then the model is named “random effects error component model” and the estimation procedure is Generalised Least Squares (GLS).

When one of the Z_{it} 's is a lagged dependent variable $(y_{i,t-1})$, the model becomes dynamic and the aforementioned problems ensue. As a solution to the latter, Anderson and Hsiao (1981) put forward an instrumental variable (IV) procedure, which is designed as follows. Considering a dynamic version of (24) in which $\eta_t = 0$, so that the model becomes

$$y_{it} = \beta y_{i,t-1} + Z_{it}\theta + \mu_i + \upsilon_{it} \qquad (26)$$

the variables are initially first differenced to obtain

$$\Delta y_{it} = \beta \Delta y_{i,t-1} + \Delta Z_{it}\theta + \Delta \upsilon_{it} \qquad (27).$$

Although the fixed effects have been removed, the errors in (27) are now correlated with $\Delta y_{i,t-1}$, so that the latter is instrumented with $y_{i,t-2}$, which is correlated with it but not with the disturbance[14]. Following Judson and Owen (1999), we will refer to this as the AH estimator[15].

Finally, using TSCS techniques implies setting $\mu_i = \eta_t = 0$, so that the model in (24) becomes

$$y_{it} = Z_{it}\beta + \upsilon_{it} \qquad (28).$$

---

[14] Anderson and Hsiao (1981) suggest $\Delta y_{i,t-2}$ as an alternative instrument, but Arellano (1989) shows that the latter leads to a significant loss of efficiency [see also Arellano and Bond (1991) and Kiviet (1995)].

[15] The first-differencing involved in the implementation of this procedure brings two additional advantages. The first is the abovementioned removal of any residual worry related to the possible presence of a unit root. The second is that, partially removing the serial correlation in the disturbances, it reduces the problems related to the parameter-heterogeneity-induced bias.

To deal with the double nature of the data, the structure of the error matrix features a high degree of flexibility and different models arise from (28) according to which of the following assumptions on the distribution of the error term are allowed for: (1) panel heteroscedasticity, (2) contemporaneously correlated errors, (3) common serially correlated errors, (4) unit-specific serially correlated errors.

If the errors in (28) meet one or more of these assumptions, OLS estimates will be consistent but inefficient. The Parks' (1967) method deals with this problem using two sequential Feasible Generalised Least Squares (FGLS) transformations to firstly purge the data from serial correlation and, subsequently, deal with cross-section heteroscedasticity and contemporaneous correlation. However, Beck and Katz (1995) show that this second correction yields downward biased standard errors and that the severity of the problem is inversely related to the time dimension of the data[16]. Monte Carlo evidence shows that, even for a ratio of T to N equal to 4, the resulting "overconfidence" of the Parks standard errors is about 30%. Since, in the case of TSCS estimations, OLS estimates are usually found to be not much less efficient than FGLS estimates, the solution they suggest is to retain OLS estimates of the regression parameters and rely on *panel-corrected standard errors* (PCSEs), which correct for contemporaneous correlation and heteroschedasticity[17].

5. Estimation of the convergence equation

All data used in the estimations are from the *Regional Accounts databank* CRENoS[18] and the period under consideration is 1970-1995. As regards the variables, we measure n as being the growth rate of labour units employed and s the investment-output ratio. Assuming a value of 0.02 for g and 0.03 for δ, MRW and Islam (1995) set $(g+\delta)=0.05$, which, in the case of the classical convergence equation, represents a slight problem with our dataset, since it leads

[16] Each off-diagonal element of the matrix of contemporaneous correlations of the errors is, on average, estimated using $2T/N$ observations. Thus, if T is close to N, as in our case, each element would be estimated using only about two observations.

[17] Note that a prerequisite for the application of PCSEs is the absence of any serial correlation in the data. For the details of the computation of PCSEs, see Beck and Katz (1995, p. 638).

[18] The database is available on line at http://www.crenos.it. The reader is referred to the CRENoS website for a description of its features.

to some negative values of $(n+g+\delta)$. Therefore, on the grounds that δ may be bigger, we set $(g+\delta)$ or $(q_i+\delta)$ equal to 0.07[19].

