Fazio G, Piacentino D.

[Convergence analysis for hierarchical longitudinal data.](#)

Economic Modelling 2018

DOI: <https://doi.org/10.1016/j.econmod.2018.03.009>

Copyright:

© 2018. This manuscript version is made available under the [CC-BY-NC-ND 4.0 license](#)

DOI link to article:

<https://doi.org/10.1016/j.econmod.2018.03.009>

Date deposited:

19/04/2018

Embargo release date:

17 April 2019

This work is licensed under a

[Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International licence](#)

Convergence analysis for hierarchical longitudinal data

Giorgio Fazio* Davide Piacentino†

Abstract

Convergence analysis is typically envisaged either from a macro or a micro perspective. However, empirical tests tend to ignore that the two levels are often “nested” in a hierarchy. Building on hierarchical growth curve modelling, we propose an approach to convergence analysis that allows contemporaneous inference on macro and micro-convergence. Compared to the classic linear convergence analysis, the suggested methodology provides a more flexible alternative to model heterogeneity and validate the results for possible Galton’s fallacy. We illustrate the approach in two empirical examples, one considering convergence across European regions and countries and the other across Italian firms and regions. In the European case, we find that the evidence of convergence depends on the choice of cross-sectional sample. Evidence on convergence in Italy applies only to part of the temporal sample and, therefore, is not robust to Galton’s fallacy. Our analysis returns more robust results on the convergence process and allows better inference for policy intervention. We can envisage that this approach will find increasing applications in the future, as disaggregated data becomes available and heterogeneity becomes an increasingly prominent feature in economic modelling.

JEL Classification: C33, O47, R11.

*Corresponding author, Newcastle University Business School, Newcastle University, 5 Barrack Road, NE1 4SE, Newcastle upon Tyne (UK), *Tel.* +44 0191 2081642, *Email* giorgio.fazio@ncl.ac.uk; *SEAS, University of Palermo*

†SEAS, University of Palermo, Edificio 13, Palermo, ITA-90128, *Tel.* +39 091 23895230, *Email* davide.piacentino@unipa.it. The authors are grateful to the Editor and two anonymous reviewers for their comments on earlier versions of the paper. The paper has also benefitted from various discussions, at various stages, with Roberta Capello, Ugo Fratesi, Yue Ma, Sergio Rey, Mariangela Sciandra, Fiona Steele and Thanasis Stengos. The usual caveat applies and all remaining errors are their own.

Keywords: Convergence analysis; Hierarchical longitudinal data; Growth curves; European convergence; Italian convergence.

1 Introduction

Understanding economic convergence is of critical importance in order to formulate growth policies at the national, regional and local level. It is not surprising, then, that convergence analysis has become one of the most important objectives of empirical macroeconomic modelling. In this direction, traditional empirical growth studies typically consider a reference level of analysis, e.g. countries or regions, and perform tests of convergence using aggregate data from national or international statistical sources.¹ Following Solow (1956), the classic linear convergence approach relies on the assumption that economies share similar preferences and technologies and, therefore, have similar aggregate production functions. However, as discussed in detail by Durlauf and Johnson (1995) and Durlauf *et al.* (2001, 2005), the data generating process underlying the Solow parameters may be substantially different across countries (or regions) leading to incorrect conclusions on the convergence process. This issue has prompted a number of contributions in the macroeconometric literature that try to model parameter heterogeneity in the convergence analysis (see, among the others, Durlauf *et al.* 2001, Phillips and Sul, 2007, Haupt *et al.*, 2017). Another issue emerging from the use of the aggregate production function is related to its ability to represent the underlying economy and, therefore, the convergence process. Actually, the very existence of aggregate production functions is part of a long-standing controversy, as illustrated by Cohen and Harcourt (2003) and more recently by Felipe and McCombie (2014). Indeed, considerable heterogeneity may exist among the underlying micro units that are part of the macro aggregate of interest (Altomonte and Colantone 2008; Fazio and Piacentino, 2010). As such, empirical growth analysis is bound to miss the underlying microeconomic processes, such as technological diffusion/catching-up and concentration/dispersion occurring at the micro level (regional, firm or industry), that according to economic theory are critical to observe

¹ See Islam (2003) for an extensive survey on the convergence debate and Martin and Sunley (1998), Rey and Janikas (2005), Le Gallo and Fingleton (2014) and Gennaioli *et al.* (2014) for a more specific regional perspective.

economic convergence or divergence at the macro-level.²

For the above reasons, a number of papers have shifted the focus from the macro to the micro level, looking for example at firms or industries (see, among the others, Hausmann *et al.*, 2007; Bartelsman *et al.*, 2008; Egger and Pfaffermayr, 2009; Huber and Pfaffermayr, 2010; Chevalier *et al.*, 2012). This literature tends to find stronger evidence of convergence at the micro rather than the macro-level.³ Yet, also a purely disaggregated and micro-level perspective has its pitfalls. First, from the policy perspective it is typically the macro level of interest - usually a certain administrative unit - that matters in order to formulate policies to boost growth and reduce inequality. Second, observational units are clearly clustered in space, a concept given particular attention in regional science and economics (e.g. Arbia *et al.*, 2010; 2012), where Garretsen and Martin (2011) and Ottaviano (2011) further highlight the need to more explicitly consider how micro-level heterogeneity and micro-level interactions among people and firms affect macro-level heterogeneity.⁴ In this context, the spatial economics literature emphasises how the choice of the appropriate level of interest and the relative statistical inference may be affected by the well-known *ecological fallacy* or *aggre-*

² For example, Bartelsman *et al.* (2008) argue for the importance to consider firm-level dynamics and allow for micro-level heterogeneity in order to properly account for the role of technological diffusion. They exploit international firm-level data to construct national and global productivity frontiers and find that UK firms tend to display greater convergence towards the national frontier rather than the global one.

³ Rodrik (2013) identifies international unconditional convergence when disaggregated sectoral data is used in place of aggregate national data. According to the Author, aggregate macro studies tend to ignore intra-sectoral heterogeneity, particularly evident in lagging countries where few industries may be close - and many others far away - from the frontier. He argues that even though strong convergence is found within manufacturing, the limited extent of manufacturing in developing and low income countries explains the failure of aggregate convergence.

⁴ For example, in this literature, a number of authors consider the role of firm heterogeneity (Ottaviano, 2011; von Ehrlich and Seidel, 2013; Fazio and Maltese, 2015) or work heterogeneity (Groot *et al.*, 2014) with respect to agglomeration economies. Venables (2011) discusses the connection between heterogeneity across workers and heterogeneity across cities. Altomonte and Colantone (2008) perform a microfounded analysis of the sources of regional economic disparities and find that such disparities are endogenous to the interaction between firm-level dynamics and the initial market conditions. Rizov and Zhang (2014) investigate regional productivity disparities in China from firm-level data, Basile *et al.* (2014) exploit firm-level data to explore the regional business cycles differentials. Campbell *et al.* (2016) discuss the role of firm heterogeneity for aggregate convergence. They construct regional indicators of firm performance and identify an important role of best performers for regional convergence within an aggregate analysis. In the opposite direction, Magrini, Gerolimetto and Duran (2015) consider the role of aggregate fluctuations, such as business cycle synchronization, in the analysis of regional convergence in the USA.

gation bias problems first discussed by Robinson (1950).⁵

Both the purely macro and the purely micro approaches, however, tend to ignore that often the economic processes under consideration, and the relative data measurements, are hierarchically nested. The importance of such hierarchical nature is evident if one considers, for example, how firms dynamics are influenced by their “ecology” or how national convergence is shaped by both within and between country forces. Acknowledging such hierarchical nature can help tackling some of the above limitations.

