Regional Convergence in Europe: A Dynamic Heterogeneous Panel Approach

Vogel, Johanna

December 2013

Regional Convergence in Europe: A Dynamic Heterogeneous Panel Approach

Johanna Vogel*†
Austrian Institute of Economic Research

This version: December 2013

Abstract

This paper studies the effects of allowing for heterogeneous slope coefficients in the Mankiw, Romer and Weil (1992) model, based on panel data for 193 EU-15 regions from 1980 to 2005. We first estimate the model using conventional pooled panel data estimators, based on data at five-year intervals, allowing at most intercepts to differ across regions. Then we relax the restriction of homogeneous slope coefficients by estimating separate time-series models for each region based on annual data, using Pesaran and Smith's (1995) mean group estimator. To account for spatial dependence, we employ the common correlated effects approach of Pesaran (2006). Our empirical analysis indicates important differences across regions in the speed of adjustment to region-specific long-run paths for the level of income per capita. Allowing for heterogeneous coefficients doubles the speed of adjustment to 22% per year on average compared to the homogenous case, which suggests downward bias in the latter. We also find a positive and significant effect of the rate of investment, although the implied capital elasticity and the estimated long-run effect of investment are smaller than expected.

Keywords: Convergence, European regions, dynamic heterogeneous panels, mean group estimation, cross-section dependence.

JEL Classification: O40, R11, C23.

---

*Correspondence: Austrian Institute of Economic Research, Arsenal Objekt 20, 1030 Vienna, Austria; johanna.vogel@wifo.ac.at.

†I am grateful to Steve Bond and Simon Wren-Lewis for helpful comments and advice on earlier versions of this paper. I would also like to thank Jonathan Stenning at Cambridge Econometrics for providing access to and help with the European Regional Database. All remaining errors are my own. Financial support by the ESRC (award number PTA-031-2004-00246) is gratefully acknowledged. The views expressed here are those of the author and do not necessarily reflect those of the Austrian Institute of Economic Research.

1 Introduction:

The evolution of income per capita across the regions of the European Union has been a key policy concern since the 1986 Single European Act, when the community's regional policy was formally established. One of the primary objectives of EU regional policy is the reduction of regional income disparities, or regional convergence. For the period 2007 to 2013 for example, 283 billion euro, or 81% of the total regional policy budget, were available to foster growth and catch-up of regions with per-capita incomes below 75% of the EU average.

Considering the EU-15, total cross-regional disparities have tended to decline since the 1980s, as indicated by measures of the dispersion of per-capita income such as the coefficient of variation (European Commission, 2010). Nevertheless, persistent differences between regions have remained. For example, the Greek region Epirus, located to the south of Albania, and the Portuguese Central region were at the bottom of the per-capita income distribution across the EU-15 regions in 1980, 1990 and 2000. On the other hand, Brussels, Hamburg and Inner London were at or near the top of the distribution at these points in time. More generally, levels of regional economic development have been much lower in the peripheral regions of Greece, Portugal, Spain and Southern Italy compared to the richer capital regions and the urban and industrial agglomerations in the centre of Europe. In the literature, the term "blue banana" is sometimes used to describe the curved area of high economic activity and high population density that stretches from the West Midlands in the UK across Belgium and parts of the Netherlands, the Rhine-Ruhr metropolitan area and south-west Germany into Northern Italy.[1]

These persistent disparities in regional per-capita incomes suggest that growth processes differ considerably between the European regions. For example, an increase in a growth driver like investment in physical capital probably has a different impact in Hamburg than in Epirus due to a number of region-specific factors including economic institutions and geographic proximity to economic centres. The speed at which regions adjust to shocks is likely to vary across Europe for similar reasons. In this paper, we investigate heterogeneity in the speed of adjustment as well as the effects of some growth determinants across the regions of the EU-15.

A substantial part of the empirical literature on regional growth in the European Union has estimated specifications for the short-run growth rate of per-capita income along a transition path to the long-run equilibrium, similar to the specification derived by Mankiw et al. (1992) from the neoclassical Solow (1956) model. Regional studies in this tradition have focused on two main themes. The first is the question of regional convergence, which has met with great empirical interest in light of the convergence objective of EU regional policy. The second theme is the development and application of spatial econometric methods

[1] The term originates from satellite images of Europe at night, where the light emanating from the banana-shaped densely populated area glows blue.

suited to the study of regional growth. Research in this direction has been partly motivated by interest in the size and sign of regional spillovers, and partly by concern about model misspecification that could arise when neglecting spatial correlation in regional data.

While the regional growth literature in the tradition of Mankiw et al. (1992) now commonly employs panel data methods that allow for region-specific intercepts, it has largely ignored heterogeneity in the slope coefficients. However, if there are good reasons to expect these coefficients to differ between cross-sectional units, Pesaran and Smith (1995) show that imposing common parameters may render standard pooled panel data estimators inconsistent. Some recent surveys of empirical growth analysis (Durlauf, Johnson and Temple 2005; Eberhardt and Teal 2011) also emphasise the importance of taking into account heterogeneous effects of explanatory variables across countries or regions.

In this paper, we address this gap in the literature by investigating the effect of parameter heterogeneity on the estimated coefficients of Mankiw et al.'s (1992) model of transitional growth in a panel data framework. Our dataset covers 193 regions from the EU-15 countries over the period 1980 to 2005. Similar to existing work, we first estimate the model using conventional pooled estimators for dynamic panel data models that allow at most intercepts to differ across regions. For this purpose, we split our data into five-year intervals to control for short-run fluctuations such as business cycles. We then allow for heterogeneous slope coefficients by estimating separate time-series models for individual regions using Pesaran and Smith's (1995) mean group estimator. This is implemented with data at annual frequency. Richer dynamics are introduced into the model to control for short-run fluctuations, and pooled annual specifications are presented for comparison.

To deal with spatial correlation in our regional data, we use the common correlated effects (CCE) approach of Pesaran (2006). His CCEMG estimator is straightforward to implement as an extension of the mean group estimator in our empirical setup. The CCE approach allows for a general form of cross-section error dependence and has been shown to remain valid in the presence of the forms of spatial error correlation typically considered in spatial econometrics. To our knowledge, it has so far not been considered in the context of the Mankiw et al. (1992) model applied to the EU regions.

Our empirical analysis highlights the importance of allowing for heterogeneity across regions in the speed of adjustment to region-specific long-run paths for the level of income per capita. When using the mean group approach, we obtain an average estimate of this parameter that is twice as large as when using conventional pooled panel data methods. This provides further support for the few empirical studies on this issue that exist at the regional level. It also suggests that much of the literature on regional convergence, which largely uses pooled panel data methods, may substantially underestimate the speed of conditional convergence.

In addition, we find a positive and significant effect of the rate of investment throughout. The size of the implied elasticity of output with respect to capital is robust across pooled and heterogeneous panel estimators, but it is smaller than we would expect based on the Mankiw et al. (1992) model and macroeconomic data on factor income shares. Our preferred estimate of the long-run elasticity of income with respect to investment is also smaller than expected. There is less robust evidence of a significant long-run effect of population growth on regional income levels.

The remainder of this paper is organised as follows. The next section reviews relevant empirical work on European regional growth in the tradition of Mankiw et al. (1992). Section 3 outlines the theoretical framework and our empirical specifications with homogeneous and heterogeneous coefficients. Sections 4 and 5 provide an overview of the estimation methods and data we use. Section 6 presents and discusses our empirical results, and section 7 concludes.

2 Review of the Empirical Literature:

The literature on European regional growth has proceeded through several stages, in parallel with developments in econometric methods and data availability. Early studies predominantly employed cross-section least-squares regressions to investigate convergence across regions. For lack of data on control variables, at most country dummies or variables describing regional economic structure were usually included as regressors. Examples are Barro and Sala-i-Martin (1991), who find a speed of convergence of about 2% per year for European regions from 1950 to 1985, as well as Armstrong (1995) and Tondl (1999), who report similar or lower rates for comparable time periods.

As longer time series on a greater number of variables became available at the EU regional level, broader models like Mankiw et al. (1992) (henceforth also MRW) were estimated, increasingly using panel data methods.[2] This review focuses primarily on results from the panel data literature, since our contribution in this paper is concerned with relaxing the restriction of homogeneous slope coefficients imposed by conventional pooled panel data estimators.

Mankiw et al. (1992) themselves use the cross-section regression framework for different samples of countries to estimate a specification for the short-run growth rate of per-capita income in transition to the long-run equilibrium. The average growth rate over the time period under consideration, 1960 to 1985, is modelled as a positive function of the rate of investment and a negative function of the growth rate of the labour force. In addition, the growth rate depends negatively on income per capita in the initial period, implying convergence to the (country-specific) long-run path for the level of income.[3] Since the model predicts that the coefficients on investment and labour force growth should be equal in

[2] We refer to the simple version of the Mankiw et al. (1992) model that does not include human capital, since regional data on this variable are not available for the full time period covered in this paper.

[3] The literature reviewed in this section is concerned with *conditional* convergence. See section 3.

magnitude but opposite in sign, it can also be estimated in restricted form, with an expected positive coefficient on the difference between these two variables. The rate of convergence to the long-run equilibrium and the elasticity of output with respect to capital can then be derived from the coefficient estimates. Together with macroeconomic data on factor income shares, the theoretical model implies a capital elasticity of about 0.33. MRW's own empirical findings substantially exceed this value unless the model is augmented with human capital.

Badinger, Müller and Tondl (2004) estimate MRW's transitional growth specification for a panel of 196 regions from the EU-15 countries, using data at five-year intervals spanning the period 1975 to 1999. They apply a spatial filtering method to control for spatial dependence in their data and employ pooled panel data methods that allow at most for region-specific intercepts ("fixed" effects). Apart from pooled OLS, these are the within-groups estimator popularised in the empirical growth literature by Islam (1995), and the first-differenced and system-GMM estimators developed by Arellano and Bond (1991) and Blundell and Bond (1998). We use all of these to estimate the MRW model with homogeneous slope coefficients in the first part of our empirical analysis.

Badinger et al. (2004) find some support for the MRW model, which they estimate in its restricted form. They obtain a positive and significant coefficient on the difference between investment and population growth rates across estimation methods. There is also significant evidence of convergence to region-specific long-run paths for income. However, the size of the implied structural parameters differs considerably across estimators. The rate of convergence ranges between 1.5% per year with pooled OLS to 25% per year using within groups, while the first-differenced and system-GMM estimates lie in-between at 18.4% and 7% respectively. The implied capital elasticity varies between 0.64 for pooled OLS and 0.18 using within groups. The first-differenced and system-GMM estimates of this parameter are 0.28 and 0.43 and thus more in line with the MRW model.

The variability of Badinger et al.'s (2004) results highlights estimation issues with pooled panel data estimators that have also been raised in studies at the country level. For example, Islam (1995) obtains a comparatively low rate of convergence and a high capital elasticity when using pooled OLS for the same samples as MRW, while within groups yields a higher rate of convergence and a lower capital elasticity. Results in Caselli, Esquivel and Lefort (1996) reveal a similar tendency for first-differenced GMM. Using Caselli et al.'s (1996) sample, Bond, Hoeffler and Temple (2001) find that the system-GMM estimate of the rate of convergence lies between pooled OLS and within groups or first-differenced GMM. It has been shown that this may be due to bias in the last three estimators, and Badinger et al. (2004) give preference to their system-GMM results for this reason.

Elhorst, Piras and Arbia (2010) evaluate the transitional growth model of MRW, also in its restricted form, as well as a spatial extension of it, where the conventional model is augmented with spatially weighted dependent and independent variables. They consider

193 EU-15 regions from 1977 to 2002. In the a-spatial case, results are obtained using cross-section OLS as well as pooled OLS and first-differenced GMM based on five-yearly panel data; they partly support the MRW model: there is significant evidence of convergence to region-specific long-run income paths, and the coefficient on the difference between investment and population growth is positive and significant for cross-section and pooled OLS. For first-differenced GMM, it is not significantly different from zero. Pooled OLS yields an implied capital elasticity of about 0.31, which is consistent with the MRW model. The implied annual speed of convergence is 0.9% with pooled OLS and 7.8% for first-differenced GMM. The latter estimate is thus considerably lower than the corresponding figure in Badinger et al. (2004). However, in contrast to these authors, Elhorst et al. (2010) do not report tests of the validity of their first-differenced GMM instruments, so the reliability of these estimates is unclear. The spatial estimates of the speed of convergence are very similar to the conventional ones, but the estimated coefficients on the investment variable are never significant and the implied capital elasticity is sometimes negative.

Bouayad-Agha and Védrine (2010) compare the MRW model in its unrestricted form with an extension that includes spatially weighted versions of the dependent and some independent variables as regressors. Their study focuses on the speed of convergence and no implied capital elasticities are presented, but it considers the same time period and employs similar estimation methods as we do in this paper. The sample consists of 191 EU-15 regions observed for five-yearly periods from 1980 to 2005. The a-spatial version of the model is estimated using pooled OLS, within groups and first-differenced GMM, but the results are mixed. There is significant evidence of convergence across estimators, and the coefficient on the investment rate - which is measured as investment divided by population rather than output - is positive and significant throughout. However, the coefficient on the population growth rate is either positive and significant, which runs counter to the model's prediction, or negative but insignificant. Further, the instruments that the authors use for first-differenced GMM are invalid. The implied speed of convergence varies from a low rate of 2% per year obtained with pooled OLS to a higher rate of 14.7% per year with first-differenced GMM. The latter figure rises to between 15.1 and 21.5% when the spatially weighted variables are added, which also improves GMM instrument validity.

In sum, the results from the pooled panel data literature that examines the MRW model using five-yearly data for the EU regions are generally consistent with the model's predictions of convergence (to region-specific long-run income paths) as well as a positive and significant effect of investment. However, the studies do not agree on the size of the implied structural parameters. Regarding the speed of convergence, the observed variability in results described above is likely due to bias in some estimators. It is noteworthy that Elhorst et al. (2010) do not use the system-GMM estimator, which may be less prone to bias in this context.[4] By contrast, we employ system GMM in the first part of our empirical

[4]Bouayad-Agha and Védrine (2010) report invalid instruments with system GMM.

analysis to estimate the MRW model with homogeneous slope coefficients, as do Badinger et al. (2004). Compared to the latter, we investigate a more recent time period.

The empirical literature reviewed so far does not allow the explanatory variables in the MRW model to have a heterogeneous impact across regions. One way of doing this is to estimate the MRW model separately for subgroups of regions that are likely to share similar parameter values (Durlauf, Johnson and Temple 2005).[5] We take the alternative route of allowing the parameters of the MRW model to vary across all European regions using Pesaran and Smith's (1995) mean group estimator.