REGION	LAGS	ADF	PERRON	KPSS
PIEMONTE	4	-3.018	-6.80023**	0.07426
VALLE D'AOSTA	4	-4.126*	-6.12126*	0.13028
LOMBARDIA	0	-2.007	-3.68035	0.39223**
TRENTINO ALTO ADIGE	0	-2.095	-4.72877	0.18431
VENETO	0	-2.731	-3.42769	0.20246*
FRIULI VENEZIA GIULIA	0	-1.789	-4.02743	0.29213**
LIGURIA	4	-2.948	-6.24038*	0.07788
EMILIA ROMAGNA	1	-3.298	-4.03206	0.11373
TOSCANA	2	-3.875*	-5.27281	0.07825
UMBRIA	0	-2.182	-4.00395	0.20899*
MARCHE	0	-1.766	-3.61252	0.27507**
LAZIO	0	-2.127	-4.31657	0.20114*
ABRUZZO	1	-3.723*	-4.91638	0.12591
MOLISE	0	-2.443	-3.69157	0.30963**
CAMPANIA	3	-4.200*	-6.04074*	0.09563
PUGLIA	1	-3.330	-5.07043	0.17269*
BASILICATA	1	-0.3004	-4.02568	0.24845**
CALABRIA	2	-3.405	-4.58471	0.09695
SICILIA	2	-4.025*	-4.24002	0.12863
SARDEGNA	0	-2.863	-3.91593	0.22040**
IPS t-bar (1)		-2.81**		
IPS t-bar (2)		-2.42^	-2.62**	-2.15

Table 1 – Unit root tests on $\ln y(t)$

Notes:

1. All tests include both an intercept and a trend. Lags selected with general-to-simple recursive procedure [see Perron (1997)];
2. "Perron" is the unit root test proposed by Perron (1997), the null hypothesis is non-stationarity;
3. "KPSS" is the unit root test proposed by Kwiatkowski et al. (1992), the null hypothesis is stationarity;
4. "IPS t-bar" is the Im et al. (2003) Panel Unit Root test. The two versions of the test are, respectively: IPS t-bar (1) reports the value of the test when applied to the entire panel of 20 regions; IPS t-bar (2) when applied to the regions for which the respective univariate test cannot reject the null of a unit root (ADF and Perron) or rejects the null of stationarity (KPSS).
5. ^ indicates rejection of the null at the 10% level, * at the 5% level and ** at the 1% level.

---

[19] Setting $(q_i+\delta)=0.05$ in equation (20) does not change significantly the estimates in quantitative terms and, what is perhaps more important, turns out to be irrelevant qualitatively (i.e. significance levels, signs, etc…).

We start our analysis by examining the results of the unit root tests on $\ln y(t)$, reported in Table 1. As expected, the ADF and Perron (1997) tests reject the null of a unit root only in a handful of cases (5 and 4, respectively), while the more powerful KPSS does not reject the null of stationarity for 10 regions out of 20. As for the Panel Unit Root test, when applied to the entire sample the IPS strongly rejects the unit root hypothesis, while the results are less clear-cut when the test is applied to the sub-samples of regions for which the univariate tests signal the presence of a unit root. In our opinion, as far as potential non-stationarity issues are concerned, the results of the unit root tests allow us to proceed to the estimation of (20) with some degree of confidence. As will be seen shortly, the latter will be further reinforced by the comparison of the AH estimates with those of the other estimators.

For comparison purposes, both the classic convergence equation (23) and the extended formulation derived in (20) are estimated and their results discussed. As mentioned, while presenting LSVD estimates throughout, the choice of the most appropriate Panel Data technique relies in each case on the insights and Monte Carlo evidence provided by Judson and Owen (1999), while the TSCS results rely on the procedure suggested by Beck and Katz (1995) (i.e. OLS with PCSEs). White's heteroscedasticity-corrected standard errors are applied whenever feasible.

We start off with the estimates from using the 5-year averages suggested by Islam, which are presented in Table 2. The routinely reported LSDV results are in this case presented together with the pooled OLS estimates, the latter technique being by far the most appropriate in this case according to Judson and Owen (1999). The difference between the two sets of results is, as expected, relevant for the coefficient on the lagged dependent variable[20]: the downward-biased LSDV estimate of β generates a convergence rate of about 14 per cent, much higher than the implied λ from OLS estimations, as well as the values usually characterising conventional convergence studies. The disparity is even greater in the case of α, the elasticity of output with respect to capital, which is about 25 per cent according to the LSDV regressions and about 55 per cent when using OLS.