Moving in this direction, this paper illustrates how the estimation of growth curves in hierarchical longitudinal data, traditionally used in the literature on education (e.g. Steele, 2008; Kremer *et al.*, 2016), can be used as an additional alternative approach to model parameter heterogeneity and overcome some of the limitations of the classic convergence analysis. The suggested methodology presents some notable advantages. First, it allows inference on convergence at the macro level of interest by exploiting the underlying information on micro-convergence.⁶ As such, it takes into account macro and micro growth heterogeneity by looking at the underlying disaggregated growth processes. Second, it obtains at the same time convergence tests at the macro and the micro levels, while accounting for interactions both within and between the levels of the hierarchy. In the process, it allows controlling for various sources of heterogeneity at both levels and their interactions. A further important advantage of the proposed approach is to allow cross-checking the convergence result for potential Galton’s fallacy, one of the main problems of the classic convergence approach.

The rest of the paper is organized as follows. The next section illustrates the proposed approach to testing for convergence over hierarchical longitudinal data. Section 3 applies the approach to two empirical examples investigating convergence among regions and countries in Europe and convergence among firms and regions in Italy. Section 4 concludes and

⁵ See Paelinck (2000) for a more extensive discussion.

⁶ Indeed, Goldstein (2011) highlights the usefulness of a hierarchical approach even when the aggregate level is the main level of interest.

discusses the implications from the analysis.

2 Hierarchical growth curves and convergence

2.1 Cross-sectional β - convergence analysis

The proposed approach is best understood in relation to the traditional β convergence analysis due to Barro and Sala-i-Martin (1991, 1992), who estimate a reduced-form equation of the neoclassical growth model due to Solow (1956, 1957). Assuming the same steady state for all economies, absolute or unconditional convergence is measured by considering the following regression equation:

$$\frac{1}{T} (y_{nT} - y_{n0}) = g_n = \beta_0 + \beta_1 y_{n0} + \varepsilon_n , \quad (1)$$

where y_{n0} and y_{nT} represent the natural log of per capita income of economy n in the initial and final period, so that g_n is the average growth of unit n over the observed period T , and ε_n is the usual error term. Equation (1) is estimated in a purely cross-sectional setting, where convergence requires economies with lower initial levels of per capita income to grow faster than economies with higher initial levels of per capital income, i.e.

$$\hat{\beta}_1 = \frac{cov(g_n, y_{n0})}{var(y_{n0})} < 0 . \quad (2)$$

As mentioned in the introduction, an important drawback of equation (1) pertains to the fact that it is based on the assumption of common production functions (parameters) across all economies under consideration. Durlauf *et al.* (2001, 2005) highlight how this assumption can be particularly stringent and lead to misleading conclusions, as different economies are likely to have different underlying production functions and corresponding growth processes.

Some of this heterogeneity can be captured in a multi-level version of equation (1).

Chasco and Lopez (2009), for example, capture country-level variability in a regional growth regression by fitting the following model, where i regions are nested within n countries and z is a vector of control variables:

$$\begin{aligned} g_{in} &= \beta_{0n} + \beta_1 y_{0in} + \theta' z_{in} + \varepsilon_{in} \\ \beta_{0n} &= \beta_0 + v_{0n} \end{aligned} \tag{3}$$

The heterogeneity in growth across countries is captured by the random intercept terms $v_{0n} \sim N(0, \sigma_{v_0})$. The authors find that the inclusion of random intercepts. i.e. cross-country differences, affects the speed of convergence across European regions.

Dapena *et al.* (2017a, 2017b) further extend equation (3) to allow not only intercepts, but also slopes β_1 to vary across the n economies. By capturing the heterogeneity in the relationship between the growth rate and the initial income, they obtain a decomposition of the overall convergence process into the national and regional parts.

However, the above approaches omit some information that is contained in the longitudinal part of the data. The importance to incorporate the time series information in convergence approaches is underlined by Bernard and Durlauf (1996, p 172) who, in comparing cross-sectional and time-series convergence approaches, assert that: *“An important advance over both approaches, though, would lie in the integration of the transition information in the cross-section approach with the steady state information in the time series approach to create a more general empirical methodology”*.

In this paper, we move in this direction and adopt a parametric approach that allows for heterogeneity in the speed of convergence across productive units and groups.⁷ As discussed in the next paragraph, this approach is also particularly useful in terms of cross-checking the robustness of the results for Galton’s fallacy.

⁷ In a recent paper, Li *et al.* (2016) consider the longitudinal part of the data in a non-parametric setting and investigate the heterogeneity in the rates of convergence across quantiles. They find evidence in favour of a varying-coefficient approach in convergence analysis or in their words: *“non-constant rates of convergence among different groups of economies”* (Li *et al.*, 2016, p. 39).

2.2 Hierarchical longitudinal convergence analysis

Here, we show how the longitudinal part of the data can be exploited in a parametric setting where, via the estimation of hierarchical growth curves, it is possible at the same time to look at aggregate macro dynamics as a consequence of micro-disaggregate dynamics, consider heterogeneity across all units in the sample and obtain convergence parameters for both levels under analysis. The approach proposed is suited for hierarchical longitudinal data. The growth curves can either be estimated as latent factors if a structural equation approach is employed or as random factors if a multilevel growth approach is employed. The convergence results are not affected by which approach is adopted and scholars frequently prefer the multilevel approach because of its computational advantages. For the same reason, we take here this second avenue.

As mentioned above, convergence analysis based on the estimation of equation 1 (and variants) is further affected by the so-called Galton's fallacy. Indeed, Sala-i-Martin (1996) underline how β convergence is a necessary, but not sufficient, condition to conclude in favor of a reduction in the cross-sectional dispersion of per capita incomes over time, also known as σ convergence. This means that under no β convergence, there cannot be σ convergence, but under β convergence, the further requirement is needed that the initial variance of incomes lies above its steady state and diminishes over time. Hence, it is in principle possible to conclude for β convergence, while the dispersion actually increases over time. In this case, concluding for convergence would result in Galton's fallacy (see Egger and Pfaffermayr, 2009, and Huber and Pfaffermayr, 2010, for further discussion). This implies that it is not possible to make inference solely on equation (1). We show here how the suggested approach can be used to cross-check whether the estimated convergence satisfies both necessary and sufficient conditions, therefore excluding potential misleading evidence based on Galton's fallacy.

To illustrate the approach, we can begin from the most simple general linear-trend model:⁸

$$y_t = \gamma_0 + \gamma_1 t + u_t, \quad (4)$$

where

$$\gamma_1 = \frac{dy}{dt} = \frac{d \ln(Y)}{dt} = \frac{1}{Y} \frac{dY}{dt} = \frac{dY/Y}{dt}$$

and u_t is the usual disturbance term. The estimate of γ_1 , $\hat{\gamma}_1$, can be interpreted as the estimated growth of y over the period $t_T - t_0$ of any generic unit of interest.⁹ This is omitted at this stage from equation (4) to keep notation simple.