At the country level, this approach is taken by Bond, Leblebicioglu and Schiantarelli (2010) to allow for a heterogeneous effect of investment on the long-run growth rate of GDP per worker. A related literature considers heterogeneous univariate panel data models of per-capita income to study convergence: Lee, Pesaran and Smith (1997) employ the mean group estimator to estimate separate time-series models for individual countries in samples similar to MRW and Islam (1995). The parameters of interest, obtained using annual data, are then averaged over countries to obtain the mean group estimates. The authors find that allowing for heterogeneous speeds of convergence in addition to heterogeneous intercepts raises the average speed of convergence across countries. When also allowing for heterogeneity in the long-run growth rate of per-capita income, captured by the coefficient on a linear time trend, the average speed of convergence increases further to values between 24% and 30% per year, depending on the sample.

Meliciani and Peracchi (2006) apply Lee et al.'s (1997) framework to annual data for 95 EU regions from 1980 to 2000 and find a similar pattern. The speed of convergence increases from 13% per annum when only intercepts are heterogeneous to 19% on average when the speed of convergence and the long-run growth rate are also allowed to be region-specific. Canova and Marcet (1995) use Bayesian methods to estimate an AR(1) model for per-capita income on annual panel data for 114 EU-15 regions from 1980 to 1992. When region-specific intercepts as well as heterogeneous speeds of convergence are allowed for, the estimated average speed of convergence across regions amounts to 23% per year, compared to only 2% when all parameters are restricted to be homogeneous.

Neither of the last two region-level studies considers heterogeneity in the effect of additional explanatory variables. Although the empirical specification in Lee et al. (1997) and thus in Meliciani and Peracchi (2006) is based on a version of the Solow model, investment and population growth rates are assumed constant over time and subsumed into country- and region-specific intercepts. By contrast, we explicitly allow for heterogeneous impacts of investment and population growth rates in our MRW specification.

[5] A related but distinct branch of the literature is concerned with identifying convergence clubs, which are groups of regions with similar structural characteristics and initial conditions that converge to similar long-run levels of income. Examples for the European regions include Corrado, Martin and Weeks (2005) and Fischer and Stirböck (2006). However, this approach is not concerned with estimating the heterogeneous effects of explanatory variables across regions.

3 Theoretical Framework and Empirical Specifications:

This section outlines the Mankiw et al. (1992) model of per-capita income growth in transition to the long-run equilibrium, which is based on the on the neoclassical model of Solow (1956) and Swan (1956). We then discuss in more detail our empirical specifications with homogeneous and heterogeneous slope coefficients across regions.

Output Y at time t is produced according to the following Cobb-Douglas production function:

$$Y(t) \;=\; K(t)^{\alpha}\,(A(t)L(t))^{1-\alpha} \qquad 0<\alpha<1. \tag{1}$$

K is physical capital, L is labour, A represents the level of technology and the constant α is the elasticity of output with respect to capital. This production function exhibits constant returns to scale and diminishing marginal returns to the inputs capital and labour. Labour and technology grow exponentially at the constant and exogenous rates n and g, respectively, so that $A(t)L(t)$ - representing "effective" units of labour - grows at rate $(n+g)$.

The capital stock accumulates over time according to

$$\dot{K}(t) \;=\; sY(t) \;-\; \delta K(t), \tag{2}$$

where the saving rate $s = S/Y$, a constant and exogenous share of output, equals the investment rate since the economy is assumed closed. $\delta > 0$ is the constant depreciation rate of capital. Rewriting equations (1) and (2) in terms of units of effective labour, where e.g. $\hat{y}(t) = \frac{Y(t)}{A(t)L(t)}$, yields

$$\hat{y}(t) = \hat{k}(t)^{\alpha} \tag{3}$$

for the production function, and similarly

$$\dot{\hat{k}}(t) \;=\; s\hat{k}(t)^{\alpha} \;-\; (n+g+\delta)\hat{k}(t) \tag{4}$$

for the evolution of the capital stock per effective worker.

In the long-run equilibrium of the model, also called the steady state or balanced growth path, all variables in units of effective labour are constant. The steady-state value of $\hat{k}(t)$, $\hat{k}^*$, can thus be solved for by setting $\dot{\hat{k}}(t) = 0$ in equation (4):

$$\hat{k}^* \;=\; \left(\frac{s}{n+g+\delta}\right)^{\frac{1}{1-\alpha}} \quad \text{and consequently,} \quad \hat{y}^* \;=\; \left(\frac{s}{n+g+\delta}\right)^{\frac{\alpha}{1-\alpha}}. \tag{5}$$

In per-capita terms, where $y(t) = Y(t)/L(t)$, an economy's steady-state log-level of income at time t is then given by

$$\ln y^*(t) \;=\; \ln A(0) \;+\; gt \;+\; \frac{\alpha}{1-\alpha}\ln(s) \;-\; \frac{\alpha}{1-\alpha}\ln(n+g+\delta). \tag{6}$$

Thus, the higher is the rate of investment for example, the higher is the economy's level of income per capita in the steady state, holding everything else constant. The same holds the lower are the growth rate of effective labour and the depreciation rate.

While the growth rates of capital and output per effective worker are zero in the steady state, it can be shown that capital and output *per capita* grow at the common exogenous rate of technological progress g - that is, they evolve over time along a path with constant slope g. Long-run per-capita economic growth in the Solow-Swan model is therefore by definition exogenous, in that it is not explained from within the system by the behavioural parameters of the model, but rather exogenously given from outside. Changes in any of the explanatory variables in equation (6) affect the steady-state level of income but not its steady-state growth rate.

For any initial values of $\hat{k}(t)$ and $\hat{y}(t)$ that lie off (but close to) the steady state path, a process of convergence takes place over time until the economy reaches the long-run equilibrium given in (5). The growth rate of income per effective worker during this transitional period, as well as the speed at which convergence takes place, can be obtained from a log-linear approximation around the steady state, which yields

$$\frac{\dot{\hat{y}}(t)}{\hat{y}(t)} = \frac{d\ln\hat{y}(t)}{dt} \approx \beta \cdot \left(\ln\hat{y}^* - \ln\hat{y}(t)\right), \tag{7}$$

where $\beta = (1-\alpha)(n+g+\delta)$ is the rate of convergence. Equation (7) illustrates that the transitional growth rate of income per effective worker is positive for values of $\ln\hat{y}(t)$ below the steady state and negative for values above it. The growth rate is also proportional to the distance between log-income per effective worker at time t and its steady-state value. The greater (more positive) this distance is, the higher is the transitional growth rate, and as $\ln\hat{y}(t)$ approaches $\ln\hat{y}^*$ over time, the growth rate declines until it reaches zero in the steady state. β is the speed at which the gap is closed, and hence this notion of convergence is usually referred to as β-convergence.

The key property of the neoclassical growth model that gives rise to convergence is that of diminishing marginal returns to capital. At low levels of capital per effective worker, the marginal product of $\hat{k}(t)$ is high, and so are the growth rates of $\hat{k}(t)$ and thus $\hat{y}(t)$. As $\hat{k}(t)$ rises over time, its marginal product declines, and the growth rate of $\hat{y}(t)$ slows down as the economy approaches the steady state.

A considerable part of the empirical literature on economic growth - particularly for the EU regions - has, based on models like Mankiw et al. (1992), focused on whether convergence has taken place *across* several countries or regions over a given time period. The original Solow model, however, is concerned with the evolution of capital and income *within* a single economy over time, as highlighted by the absence of unit-specific i-subscripts in our outline of the model so far. Islam (1998, p. 327) refers to this as the "tension between the *within* and *across* dimensions of the concept of convergence". To reconcile the cross-country or cross-regional approach with the Solow model, Mankiw et al. (1992) and Barro and Sala-i-Martin (1992) distinguish between the notions of absolute and conditional β-convergence, which can also be illustrated by means of equations (6) and (7).

Absolute convergence refers to a situation where a group of economies converges to the same steady-state level of per-capita income, so that initially poorer regions should grow faster than richer ones along the transition path until they eventually catch up with the rich. For this outcome to be consistent with the Solow model however, the group of regions in question would have to be structurally identical in terms of their production functions and the characteristics $A(0)$, s, n, g and δ that determine the (common) steady-state income $\hat{y}^*$. Conditional convergence, on the other hand, implies that each economy converges to its *own* steady-state level of income, with growth faster the further it is away from it. Then, initially poorer regions should grow faster than richer ones only after controlling for differences in the determinants of their steady states. This concept of convergence may be more realistic for a heterogeneous set of regions with diverse production functions, saving rates and other characteristics.

From equation (7), an expression can be derived for the transitional growth rate of per-capita income between two points in time that are τ years apart:[6]

$$\begin{aligned} \ln y_t - \ln y_{t-\tau} \;=\; & \left(1 - e^{-\beta\tau}\right) \frac{\alpha}{1-\alpha} \ln(s) \;-\; \left(1 - e^{-\beta\tau}\right) \frac{\alpha}{1-\alpha} \ln(n+g+\delta) \;- \\ & -\; \left(1 - e^{-\beta\tau}\right) \ln y_{t-\tau} \;+\; \left(1 - e^{-\beta\tau}\right) \ln A(0) \;+ \\ & +\; g\left(t - e^{-\beta\tau}(t-\tau)\right) \end{aligned} \tag{8}$$

MRW apply this model to time-averaged data on y, s and n for a single cross-section of countries, so that all regression coefficients are restricted to be homogeneous across countries. g and δ are assumed to be common to all countries and constant over time, with their sum equal to 0.05. MRW recognise that the unobservable initial level of technology $\ln A(0)$, which they define more broadly to include also physical geography, climate and economic institutions, may differ across countries. To estimate equation (8) using OLS, they therefore replace $\ln A(0)$ with a constant and a country-specific error term that is assumed to be uncorrelated with the other regressors in the econometric model. This last assumption is not entirely consistent with the theoretical model, however. From equation (6), the country-specific component of $\ln A(0)$ is a determinant of steady-state income $\ln y_t^*$ and hence also of $\ln y_{t-\tau}^*$, so that it is by construction correlated with $\ln y_{t-\tau}$ on the right-hand side of equation (8). Given MRW's own broad interpretation of $\ln A(0)$,

[6] An intermediate step between equations (7) and (8) is solving the former for $\ln \hat{y}(t)$, which yields $\ln \hat{y}(t) = \left(1 - e^{-\beta t}\right) \ln \hat{y}^* + e^{-\beta t} \ln \hat{y}(0)$. Barro and Sala-i-Martin (1992) derive the same equation from a model where the saving rate is determined endogenously. However, instead of next substituting in for $\hat{y}^*$ from (5) as MRW do, these authors proxy for $\hat{y}^*$ in a more *ad hoc* manner with variables they deem important determinants of steady-state income. Empirical specifications based on Barro and Sala-i-Martin (1992) are also known in the literature as "Barro regressions".

its country-specific component may also be correlated with country-specific saving and population growth rates.

3.1 Empirical Specification with Homogeneous Coefficients:

To estimate equation (8) consistently, the panel data framework has been advocated by Islam (1995) among others. This approach permits treating the initial level of technology as a country- or region-specific "fixed" effect that remains constant over time and may be correlated with the other explanatory variables. Allowing the regressors to vary both across regions i and over time t, and defining time period $t-1$ as a point in time τ years prior to t, equation (8) can be rewritten as a dynamic panel data model for the log-level of income per capita:

$$\begin{aligned} \ln y_{it} &= \gamma \ln y_{i,t-1} \ + \ \theta_1 \ln s_{it} \ + \ \theta_2 \ln(n_{it} + g + \delta) \ + \ u_{it} \\ u_{it} &= \mu_i \ + \ \eta_t \ + \ v_{it}, \end{aligned} \tag{9}$$

or equivalently, parameterised with the growth rate of per-capita income as the dependent variable:

$$\Delta \ln y_{it} \ = \ (\gamma - 1) \ln y_{i,t-1} \ + \ \theta_1 \ln s_{it} \ + \ \theta_2 \ln(n_{it} + g + \delta) \ + \ \mu_i \ + \ \eta_t \ + \ v_{it},$$

where $\gamma = e^{-\beta\tau}$,
$\theta_1 = \left(1 - e^{-\beta\tau}\right) \frac{\alpha}{1-\alpha}$ and $\theta_2 = -\left(1 - e^{-\beta\tau}\right) \frac{\alpha}{1-\alpha}$,
$\mu_i = \left(1 - e^{-\beta\tau}\right) \ln A(0)$ are unobservable region-specific fixed effects,
$\eta_t = g\left(t - e^{-\beta\tau}(t-\tau)\right)$ captures unobservable period-specific effects that are common to all regions but vary over time, such as technology or other macroeconomic shocks; and
v_{it} is a mean-zero error term.

From γ and θ_1, the annual rate of convergence and the elasticity of output with respect to capital can be recovered as $\beta = -(\ln \gamma)/\tau$ and $\alpha = \theta_1/(1 - \gamma + \theta_1)$. Because markets are assumed competitive in the Solow model, α equals the share of capital in total income, which has been found to be roughly one-third from macroeconomic data on factor income shares, especially in developed countries (Gollin 2002). s_{it} and n_{it} are measured as averages over the time period between year $t - \tau$ and year t, and we set $g + \delta = 0.05$ for all regions and years. The unobserved common time effects η_t are accounted for by including time dummies, which permits a flexible, possibly nonlinear, evolution of technology and long-run income over time.

We use panel data at five-year intervals to estimate equation (9), so that $\tau = 5$. This is one approach chosen in the empirical literature to minimise the influence of high-frequency fluctuations such as business cycles when trying to isolate patterns of transitional or long-run growth.[7] Splitting our sample period into subperiods of five years each leaves us with a total number of time periods $T = 5$.[8]

Equation (9) allows for heterogeneity in regional long-run levels of income per capita since some of the determinants of steady-state income - the unobserved $\ln A(0)$ and the observables s and n - are allowed to be region-specific. On the other hand, the long-run growth rate of per-capita income g, captured by the time effects η_t, as well as the slope coefficients in equation (9) are restricted to be common across regions. In time-income space, region-specific long-run growth paths can therefore be visualised as parallel lines with slope g, to which the regions converge at common speed β. The notion of convergence implicit in equation (9) is conditional, since each region converges to its own steady-state level of income, albeit at a speed that it shares with the other regions.

Thus, an estimate of γ that is significant and positive but strictly less than one would be consistent with conditional convergence taking place in our sample of EU-15 regions. Further, the model predicts that θ_1 and θ_2 should be of equal magnitude, with a positive and a negative sign respectively: since a permanent increase in the rate of capital accumulation raises a region's steady-state level of income per capita, the growth rate in transition to the (now higher) steady state also rises for a given level of initial income. The opposite is the case for an increase in the growth rate of labour, which should lower the transitional income growth rate. The long-run coefficients $\frac{\theta_1}{1-\gamma}$ and $\frac{\theta_2}{1-\gamma}$ can be estimated to gauge the effects of permanent changes in investment or population growth on the steady-state level of income. If the capital elasticity α implied by the estimate of θ_1 is about one-third, equation (6) suggests that $\frac{\theta_1}{1-\gamma}$ should roughly equal 0.5.