Turning our attention to the comparison between the classic and the extended convergence equations, we firstly note that the coefficient of $\dot{X}(t)$ turns out to be significant for both the LSDV and the OLS regressions, taking a value of, respectively, about 0.47 and 0.33. In the

[20] The magnitude of the difference is very close to that predicted by the Monte Carlo evidence provided by Judson and Owen (1999). It is reassuring that, generally, this turns out to be true for the other estimations as well.

context of our approach, this implies that, in the short-run, technology diffusion was on average at least one third (nearly 50 per cent according to the LSDV estimates) or, equivalently, that each region's productivity growth rate rose by about 0.3 per cent with every additional percentage point of the structurally-weighted productivity growth rate in the rest of the country.

VARIABLES	PANEL ESTIMATION			
	LSDV		*OLS*	
	Classical	*Extended*	*Classical*	*Extended*
Constant	1.5369* (0.3835)	1.8789 (2.9757)	0.0682 (0.1047)	-0.1061 (0.1536)
$\ln y(t-1)$	0.4949* (0.0952)	0.4920* (0.0985)	0.8842** (0.031)	0.8558** (0.0476)
$\ln s(t)$	-0.0112 (0.0343)	-0.0130 (0.0356)	-0.0428 (0.0305)	-0.0378 (0.0291)
$\ln(n+g+\delta)$	-0.1674* (0.0388)	-	0.1577** (0.0470)	-
$\ln(n+q+X+\delta)$	-	-0.1808* (0.0515)	-	-0.1691** (0.0629)
$\ln X(t-1)$	-	-0.1038 (0.8072)	-	0.0732 (0.0675)
$\dot{X}(t)$	-	0.4683* (0.1247)	-	0.3267* (0.1315)
Adjusted R^2	0.9721	0.9705	0.9315	0.9298
Implied λ	0.1407* (0.0385)	0.1418* (0.0400)	0.0246** (0.0069)	0.0311** (0.0111)
Implied α	0.2489* (0.0608)	0.2625* (0.0732)	0.5766** (0.0663)	0.5398** (0.1182)
Wald test, *p*-value				
$H_0: \theta_2+\theta_3=0$	0.0001	0.0008	0.001	0.0045
$H_0: (1-\beta)\theta_4-\theta_3=0$	-	0.6719	-	0.6809

Table 2 - 5-year averages, dependent variable is $\ln y(t)$ **- (** significant at the 1% level, * at the 5%).**

On the contrary, θ_3, the coefficient of the $\ln X(t-1)$ variable, turns out to be not significant, a result that, as already mentioned, indicates that the absolute technology gap of the average Italian region from the others is not substantial. The formulation in (20), according to which $\theta_3 = (1-\beta)\theta_4$, allows us to say something more as regards the value of θ_3. Specifically, it is possible to perform a Wald test on $\mathrm{H}_0: (1-\beta)\theta_4 - \theta_3 = 0$ to check whether the estimated relation between the values of θ_3 and the other parameters differs significantly from the one arrived at theoretically. The results of the test, given in the last row of the table, strongly suggest that the null hypothesis cannot be rejected, lending empirical support to the theoretical fundamentals of our model. On the other hand, the linear restriction on the parameters of $\ln s(t)$ and $\ln(n+g+\delta)$ or $\ln(n+q+X+\delta)$, i.e. $\mathrm{H}_0: \theta_2 + \theta_3 = 0$, is always rejected, an outcome that is likely to be driven by the insignificance of $\ln s(t)$, which also enters with the wrong sign. As for the implied convergence rate, the extended version estimate is slightly faster (3 per cent, against 2.5) but this can hardly change the overall impression that the estimation of the two convergence equations depicts a very similar picture.

The results remain remarkably consistent when we turn to the yearly data regressions, whose estimates, both for the classic and extended versions, are reported in Table 3. The AH estimator is in this case the most reliable[21], while, together with the LSDV ones, the TSCS results are also provided as a further robustness check for $\dot{X}(t)$ and $\ln X(t-1)$. Further, notice that the AH regressions are also presented in "*Restricted*" version, since the Wald test could not reject $\mathrm{H}_0: \theta_2 + \theta_3 = 0$ in either case. Starting our comment with a comparison of the results in Table 2 with those in Table 3, what can immediately be noticed is their remarkable similarity, which suggests that the aforementioned concerns about the effects of business cycle volatility on yearly data regressions may be largely misplaced in our case.