Using the notation familiar to multilevel models, if we consider y_{ti} as the per capita income of unit i recorded at time t , we can estimate equation (4) as follows:¹⁰

$$y_{ti} = \gamma_{0i} + \gamma_{1i} t + \varepsilon_{ti} \quad (5)$$

$$\gamma_{0i} = \gamma_0 + \eta_{0i}$$

$$\gamma_{1i} = \gamma_1 + \eta_{1i}$$

where γ_{0i} is an individual-specific intercept composed of a fixed part, γ_0 , and a random term, $\eta_{0i} \sim N(0, \sigma_{\eta_0}^2)$; γ_{1i} is an individual-specific slope with respect to time, again composed of the fixed part γ_1 and the random part $\eta_{1i} \sim N(0, \sigma_{\eta_1}^2)$. The final term $\varepsilon_{ti} \sim N(0, \sigma_\varepsilon^2)$ is the

⁸ From the simple compound growth process $Y_t = Y_0(1+g)^t$, with Y_t and Y_0 per capita incomes at times t and 0, and g the growth rate, take natural logs to obtain the log-linear growth process $\ln(Y_t) = \ln(Y_0) + t \cdot \ln(1+g)$. Defining $\ln(Y_t) = y_t$, $\ln(Y_0) = \gamma_0$ and $\ln(1+g) = \gamma_1$, equation (4) is obtained.

⁹ Note that $\hat{\gamma}_1$ is similar but not identical to the $\Delta \bar{y}$ of equation (1).

¹⁰ Note that in the multilevel notation, t_{ti} is the point in time when an observation i is recorded. In longitudinal macroeconomic data, responses are typically recorded continuously and contemporaneously for all units, so that $t_{ti} = t_t$.

random component related to time. In the hierarchical growth curve setting, the system of equations (5) is a simple two-level hierarchical structure, where the recording time represents the *first level* and the realization of y for unit i is the *second level* (see Steele, 2008). Then, the above system of equations can be represented in reduced form, as follows:

$$y_{ti} = \underbrace{\gamma_0 + \gamma_1 t_t}_{\text{deterministic}} + \underbrace{\eta_{0i} + \eta_{1i} t_t + \varepsilon_{ti}}_{\text{stochastic}}. \quad (6)$$

While the term $\gamma_0 + \gamma_1 t_t$ represents the common growth trajectory, $\eta_{0i} + \eta_{1i} t_t$ denotes the individual growth trajectory resulting from an individual intercept (e.g. an individual initial level of per capita income) and an individual slope (e.g. an individual growth rate of per capita income). Residuals may be level-correlated and the covariance structure of this 2-level model is given by:

$$\Omega_\eta = \begin{pmatrix} \sigma_{\eta_0}^2 & \\ \sigma_{\eta_{01}} & \sigma_{\eta_1}^2 \end{pmatrix}, \quad (7)$$

where $\sigma_{\eta_0}^2$ and $\sigma_{\eta_1}^2$ are respectively the variance of individual intercepts and slopes (growth rates)¹¹ and $\sigma_{\eta_{01}}$ is the covariance between intercepts and slopes, giving information on convergence. A statistically significant negative (positive) covariance will imply convergence (divergence): individual units with lower (higher) values of y at the initial period experience higher (lower) growth rates over the observed period.

Since $\sigma_{\eta_{01}} = \text{cov}(\hat{\gamma}_{0i}, \hat{\gamma}_{1i}) \simeq \text{cov}(y_{i0}, g_i)$, a measure similar to the β -convergence parameter of standard growth regressions can be obtained by dividing the covariance $\sigma_{\eta_{01}}$ by the variance of the intercepts:

¹¹ Here, t is centered around the first observed year so that the intercepts represent the initial period individual per capital incomes and $\sigma_{\eta_0}^2$ their variance. In the multilevel literature, the time variable t is usually centered around the mid-point and $\sigma_{\eta_0}^2$ is interpreted as the between-individual variance in y at the mid-point.

$$\hat{\mu} = \frac{cov(\hat{\gamma}_{0i}, \hat{\gamma}_{1i})}{var(\hat{\gamma}_{0i})} \simeq \frac{cov(y_{i0}, g_i)}{var(y_{i0})} = \hat{\beta} \quad (8)$$

However, while a positive estimate of $\hat{\mu}$ is clearly evidence of divergence, a negative $\hat{\mu}$ is a necessary, but not sufficient condition to conclude in favor of convergence over the entire period. Indeed, the between-units variance in the expected value of y is a quadratic function of time:

$$V_y = E[(y_{ti} - \gamma_0 - \gamma_1 t_i)(y_{ti} - \gamma_0 - \gamma_1 t_i)] = \sigma_{\eta_0}^2 + \sigma_{\eta_1}^2 t_i^2 + \sigma_{\eta_{01}} 2t_i + \sigma_{\varepsilon}^2. \quad (9)$$

V_y in eq. (9) can increase or decrease with respect to time depending on the sign of the covariance $\sigma_{\eta_{01}}$ and on the size of t , as

$$\frac{dV_y}{dt} = 2\sigma_{\eta_1}^2 t_i + 2\sigma_{\eta_{01}}; \quad \frac{d^2V_y}{dt^2} = 2\sigma_{\eta_1}^2 \geq 0. \quad (10)$$

V_y has its minimum for $t^* = \frac{-\sigma_{\eta_{01}}}{\sigma_{\eta_1}^2}$. From eq. (10), when the covariance $\sigma_{\eta_{01}} > 0$ - so that V_y is a strictly positive function of time - the result is that of divergence over the entire period T . When the covariance $\sigma_{\eta_{01}} < 0$, V_y evolves over time depending on the size of $\sigma_{\eta_1}^2$ relative to $\sigma_{\eta_{01}}$. As $\sigma_{\eta_1}^2 \rightarrow 0$, $t^* \rightarrow \infty$, providing evidence of convergence over the entire period T . As the variance of the slopes increases relative to the covariance, t^* becomes smaller and it may lie within the observed period T . In this case, V_y will decrease for $t < t^*$, providing evidence of convergence and will increase for $t > t^*$, providing evidence of divergence. This could be potentially indicating evidence of catching-up and overtaking. Figure A1 in Appendix A provides a graphical illustration of the possible expected values and variances of y relative to alternative convergence/divergence paths.

In summary, two conditions must be met to conclude in favor of convergence over the entire period:

Condition i)

The convergence parameter has to be significantly negative, i.e. $\mu < 0$;

Condition ii)

The minimum value of V_y has to fall outside the observed period, i.e. $t^* > T$.

Importantly, and differently from the multilevel versions of the classic convergence approach proposed in the literature, the second condition gives us information similar to the one obtained in a σ convergence analysis, so that the proposed approach is also validated in terms of the above mentioned Galton's fallacy.