In country-level regressions based on equations (8) or (9), a host of additional explanatory variables have been included, under the assumption that they also proxy for differences in initial technology or steady-state income levels across countries. Examples are indicators of human capital, trade openness, institutional or geographical variables. At the regional level however, data availability is limited for many of these, as data either do not exist for the EU regions (trade) or have only been collected for more recent years (human capital). The role of human capital is investigated in Vogel (2013), where we consider a shorter sample period. In addition, the region-specific fixed effects μ_i enable us to control for any permanent differences in human capital or trade openness across the regions in our sample, as well as for physical geography and to some extent institutions.

[7] This approach is taken by all pooled panel data studies reviewed in section 2, at both country and regional levels.

[8] The subperiods are 1980-1985, 1985-1990, 1990-1995, 1995-2000, 2000-2005. The investment share s_{it} is calculated over the five-year periods 1981-1985, 1986-1990 etc., so that the subperiods do not overlap.

3.2 Empirical Specifications with Heterogeneous Coefficients:

To allow for heterogeneity in the slope coefficients across regions, equation (9) can be modified as follows:

$$\ln y_{it} \;=\; \gamma_i \ln y_{i,t-1} \;+\; \theta_{1i} \ln s_{it} \;+\; \theta_{2i} \ln(n_{it} + g + \delta) \;+\; \mu_i \;+\; \eta_t \;+\; v_{it} \tag{10}$$

This model implies that the short- and long-run effects of investment and population growth on the level of income are now also free to differ across regions, as is the speed of adjustment to the steady state, β_i. The underlying long-run growth rate g is still assumed to be common to all regions, so that regional long-run growth paths are again parallel lines with slope g. However, by allowing for region-specific speeds of convergence, we have taken a step away from the notion of convergence embodied by equation (9), which envisages region-specific steady-state levels of income but assumes that β is homogeneous for all i.

To implement equation (10) empirically, we estimate separate time-series models for individual regions. For this purpose, we use data at annual frequency ($\tau = 1$), so that the number of time periods ($T = 25$) available is large enough. Since business cycles and other short-run fluctuations may loom large in annual data, richer dynamics are introduced into our empirical specification to control for their influence. This approach has been taken recently by Bond et al. (2010), for example. Given our theoretical framework, these additional dynamic terms may capture not only high-frequency business cycle fluctuations, but also transitional growth as regions adjust to their long-run growth paths.

Instead of including time dummies, we control for the common time effects η_t by measuring the data in deviations from cross-section averages, that is, by subtracting from each observation the average across regions for that time period. For example, $\ln y_{it}$ demeaned in this manner yields $\ln \widetilde{y}_{it} = \ln y_{it} - \frac{1}{N}\sum_{j=1}^{N} \ln y_{jt}$. This transformation is equivalent to using time dummies when all slope coefficients are homogeneous, as in equation (9), but only approximately so when coefficients are heterogeneous across regions.[9]

Overall, our estimating equation allowing for heterogeneity in the slope coefficients is based on the following autoregressive-distributed lag (ADL) model of order p:

$$\begin{aligned}\ln \widetilde{y}_{it} \;=\; & \gamma_{1i} \ln \widetilde{y}_{i,t-1} \;+\; \ldots \;+\; \gamma_{pi} \ln \widetilde{y}_{i,t-p} \;+\; \theta_{1i}^{0} \ln \widetilde{s}_{it} \;+\; \ldots \;+\; \theta_{1i}^{p} \ln \widetilde{s}_{i,t-p} \;+ \\ & +\; \theta_{2i}^{0} \ln(\widetilde{n}_{it} + g + \delta) \;+\; \ldots \;+\; \theta_{2i}^{p} \ln(\widetilde{n}_{i,t-p} + g + \delta) \;+\; \widetilde{\mu}_i \;+\; \widetilde{v}_{it}.\end{aligned} \tag{11}$$

From this specification, the region-specific speed of adjustment can be derived as $\beta_i = -\ln(\gamma_{1i} + \ldots + \gamma_{pi})$, while the long-run effect of a change in the investment rate on region i's level of per-capita income is given by $\frac{\theta_{1i}^{0} + \ldots + \theta_{1i}^{p}}{1 - \gamma_{1i} - \ldots - \gamma_{pi}}$. We reparameterise equation

[9] When slope coefficients are heterogeneous, demeaning introduces additional terms into the error term of equation (10), which may increase standard errors in estimation. See Smith and Fuertes (2004), p. 62-63.

(11) in terms of first differences and levels, which is more convenient for estimation:

$$\begin{aligned}
\Delta \ln \widetilde{y}_{it} \;=\;& (\gamma_{1i}-1)\Delta \ln \widetilde{y}_{i,t-1} \;+\; (\gamma_{1i}+\gamma_{2i}-1)\Delta \ln \widetilde{y}_{i,t-2} \;+\; \dots \;+\; (\sum_{j=1}^{p}\gamma_{ji}-1)\ln \widetilde{y}_{i,t-p} \;+\\
&+\; \theta^{0}_{1i}\Delta \ln \widetilde{s}_{it} \;+\; (\theta^{0}_{1i}+\theta^{1}_{1i})\Delta \ln \widetilde{s}_{i,t-1} \;+\; \dots \;+\; \sum_{j=0}^{p}\theta^{j}_{1i}\ln \widetilde{s}_{i,t-p} \;+\\
&+\; \theta^{0}_{2i}\Delta \ln(\widetilde{n}_{it}+g+\delta) \;+\; (\theta^{0}_{2i}+\theta^{1}_{2i})\Delta \ln(\widetilde{n}_{i,t-1}+g+\delta) \;+\; \dots \;+\\
&+\; \sum_{j=0}^{p}\theta^{j}_{2i}\ln(\widetilde{n}_{i,t-p}+g+\delta) \;+\; \widetilde{\mu}_i \;+\; \widetilde{v}_{it}
\end{aligned} \qquad (12)$$

Thus the quantities of interest in equation (11) - the sums of the coefficients on the individual (lagged) levels of each variable - can be directly obtained from the lagged levels dated $t-p$ in equation (12).

Finally, cross-section dependence in the error terms can be modelled by introducing a region-specific coefficient on the common time effect η_t in equation (10):

$$\ln y_{it} \;=\; \gamma_i \ln y_{i,t-1} \;+\; \theta_{1i}\ln s_{it} \;+\; \theta_{2i}\ln(n_{it}+g+\delta) \;+\; \mu_i \;+\; \phi_i\eta_t \;+\; v_{it} \qquad (13)$$

In contrast to the previous models, the impact of the unobserved common time effect η_t is now no longer restricted to be identical for all regions but may differ according to ϕ_i. In effect, this allows the steady-state growth rate to be region-specific, since each region may now deviate from g each period by the coefficient ϕ_i. Regional long-run growth paths need therefore no longer be parallel and may even diverge. In terms of the two notions of convergence within and across regions highlighted by Islam (1998), we have now completely eliminated the cross-regional dimension: each region converges at its own speed to its own steady-state income level, which may grow at a region-specific rate over time.[10]

Equation (13) is also implemented using annual data to estimate separate regressions for each region, so we assume an ADL model of order p again to capture business cycle effects, as in equation (11). Since the presence of the term $\phi_i\eta_t$ is controlled for using Pesaran's (2006) CCEMG estimator (see section 4.3), the data are not transformed into deviations from cross-section averages. Hence, our empirical specification allowing for cross-section dependence is

$$\begin{aligned}
\ln y_{it} \;=\;& \gamma_{1i}\ln y_{i,t-1} \;+\; \dots \;+\; \gamma_{pi}\ln y_{i,t-p} \;+\; \theta^{0}_{1i}\ln s_{it} \;+\; \dots \;+\; \theta^{p}_{1i}\ln s_{i,t-p} \;+\\
&+\; \theta^{0}_{2i}\ln(n_{it}+g+\delta) \;+\; \dots \;+\; \theta^{p}_{2i}\ln(n_{i,t-p}+g+\delta) \;+\; \mu_i \;+\; \phi_i\eta_t \;+\; v_{it},
\end{aligned}$$

[10] We therefore use the term "speed of convergence" only in the context of model (9) and otherwise refer to the "speed of adjustment".

which we reparameterise in differences-and-levels form, analogous to equation (12):

$$\begin{aligned}\Delta \ln y_{it} \;=\; & (\gamma_{1i}-1)\Delta \ln y_{i,t-1} \;+\; (\gamma_{1i}+\gamma_{2i}-1)\Delta \ln y_{i,t-2} \;+\; ... \;+\; (\sum_{j=1}^{p}\gamma_{ji}-1)\ln y_{i,t-p} \;+ \\ & +\; \theta_{1i}^{0}\Delta \ln s_{it} \;+\; (\theta_{1i}^{0}+\theta_{1i}^{1})\Delta \ln s_{i,t-1} \;+\; ... \;+\; \sum_{j=0}^{p}\theta_{1i}^{j}\ln s_{i,t-p} \;+ \\ & +\; \theta_{2i}^{0}\Delta \ln(n_{it}+g+\delta) \;+\; (\theta_{2i}^{0}+\theta_{2i}^{1})\Delta \ln(n_{i,t-1}+g+\delta) \;+\; ... \;+ \\ & +\; \sum_{j=0}^{p}\theta_{2i}^{j}\ln(n_{i,t-p}+g+\delta) \;+\; \mu_i \;+\; \phi_i\eta_t \;+\; v_{it}.\end{aligned} \tag{14}$$

4 Estimation Issues and Methods:

4.1 Pooled Dynamic Panel Data Estimators:

To estimate equation (9), two econometric issues need to be considered: first, the joint presence of region-specific fixed effects and a lagged dependent variable, which renders some estimators inconsistent, particularly when the number of time periods is small; and second, the potential endogeneity of some explanatory variables due to e.g. simultaneity. The first-differenced and system-GMM estimators developed in Arellano and Bond (1991), Arellano and Bover (1995) and Blundell and Bond (1998) are, under certain conditions, able to deal with both issues. Below, we review these and other available estimators. Let x_{it} be a row vector containing the explanatory variables $\ln s_{it}$ and $\ln(n_{it}+g+\delta)$ for this purpose.

The standard pooled OLS estimator is inconsistent for equation (9) because the lagged dependent variable $\ln y_{i,t-1}$ is by construction correlated with the region-specific fixed effects μ_i, which may also be correlated with other explanatory variables in the model. This is a form of endogeneity resulting from omitted variables that are correlated with included explanatory variables. For the coefficient γ, it can be shown that the large-sample bias in the OLS estimator is likely to be upward. Given that $\beta = -(\ln\gamma)/\tau$, this implies that the speed of convergence β may be underestimated.

The within-groups (WG) or fixed-effects estimator is also inconsistent if the number of time periods in the panel is small (Nickell 1981), as in our case. Within-groups estimation consists of applying OLS to the data transformed into deviations from region-specific means, whereby time-invariant variables like μ_i are removed ("within transformation").[11] Equivalently, this estimator can be obtained by including a set of N region-specific dummy variables in the OLS regression, so that it is also called Least-Squares Dummy Variables

[11]Within-transforming any variable y_{it} yields $y_{it} - \frac{1}{T}\sum_{s=1}^{T} y_{is}$.

(LSDV). The WG estimate of γ is likely to be biased downwards in large-N samples, in the opposite direction to OLS, so that the speed of convergence may be overestimated. When assessing the empirical outcomes of other estimators, it is therefore advisable to be sceptical of those yielding an estimate $\hat{\gamma}$ that is considerably lower (higher) than the corresponding WG (pooled OLS) estimate (Bond 2002). Consequently, we provide OLS and WG estimates primarily as benchmarks.

If one or more of the explanatory variables in equation (9) are endogenous because they are contemporaneously correlated with v_{it}, both OLS and WG estimators are inconsistent. One possible cause for this is simultaneity, which may be the case for the share of output that is invested, s_{it}, as well as the growth rate of labour, n_{it}, if we allow for inward or outward migration to take place in response to shocks to regional output.

The first-differenced and system-GMM estimators use model (9) in first differences, which, like the within transformation, removes the time-invariant fixed effect μ_i. Since the first-differenced lagged dependent variable, $\Delta \ln y_{i,t-1}$, is now correlated with the first-differenced error term, Δv_{it}, both GMM estimators employ lagged levels of the dependent variable dated $t-2$ and earlier as instruments. Similarly, lagged levels dated $t-2$ and earlier of those explanatory variables x_{it} that are endogenous in equation (9) due to correlation with v_{it} can be used as instruments for the first-differenced equations. While first-differenced GMM (FD-GMM) is based on equation (9) in first differences only, the system-GMM estimator (S-GMM) uses a combination of this equation in first differences and in levels. The instruments employed for the first-differenced equations are as described above. For the levels equations, first-differences dated $t-1$ of the dependent variable and the endogenous explanatory variables may serve as instruments.

Both GMM estimators are consistent for model (9) for large N and small T if the instruments they employ are valid - that is, uncorrelated with the error term of the relevant equation - and informative, that is, correlated with the variables that are treated as endogenous. If these conditions are satisfied, the GMM estimators offer a solution to the two estimation issues discussed above by first-differencing and making use of instruments "internal" to the model that are available in the panel framework.

To ensure that the instruments employed for the first-differenced equations are valid - that is, uncorrelated with Δv_{it} - v_{it} must be serially uncorrelated.[12] This condition can be tested using Arellano and Bond's (1991) test for serial correlation in the first-differenced regression residuals. If the number of available instruments is greater than the number of explanatory variables, the Sargan (1958)/Hansen (1982) test of overidentifying restrictions provides an additional tool for assessing the validity of the instruments.

The informativeness of the instruments for the first-differenced equations may be weak, however, when the series used in estimation are very persistent. In this case, lagged levels

[12] An additional requirement for both GMM estimators considered is the absence of feedback from the current shock v_{it} to past values of the dependent and explanatory variables. This is a standard and not unreasonable assumption.

of the variables are only weakly correlated with the subsequent first differences they serve as instruments for. This could lead to considerable finite-sample bias and imprecision in the FD-GMM estimates, particularly when the time dimension of the panel is short. The bias in $\hat{\gamma}$ is likely to be downward, in the direction of WG. On the other hand, the S-GMM estimator addresses the weak instruments problem in FD-GMM by introducing the additional instruments for the levels equations.[13] Since in this paper, we make use of series that may be quite persistent over time, such as output per capita and the output share of investment, S-GMM may be the preferred estimator.[14]

For the additional instruments for the equations in levels employed by S-GMM to be valid - that is, uncorrelated with u_{it} - one requirement is again the absence of serial correlation in v_{it}. Further, the additional instruments for the levels equations must be uncorrelated with the region-specific fixed effects μ_i. In the context of a multivariate model similar to equation (9), Blundell and Bond (2000) show that for the lagged first differences of the dependent variable, the latter condition requires a restriction on the initial conditions of the dependent variable, $\ln y_{it}$ in our case. Provided that the first-differenced explanatory variables Δx_{it} are uncorrelated with μ_i, this restriction is satisfied if equation (9) has generated $\ln y_{it}$ for long enough before the start of the sample period for the influence of the true initial conditions to vanish. In our application to data for Western European regions since 1980, it does not seem unreasonable to make this assumption. The validity of the additional instruments for the equations in levels can be formally investigated using the Difference Sargan/Hansen test.