The differences between the LSDV estimates and the results of the other estimators display the same pattern as that already described for the 5-year averages estimations, so that we avoid any further comment on them and, focussing primarily on the AH results, carry on with the evaluation of the classic and extended versions. The convergence rate estimated in both cases is again fairly similar, around 1.65 per cent and, thus, somewhat slower than what appears to be when using 5-year averages.

[21] As mentioned, the LSDV regression revealed the presence of significant time effects, so that, before proceeding with the application of the AH estimator, the between-group transformation was applied to the data in order to ensure the suitability of the one-way model specified in (26).

VARIABLES	CLASSICAL MODEL				EXTENDED MODEL			
	Unrestricted			*Restricted*	*Unrestricted*			*Restricted*
	LSDV	**AH**	**TSCS**	**AH**	**LSDV**	**AH**	**TSCS**	**AH**
Constant	0.4310** (0.0823)	-	0.0385 (0.0523)	-	-0.4693 (0.5941)	-	0.0165 (0.0761)	-
$\ln y(t-1)$	0.8721** (0.0214)	0.9834* (0.0074)	0.9809** (0.0142)	0.9835* (0.0074)	0.8725** (0.0216)	0.9837* (0.0073)	0.9780** (0.0124)	0.9838* (0.0073)
$\ln s(t)$	-0.0030 (0.0064)	-0.0007 (0.0119)	-0.0036 (0.0078)	-	-0.0026 (0.0064)	0.0003 (0.0120)	-0.0033 (0.0077)	-
$\ln(n+g+\delta)$	-0.0256** (0.0282)	-0.0130* (0.0023)	-0.0190** (0.0051)	-	-	-	-	-
$\ln s-\ln(n+g+\delta)$	-	-	-	-0.0124** (0.0022)	-	-	-	-
$\ln(n+q+X+\delta)$	-	-	-	-	-0.0355** (0.0040)	-0.0186** (0.0034)	-0.0267** (0.0074)	-
$\ln s-(n+q+X+\delta)$	-	-	-	-	-	-	-	-0.0170** (0.0073)
$\ln X(t-1)$	-	-	-	-	0.2322 (0.1552)	0.2882 (0.2856)	0.0036 (0.0236)	0.2754 (0.2793)
$\dot{X}(t)$	-	-	-	-	0.5086** (0.0961)	0.3180* (0.1608)	0.3018** (0.0747)	0.2817^ (0.1604)
Adjusted R^2	0.9925	0.0900	-	0.0894	0.9926	0.0919	-	0.0883
Implied λ	0.1368** (0.0245)	0.0168* (0.0076)	0.0193 (0.0144)	0.0162* (0.0075)	0.1364** (0.0248)	0.0164* (0.0074)	0.0223^ (0.0127)	0.0163* (0.0074)
Implied α	0.1669** (0.0291)	0.4388** (0.1136)	0.4975* (0.2051)	0.4293** (0.1144)	0.2178** (0.0366)	0.5322** (0.1151)	0.5475** (0.1659)	0.5117** (0.1184)
	Wald test, *p*-value							
$H_0:\theta_2+\theta_3=0$	0.0000	0.2490	0.0177	-	0.0000	0.1370	0.0056	-
$H_0:(1-\beta)\theta_4-\theta_3=0$	-	-	-	-	0.2896	0.3215	0.8927	0.3320

Table 3 – Yearly data, dependent variable is $\ln y(t)$ - (significant at the 1% level, * at the 5%, ^ at the 10%)**

For the extended model, the value of the coefficient on $\dot{X}(t)$ is again found to be about 0.3 (a result confirmed by the TSCS regression) and, notwithstanding the loss of efficiency implied by the AH estimator, it maintains its significance, although only at the 90 per cent level of confidence for the "*Restricted*" version. As for $\ln X(t-1)$, once again, the previous comments remain valid: the variable does not appear to have a significant effect on $\ln y(t)$, but it enters with the correct sign and the Wald test suggests that its size, relative to the other estimated parameters, is not significantly different from what it is expected to be. Finally, apart from the ever-present and puzzling result of an insignificant coefficient for $\ln s(t)$, it may be noted that the estimate of α varies somewhat between the two versions, being just above 50 per cent for the extended model and about 10 percentage points lower for the classic version.