Equation (6) still has to be extended to allow for the estimation of μ also for higher levels in the hierarchy with disaggregated units i nested within aggregate units n (e.g. firms within regions or regions within countries). In hierarchical growth curve notation, this becomes the three level model:¹²

$$\begin{aligned} y_{tin} &= \gamma_{0in} + \gamma_{1in}t_t + \varepsilon_{tin} \\ \gamma_{0in} &= \gamma_0 + v_{0n} + \eta_{0in} \\ \gamma_{1in} &= \gamma_1 + v_{1n} + \eta_{1in} \end{aligned} \tag{11}$$

or in reduced form:

$$y_{tin} = \gamma_0 + \gamma_1 t_t + \eta_{0in} + v_{0n} + \eta_{1in} t_t + v_{1n} t_t + \varepsilon_{tin} . \tag{12}$$

In equation (12), the growth rate is now allowed to vary both across the disaggregated units, i , and the aggregate macro-level units, n . Via equation 12, and exploiting the greater dimensionality of data at the first level, this approach now has the flexibility needed to model

¹² Since multilevel analysis uses random effects estimation, it controls for unobserved effects across units and groups. Also, it can easily accommodate for the introduction of conditioning variables to disentangle the unobserved effects from potentially observable factors. Moreover, the model could be easily adapted to allow for non-linear trends in the data.

departures from common growth at both the individual level and at the group or macro level. Estimation of equation (12) yields the following two variance-covariance matrices:

$$\Omega_{\eta} = \begin{pmatrix} \sigma_{\eta_0}^2 & \\ \sigma_{\eta_{01}} & \sigma_{\eta_1}^2 \end{pmatrix}; \quad \Omega_v = \begin{pmatrix} \sigma_{v_0}^2 & \\ \sigma_{v_{01}} & \sigma_{v_1}^2 \end{pmatrix} \quad (13)$$

where $\sigma_{\eta_{01}}$ can be used, as before, to obtain a measure of convergence among the i units and $\sigma_{v_{01}}$ allows now measuring convergence among the n units. While in the β convergence framework, we would have run separate regressions on the different levels to obtain the parameters β_i and β_n , using this approach we are able to simultaneously obtain two μ parameters and the relative t^* for each of the levels involved in the analysis:

$$\mu_i = \frac{\sigma_{\eta_{01}}}{\sigma_{\eta_0}^2}, \quad t_i^* = \frac{-\sigma_{\eta_{01}}}{\sigma_{\eta_1}^2}; \quad \mu_n = \frac{\sigma_{v_{01}}}{\sigma_{v_0}^2}, \quad t_n^* = \frac{-\sigma_{v_{01}}}{\sigma_{v_1}^2}. \quad (14)$$

This is a particularly useful feature, since it allows assessing whether the rates of convergence differ across the two levels, as envisaged in recent studies. Standard errors for the μ statistics can be obtained using the Delta method. The above models are estimated using Restricted Maximum Likelihood (REML).¹³

In the next section, we illustrate the above approach using two empirical examples, the first based on European regions naturally hierarchically nested within countries and the sec-

¹³ The choice of REML in place of Maximum Likelihood (ML) estimation is due to the fact that ML would lead to bias in the estimation of the variance components (Nakagawa and Schielzeth, 2013). This, however, does not allow comparing likelihoods across models and, therefore, the use of Likelihood Ratio or Wald tests. A further issue is the difficulty to calculate goodness-of-fit statistics. As discussed extensively by Nakagawa and Schielzeth (2013) and Johnson (2014), there are both theoretical and practical problems in the estimation of these measures for linear mixed effects models, and there are no solutions available when more than two levels and both random intercepts and slopes are considered (in our case there are three levels: time, micro, macro). As a consequence, R^2 -like measures are rarely presented and information criteria are used as a diagnostic tool. A possible alternative to obtain goodness-of-fit measures would be the Bayesian framework suggested by Gelman and Pardoe (2006). While this method provides a theoretical solution, it is empirically cumbersome and, as noted by Nakagawa and Schielzeth (2013), the “implementation of this method is rather difficult” (p.136). Further, its computational difficulty increases with the complexity of the specification adopted. This method, moreover, would be useful only if one were interested in the goodness-of-fit of single levels, rather than the overall model. For the above reasons, in the empirical analysis in section 3, we compare models using the Akaike Information Criterion (AIC).

and based on Italian firms each belonging to their NUTS 3 regions.

3 Two empirical examples

3.1 Convergence of Regions and Countries in Europe

In the first empirical exercise, along the lines of the seminal contribution of Sala-i-Martin (1996), we consider convergence in Europe across NUTS 2 regions and countries. Data on GDP per capita is obtained from Cambridge Econometrics European Regional Database¹⁴ (see Appendix B for some simple descriptive statistics). In our analysis, we also look for robustness of the results across three different samples. The first sample includes the same 8 countries (Belgium, Denmark, France, Germany, Italy, the Netherlands, Spain, United Kingdom) used by Sala-i-Martin (1996) in his pioneering work, but for the more recent period 1980-2012. The second sample uses the same time span, but includes 16 countries and 213 regions. The final sample extends the cross section of countries and regions even further to 26 and 277, respectively, to include also the Central-Eastern European economies. This, however, requires a compromise on the time dimension, by limiting the time span to 1991-2012, because of the shorter data availability for this set of countries. A larger cross-section of countries and regions introduces greater heterogeneity, but it also increases the number of second-level units and, therefore, the statistical power of the multilevel analysis in estimating the second-level effects of the hierarchy.

As a preliminary analysis, table 1 reports the simple β convergence regressions as per equation (1) when regions and countries are taken separately. These results indicate evidence of convergence at the regional level and divergence at the country level irrespective of the sample considered.

¹⁴ <http://www.camecon.com/>

<insert table 1 about here>

Then, we perform the proposed hierarchical analysis over the three above samples. In the first three specifications of table 2, we show how hierarchical growth convergence can be modelled in reference to the smaller sample of 8 countries included in Sala-i-Martin (1996). For this sample, we begin our estimations by considering the simplest model with only random intercepts at the regional level in column 1 and then with random intercepts and slopes at the regional level to model regional convergence only (column 2) and with random intercepts and slopes at both the regional and the country-level to model simultaneously regional and country-level convergence (column 3). The AIC criterion¹⁵ confirms that the specification in column 3, where both regional and country-level growth trajectories are considered, is indeed the most informative. Results in columns 4 and 5 report results for the larger sample directly using the specification of column 3.

Statistically significant evidence of regional convergence is found for the samples of 171 regions (8 countries) and 213 regions (16 countries), but not for the sample of 277 regions (26 countries). In order to validate these results in terms of Galton's fallacy, we also consider the evolution of the expected variance of GDP per capita over time and, in particular, we calculate the t_i^* discussed above (see condition (ii) in table 2) for those instances where there is evidence of statistically significant convergence according to condition (i). Importantly, t_i^* is smaller than T in columns 2, 3 and 4, indicating a decrease in cross-sectional variance only for part of the sample. In particular, for the sample of 171 regions, evidence of convergence is limited to the first 3 or 6 years depending on whether country-level convergence is also modelled. Convergence is present for the first eight years between the 213 regions of the second sample that excludes the Central Eastern European block. Therefore, while we would be tempted to conclude in favor of convergence by looking at the μ_i parameter only (or on

¹⁵ The AIC, equal to two times the number of parameters minus two times the log-likelihood, is here -10949, -14493, -14577 in the each of the first three regressions.

the grounds of the classic convergence approach in table 1), the second condition leads us to conclude against regional convergence over the entire period (i.e. Galton's fallacy).