4.2 Parameter Heterogeneity:

To estimate equation (12), we employ the mean group (MG) approach of Pesaran and Smith (1995), which consists of estimating separate time-series regressions for each region in the sample and averaging the estimated coefficients across regions. Pesaran and Smith (1995) show that neglecting parameter heterogeneity in a dynamic panel data model like (9) renders the pooled estimators considered in section 4.1 inconsistent. This is because the unmodelled heterogeneous components of the terms on the right-hand side of equation (9) - say $(\gamma_i - \gamma)\ln y_{i,t-1}$ and $(\theta_i - \theta)x_{it}$, where $\theta = (\theta_1\ \theta_2)'$[15] - become part of the error

[13]Monte Carlo simulations conducted by Blundell and Bond (1998) for the univariate AR(1) case with persistent series indicate that the gains in terms of bias and precision from using the S-GMM estimator are substantial for T as large as 11. Simulations by Blundell, Bond and Windmeijer (2000) for the multivariate case suggest that this result extends to a model with additional explanatory variables, which is more relevant to equation (9).

[14]Estimating simple AR(1) models for $\ln y_{it}$ and $\ln s_{it}$ based on five-yearly data provides evidence of persistence in both series. For $\ln y_{it}$, the autoregressive coefficient lies above 0.95 when using pooled OLS and both GMM estimators, with a lower estimate of 0.55 obtained using WG. For $\ln s_{it}$, all estimators except WG yield an autoregressive coefficient above 0.80.

[15]These expressions result from assuming a random coefficients model for the heterogeneous parameters: $\gamma_i = \gamma + \xi_{1i}$ and $\theta_i = \theta + \xi_{2i}$, where ξ_{1i} and ξ_{2i} are zero-mean random variables. This setup is considered

term, which is then correlated with the included regressors if these are serially correlated (at least as long as $\gamma_i \neq 0$).[16] An instrumental variables approach such as the GMM estimators described above is probably also infeasible, since any candidate instrument that is correlated with $\ln y_{it}$ and x_{it} will also be correlated with the error term and therefore invalid.

The pooled OLS and WG estimates of γ are likely to be biased upwards asymptotically (for both N and T large) when parameter heterogeneity is ignored, thus leading to potential underestimation of the speed of adjustment. The long-run coefficients $\frac{\theta}{1-\gamma}$ are also likely biased upwards, while θ tends to be underestimated.[17] These large-sample biases have been found to increase with the degree of parameter heterogeneity and the degree of serial correlation in the explanatory variables.

If both N and T are large, the MG approach provides consistent estimates of the averages of the regression coefficients when these are heterogeneous across regions, as in model (12). The MG estimator of the average speed of adjustment, for example, is given by $\hat{\beta}_{MG} = N^{-1}\sum_{i=1}^{N}\hat{\beta}_i = N^{-1}\sum_{i=1}^{N} -\ln(\hat{\gamma}_{1i} + ... + \hat{\gamma}_{pi})$, where $\hat{\gamma}_{1i}$ to $\hat{\gamma}_{pi}$ are region-specific (e.g. OLS) estimates obtained from the individual time-series regressions of equation (12). The average effect of investment on long-run income across regions can be computed as $N^{-1}\sum_{i=1}^{N}\left(\frac{\hat{\theta}_{1i}^{0} + ... + \hat{\theta}_{1i}^{p}}{1 - \hat{\gamma}_{1i} - ... - \hat{\gamma}_{pi}}\right)$. The region-specific parameters in equation (12) can also be estimated by instrumental variables (IV) methods to allow for the possibility that the explanatory variables x_{it} are endogenous, for example using their own lags as instruments in a time-series model for each region.

The MG estimator is consistent both when the variables in the model are covariance stationary, and when $\ln y_{it}$ and x_{it} are integrated of order 1, or $I(1)$, with a cointegrating relationship between them for each region. In this case, the long-run equilibrium relationship between $\ln y_{it}$ and x_{it} can be estimated by OLS, which is super-consistent for θ_i, and the average long-run coefficients across regions can be obtained as above.

Monte Carlo studies conducted by Pesaran, Smith and Im (1996) and Hsiao, Pesaran and Tahmiscioglu (1999) suggest that for fairly short time series ($T = 20$), the MG estimator may be biased downwards for γ and θ. For the long-run coefficients, however, the MG estimator performs much better even for small panels ($N = T = 20$). Since we use 25 time periods in estimation, we will therefore not focus on the short-run effects of $\ln s_{it}$ and $\ln(n_{it} + g + \delta)$ on income levels.

for convenience, and Pesaran and Smith (1995) find that their main results also hold when the coefficients are assumed fixed but different across i.

[16] Using annual data to estimate AR(1) models for $\ln y_{it}$ and $\ln s_{it}$ yields strong evidence of autocorrelation in the series. For per-capita income, the autoregressive coefficient is 0.99 and 0.90 when using pooled OLS and WG respectively. For the investment share, we obtain 0.94 with OLS and 0.85 with WG.

[17] The direction of the bias for γ and $\frac{\theta}{1-\gamma}$ depends on the assumption that the x_{it} variables follow AR(1) processes with positive autocorrelation coefficients. For negative autocorrelation, the pooled OLS and WG estimates of γ and $\frac{\theta}{1-\gamma}$ are biased downwards. See Pesaran and Smith (1995).

4.3 Cross-Section Dependence:

To estimate equation (14), we use the common correlated effects mean group estimator (CCEMG) proposed by Pesaran (2006). This estimator controls for the term $\phi_i \eta_t$ by augmenting the equation with the cross-section averages for each time period of the dependent variable and each of the explanatory variables. In equation (14), the terms $\Delta \ln \bar{y}_t$, $\Delta \bar{x}_t$ and their lags as well as $\ln \bar{y}_{t-p}$ and $\bar{x}_{t-p}$ are thus added to the regressors on the right-hand side, where e.g. $\Delta \ln \bar{y}_t = \frac{1}{N} \sum_{i=1}^{N} \Delta \ln y_{it}$. The augmented model is then estimated by mean groups, based on individual time-series regressions for each region.

Pesaran's (2006) approach belongs to a strand of research that assumes a factor structure for the error term in order to model cross-sectional error correlation, or cross-section dependence. In this framework, the error term $\varepsilon_{it} = \phi_i \eta_t + v_{it}$ in equation (14) can be interpreted as a single factor model, where η_t is an unobserved common factor that may influence each region differently, with "factor loading" ϕ_i. Assuming that ϕ_i is non-stochastic and $\eta_t \sim iid\ N(0,1)$ over t, the covariance between the errors for two regions $i \neq j$ is given by $E(\varepsilon_{it}\varepsilon_{jt}) = \phi_i \phi_j$.[18] This approach does not assume that the structure of the error correlation is known and therefore allows for a general form of cross-section dependence. By contrast, the spatial econometrics literature models cross-section correlation by means of a pre-selected weight matrix, so that the nature of the correlation between regions is specified in advance. The CCE approach, on the other hand, remains valid in the presence of the forms of spatial error correlation typically considered in spatial econometrics (Pesaran and Tosetti 2011). It also remains valid in the presence of a finite number of unobserved common factors with heterogeneous coefficients, not just the single common factor η_t.

Therefore, the CCEMG estimator is a feasible way to control for a variety of possible causes of cross-sectional error correlation, including spatial dependence, in our annual panel setup where T may be considered sufficiently long to estimate separate time-series regressions for each region. However, the estimated parameters on the included cross-section averages cannot be directly interpreted, so that we cannot make statements on the sign and strength of the cross-section dependence.

The expression for the error covariance above illustrates that cross-section dependence that remains unaccounted for may induce inefficiency in the MG estimator. It will also be biased and inconsistent if the common factor η_t affects not only the dependent variable, but also one or more of the explanatory variables in our model (possibly again in a region-specific manner). An example of such a situation would be the heterogeneous impact of the recent "great recession" on EU regional incomes and investment, depending for instance on regional industrial structure.

Pesaran (2006) shows that when N is large, the CCEMG estimator yields consistent estimates for model (14). Intuitively, this is because the cross-sectional averages that are included as additional regressors approximate η_t - in general, a finite number of unobserved

[18] This example is adapted from Phillips and Sul (2003), p. 225.

common factors - as N becomes large. Kapetanios, Pesaran and Yamagata (2011) establish that this result holds not only when variables and common factors are covariance stationary, as in Pesaran (2006), but also when the common factors contain unit roots, inducing nonstationarity in the observed variables $\ln y_{it}$ and x_{it}. A requirement is that the model errors v_{it} are stationary, however, so that given η_t, $\ln y_{it}$ and x_{it} must be cointegrated.

Monte Carlo simulations by Coakley, Fuertes and Smith (2006), Pesaran (2006) and Kapetanios et al. (2011) are supportive of the CCEMG estimator in samples of the size we consider. These studies find that it outperforms the MG estimator in terms of bias and precision under different scenarios regarding the factor structure of the errors, correlation between factors and explanatory variables, and stationarity properties of the common factors. In particular, the simulations in Pesaran (2006) indicate that not accounting for common factors with heterogeneous coefficients may lead to substantial bias and imprecision in the conventional MG estimator.

5 Data and Variables:

We use data from the 2007 edition of the Cambridge Econometrics (CE) European Regional Database, which covers the period from 1980 to 2005. Regional per-capita income y_{it} is constructed as regional gross value added (GVA) divided by population. In the CE database, GVA is measured at constant prices in 2000 euros, which we adjust for cross-country differences in price levels via national purchasing power parities (PPPs) defined relative to the EU average.[19] We measure the saving rate s_{it} as the share of gross fixed capital formation, also expressed in PPP-adjusted 2000 euros, in total regional GVA. The growth rate of the total population is used to proxy the growth rate of labour input n_{it}.

Our sample consists of 193 NUTS-2 regions from the EU-15 countries.[20] It does not include the French, Spanish and Portuguese small island territories in the Atlantic, Caribbean and Indian Oceans, nor Ceuta and Melilla, two Spanish exclaves on the Moroccan coast. In addition, we exclude the East German regions, for which data are only available from 1991 onwards, and the Dutch region Flevoland, with data only from 1986. We also drop Groningen from the Netherlands, whose per-capita income series exhibits great volatility over our sample period; output in this region is strongly dependent on natural gas production. For a list of all regions in the sample, see Appendix A. Appendix B provides summary statistics of our variables.

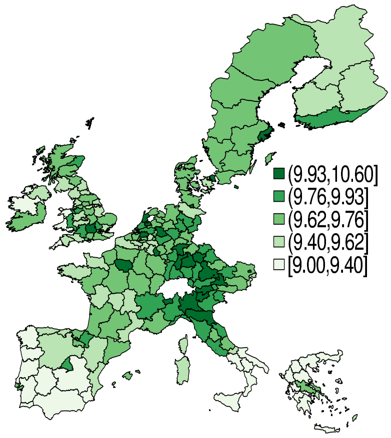
Figure 1 below illustrates the spatial distribution of the variables across all regions in the sample. In the top left-hand map, a banana-shaped arc of regions with high GVA per

[19] The resulting artificial common European currency unit is the Purchasing Power Standard (PPS), where one PPS equals the average purchasing power of one euro across the European Union.

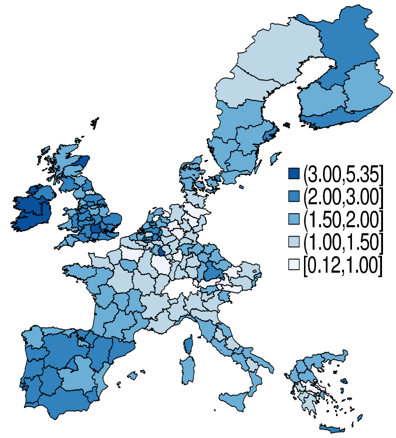
[20] The sample covers regions from Austria (9 regions), Belgium (11), Denmark (3), Finland (5), France (22), Germany (30), Greece (13), Ireland (2), Italy (21), Luxembourg (1), Netherlands (10), Portugal (5), Spain (16), Sweden (8) and the United Kingdom (37).

capita is clearly visible, stretching from Southern England across the Benelux countries and south-western Germany into Northern Italy. The richest region over our sample period is Inner London, followed by Brussels, Hamburg and Luxembourg. Average levels of income per capita are substantially lower in Portugal, the south of Spain, Southern Italy and Greece. The poorest regions in per-capita terms are Epirus in Greece, the Portuguese Central region and Extremadura in Spain.

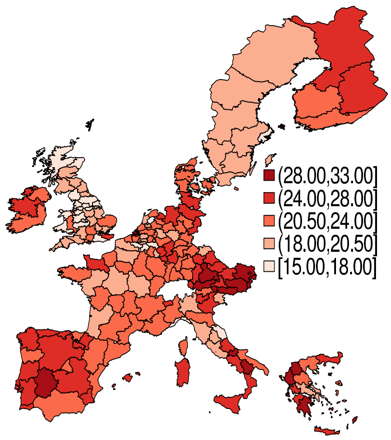
Figure 1: Variable Distribution across EU-15 NUTS-2 Regions, 1980-2005



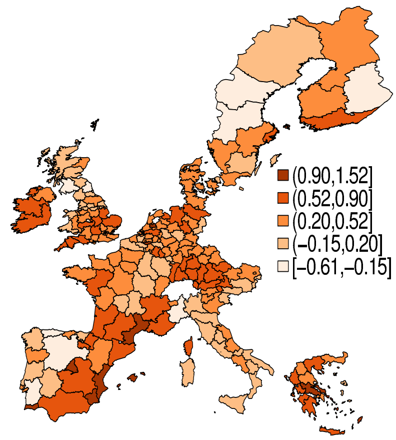
Average log-level of GVA p.c.



Average annual growth of GVA p.c. (%)



Average investment share (%)



Average population growth rate (%)

In the top right-hand map of Figure 1, the fastest-growing regions between 1980 and 2005 are Luxembourg and the two Irish NUTS-2 regions. The maximum is attained by Southern and Eastern Ireland, where per-capita GVA grew at an annualised rate of 5.4%. The average growth rate in our sample is 1.8% per year. British and Finnish regions also registered high growth rates, as did some regions with low levels of per-capita GVA: Extremadura, Crete in Greece and the Algarve in southern Portugal. On the other hand, Central Greece registered by far the lowest average annual growth rate (0.12%), followed by Drenthe in the Netherlands and several Austrian and German regions.