6. Conclusion

Using Italy as a case study, this paper investigates the link between economies' structural characteristics and their growth performances. We assume a structural channel for technological spillovers, derive an "extended" convergence equation from a modified version of the Solow model and estimate it by means of Panel Data procedures and data on the Italian regions over the 1970-1995 period. The results are remarkably robust to different techniques and provide empirical support for the validity of our approach.

From the theoretical viewpoint, our model implies that the effects of technology diffusion on the convergence process are twofold. Firstly, if technological progress is partly dependent on external innovations, the temporal evolution of an economy's productivity level, and its speed of convergence to the steady state value, cannot be ascribed solely to the existence of diminishing returns to capital, as suggested by Neoclassical Growth Theory, but is affected by technology diffusion as well. The difficulty in disentangling the effects of the two factors on the convergence rate remains, but the "extended" convergence equation reveals that the size of technological spillovers will have a *level effect* on productivity. With innovations flowing through a structural channel, the degree to which economies can enjoy such an effect will depend on their structural features. Secondly, treating technological progress as sector dependent, our model implies potential growth rate heterogeneity. Thus, the estimated convergence rate can be ascribed to the concept of "Weak Conditional Convergence", with each economy converging to *its own* steady state growth rate, which is the more likely to be different from the others the more diverse the steady state production structures.

As pointed out by PS and LPS, this potential parameter heterogeneity entails some serious econometric problems for dynamic Panel Data estimations, which may lead to downward biased estimates of the convergence rate and the parameters of the independent variables. As a result, our estimates should be treated with some care, since they may be underestimating the importance of structurally-weighted technology diffusion between the Italian regions which, nonetheless, remains significant.

Bibliography

Abramovitz, M. (1986) – "Catching-Up, Forging Ahead and Falling Behind", *Journal of Economic History*, 46, 386-406.

Abramovitz, M. (1993) – "The Origins of the Post-War Catch-Up and Convergence Boom", in J. Fagerberg, B. Verspagen and N. von Tunzelmann (eds), *The Dynamics of Technology, Trade and Growth*, Aldershot, Edward Elgar, 21–52.

Abramovitz, M. (1994) –"Catch-Up and Convergence in the Post-war Boom and After", in W. J. Baumol, R. R. Nelson and E. N. Wolff (eds.), *Convergence of Productivity: Cross-Country Studies and Historical Evidence*, Oxford, Oxford University Press, 86-125.

Anderson, T. W. and Hsiao, C. (1981) – "Estimation of Dynamic Models with Error Components", *Journal of American Statistical Association*, 589-606.

Arellano, M. and Bond, S. (1991) – "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations", *Review of Economic Studies*, 58, 277-297.

Arellano, M. and Bond, S. (1995) – "Another Look at the Instrumental Variable Estimation of Error-Components Models", *Journal of Econometrics*, 68, 29-51.

Beck, N. and Katz, J. N. (1995) – "What to Do and not to Do with Time-Series Cross-Section Data", *American Political Science Review*, 89, 634-647.

Carmeci, G. and Mauro, L. (2002) – "The Convergence of the Italian Regions and Unemployment: Theory and Evidence", *Journal of Regional Science*, 42, 509-532.

Dollar, D. and Wolff, E. N. (1988) – "Convergence of Industry Labor Productivity among Advanced Economies, 1963-1982", *The Review of Economics and Statistics*, 70, 549-558.

Fagerberg, J. (2000) – "Technological Progress, Structural Change and Productivity Growth: a Comparative Study", *Structural Change and Economic Dynamics*, 11, 193-411.

Fingleton, B. (2001) – "Theoretical Economic Geography and Spatial Econometrics: Dynamic Perspectives", *Journal of Economic Geography*, 1, 201-225.

Fingleton, B. and McCombie, J. S. L. (1998) – "Increasing Returns and Economic Growth: Some Evidence for Manufacturing from the European Union Regions", *Oxford Economic Papers*, 50, 89-105.