At the country level, where the β convergence tests always concluded for no convergence, we now find evidence of statistically significant convergence only for the sample including 26 countries. In this case t_n^* falls well beyond T , indicating a monotonic decrease over time of the expected variance of GDP per capita (see panel a in figure 1), so that the result is not subject to problems of Galton's fallacy. Interestingly, this result shows that, once a richer dataset is used within a more flexible setting that accounts for heterogeneity of the growth processes across regions and countries, convergence can be detected across European countries. This convergence is only partially visible in the empirical values of GDP per capita (panel b of figure 1), where there is a tendency for the high-income countries to gather around a common steady state and for the dispersion of income between the lower-income countries and the middle-income countries to go down over time. This example highlights how hierarchical convergence analysis can overcome some of the limitations of a purely cross-sectional macro convergence approach that looks just at either regions or countries and can help shed further light on the European convergence process.¹⁶

In the next paragraph, we present a second example based on Italian firms and regions.

<insert table 2 about here>

<insert figure 1 about here>

3.2 Convergence of Firms and Regions in Italy

In our second example we look at convergence in Italy, a country well-known for its spatial disparities (see, among the others, Aiello and Scoppa, 2000; Byrne *et al.*, 2009; Iuzzolino

¹⁶ Poor evidence of convergence across European regions is also obtained by Azomahou *et al.* (2011) using an alternative semi-parametric approach. The authors find evidence of convergence only for some group of regions.

et al., 2013). Specifically, here, we investigate labor productivity convergence over the period 1999-2007 across a sample of Italian firms in the manufacturing and services sectors (for obvious reasons, public utilities and state monopolies have been excluded from the sample). Labor productivity at the firm-level is calculated as the ratio between the value added and the number of employees, both drawn from AIDA, the Italian section of the Bureau Van Dijk Database, a dataset reporting balance-sheet data of Italian companies with a value of production above 100.000 Euros.¹⁷ For both theoretical and statistical considerations we exclude micro firms, i.e. firms with less than 10 employees, from the sample.¹⁸ After some data housekeeping, our sample consists of 21295 firms (see Appendix B for some usual summary statistics). As macro-level of analysis, we consider the 110 Italian NUTS-3 regions (also associated to the administrative level of Provinces). This level of aggregation ensures sufficient degrees of freedom to robustly estimate variability at the second level of our hierarchical structure.

Again, in table 3 we first present a preliminary analysis based on the β convergence regressions, where tests at the disaggregated (firm) and aggregate (regional) levels are performed separately. In these tests, while convergence is identified at the firm-level, divergence is found for the region-level aggregation obtained by summing over the firm-level data (see column 1). However, this evidence depends on the region (province) of Milan. Excluding Milan from the sample leads to concluding in favor of convergence at both the disaggregated micro and the aggregate macro level (column 2). These results highlight the sensitivity of the simple single-level β convergence tests to the influence of single data points. As we will see below, our approach reduces such risk because convergence is not tested on single points, but on estimated growth trajectories, and exploits the full hierarchy of the data, i.e. both micro

¹⁷ Compared to the population of firms with these features, AIDA presents a negligible missing rate, minimizing sample bias issues. For greater detail please visit <http://www.bvdinfo.com>. Value Added has been deflated using sectoral prices from the OECD.

¹⁸ Data on micro and nano firms are usually less reliable and their labour productivity dynamics are obviously more volatile (e.g. the incremental contribution of one additional employee to labour productivity is more relevant in a firm with few employees than in a firm with many employees).

and macro-levels.

<insert table 3 about here>

Table 4 reports on the results of the hierarchical μ convergence analysis. As in the previous example, we first show the results of estimating the models with only random intercepts and with random intercepts and slopes at the lowest level of hierarchy (columns 1 and 2). Then, in column 3 the model with random intercepts and slopes at both levels is presented. Given the preliminary evidence obtained in the β convergence analysis on the influence of a specific region on the results, we also estimate a specification that excludes the region of Milan (column 4).

<insert Table 4 about here>

As in the previous example, the AIC suggests that the specification including random intercepts and slopes at both first and second levels is the most informative.¹⁹ In this specification, an overall negative trend for productivity emerges (the negative parameter γ_1 in columns 3 and 4). This reflects the well-documented Italian productivity puzzle (see Calligaris *et al.*, 2016; Pellegrino and Zingales, 2017). In terms of condition (i), we find evidence of convergence both at the firm and at the regional level. However, condition (ii) is never respected for any of the two levels so that there is evidence of convergence for only part of the period, as we expect an inversion from convergence to divergence during the period. Specifically, there is evidence of firm-level convergence only for the first year and a half and regional convergence for the first five of the seven years. These results imply that during a period of overall decreasing national productivity Italian firms and regions have tended to converge for only part of the period and diverge afterwards, leading us to conclude against

¹⁹ AIC values are 161657, 121395, and 120967 for the first three specifications, respectively

the overall hypothesis of convergence. In figure 2 we can observe this partial convergence at the regional level ($t_n^* = 5$) and at the firm-level ($t^* = 1.5$).

Finally, differently from the preliminary β convergence approach, our results are now unaffected by the inclusion/exclusion of Milan from the sample indicating that this approach is less sensitive to single data points.

<insert figure 2 about here>

4 Conclusions

Most approaches to convergence analysis rely on either the macro or the micro level of observation (e.g. countries, regions, firms). Approaches looking at each of these levels separately, however, can suffer from a number of problems: insufficient flexibility to fully account for parameter heterogeneity, inability to capture micro-macro interactions and lack of consideration for the hierarchical structure of the data.

In this paper, we have illustrated how hierarchical growth curve estimation can be implemented to investigate convergence across different levels of the hierarchy under consideration. Compared to the classic convergence analysis, this approach exploits the growth trajectories of individual units (e.g. regions/firms nested within countries/regions) and simultaneously obtains convergence parameters for both levels of the hierarchy, i.e. the micro-disaggregate and the macro-aggregate level of interest. Additionally, the result of convergence can be validated in terms of the well-known Galton’s fallacy problem, by identifying whether a reduction in dispersion across units occurs for only part of or for the entire period under observation. From the policy standpoint, disentangling convergence across the levels of the hierarchy can be useful to determine at which level (national, regional or both) governments should intervene.

We have performed two applications where we have looked at convergence for both NUTS-2 regions and countries in Europe and firms and NUTS-3 regions in Italy. In both cases, this approach returns evidence that is more robust and more informative for policy purposes. The first example, looking at European regions and countries, gives a number of insights into the process of European convergence, in particular with respect to Eastern enlargement. Specifically, we find statistically significant evidence of convergence at the regional level for the smaller, but longer, sub-samples of 171 or 213 Western European regions over the period 1980-2012. However, this evidence cannot be considered as conclusive since it only applies to the first part of the period under consideration (first three to eight years depending on the sample and specification used). On the other hand, there is no statistically significant evidence of regional convergence, even for just part of the period, in the larger but shorter sample that also includes Central-Eastern European regions. Across countries, we cannot find evidence of convergence for the sub-sample of Western countries over the longer time span. Evidence of convergence emerges when the larger sample, inclusive of Central Eastern countries, is considered over the shorter post-1990 period.