Overall, the two top panels of the Figure suggest that the growth and convergence experiences of the EU-15 regions differed widely across and within countries: both high- and low-income regions grew fast between 1980 and 2005, and both high- and low-income regions grew slowly. Hence, this first look at the data indicates that allowing for heterogeneity in speeds of adjustment is sensible.

As the bottom left-hand panel of Figure 1 shows, the share of total GVA invested (GFCF) was highest on average in regions from Greece and Austria, Germany, Southern Italy and Spain (around 30%). At the other end of the spectrum, average investment shares amounted to only about half of this figure (15%) in regions across the UK, in Brussels and in Stockholm.

Rates of population growth (bottom right-hand panel) are rather heterogeneous across the EU-15. In Greece and Ireland, the population grew rather strongly across all regions, while most regions grew moderately in Germany, Belgium, the Netherlands and the south and centre of the UK. On the other hand, the population barely changed on average in Italy: Liguria registered the strongest population decline of all EU-15 regions (-0.62% per year on average), while growth rates were modest or zero in all other Italian regions. In Sweden, France, Spain and Portugal, structurally weaker peripheral or inland regions lost population, while economic centres gained. The sample maximum average population growth rate of 1.52% per year was achieved by the Spanish Balearic Islands.

5.1 Time-Series Properties:

When the time dimension of the data used in estimation is moderate or large, the time-series properties of the data matter. Many macroeconomic variables are nonstationary; if this is the case for $\ln y_{it}$ and some of the x_{it} in this paper, the MG and CCEMG approaches require cointegration between them for consistency. Panel A of Table 1 reports the results of applying the Im, Pesaran and Shin (2003) panel unit root test to each variable. Their W test statistic is a standardised average of N region-specific augmented Dickey-Fuller (ADF) statistics, so the test allows for heterogeneous parameters in the individual ADF regressions. The ADF lag length of greatest interest to us is lag three, as we set $p = 3$ in our empirical implementation of the region-specific models (12) and (14). We also present results for shorter lag lengths for robustness. The tests are carried out on data in deviations from cross-section averages, which removes trends that are common across regions.

The results of the Im et al. (2003) test suggest that $\ln \widetilde{y}_{it}$ may be considered integrated of order 1, or $I(1)$: first, the test cannot reject the null hypothesis that each regional $\ln \widetilde{y}_{it}$ series is nonstationary, irrespective of whether region-specific time trends are included in the individual ADF regressions. Second, nonstationarity is rejected for the variable in first-differences, both with and without trends.

The investment share $\ln \widetilde{s}_{it}$ also appears to be $I(1)$ at all lag lengths in the version of the test that does not allow for region-specific trends. On the other hand, when trends are included, the test indicates that $\ln \widetilde{s}_{it}$ may be stationary around these trends, except at lag length three, for which it seems to be nonstationary even given the trends. In the long term, it does not seem plausible that the investment share contains stochastic or deterministic trends, given that it is bounded by zero and one. However, over shorter time spans, the series may be trending and thus look $I(1)$. We therefore proceed on the premise that the investment share behaves like an $I(1)$ series over our sample period.

Table 1: Variable Time-Series Properties

Panel A: Im, Pesaran and Shin (2003) Test for Unit Roots in Heterogeneous Panels

	Constant			Constant and trend		
lags	$\ln \widetilde{y}_{it}$	$\ln \widetilde{s}_{it}$	$\ln(\widetilde{n}_{i,t}+g+\delta)$	$\ln \widetilde{y}_{it}$	$\ln \widetilde{s}_{it}$	$\ln(\widetilde{n}_{i,t}+g+\delta)$
0	1.73 (0.96)	1.77 (0.97)	-18.00 (0.00)	2.78 (1.00)	-5.16 (0.00)	-12.83 (0.00)
1	-0.93 (0.18)	0.26 (0.60)	-10.83 (0.00)	-0.64 (0.26)	-6.07 (0.00)	-4.65 (0.00)
2	1.07 (0.86)	1.35 (0.91)	-8.08 (0.00)	0.76 (0.78)	-3.84 (0.00)	-1.16 (0.12)
3	1.27 (0.90)	2.58 (1.00)	-6.53 (0.00)	1.20 (0.88)	-0.69 (0.25)	0.45 (0.67)
	Constant			Constant and trend		
lags	$\Delta \ln \widetilde{y}_{it}$	$\Delta \ln \widetilde{s}_{it}$	$\Delta\ln(\widetilde{n}_{i,t}+g+\delta)$	$\Delta \ln \widetilde{y}_{it}$	$\Delta \ln \widetilde{s}_{it}$	$\Delta\ln(\widetilde{n}_{i,t}+g+\delta)$
0	-47.19 (0.00)	-54.11 (0.00)	-70.92 (0.00)	-42.07 (0.00)	-47.45 (0.00)	-64.72 (0.00)
1	-30.66 (0.00)	-32.67 (0.00)	-41.54 (0.00)	-25.78 (0.00)	-24.88 (0.00)	-34.68 (0.00)
2	-19.43 (0.00)	-24.26 (0.00)	-25.05 (0.00)	-15.12 (0.00)	-17.26 (0.00)	-18.99 (0.00)
3	-14.97 (0.00)	-17.70 (0.00)	-16.26 (0.00)	-11.00 (0.00)	-10.92 (0.00)	-10.31 (0.00)

Panel B: Pedroni (1999, 2004) Test for Cointegration between $\ln \widetilde{y}_{it}$ and $\ln \widetilde{s}_{it}$

	Constant		Constant and trend	
lags	Panel ADF	Group ADF	Panel ADF	Group ADF
0	-1.50 (0.07)	-2.46 (0.01)	-0.80 (0.21)	-0.92 (0.18)
1	-1.43 (0.08)	-2.85 (0.00)	-0.69 (0.25)	-1.90 (0.03)
2	-1.64 (0.05)	-3.58 (0.00)	-0.99 (0.16)	-5.55 (0.00)
3	-1.69 (0.04)	-4.58 (0.00)	-1.67 (0.05)	-8.67 (0.00)

Notes: Im et al. (2003) test: W statistic reported, p-values in parentheses; W statistic is asymptotically standard normal under the null hypothesis that all region-specific series are nonstationary; alternative hypothesis (one-sided) is that *some* region-specific series are stationary.
Pedroni (1999, 2004) test: p-values in parentheses; panel and group ADF statistics are both asymptotically standard normal under the null hypothesis of no cointegration; alternative hypothesis is one-sided.

Lastly, for the population growth rate $\ln(n_{it} + g + \delta)$, the Im et al. (2003) test rejects nonstationarity for both level and first-difference when no region-specific trends are included. With trends, there is some evidence of nonstationarity at lag lengths two and three. However, European population levels over the last three decades have clearly not evolved in a way that would be consistent with a unit root in the population growth rate. We therefore treat $\ln(n_{it} + g + \delta)$ as stationary or $I(0)$.

For all variables, we also carried out Pesaran's (2007) unit root test for cross-sectionally dependent data, which mostly confirms our conclusions from the Im et al. (2003) test. The results are given in Table C.1 in the Appendix.

Given the evidence for $I(1)$ behaviour in $\ln \widetilde{y}_{it}$ and $\ln \widetilde{s}_{it}$ presented in Panel A of Table 1, Panel B reports results of the Pedroni (1999, 2004) test for cointegration between the two variables. This test is applied to the residuals of region-specific cointegrating regressions, thereby allowing for heterogeneity in the long-run cointegrating relationships between $\ln \widetilde{y}_{it}$ and $\ln \widetilde{s}_{it}$. The group ADF statistic further allows for heterogeneous coefficients in the ADF regressions, while the panel ADF statistic is based on a pooled specification.

The group ADF statistic rejects the null hypothesis of no cointegration, both with and without region-specific time trends, at almost all lag lengths. At lag three, the panel ADF statistic also rejects at the 5% level for both versions of the test. Since this is the relevant lag length for us, we assume that $\ln \widetilde{y}_{it}$ and $\ln \widetilde{s}_{it}$ are cointegrated for the remainder of this paper. We note however that the panel ADF statistic provides less evidence of cointegration in the model with trends at shorter lag lengths.

Finally, Table C.2 in the Appendix contains the results of Pesaran's (2004) CD test for cross-section dependence applied to each of the series used in estimation. The test statistic is based on the pairwise correlation coefficients between the regional observations at each point in time. For five-yearly and annual data, the CD test provides significant evidence of cross-section dependence for all variables (panels A and B). For annual data in deviations from cross-section averages in panel C, this is the case only for some series.

6 Estimation Results and Discussion:

We begin our empirical analysis with equation (9), where all parameters except region-specific fixed effects are assumed to be homogeneous across regions. Table 2 presents the results of estimating this model using conventional pooled estimators based on data at five-year intervals. For ease of comparison with the annual results below, we report results for equation (9) parameterised with $\Delta \ln y_{it}$ as the dependent variable, so that the coefficient on $\ln y_{i,t-1}$ given in the table is $(\hat{\gamma} - 1)$. A full set of time dummies is included to account for the unobserved common time effects η_t.

The pooled OLS specification in column (i) does not appear to lend much support to the Mankiw et al. (1992) model, since only the estimated coefficient on the lagged dependent variable is significantly different from zero. The size of this estimate implies that convergence - which occurs at a common speed to region-specific but parallel steady-state income growth paths in the context of equation (9) - takes place across regions at the slow rate of 1% per year. The associated half-life of convergence, which measures the time required to close half of the gap to the steady state, is 69.8 years.[21] The finding of

[21]The annual speed of convergence is computed as $\hat{\beta} = -(\ln \hat{\gamma})/5$; the half-life of convergence is $(\ln 2)/\hat{\beta}$.

slow convergence when using pooled OLS is consistent with the results of both country- and region-level studies reviewed in section 2 (Islam 1995; Badinger et al. 2004). If there are important unobserved region-specific effects however, the OLS estimate of $(\gamma - 1)$ in column (i) is likely to be biased upwards towards zero, so that the speed of convergence may be underestimated.

The within-groups results in column (ii) differ markedly from OLS, which suggests that allowing for region-specific fixed effects matters. The WG results are also more promising for the MRW model. All coefficients are highly significant and have the expected signs. $\hat{\theta}_1$ and $\hat{\theta}_2$ are not significantly different in absolute value, as the model would predict.[22] However, at 0.154, the implied elasticity of output with respect to capital $\hat{\alpha}$ is only half its usual size found from macroeconomic data on factor income shares (0.33). The long-run elasticities of per-capita income with respect to investment and population growth are denoted by LR $\ln s_{it}$ and LR $\ln(n_{it} + g + \delta)$ in the table:[23] our estimates indicate that a permanent increase in the rate of investment by 1% raises the long-run level of income per capita for the average region by about 0.18%, while an equal increase in the rate of population growth lowers it by 0.26%. The first of these numbers is considerably lower than the value of 0.5 that would be consistent with the MRW model if our estimate of α were equal to 0.33. The speed of convergence implied by the WG estimate of $(\gamma - 1)$, 11.7% per annum, is much higher than the corresponding OLS estimate, a finding that is again in line with other empirical studies both at the level of countries and EU regions. At a convergence speed of 11.7%, the half-life amounts to just 5.9 years. The large discrepancy between the OLS and WG estimates of $(\gamma - 1)$ could however reflect the likely downward bias in the WG estimate, towards -1, so that β may be overestimated.

Arellano and Bond's (1991) tests of serial correlation, AB-AR(1) and AB-AR(2), indicate significant second-order serial correlation in the residuals of both OLS and within-groups specifications in columns (i) and (ii). There is also some evidence of first-order serial correlation for WG but none for OLS.[24] Since both OLS and WG estimators are likely to be biased, diagnostic tests on the residuals in columns (i) and (ii) need not be reliable.

The first-differenced and system-GMM results in columns (iii) and (v) make use of all available lagged levels, dated $t - 2$ and earlier, of $\ln y_{it}$, $\ln s_{it}$ and $\ln(n_{it} + g + \delta)$ as instruments for the first-differenced equations.[25] That is, we treat both the saving rate

[22] A t-test of the null hypothesis $\hat{\theta}_1 = -\hat{\theta}_2$ produces a p-value of 0.390.

[23] The implied capital elasticity is computed according to $\hat{\alpha} = \hat{\theta}_1/(1 - \hat{\gamma} + \hat{\theta}_1)$, and the long-run coefficient on investment is calculated as $\hat{\theta}_1/(1 - \hat{\gamma})$. Standard errors for $\hat{\alpha}$, $\hat{\beta}$ and the long-run coefficients are obtained using the delta method.

[24] Given that the OLS estimator does not account for the region-specific fixed effects μ_i, significant serial correlation in the OLS residuals may be expected. For within groups however, the tests are obtained from its least-squares dummy-variables variant and are thus carried out on v_{it}, which we would hope to be serially uncorrelated.

[25] All GMM estimators are implemented as two-step estimators with standard errors corrected for small-sample bias as suggested by Windmeijer (2005). The GMM estimates in this paper are computed using

and the population growth rate as potentially endogenous variables. In columns (iv) and (vi), we use only levels dated $t-3$ and earlier as a robustness check. Overall, the GMM

Table 2: Pooled MRW Model Estimates - Model (9) using Five-Yearly Data

Dependent variable:	(i)	(ii)	(iii)	(iv)	(v)	(vi)
$\Delta \ln y_{it}$	POLS	WG	FD-GMM	FD-GMM	S-GMM	S-GMM
FD IV-set first lag:			$t-2$	$t-3$	$t-2$	$t-3$
$\ln y_{i,t-1}$	-0.049***	-0.443***	-0.211*	-0.384	-0.089***	-0.207***
	(0.010)	(0.061)	(0.114)	(0.273)	(0.031)	(0.061)
$\ln s_{it}$	-0.016	0.080***	0.110*	0.392*	0.080**	-0.127
	(0.015)	(0.019)	(0.056)	(0.234)	(0.034)	(0.083)
$\ln(n_{it}+g+\delta)$	-0.011	-0.116***	-0.292***	-0.711*	-0.117	0.172
	(0.036)	(0.039)	(0.087)	(0.425)	(0.103)	(0.193)
Implied $\hat{\beta}$	0.010***	0.117***	0.047	0.097	0.019***	0.047***
	(0.002)	(0.022)	(0.029)	(0.089)	(0.007)	(0.015)
Implied $\hat{\alpha}$	-0.492	0.154***	0.342***	0.505***	0.475***	-1.572
	(0.675)	(0.038)	(0.112)	(0.155)	(0.151)	(2.034)
LR $\ln s_{it}$	-0.330	0.181***	0.520***	1.020	0.903*	-0.611**
	(0.303)	(0.053)	(0.258)	(0.633)	(0.548)	(0.307)
LR $\ln(n_{it}+g+\delta)$	-0.216	-0.262***	-1.385	-1.852	-1.315	0.831
	(0.741)	(0.087)	(0.873)	(1.541)	(1.369)	(0.818)
AB-AR(1)	1.35	-1.82	-4.77	-2.76	-6.30	-5.80
	(0.178)	(0.069)	(0.000)	(0.006)	(0.000)	(0.000)
AB-AR(2)	2.09	-3.55	1.15	-0.57	1.58	1.26
	(0.037)	(0.000)	(0.249)	(0.568)	(0.113)	(0.207)
Hansen J			55.92	34.18	84.30	43.92
			(0.000)	(0.000)	(0.000)	(0.000)
Dif-Hansen					35.28	17.04
					(0.000)	(0.017)
Time Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	965	965	772	772	965	965
Number of Regions	193	193	193	193	193	193
Number of Instruments			26	16	37	24

Notes: Standard errors, reported in parentheses, are robust to heteroskedasticity and serial correlation; for POLS and WG, they are Huber-White standard errors clustered on regions; GMM estimators are two-step estimators with standard errors corrected for small-sample bias as suggested by Windmeijer (2005); ***, **, and * indicate significance at 1%, 5% and 10% levels; AB-AR(1) and AB-AR(2) are Arellano and Bond's (1991) tests of first- and second-order residual serial correlation, asymptotically standard normal under the null of no serial correlation, p-values in parentheses; Hansen J and Dif-Hansen are the Hansen (1982) and Difference Hansen tests of m overidentifying restrictions, asymptotically $\chi^2(m)$ under the null that the overidentifying restrictions are valid, p-values in parentheses.