Gerschenkron, A. (1962) – "Economic Backwardness in Historical Perspective", Cambridge, MA, Belknap Press.

Islam, N. (1995) – “Growth Empirics: a Panel Data Approach”, *Quarterly Journal of Economics*, 110, 1127-1170.

Islam, N. (1998) – “Growth Empirics: A Panel Data Approach – A Reply”, *Quarterly Journal of Economics*, 113, 325-329.

Islam, N. (2003) – “What Have We Learnt from the Convergence Debate?”, *Journal of Economic Surveys*, 17, 309-362.

Judson, R. A. and Owen, A. L. (1999) – “Estimating Dynamic Panel Data Models: a Guide for Macroeconomists”, *Economics Letters*, 65, 9-15.

Kiviet, J. F. (1995) – “On Bias, Inconsistency and Inefficiency of Various Estimators in Dynamic Panel Data Models”, *Journal of Econometrics,* 68, 53-78.

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

Lundvall, B. (1992) – “National Systems of Innovation: Towards a Theory of innovation and Interactive Learning”, London, Pinter Pub.

Mankiw, N. G., Romer, D. and Weil, D. (1992) – “A Contribution to the Empirics of Economic Growth”, *Quarterly Journal of Economics,* 107, 407-437.

Mauro, L. and Podrecca, E. (1994) – “The Case of Italian Regions: Convergence or Dualism?”, *Economics Letters*, 23, 447-468.

Maurseth, P. (2001) – “Convergence, Geography and Technology”, *Structural Change and Economic Dynamics,* 12, 247-276.

Midelfart-Knarvik, K. H., Overman, H.G., Redding, S. J. and Venables, A. J. (2000) – “The Location of European Industry”, *Economic Papers 142*, *European Commission Directorate-General for Economic and Financial Affairs*.

Nelson, R.R. (1981) – “Research on Productivity Growth and Productivity Differences: Dead Ends and New Departures”, *Journal of Economic Literature*, 19, 1029-1064.

Nelson, R.R. (1993) – “National Innovation Systems: A Comparative Analysis”, Oxford, Oxford University Press.

Nelson, R.R. and Wright, G. (1992) – “The Rise and Fall of American Technological Leadership: The Post-war Era in Historical Perspective”, *Journal of Economic Literature*, 30, 1931-1964.

Nickell, S. (1981) – “Biases in Dynamic Models with Fixed Effects”, *Econometrica*, 49, 1417-1426.

Ohkawa, K. and Rosovsky, H. (1973) – “Japanese Economic Growth”, Stanford, Stanford University Press.

Paci, R. and Pigliaru, F. (1997) – “Structural Change and Convergence: an Italian Regional Perspective”, *Structural Change and Economic Dynamics*, 8, 297-318.

Paci, R. and Pigliaru, F. (1998). “Growth and Sectoral Dynamics in the Italian Regions.” In J. Adams and F. Pigliaru, eds. (1998): *Beyond Convergence. National and Regional Patterns of Growth and Chang*e. Cheltenham: Edward Elgar.

Parks, R. (1967) – “Efficient Estimation of a System of Regressions Equations when Disturbances Are Both Serially and Contemporaneously Correlated”, *Journal of the American Statistical Association*, 62, 500-509.

Pesaran, M. H., and Smith, R. (1995) – “Estimating Long-Run Relationships from Heterogeneous Panels”, *Journal of Econometrics*, 68, 79-113.

Rey, S. and Montuory, B. (1999) – “US Regional Income Convergence: A Spatial Econometric Perspective”, *Regional Studies*, 33, 143-156.

Sala-i-Martin, X. (1996) – “Regional Cohesion: Evidence and Theories of Regional Growth and Convergence”, *European Economic Review*, 40, 1325-1352.

Salter, W. E. G. (1960) – “Productivity and Technical Change”, Cambridge University Press, Cambridge.

Solow, R. M. (1956) – “A Contribution to the Theory of Economic Growth”, *Quarterly Journal of Economics*, 70, 65-94.

Temple, J. (1999) – “The New Growth Evidence”, *Journal of Economic Literature*, 37, 112-156.

Terrasi, M. (1999) – “Convergence and Divergence across Italian Regions” – *The Annals of Regional Science*, 33, 491-510.