In the Italian example, there seems to be evidence of convergence at both the firm and regional levels (albeit around the well-known national negative trend of productivity), but this evidence is restricted to only part of the sample (the first year and a half for firms and five years for regions) with divergence following after this period. In conclusion, both examples show how the hierarchical convergence approach provides a clearer, richer and more robust picture compared to the classic convergence analysis and it can, therefore, be an additional informative tool for economic policies at the international, national and regional levels to boost growth and reduce spatial disparities. In the European case, our results suggest the need for both EU and national policies that promote cohesion across the regions. In the Italian case, they suggest that sustained convergence requires intervention not only at the usual macro-regional level, where we document the persistence of the well-known disparities, but the

need to mix place-based policies with micro or firm-level policies (for example, reducing the firm-level misallocation of investment documented in recent studies, especially in lagging regions).

From the methodological standpoint, the flexibility of the approach also allows potentially interesting extensions that have been omitted here for brevity. For example, conditioning variables could be easily added at both the macro and micro levels of the hierarchy to look at conditional convergence. The model could be extended to a polynomial specification to look at potential non-linearities in the dynamics of productivity. Also, the approach could be extended along the lines suggested by Corrado and Fingleton (2012) to include a spatial lag structure.

We believe this approach will find further interesting applications in the future, as disaggregated data become increasingly available and the role of heterogeneity becomes an increasingly prominent feature in economic modelling.

References

- Aiello, F., Scoppa, V. (2000). Uneven regional development in Italy: explaining differences in productivity levels. *Giornale degli Economisti e Annali di Economia*, **60**, 270–298.
- Altomonte, C., Colantone, I. (2008). Firm heterogeneity and endogenous regional disparities. *Journal of Economic Geography*, **8**, 779–810.
- Arbia, G., Espa, G., Giuliani, D., Mazzitelli, A. (2010). Detecting the existence of space–time clustering of firms. *Regional Science and Urban Economics*, **40**, 311–323.
- Arbia, G., Espa, G., Giuliani, D., Mazzitelli, A. (2012). Clusters of firms in an inhomogeneous space: The high–tech industries in Milan. *Economic Modelling*, **29**, 3–11.
- Azomahou, T.T., El Ouardighi, J., Nguyen–Van, P., Pham, T.K.C. (2011). Testing convergence of European regions: A semiparametric approach. *Economic Modelling*, **28**, 1202–1210.
- Barro, R.J., Sala-i-Martin, X. (1991). Convergence across States and Regions. *Brookings Papers on Economic Activity*, **22**, 107–182.
- Barro, R.J., Sala-i-Martin, X. (1992). Convergence. *Journal of Political Economy*, **100**, 223–51.
- Bartelsman, E., Haskel, J., Martin, R. (2008). Distance to which frontier? Evidence on productivity convergence from international firm-level data. *CEPR Discussion Paper* **DP7032**.
- Basile, R., de Nardis, S., Pappalardo, C. (2014). Firm heterogeneity and regional business cycles differentials. *Journal of Economic Geography*, **14**, 1087–1115.
- Bernard, A.B., Durlauf, S.N. (1996). Interpreting tests of the convergence hypothesis *Journal of Applied Econometrics*, **71**, 161–173.

- Byrne, J.P., Fazio, G., Piacentino, D. (2009). TFP convergence among Italian regions: some evidence from Panel Unit Roots. *Regional Studies*, **43**, 63–76.
- Calligaris, S., Del Gatto, M., Fadi, H., Ottaviano, G.I.P., Schivardi, F. (2016). Italy's Productivity Conundrum. A Study on Resource Misallocation in Italy. European Commission, Economic and Financial Affairs Discussion Paper 30
- Campbell, H.S., James, R.D., Kunkle, G. (2016). Sustained Firm Growth and Regional Income Convergence. *Papers in Applied Geography*, **2**, 49–65.
- Chasco, C., Lopez, A.M. (2009). Multilevel models: an application to the beta-convergence model. *Region et Developpement*, **30**, 35–58.
- Chevalier, P.A., Lecat, R., Oultn, N. (2012). Convergence of firm-level productivity, globalisation and information technology: Evidence from France. *Economics Letters*, **116**, 244–246.
- Cohen, A.J., Harcourt, G.C. (2003). Whatever happened to the Cambridge capital theory controversy?. *Journal of Economic Perspectives*, **17**, 199–214.
- Corrado, L., Fingleton, B. (2012). Where is the economics in spatial econometrics? *Journal of Regional Science*, **52**, 2, 210–239.
- Dapena, A.D., Morolliçæn, F.R., Paredes, A. (2017a). Are there different local patterns of convergence concealed beneath the regional level? An analysis for US states and counties using a multilevel approach. *Annals of Regional Science*, **58**, 3, 623-640.
- Dapena, A.D., Morolliçæn, F. R., de Moura Pires, M., da Silva Gomes, A. (2017b). Convergence in Brazil: new evidence using a multilevel approach. *Applied Economics*, **49**, 50, 5050-5062.

- Durlauf, S.N., Johnson, P. (1995). Multiple regimes and cross-country growth behaviour. *Journal of Applied Econometrics*, **10**, 4, 365–384.
- Durlauf, S.N., Kourtellos, A., Minkin, A. (2001). The local Solow growth model. *European Economic Review*, **45**, 4-6, 928–940.
- Durlauf, S.N., Johnson, P.A., Temple, J.R.W. (2005). Growth Econometrics. *Handbook of Economic Growth*, Chapter 8, 556–677.
- Egger, P., Pfaffermayr, M. (2009). On testing conditional sigma-convergence. *Oxford Bulletin of Economics and Statistics*, **71**, 453–473.
- Fazio, G., Maltese, E. (2015). Agglomeration Externalities and the Productivity of Italian Firms. *Growth and Change*, **46**, 354–378.
- Fazio, G., Piacentino, D. (2010). A spatial multilevel analysis of Italian SMEs productivity. *Spatial Economic Analysis*, **5**, 299–316.
- Felipe, J., McCombie, J.S.L. (2014). The aggregate production function: 'Not even wrong'. *Review of Political Economy*, **26**, 1, 60–84.
- Garretsen, H., Martin, R. (2011). The Journal of Economic Geography a decade on: where do we go from here?. *Journal of Economic Geography*, **11**, 207–213.
- Gelman, A., Pardoe, I. (2006). Bayesian measures of explained variance and pooling in multilevel (hierarchical) models. *Technometrics*, **48**, 2, 241–251
- Gennaioli, N., La Porta, R., De Silanes, F.L., Shleifer, A. (2014). Growth in regions. *Journal of Economic Growth*, **19**, 259–309.
- Goldstein H. (2011). *Multilevel Statistical Models*. Wiley Series in Probability and Statistics, Wiley.

- Groot, S.P.T., de Groot, H.L.F., Smit, M.J. (2014). Regional wage differences in the Netherlands: micro evidence on agglomeration externalities. *Journal of Regional Science*, **54**, 503–523.
- Haupt, H., Schnurbus, J., Semmler, W. (2017). Estimation of grouped, time-varying convergence in economic growth Author links open overlay panel. *Economics and Econometrics*, in press
- Hausmann, R., Hwang, J., Rodrik, D. (2007). What you export matters. *Journal of Economic Growth*, **12**, 1–25.
- Huber, P., Pfaffermayr, M. (2010). Testing for conditional convergence in variance and skewness: The firm size distribution revisited. *Oxford Bulletin of Economics and Statistics*, **72**, 648–688.
- Islam, N. (2003). What have we learnt from the convergence debate?. *Journal of Economic Surveys*, **17**, 309–362.
- Iuzzolino, G., Pellegrini, G., Viesti, G. (2013). Regional Convergence. In: *The Oxford Handbook of the Italian Economy Since Unification*, G. Toniolo (eds), 571–598.
- Johnson, P.C.D. (2014). Extension of Nakagawa & Schielzeth's R^2_{GLMM} to random slopes models. *Methods in Ecology and Evolution*, **5** 944–946.
- Kremer, K.P., Flower, A., Huang, J., Vaughn, M.G. (2016). Behavior problems and children's academic achievement: A test of growth-curve models with gender and racial differences. *Children and Youth Services Review*, **67**, 95–104.
- Le Gallo, J., Fingleton, B. (2014). Regional growth and convergence empirics. *Handbook of Regional Science*, **1**, 291–315.