Roodman's (2009) `xtabond2` command for Stata.

estimators using lags $t-2$ and earlier perform better, and the parameter estimates in column (iii) are quite sensible. The AB-AR tests in columns (iii) to (vi) provide strong evidence of first- but not of second-order residual autocorrelation. While negative first-order serial correlation is to be expected in the residuals of the first-differenced equations, the absence of significant second-order residual autocorrelation is consistent with serially uncorrelated v_{it}, a condition for instrument validity in FD- and S-GMM. However, the Hansen and Difference Hansen tests in all columns reject the null hypothesis that the instruments are valid. One reason for this may be that the restriction of homogeneous parameters imposed by the pooled panel data estimators does not hold, as noted in section 4.2. Examining the validity of instrument subsets by individual variables and lag length did not reveal a solution to the problem, so that the GMM results must be interpreted with caution.

The FD-GMM parameter estimates in column (iii) improve in some respects compared to the WG results. The capital elasticity $\hat{\alpha}$ is now very close to the value for capital's share in income generally found from data on factor income shares. $(\hat{\gamma}-1)$ lies approximately in the middle between the OLS and WG point estimates, which is reassuring given the likely direction of bias in these estimators; it could be that any downward bias arising from weak instruments in the FD-GMM estimate of $(\gamma-1)$ is at least partly offset by upward bias due to invalid instruments. The implied speed of convergence of 4.7% per year is marginally significant at the 10.3% level, and the long-run effect on per-capita income of a 1% increase in the investment rate rises to 0.5%, as the MRW model would imply given our estimate of α. However, the standard errors of the short-run coefficients have increased, so that the long-run effects are not very precisely estimated. There is also some indication that $\hat{\theta}_1$ and $\hat{\theta}_2$ are not equal in absolute value: a t-test rejects this hypothesis with a p-value of 0.059.

To investigate further the problem with instrument validity highlighted by the Hansen test, we re-estimated column (iii) using different instrument subsets by variables and lag length, albeit unsuccessfully. In column (iv), we report FD-GMM results using only lagged levels dated $t-3$ and earlier as instruments, which the Hansen test still rejects. Moreover, most coefficients are now very imprecisely estimated, and $(\hat{\gamma}-1)$ is lower than in column (iii) and closer to the WG estimate in column (ii). Both these factors could indicate weak instruments, and we found that restricting the instrument set to even longer lags tended to amplify this problem without yielding significant improvements in validity.

In column (v), the S-GMM estimate of $(\gamma-1)$ is higher than for WG and FD-GMM and closer to the OLS estimate in column (i). Since the Hansen test continues to reject the validity of our instrument set for the first-differenced equations, and the Difference Hansen test suggests that the additional instruments for the levels equations are also invalid, we have reason to suspect that this may reflect upward bias in $(\hat{\gamma}-1)$. The implied estimates of α and the long-run coefficient on investment which, at 0.475 and 0.903 respectively, appear unreasonably high, are consistent with this suspicion.

In column (vi), excluding instruments dated $t-2$ and $t-1$ from the instrument sets for the first-differenced and the levels equations respectively improves overall instrument validity only very marginally. The Hansen test rejects the null hypothesis at the 1% level as in the previous columns, and the Difference Hansen test rejects at the 5% level. In addition, while $\hat{\theta}_1$ and $\hat{\theta}_2$ switch signs but are insignificant, the long-run coefficient on the investment rate is negative and significant, which runs counter to the prediction of the Solow model.

Overall, all estimators available when using data at five-year intervals seem subject to considerable bias, so that in this framework, we cannot hope to obtain reliable estimates of the structural and long-run coefficients implied by the MRW model for our sample of EU regions. In particular, the assumption of homogeneous slope coefficients underlying the pooled estimators could be behind the problems with instrument validity in FD- and S-GMM. Therefore, we now turn to annual data, which allows us to investigate parameter heterogeneity and cross-section dependence.

In equation (12), all parameters except that on the unobserved common time effect η_t are free to differ across regions. Before estimating this model by means of separate time-series regressions for each region, in Table 3 we first present results obtained using pooled annual data. This intermediate step provides a benchmark against which the effects of allowing all parameters to vary across regions can be evaluated. For comparability with the region-specific models estimated by MG and CCEMG later on, where each individual regression contains a region-specific intercept, we include a full set of region dummies in Table 3. Considering constraints on degrees of freedom in the region-specific regressions, we aim for parsimony in the dynamics of equation (12) and thus set the maximum lag length p equal to three. The coefficient on $\ln \widetilde{y}_{i,t-3}$ reported in Table 3 is therefore $(\hat{\gamma}_1 + \hat{\gamma}_2 + \hat{\gamma}_3 - 1)$, and all data are measured in deviations from cross-section averages to control for η_t. In terms of estimation, we pursue a general-to-specific strategy eliminating insignificant variables one at a time to arrive at a parsimonious specification. Since the GMM results in Table 2 were broadly similar when treating s_{it} and n_{it} as predetermined rather than endogenous, we employ OLS to estimate the pooled annual model. Eliminating insignificant lags also leaves us with a specification that involves no levels dated t of the investment rate, thus reducing simultaneity concerns for this variable.

Column (i) of Table 3 contains estimates for the full model (12). Dropping statistically insignificant lags yields the results in column (ii). Given these, we impose the restriction in column (iii) that the coefficients on all included first-differences of $\ln(\widetilde{n}_{i,t} + g + \delta)$ and its lags are equal, which minimises the number of explanatory variables and provides the basis for the region-specific regressions below. An F-test of the restriction, shown in column (ii), indicates that it cannot be rejected.

Across all specifications, the estimated coefficients on the lagged levels dated $t-3$ are highly significant and have the signs that the MRW model predicts. The coefficient on $\ln \widetilde{y}_{i,t-3}$ implies a speed of adjustment - which, given that we impose homogeneous coefficients on model (12) in Table 3, is still common across regions and occurs to region-specific but parallel steady-state income growth paths - of 11.2% per year.[26] This estimate is very similar to the WG estimate of 11.7% that we found in Table 2 based on five-yearly data. If model parameters are heterogeneous across regions however, it may be biased downwards.

Table 3: Pooled MRW Model Estimates
Model (12) with Homogeneous Coefficients using Annual Data

Dependent variable: $\Delta \ln \widetilde{y}_{it}$	(i) LSDV Unrestricted coeff.	(s.e.)	(ii) LSDV Parsimonious coeff.	(s.e.)	(iii) LSDV Restricted coeff.	(s.e.)
$\Delta \ln \widetilde{y}_{i,t-1}$	-0.006	(0.032)				
$\Delta \ln \widetilde{y}_{i,t-2}$	-0.121***	(0.044)	-0.121***	(0.044)	-0.122***	(0.043)
$\ln \widetilde{y}_{i,t-3}$	-0.106***	(0.023)	-0.106***	(0.022)	-0.106***	(0.022)
$\Delta \ln \widetilde{s}_{it}$	-0.002	(0.011)				
$\Delta \ln \widetilde{s}_{i,t-1}$	0.021***	(0.007)	0.020***	(0.007)	0.020***	(0.007)
$\Delta \ln \widetilde{s}_{i,t-2}$	0.008	(0.008)				
$\ln \widetilde{s}_{i,t-3}$	0.014***	(0.005)	0.013***	(0.005)	0.013***	(0.005)
$\Delta \ln(\widetilde{n}_{it} + g + \delta)$	-0.022***	(0.004)	-0.022***	(0.004)		
$\Delta \ln(\widetilde{n}_{i,t-1} + g + \delta)$	-0.022***	(0.004)	-0.023***	(0.004)	-0.022***	(0.004)
$\Delta \ln(\widetilde{n}_{i,t-2} + g + \delta)$	-0.022***	(0.005)	-0.022***	(0.005)		
$\ln(\widetilde{n}_{i,t-3} + g + \delta)$	-0.035***	(0.005)	-0.035***	(0.004)	-0.035***	(0.004)
Implied $\hat{\beta}$	0.112***	(0.026)	0.112***	(0.024)	0.112***	(0.024)
Implied $\hat{\alpha}$	0.117***	(0.041)	0.107***	(0.036)	0.107***	(0.036)
LR $\ln s_{it}$	0.132**	(0.052)	0.119***	(0.045)	0.120***	(0.045)
LR $\ln(n_{it} + g + \delta)$	-0.325***	(0.075)	-0.327***	(0.075)	-0.330***	(0.073)
F-Test Restriction			0.02 (0.979)			
AB-AR(1)	-0.99 (0.324)		-0.15 (0.878)		-0.15 (0.879)	
AB-AR(2)	-0.93 (0.351)		-0.98 (0.328)		-0.97 (0.333)	
CD Test	-0.10 (0.917)		-0.15 (0.879)		-0.15 (0.879)	
Region Dummies	Yes		Yes		Yes	
Observations	4246		4246		4246	
Number of Regions	193		193		193	

Notes: Standard errors, reported in parentheses, are robust to heteroskedasticity and serial correlation (Huber-White standard errors clustered on regions); CD is the Pesaran (2004) test for cross-section dependence in the model residuals, asymptotically standard normal under the null hypothesis of no cross-section dependence, p-values in parentheses.

[26] The annual speed of adjustment is computed from the coefficient on $\ln \widetilde{y}_{i,t-3}$ as $\hat{\beta} = -\ln(\hat{\gamma}_1 + \hat{\gamma}_2 + \hat{\gamma}_3)$.

The coefficients on $\ln \widetilde{s}_{i,t-3}$ and $\ln(\widetilde{n}_{i,t-3} + g + \delta)$ differ significantly in size: in every column, a t-test rejects their equality in absolute value at the 1% level. Moreover, the point estimates of α and the long-run effect of investment in columns (ii) and (iii), 0.107 and about 0.12, are even lower than the corresponding WG estimates in Table 2, which are themselves only half the size suggested by MRW. On the other hand, the long-run coefficient on the population growth rate, about -0.33, is larger in absolute value and more precisely estimated than in Table 2. From the AB-AR and CD tests, there is no evidence that the residuals in Table 3 are serially correlated or cross-sectionally dependent. Demeaning the annual data therefore seems to capture any cross-sectional dependence present. Overall however, it seems that moving from pooled five-yearly to pooled annual data does little to improve the fit of the Solow model to our EU regional data.

In Table 4, we turn to the results of estimating equations (12) and (14) by means of separate time-series regressions for each region. For this purpose, we use the restricted specification in column (iii) of Table 3 and focus on the lagged levels dated $t-3$ of each variable as well as on the implied structural and long-run parameters. For each coefficient, we report its mean, the standard error of the mean, and the median across the region-specific regression estimates, where the mean is an estimate that is robust to outliers.[27]

For the MG estimator, we present results for two different ways of controlling for the unobserved common time effects η_t. In column (i), we use data in deviations from cross-section averages, which allows for an unconstrained, possibly nonlinear, evolution of the unobservable factors over time but restricts their impact to be identical across regions. As an alternative and a transition to CCEMG, in column (ii) we use the original data but add linear trends to the region-specific regressions. Thus the unobservables may differ in impact across regions but are constrained to evolve linearly. Column (iii) presents the CCEMG results - also based on the original, untransformed data - and in the final column, we include region-specific trends in CCEMG estimation.

On the whole, the MG estimator based on demeaned data in column (i) appears to be most successful at identifying the model coefficients in Table 4. The estimated mean coefficient on $\ln y_{i,t-3}$ lies below its pooled counterpart in the previous table and implies that on average, regions adjust to their (still parallel) long-run income growth paths at a rate of almost 22% per year. Put differently, the regions are able to close half of the gap to their steady-state paths in just 3.2 years. Allowing for heterogeneous coefficients across regions thus yields a much higher average speed of adjustment than the pooled estimators in Tables 2 and 3, which is consistent with important regional differences in this parameter. This finding is also in line with the existing regional studies that have allowed for heterogeneous speeds of adjustment in univariate models, Meliciani and Peracchi (2006)

[27] The robust estimate of the mean is obtained from a robust regression of the 193 region-specific coefficient estimates on a constant only, using the `rreg` command in Stata. The standard error of the robust mean is then computed using the observation-specific weights (with outliers downweighted) generated by `rreg`.

and Canova and Marcet (1995).