- Li, K., Zhou, X., Pan, Z. (2016). Cross-country output convergence and growth: Evidence from varying coefficient nonparametric method. *Economic Modelling*, **55**, 32-41.
- Magrini, S., Gerolimetto, M., Duran, H.E.(2015). Regional Convergence and Aggregate Business Cycle in the United States. *Regional Studies*, **49**, 251–272.
- Martin, R., Sunley, P. (1998). Slow convergence? The new endogenous growth theory and regional development. *Economic Geography*, **74**, 201-227.
- Nakagawa, S., Schielzeth, H. (2013). A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, **4**, 2, 133–142
- Ottaviano, G.I.P. (2011). 'New' new economic geography: firm heterogeneity and agglomeration economies. *Journal of Economic Geography*, **11**, 231–240.
- Paelinck, J.H.P. (2000). On aggregation in spatial econometric modelling. *Journal of Geographical Systems*, **2**, 157–165.
- Pellegrino, B., Zingales, L. (2017). Diagnosing the Italian disease. NBER Working Paper No. 23964
- Phillips, P.C.B., Sul, D. (2007). Some empirics on economic growth under heterogeneous technology. *Journal of Macroeconomics*, **29**, 455–469.
- Rey, S.J., Janikas M.V. (2005). Regional convergence, inequality, and space. *Journal of Economic Geography*, **5**, 155–176.
- Rizov, M., Zhang, X. (2014). Regional disparities and productivity in China: Evidence from manufacturing micro data. *Papers in Regional Science*, **93**, 321–340.
- Robinson, W.S. (1950). Ecological correlations and the behaviour of individuals. *American Sociological Review*, **15**, 351–357.

- Rodrik, D. (2013). Unconditional convergence in manufacturing. *Quarterly Journal of Economics*, 128, 165–204.
- Sala-i-Martin, X. (1996). Regional cohesion: Evidence and theories of regional growth and convergence. *European Economic Review*, 40, 1325–1352.
- Solow, R. (1956). A contribution to the theory of economic growth. *Quarterly Journal of Economics*, 70, 65–94.
- Solow, R. (1957). Technical change and aggregate production function. *Review of Economics and Statistics*, 39, 312–320.
- Steele, F. (2008). Multilevel models for longitudinal data. *Journal of The Royal Statistical Society Series A*, 171, 5–19.
- Venables, A.J. (2011). Productivity in cities: self-selection and sorting. *Journal of Economic Geography*, 11(2), 241–251.
- von Ehrlich, M., Seidel, T. (2013). More similar firms – More similar regions? On the role of firm heterogeneity for agglomeration. *Regional Science and Urban Economics*, 43, 539–548.

APPENDIX A: Hierarchical longitudinal growth and convergence

Figure A1 illustrates possible alternative convergence results depending on different starting and ending configurations of individual units.

APPENDIX B: Summary statistics

Table B1 - Summary Statistics. European data (Sample: 1980-2012 - 8 countries)

<i>GDP per capita (Y/L)</i>			
<i>Year</i>	<i>Regions</i>	<i>Mean</i>	<i>Std. Dev.</i>
1980	160	17317.97	5840.285
1981	160	17483.98	5937.41
1982	160	17611.15	5780.374
1983	160	17924.3	5811.176
1984	160	18433.58	6108.93
1985	160	18911.85	6268.688
1986	161	19207.07	6127.675
1987	161	19652.7	6004.12
1988	161	20386.42	6089.451
1989	161	20975.87	6145.389
1990	161	21442.92	6362.012
1991	171	21339.98	6190.795
1992	171	21535.24	6400.905
1993	171	21424.19	6577.554
1994	171	22048.2	6835.531
1995	171	22545.18	6928.921
1996	171	22885.57	7035.933
1997	171	23471.56	7229.864
1998	171	24062.2	7421.457
1999	171	24640.5	7647.93
2000	171	25424.55	7942.647
2001	171	25836.3	8077.892
2002	171	25991.24	8068.308
2003	171	26207.69	8201.316
2004	171	26646.75	8309.248
2005	171	26991.78	8557.879
2006	171	27744.73	8856.309
2007	171	28309.89	9114.837
2008	171	28247.66	9286.427
2009	171	26907.39	8775.933
2010	171	27389.04	8943.119
2011	171	27798.8	9039.89
2012	171	27552.01	8986.887

Table B2 - Summary Statistics. European data (Sample: 1991-2012 - 26 countries)

<i>Year</i>	<i>GDP per capita (Y/L)</i>		
	<i>Obs</i>	<i>Mean</i>	<i>Std. Dev.</i>
1991	277	18018.58	9315.5
1992	277	18133.27	9513.878
1993	277	18043.69	9601.158
1994	277	18572.65	9935.118
1995	277	19031.48	10137.17
1996	277	19395.9	10343.3
1997	277	19961.22	10662.07
1998	277	20512.57	10959.41
1999	277	21042.01	11275.95
2000	277	21701.92	11661.59
2001	277	22081.83	11853.65
2002	277	22298.2	11892.93
2003	277	22569.25	11919.67
2004	277	23072.07	12124.53
2005	277	23419.37	12340.67
2006	277	24129.28	12618.61
2007	277	24703.25	12870.38
2008	277	24678.49	12794.58
2009	277	23504.39	12192.23
2010	277	23875.73	12401.19
2011	277	24160.93	12544.73
2012	277	23956.44	12539.28

Table B3 - Summary Statistics. European data (Sample: 1980-2012 - 16 countries)

<i>GDP per capita (Y/L)</i>			
<i>Year</i>	<i>Obs</i>	<i>Mean</i>	<i>Std. Dev.</i>
1980	213	17324.25	6331.258
1981	213	17469.21	6399.063
1982	213	17599.94	6305.708
1983	213	17914.95	6447.85
1984	213	18421.25	6811.057
1985	213	18933.01	6967.112
1986	213	19258.36	7033.598
1987	213	19681.99	7015.485
1988	213	20352.45	6993.724
1989	213	20922.62	7054.36
1990	213	21352.34	7212.405
1991	213	21354.91	7219.369
1992	213	21635.13	7406.051
1993	213	21565.37	7601.704
1994	213	22213.64	7927.342
1995	213	22724.19	8121.372
1996	213	23133.23	8343.463
1997	213	23816.06	8621.058
1998	213	24486.65	8860.384
1999	213	25122.75	9133.1
2000	213	25922.37	9465.158
2001	213	26358.27	9648.337
2002	213	26583.03	9688.081
2003	213	26871.06	9705.566
2004	213	27421.94	9926.204
2005	213	27778.42	10228.39
2006	213	28554.49	10496.42
2007	213	29153.86	10799.32
2008	213	29036.66	10824.95
2009	213	27636.17	10339.89
2010	213	28043.42	10550.97
2011	213	28323.82	10749.29
2012	213	28068.04	10850.07

Table B4 - Summary Statistics. Italian data

<i>Labor Productivity (Y/L)</i>			
<i>Year</i>	<i>Obs</i>	<i>Mean</i>	<i>Std. Dev.</i>
1999	21295	49322.97	71831.16
2000	21295	50185.37	51063.03
2001	21295	46429.36	43998.79
2002	21295	45706.09	35593.47
2003	21295	46170.07	32321.84
2004	21295	53822.41	47211.46
2005	21295	54372.68	51443.84
2006	21295	57400.07	57474.18
2007	21295	59793.69	63856.86

Figure 1: Convergence in GDP per capita - 26 European countries
 (a) Expected variance (b) Empirical data

Figure 2: Expected variance of labor productivity in Italian firms and regions
 (a) Firms (b) NUTS-3 Regions

Notes: the region of Milan is excluded.

Notes: the region of Milan is excluded.

Table 1: β convergence - European regions and countries

		(1)	(2)	(3)
Regions	β	-0.0060 (0.0013)	-0.0034 (0.0013)	-0.0069 (0.0007)
	$Adj-R^2$	0.11	0.02	0.23
	<i>Period</i>	1980-2012	1980-2012	1991-2012
	<i>Obs.</i>	171	213	277
		convergence	convergence	convergence
Countries	β	0.0054 (0.0037)	0.0015 (0.0002)	0.0011 (0.0002)
	$Adj-R^2$	0.13	0.69	0.51
	<i>Period</i>	1980-2012	1980-2012	1991-2012
	<i>Obs.</i>	8	16	26
		not significant	divergence	divergence

Notes: Estimation by OLS; Standard Errors in parentheses. All variables are in logs

Table 2: Hierarchical μ convergence - European regions and countries

		(1)	(2)	(3)	(4)	(5)
Regions	γ_0	9.7216 (0.0202)	9.7245 (0.0216)	9.7471 (0.070)	9.7560 (0.0727)	9.3150 (0.1858)
	γ_1	0.0170 (0.0001)	0.0169 (0.0005)	0.0169 (0.0014)	0.0188 (0.0014)	0.0239 (0.0025)
	σ_ε^2	0.0067 (0.0001)	0.0031 (0.0001)	0.0031 (0.0001)	0.0039 (0.0001)	0.0031 (0.0000)
	$\sigma_{\eta 0}^2$	0.0689 (0.0075)	0.0801 (0.0087)	0.0580 (0.0065)	0.0631 (0.0064)	0.0563 (0.0051)
	$\sigma_{\eta 01}$		-0.0006 (0.0001)	-0.0003 (0.0001)	-0.0008 (0.0001)	-0.0001 (0.0001)
	$\sigma_{\eta 1}^2$		0.0001 (0.0000)	0.0001 (0.000)	0.0001 (0.000)	0.0001 (0.000)
	μ_i		-0.0082 (0.0016)	-0.0061 (0.0016)	-0.0124 (0.0016)	-0.0021 (0.0016)
	condition (i)		yes	yes	yes	no
	t_i^*	-	6	3	8	-
	condition (ii)		no	no	no	-
Countries	σ_{v0}^2		-	0.0352 (0.0217)	0.0755 (0.0320)	0.8825 (0.2558)
	σ_{v01}		-	-0.0003 (0.0004)	0.0004 (0.0004)	-0.0082 (0.0030)
	σ_{v1}^2		-	0.0001 (0.000)	0.0001 (0.000)	0.0001 (0.000)
	μ_n		-	-0.0094 (0.0076)	0.0055 (0.0059)	-0.0093 (0.0020)
	condition (i)			no	no	yes
	t_n^*	-	-	-	-	82
	condition (ii)		-	-	-	yes
	$\log(L)$	5478.6	7252.5	7297.5	8456.9	7659.1
	<i>Regions</i>	171	171	171	213	277
	<i>Countries</i>	8	8	8	16	26
	<i>Period</i>	1980-2012	1980-2012	1980-2012	1980-2012	1991-2012
	<i>Obs.</i>	5527	5527	5527	7023	6094

Notes: Estimation by Restricted Maximum Likelihood; Standard Errors in parentheses. All variables are in logs

Table 3: β convergence - Italian firms and regions example

		(1)	(2)
Firms	β	-00156	-0.0152
		(0.0006)	(0.0006)
	$Adj-R^2$	0.03	0.03
	<i>Period</i>	1999-2007	1999-2007
	<i>Obs.</i>	21295	18539
		convergence	convergence
Regions	β	0.0002	-0.0002
		(0.0000)	(0.0000)
	$Adj-R^2$	0.17	0.18
	<i>Period</i>	1999-2007	1999-2007
	<i>Obs.</i>	110	109
		divergence	convergence

Notes: Estimation by OLS; Standard Errors in parentheses. All variables are in logs. In columns (2), the region of Milan is excluded.

Table 4: Hierarchical μ convergence - Italian firms and regions

		(1)	(2)	(3)	(4)
Firms	γ_0	12.7941 (0.0050)	12.7941 (0.0049)	12.8460 (0.0137)	12.8476 (0.0137)
	γ_1	0.0010 (0.0003)	0.0010 (0.0005)	-0.0037 (0.0011)	-0.0040 (0.0010)
	σ_ε^2	0.0876 (0.0003)	0.0562 (0.0002)	0.05262 (0.0001)	0.0521 (0.0002)
	$\sigma_{\eta 0}^2$	0.05027 (0.0049)	0.4924 (0.0050)	0.4835 (0.0049)	0.4783 (0.0052)
	$\sigma_{\eta 01}$		-0.0066 (0.0003)	-0.0063 (0.0003)	-0.0060 (0.0004)
	$\sigma_{\eta 1}^2$		0.0042 (0.0000)	0.0041 (0.0000)	0.0041 (0.0000)
	μ_i	-	-0.0135 (0.0007)	-0.0131 (0.0007)	-0.0125 (0.0007)
	condition (i)		yes	yes	yes
	t_i^*	-	1.58	1.54	1.47
	condition (ii)	-	no	no	no
Regions	σ_{v0}^2			0.0133 (0.0029)	0.0133 (0.0029)
	σ_{v01}			-0.0003 (0.0002)	-0.0003 (0.0002)
	σ_{v1}^2			0.0001 (0.0000)	0.0001 (0.0000)
	μ_n	-	-	-0.0264 (0.0112)	-0.0222 (0.0110)
	condition (i)		-	yes	yes
	t_n^*	-	-	5.14	5.16
	condition (ii)	-	-	no	no
	$\text{Log}(L)$	-80824.31	-60691.32	-60474.40	-47453.91
	<i>Firms</i>	21295	21295	21295	18539
	<i>Regions</i>	110	110	110	109
	<i>Period</i>	1999-2007	1999-2007	1999-2007	1999-2007
	<i>Obs.</i>	191655	191655	191655	166851

Notes: Estimation by Restricted Maximum Likelihood; Standard Errors in parentheses.

All variables are in logs. In column (4), the region of Milan is excluded.