Table 4: Heterogeneous MRW Model Estimates
Restricted Models (12) and (14) using Annual Data

Dependent variable: $\Delta \ln y_{it}$		(i) MG Demeaned	(ii) MG +Trends	(iii) CCEMG	(iv) CCEMG +Trends
$\ln y_{i,t-3}$	Mean	-0.195***	-0.304***	-0.253***	-0.324***
		(0.026)	(0.029)	(0.033)	(0.034)
	Median	-0.194	-0.315	-0.253	-0.321
$\ln s_{i,t-3}$	Mean	0.017	-0.027	0.003	-0.007
		(0.017)	(0.020)	(0.022)	(0.022)
	Median	0.017	-0.026	0.003	-0.004
$\ln(n_{i,t-3} + g + \delta)$	Mean	-0.018	-0.026	-0.013	0.007
		(0.019)	(0.020)	(0.021)	(0.021)
	Median	-0.013	-0.017	-0.012	0.002
Implied $\hat{\beta}$	Mean	0.217***	0.363***	0.292***	0.391***
		(0.027)	(0.031)	(0.035)	(0.043)
	Median	0.216	0.378	0.291	0.390
Implied $\hat{\alpha}$	Mean	0.159***	0.051	0.115**	0.079
		(0.036)	(0.032)	(0.057)	(0.085)
	Median	0.159	0.023	0.099	0.058
LR $\ln s_{it}$	Mean	0.120**	-0.038	0.016	-0.005
		(0.057)	(0.039)	(0.077)	(0.072)
	Median	0.099	-0.003	0.020	-0.015
LR $\ln(n_{it} + g + \delta)$	Mean	-0.087	-0.021	-0.007	0.029
		(0.059)	(0.107)	(0.059)	(0.059)
	Median	-0.054	-0.008	-0.028	0.017
CD Test		0.72 (0.472)	72.26 (0.000)	-0.12 (0.906)	-1.28 (0.201)
Observations		4246	4246	4246	4246
Number of Regions		193	193	193	193

Notes: Outlier-robust means reported; standard errors of robust means in parentheses; standard errors are also robust to heteroskedasticity and autocorrelation.
The specifications in columns (i) and (ii) also include the following explanatory variables: $\Delta \ln y_{i,t-2}$, $\Delta \ln s_{i,t-1}$ and $\Sigma_{q=0}^{2} \Delta \ln(n_{i,t-q} + g + \delta)$.
The specifications in columns (iii) and (iv) also include the following explanatory variables: $\Delta \ln y_{i,t-2}$, $\Delta \ln s_{i,t-1}$ and $\Sigma_{q=0}^{2} \Delta \ln(n_{i,t-q} + g + \delta)$; in addition, they include the sample means of these variables, of $\ln y_{i,t-3}$, $\ln s_{i,t-3}$ and $\ln(n_{i,t-3} + g + \delta)$, and of the dependent variable $\Delta \ln y_{it}$.

The other average short-run coefficients in column (i) are not significantly different from zero, a feature that recurs in the remaining columns of the table and is not entirely surprising given that at most 25 time-series observations are used in the region-specific regressions. The long-run coefficient on population growth is also insignificant throughout. However, the average estimates of α and the long-run coefficient on investment in column (i) are significant, positive and quite similar to the pooled WG and LSDV estimates in Tables 2 and 3.[28] The CD test detects no evidence of residual cross-section dependence.

In column (ii) by contrast, the CD test clearly indicates that introducing region-specific trends instead of demeaning does not adequately capture the cross-section dependence in the data. The average speed of adjustment increases to 36% but none of the other implied and long-run parameters are significant. Given the evidence for residual cross-section dependence, we do not regard these results as very reliable.

The CCEMG estimator in column (iii) allows for unobservable common time effects that may both be nonlinear and heterogeneous in their regional impact. From the CD test, this strategy indeed seems to deal with cross-section dependence quite comprehensively. Compared to column (i), the smaller mean coefficient on $\ln \widetilde{y}_{i,t-3}$ raises the implied average speed of adjustment further, to almost 30% per year. At this rate, the half-life of convergence is only 2.4 years. This may be explained by the fact that conceptually, all homogeneity constraints on regional long-run adjustment processes have been eliminated: steady-state income paths may evolve at region-specific growth rates over time and may even diverge.

The average short-run coefficients in column (iii) are very small, and of the implied and long-run coefficients, only the output elasticity with respect to capital is significant although it is also rather small. It should be borne in mind however that the number of explanatory variables in the region-specific regressions underlying column (iii) is 13 due to the inclusion of cross-section average terms for CCEMG estimation. Identifying individual coefficients may thus be quite challenging given our moderate time dimension. In column (iv), the addition of region-specific linear trends does not produce a well-specified model, as both average short- and long-run effects of investment and population growth switch signs and appear to be dominated by the autoregressive income process. Moreover, the CD test statistic has increased in absolute value compared to the previous column.

Overall, our preferred results in Table 4 are those in column (i), given that MG based on demeaned data leaves no cross-section dependence in the residuals and the results are the most informative about the role played by investment and population growth in driving regional long-run income levels.

In summary, throughout our empirical analysis in this paper, we have found the role of investment to be positive and statistically significant but generally smaller than predicted by the MRW model together with macroeconomic data on factor shares. From the pooled fixed-effects results in Tables 2 and 3 to the region-specific mean-group estimates in Table 4, the output elasticity of capital α lies between 0.107 and 0.159, which is substantially lower than the 0.33 we would expect. The long-run effect of a 1% increase in investment

[28]The estimates of the implied structural and long-run parameters in Table 4 are obtained as nonlinear combinations of the estimated coefficients on $\ln y_{i,t-3}$ and $\ln s_{i,t-3}$. These nonlinear combinations are constructed separately for each region and then averaged over all regions. See the discussion between equations (11) and (12) for the formulae of the region-specific implied parameters.

on income ranges between approximately 0.12% and 0.18%, which is also smaller than expected. The only exception are the first-differenced GMM estimates in Table 2, which are much more consistent with the Solow model but should not be overemphasised because of the evidence of invalid instruments. The prediction that the short-run coefficients on investment and population growth should be equal in magnitude but opposite in sign is only found to hold in the within-groups estimates of the pooled model for five-year periods (Table 2, column ii). The long-run effect of a 1% increase in population growth, where significant, lies between -0.26% and -0.33%.

One implication of these results is that factors other than capital accumulation and population growth need to be considered to explain the evolution of regional incomes across the EU-15 since 1980. In Vogel (2013), we examine the role of human capital and research and development, which are emphasised by endogenous growth theories.

Regarding the speed of convergence to the steady state, an issue that has attracted much attention in research on EU regional growth, we find that it occurs at a speed of about 11% as long as only long-run income levels are free to differ between regions. Allowing the speed of adjustment itself to be heterogeneous while maintaining the assumption of a common long-run growth rate raises its average value across regions to 22%. This figure suggests that the speed of adjustment may have been substantially underestimated in the empirical literature using pooled panel data estimators, which impose homogeneity in slope coefficients across regions. In our view, allowing for region-specific speeds of adjustment is therefore an important and sensible extension to allowing for region-specific long-run income levels only, as is currently common practise.

7 Conclusion:

Observed patterns of regional economic development across the European Union since the 1980s suggest that regional growth processes are characterised by substantial heterogeneity. The existing empirical literature on European regional growth has so far not emphasised heterogeneity in estimated regression coefficients, although it has been shown that ignoring this may lead to bias (Pesaran and Smith 1995). In this paper, we contribute to the literature by investigating the effects of allowing for heterogeneous slope coefficients in the model of Mankiw et al. (1992), using panel data for 193 EU-15 regions from 1980 to 2005.

Similar to previous studies, we first use pooled panel data estimators, based on data at five-year intervals, restricting all slope coefficients to be homogeneous across regions. Then we relax this restriction and estimate separate time-series models for each region, based on annual data, using Pesaran and Smith's (1995) mean group estimator. To address possible spatial correlation in our data, we employ the common correlated effects approach of Pesaran (2006).

Our empirical analysis indicates important differences across regions in the speed of adjustment to region-specific long-run paths for the level of income per capita. Restricting this parameter to be homogeneous results in an estimate of 11% per year, while allowing for heterogeneity raises it to 22% per year on average across regions. This finding is consistent with previous studies at the EU regional level, which have not considered heterogeneity in the effect of additional explanatory variables.

Throughout most of our empirical analysis, we find a positive and statistically significant coefficient on the rate of investment. The size of the implied elasticity of output with respect to capital is fairly robust across pooled and heterogeneous panel estimators and lies between 0.107 and 0.159. We thus find this parameter to be at most half as large as the value of 0.33 implied by the Mankiw et al. (1992) model together with macroeconomic data on factor income shares. Our preferred mean-group estimate of the long-run elasticity of per-capita income with respect to investment is, at 0.12%, also smaller than what the Mankiw et al. (1992) model would predict. We find less robust evidence regarding the effect of population growth on long-run income levels. Only the pooled estimators indicate a significant long-run elasticity, which ranges between -0.26% and -0.33%.

Overall, this paper highlights the importance of allowing for heterogeneous parameters in empirical growth models. We illustrate this with our finding that regions' average speed of adjustment to their long-run paths for the level of income per capita doubles when considering heterogeneity in this parameter compared to the case where only long-run income levels are allowed to differ between regions. Therefore, the empirical literature on regional income convergence, which predominantly considers the latter case, may substantially underestimate regional speeds of adjustment to equilibrium.

References

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations, *The Review of Economic Studies* **58**(2): 277–297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error-Components Models, *Journal of Econometrics* **68**(1): 29–51.

Armstrong, H. W. (1995). Convergence among Regions of the European Union, 1950-1990, *Papers in Regional Science* **74**(2): 143–152.

Badinger, H., Müller, W. G. and Tondl, G. (2004). Regional Convergence in the European Union, 1985-1999: A Spatial Dynamic Panel Analysis, *Regional Studies* **38**(3): 241–253.

Barro, R. J. and Sala-i-Martin, X. (1991). Convergence across States and Regions, *Brookings Papers on Economic Activity* **22**(1): 107–182.

Barro, R. J. and Sala-i-Martin, X. (1992). Convergence, *Journal of Political Economy* **100**(2): 223–251.

Blundell, R. and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel Data Models, *Journal of Econometrics* **87**(1): 115–143.

Blundell, R. and Bond, S. (2000). GMM Estimation with Persistent Panel Data: An Application to Production Functions, *Econometric Reviews* **19**(3): 321–340.

Blundell, R., Bond, S. and Windmeijer, F. (2000). Estimation in Dynamic Panel Data Models: Improving on the Performance of the Standard GMM Estimator, *in* B. H. Baltagi (ed.), *Advances in Econometrics, Volume 15: Nonstationary Panels, Panel Cointegration and Dynamic Panels*, Amsterdam: JAI Elsevier Science.

Bond, S., Leblebicioglu, A. and Schiantarelli, F. (2010). Capital Accumulation and Growth: A New Look at the Empirical Evidence, *Journal of Applied Econometrics* **25**(7): 1073–1099.

Bond, S. R. (2002). Dynamic Panel Data Models: A Guide to Micro Data Methods and Practice, *Portuguese Economic Journal* **1**(2): 141–162.

Bond, S. R., Hoeffler, A. and Temple, J. (2001). GMM Estimation of Empirical Growth Models, *CEPR Discussion Paper 3048*.

Bouayad-Agha, S. and Védrine, L. (2010). Estimation Strategies for a Spatial Dynamic Panel using GMM. A New Approach to the Convergence Issue of European Regions, *Spatial Economic Analysis* **5**(2): 205–227.

Canova, F. and Marcet, A. (1995). The Poor Stay Poor: Non-Convergence across Countries and Regions, *CEPR Discussion Paper 1265.*

Caselli, F., Esquivel, G. and Lefort, F. (1996). Reopening the Convergence Debate: A New Look at Cross-Country Growth Empirics, *Journal of Economic Growth* **1**(3): 363–389.

Coakley, J., Fuertes, A.-M. and Smith, R. (2006). Unobserved Heterogeneity in Panel Time Series Models, *Computational Statistics and Data Analysis* **50**(9): 2361–2380.

Corrado, L., Martin, R. and Weeks, M. (2005). Identifying and Interpreting Regional Convergence Clusters across Europe, *Economic Journal* **115**(502): C133–C160.

Durlauf, S. N., Johnson, P. A. and Temple, J. R. (2005). Growth Econometrics, *in* P. Aghion and S. N. Durlauf (eds), *Handbook of Economic Growth*, Vol. 1A, Amsterdam: North Holland, pp. 555–677.

Eberhardt, M. and Teal, F. (2011). Econometrics for Grumblers: A New Look at the Literature on Cross-Country Growth Empirics, *Journal of Economic Surveys* **25**(1): 99–155.

Elhorst, J. P., Piras, G. and Arbia, G. (2010). Growth and Convergence in a Multi-Regional Model with Space-Time Dynamics, *Geographical Analysis* **42**(3): 338–355.

European Commission (2010). *Investing in Europe's Future: Fifth Report on Economic, Social and Territorial Cohesion*, Luxembourg: Publications Office of the European Union.

Fischer, M. M. and Stirböck, C. (2006). Pan-European Regional Income Growth and Club-Convergence, *Annals of Regional Science* **40**(4): 693–721.

Gollin, D. (2002). Getting Income Shares Right, *Journal of Political Economy* **110**(2): 458–474.

Hansen, L. P. (1982). Large Sample Properties of Generalized Method of Moments Estimators, *Econometrica* **50**(4): 1029–1054.

Hsiao, C., Pesaran, M. H. and Tahmiscioglu, A. K. (1999). Bayes Estimation of Short-Run Coefficients in Dynamic Panel Data Models, *in* C. Hsiao, M. H. Pesaran, K. Lahiri and L. F. Lee (eds), *Analysis of Panels and Limited Dependent Variable Models: In Honour of G.S. Maddala*, Cambridge: Cambridge University Press.

Im, K.-S., Pesaran, M. H. and Shin, Y. (2003). Testing for Unit Roots in Heterogeneous Panels, *Journal of Econometrics* **115**(1): 53–74.

Islam, N. (1995). Growth Empirics: A Panel Data Approach, *Quarterly Journal of Economics* **110**(4): 1127–1170.

Islam, N. (1998). Growth Empirics: A Panel Data Approach - A Reply, *Quarterly Journal of Economics* **113**(1): 325–329.

Kapetanios, G., Pesaran, M. H. and Yamagata, T. (2011). Panels with Non-Stationary Multifactor Error Structures, *Journal of Econometrics* **160**(2): 326–348.

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**(4): 357–392.

Mankiw, N. G., Romer, D. and Weil, D. N. (1992). A Contribution to the Empirics of Economic Growth, *Quarterly Journal of Economics* **107**(2): 407–437.

Meliciani, V. and Peracchi, F. (2006). Convergence in Per-Capita GDP across European Regions: A Reappraisal, *Empirical Economics* **31**(3): 549–568.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects, *Econometrica* **49**(6): 1417–1426.

Pedroni, P. (1999). Critical Values for Cointegration Tests in Heterogeneous Panels with Multiple Regressors, *Oxford Bulletin of Economics and Statistics* **61**(S1): 653–670.

Pedroni, P. (2004). Panel Cointegration: Asymptotic and Finite Sample Properties of Pooled Time Series Tests with an Application to the PPP Hypothesis, *Econometric Theory* **20**(3): 597–625.

Pesaran, M. H. (2004). General Diagnostic Tests for Cross Section Dependence in Panels, *IZA Discussion Paper 1240.*

Pesaran, M. H. (2006). Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure, *Econometrica* **74**(4): 967–1012.

Pesaran, M. H. (2007). A Simple Panel Unit Root Test in the Presence of Cross-Section Dependence, *Journal of Applied Econometrics* **22**(2): 265–312.

Pesaran, M. H. and Smith, R. (1995). Estimating Long-Run Relationships from Dynamic Heterogeneous Panels, *Journal of Econometrics* **68**(1): 79–113.

Pesaran, M. H., Smith, R. and Im, K.-S. (1996). Dynamic Linear Models for Heterogeneous Panels, *in* L. Matyas and P. Sevestre (eds), *The Econometrics of Panel Data: A Handbook of Theory with Applications*, 2nd rev. ed., Dordrecht: Kluwer Academic.

Pesaran, M. H. and Tosetti, E. (2011). Large Panels with Common Factors and Spatial Correlation, *Journal of Econometrics* **161**(2): 182–202.

Phillips, P. C. B. and Sul, D. (2003). Dynamic Panel Estimation and Homogeneity Testing Under Cross Section Dependence, *Econometrics Journal* **6**(1): 217–259.

Roodman, D. (2009). How To Do `xtabond2`: An Introduction to Difference and System GMM in Stata, *Stata Journal* **9**(1): 86–136.

Sargan, J. D. (1958). The Estimation of Economic Relationships using Instrumental Variables, *Econometrica* **26**(3): 393–415.

Smith, R. P. and Fuertes, A.-M. (2004). Panel Time-Series, *mimeo*, Centre for Microdata Methods and Practice, April 2004.

Solow, R. M. (1956). A Contribution to the Theory of Economic Growth, *Quarterly Journal of Economics* **70**(1): 65–94.

Swan, T. W. (1956). Economic Growth and Capital Accumulation, *Economic Record* **32**(2): 334–361.

Tondl, G. (1999). The Changing Pattern of Regional Convergence in Europe, *Review of Regional Research* **19**(1): 1–33.

Vogel, J. (2013). The Two Faces of R&D and Human Capital: Evidence from Western European Regions, *Papers in Regional Science,* **doi: 10.1111/pirs.12084**.

Windmeijer, F. (2005). A Finite Sample Correction for the Variance of Linear Efficient Two-Step GMM Estimators, *Journal of Econometrics* **126**(1): 25–51.

Appendices

A List of Regions

Country	Code	Region Name	Country	Code	Region Name
Austria	AT11	Burgenland	France	FR51	Loire Counties
	AT12	Lower Austria		FR52	Brittany
	AT13	Vienna		FR53	Poitou-Charentes
	AT21	Carinthia		FR61	Aquitaine
	AT22	Styria		FR62	South Pyrenées
	AT31	Upper Austria		FR63	Limousin
	AT32	Salzburg		FR71	Rhône-Alpes
	AT33	Tyrol		FR72	Auvergne
	AT34	Vorarlberg		FR81	Languedoc-Roussillion
				FR82	Provence-Alpes-Côte d'Azur
Belgium	BE10	Brussels		FR83	Corsica
	BE21	Antwerp			
	BE22	Limburg	Germany	DE11	Stuttgart
	BE23	East Flanders		DE12	Karlsruhe
	BE24	Flemish Brabant		DE13	Freiburg
	BE25	West Flanders		DE14	Tübingen
	BE31	Walloon Brabant		DE21	Upper Bavaria
	BE32	Hainault		DE22	Lower Bavaria
	BE33	Liège		DE23	Upper Palatinate
	BE24	Luxembourg (BE)		DE24	Upper Franconia
	BE25	Namur		DE25	Middle Franconia
				DE26	Lower Franconia
Denmark	DK01	Capital Region		DE27	Swabia
	DK02	East of the Great Belt		DE50	Bremen
	DK03	West of the Great Belt		DE60	Hamburg
				DE71	Darmstadt
Finland	FI13	East Finland		DE72	Gießen
	FI18	South Finland		DE73	Kassel
	FI19	West Finland		DE91	Braunschweig
	FI1A	North Finland		DE92	Hannover
	FI20	Åland Islands		DE93	Lunenburg
				DE94	Weser-Ems
France	FR10	Île de France		DEA1	Düsseldorf
	FR21	Champagne-Ardenne		DEA2	Cologne
	FR22	Picardy		DEA3	Münster
	FR23	Upper Normandy		DEA4	Detmold
	FR24	Centre		DEA5	Arnsberg
	FR25	Lower Normandy		DEB1	Koblenz
	FR26	Burgundy		DEB2	Trier
	FR30	North-Pas de Calais		DEB3	Rhine-Hesse-Palatinate
	FR41	Lorraine		DEC0	Saarland
	FR42	Alsace		DEF0	Schleswig-Holstein
	FR43	Franche-Comté			

Country	Code	Region Name	Country	Code	Region Name
Greece	GR11	East Macedonia	Netherlands	NL34	Zeeland
	GR12	Central Macedonia		NL41	North Brabant
	GR13	West Macedonia		NL42	Limburg
	GR14	Thessaly			
	GR21	Epirus	Portugal	PT11	North
	GR22	Ionian Islands		PT15	Algarve
	GR23	West Greece		PT16	Centre
	GR24	Central Greece		PT17	Lisbon
	GR25	Peloponnese		PT18	Alentejo
	GR30	Attica			
	GR41	North Aegean	Spain	ES11	Galicia
	GR42	South Aegean		ES12	Principality of Asturias
	GR42	Crete		ES13	Cantabria
				ES21	Basque Country
Ireland	IE01	Border, Midlands and Western		ES22	Navarre
				ES23	La Rioja
	IE02	Southern and Eastern		ES24	Aragón
				ES30	Community of Madrid
Italy	ITC1	Piedmont		ES41	Castile and León
	ITC2	Aosta Valley		ES42	Castile-La Mancha
	ITC3	Liguria		ES43	Extremadura
	ITC4	Lombardy		ES51	Catalonia
	ITD1	Bolzano-Bozen		ES52	Community of Valencia
	ITD2	Trento		ES53	Balearic Islands
	ITD3	Veneto		ES61	Andalusia
	ITD4	Friuli-Venezia Giulia		ES62	Region of Murcia
	ITD5	Emilia-Romagna			
	ITE1	Tuscany	Sweden	SE01	Stockholm
	ITE2	Umbria		SE02	Eastern Central Sweden
	ITE3	Marche		SE04	South Sweden
	ITE4	Latium		SE06	Northern Central Sweden
	ITF1	Abruzzo		SE07	Central Norrland
	ITF2	Molise		SE08	Upper Norrland
	ITF3	Campania		SE09	Småland and Islands
	ITF4	Apulia		SE0A	West Sweden
	ITF5	Basilicata			
	ITF6	Calabria	UK	UKC1	Tees Valley and Durham
	ITG1	Sicily		UKC2	Northumberland and Tyne and Wear
	ITG2	Sardinia			
				UKD1	Cumbria
Luxembourg	LU00	Luxembourg		UKD2	Cheshire
				UKD3	Greater Manchester
Netherlands	NL12	Friesland		UKD4	Lancashire
	NL13	Drenthe		UKD5	Merseyside
	NL21	Overijssel		UKE1	East Riding and North Lincolnshire
	NL22	Gelderland			
	NL31	Utrecht		UKE2	North Yorkshire
	NL32	North Holland		UKE3	South Yorkshire
	NL33	South Holland		UKE4	West Yorkshire

Country	Code	Region Name	Country	Code	Region Name
UK	UKF1	Derbyshire and Nottinghamshire	UK	UKJ1	Berkshire, Buckinghamshire and Oxfordshire
	UKF2	Leicestershire, Rutland and Northamptonshire		UKJ2	Surrey, East and West Sussex
				UKJ3	Hampshire and Isle of Wight
	UKF3	Lincolnshire		UKJ4	Kent
	UKG1	Herefordshire, Worcestershire and Warwickshire		UKK1	Gloucestershire, Wiltshire and North Somerset
				UKK2	Dorset and Somerset
	UKG2	Shropshire and Staffordshire		UKK3	Cornwall and Isles of Scilly
	UKG3	West Midlands		UKK4	Devon
	UKH1	East Anglia		UKL1	West Wales and The Valleys
	UKH2	Bedfordshire and Hertfordshire		UKL2	East Wales
				UKM1	North Eastern Scotland
	UKH3	Essex		UKM2	Eastern Scotland
	UKI1	Inner London		UKM3	South Western Scotland
	UKI2	Outer London		UKM4	Highlands and Islands
				UKN0	Northern Ireland

NUTS version: 2003.

B Summary Statistics

Table B.1: Summary Statistics

Variable		Mean	Std. Dev.	Min.	Max.	Observations
y_{it}	overall	16,536	5,161	5,871	54,367	$N = 5018$
	between		4,468	8,293	38,505	$n = 193$
	within		2,601	2,066	32,406	$T = 26$
$\Delta \ln y_{it}$	overall	0.018	0.029	-0.159	0.479	$N = 4825$
	between		0.007	0.001	0.054	$n = 193$
	within		0.027	-0.162	0.476	$T = 25$
s_{it}	overall	0.223	0.047	0.088	0.500	$N = 5018$
	between		0.037	0.151	0.327	$n = 193$
	within		0.030	0.023	0.431	$T = 26$
n_{it}	overall	0.004	0.006	-0.046	0.060	$N = 4825$
	between		0.003	-0.006	0.015	$n = 193$
	within		0.005	-0.052	0.061	$T = 25$

Table B.2: Correlation Matrix

Panel A: Five-Yearly Data	$\Delta \ln y_{it}$	$\ln y_{i,t-1}$	$\ln s_{it}$	$\ln(n_{it} + g + \delta)$
$\Delta \ln y_{it}$	1			
$\ln y_{i,t-1}$	-0.1889***	1		
$\ln s_{it}$	0.0475	-0.1793***	1	
$\ln(n_{it} + g + \delta)$	-0.0774**	0.1130***	0.1496***	1
Panel B: Annual Data	$\Delta \ln y_{it}$	$\ln y_{i,t-1}$	$\ln s_{it}$	$\ln(n_{it} + g + \delta)$
$\Delta \ln y_{it}$	1			
$\ln y_{i,t-1}$	-0.0708***	1		
$\ln s_{it}$	0.0126	-0.1538***	1	
$\ln(n_{it} + g + \delta)$	-0.1313***	0.0997***	0.0705***	1
Panel C: Annual Data, Cross-Sectionally Demeaned	$\Delta \ln \widetilde{y}_{it}$	$\ln \widetilde{y}_{i,t-1}$	$\ln \widetilde{s}_{it}$	$\ln(\widetilde{n}_{it} + g + \delta)$
$\Delta \ln \widetilde{y}_{it}$	1			
$\ln \widetilde{y}_{i,t-1}$	-0.0861***	1		
$\ln \widetilde{s}_{it}$	-0.0187	-0.1759***	1	
$\ln(\widetilde{n}_{it} + g + \delta)$	-0.1072***	0.0667***	0.0754***	1

Notes: *** and ** indicate significance at the 1% and 5% levels respectively.

C Data Properties

Table C.1: Pesaran (2007) Test for Unit Roots under Cross-Section Dependence

	Constant			Constant and trend		
lags	$\ln y_{it}$	$\ln s_{it}$	$\ln(n_{i,t}+g+\delta)$	$\ln y_{it}$	$\ln s_{it}$	$\ln(n_{i,t}+g+\delta)$
0	0.17 (0.57)	4.63 (1.00)	-15.02 (0.00)	3.57 (1.00)	-5.15 (0.00)	-14.43 (0.00)
1	-1.47 (0.93)	6.26 (1.00)	-4.89 (0.00)	5.92 (1.00)	-1.76 (0.04)	-4.40 (0.00)
2	5.00 (1.00)	9.71 (1.00)	2.51 (0.99)	9.90 (1.00)	2.89 (0.99)	3.27 (0.99)
3	8.05 (1.00)	11.41 (1.00)	9.29 (1.00)	14.40 (1.00)	-8.39 (1.00)	11.48 (1.00)
	Constant			Constant and trend		
lags	$\Delta \ln y_{it}$	$\Delta \ln s_{it}$	$\Delta\ln(n_{i,t}+g+\delta)$	$\Delta \ln y_{it}$	$\Delta \ln s_{it}$	$\Delta\ln(n_{i,t}+g+\delta)$
0	-36.97 (0.00)	-46.89 (0.00)	-52.03 (0.00)	-33.07 (0.00)	-40.71 (0.00)	-48.44 (0.00)
1	-20.34 (0.00)	-26.66 (0.00)	-32.72 (0.00)	-16.79 (0.00)	-20.65 (0.00)	-28.05 (0.00)
2	-8.95 (0.00)	-15.65 (0.00)	-19.43 (0.00)	-5.64 (0.00)	-9.26 (0.00)	-15.50 (0.00)
3	-2.49 (0.00)	-6.54 (0.00)	-4.58 (0.00)	-0.60 (0.28)	-1.04 (0.15)	-2.70 (0.00)

Notes: Z-statistic reported, p-values in parentheses; Z statistic is asymptotically standard normal under the null hypothesis that all region-specific series are nonstationary; alternative hypothesis (one-sided) is that *some* region-specific series are stationary; the test is based on a standardised average of N region-specific ADF test statistics, where the ADF regressions are augmented with cross-section averages of dependent and all explanatory variables (lagged levels and first-differences of the individual series) to control for unobserved common factors that may affect each region differently (cross-section dependence).

Table C.2: Pesaran (2004) Test for Cross-Section Dependence

Panel A: Five-Yearly Data

	$\ln y_{it}$	$\ln s_{it}$	$\ln(n_{i,t}+g+\delta)$	$\Delta \ln y_{it}$	$\Delta \ln s_{it}$	$\Delta\ln(n_{i,t}+g+\delta)$
CD test	282.84	7.41	33.96	108.19	30.06	33.17
p-value	0.000	0.000	0.000	0.000	0.000	0.000
Avg. ρ	0.929	0.024	0.112	0.397	0.110	0.122
Avg. $\lvert\rho\rvert$	0.929	0.568	0.447	0.572	0.501	0.506

Panel B: Annual Data

	$\ln y_{it}$	$\ln s_{it}$	$\ln(n_{i,t}+g+\delta)$	$\Delta \ln y_{it}$	$\Delta \ln s_{it}$	$\Delta\ln(n_{i,t}+g+\delta)$
CD test	625.55	66.25	64.09	135.24	108.32	30.34
p-value	0.000	0.000	0.000	0.000	0.000	0.000
Avg. ρ	0.919	0.097	0.094	0.203	0.162	0.046
Avg. $\lvert\rho\rvert$	0.919	0.425	0.278	0.271	0.242	0.203

Panel C: Annual Data, Cross-Sectionally Demeaned

	$\ln \widetilde{y}_{it}$	$\ln \widetilde{s}_{it}$	$\ln(\widetilde{n}_{i,t}+g+\delta)$	$\Delta \ln \widetilde{y}_{it}$	$\Delta \ln \widetilde{s}_{it}$	$\Delta\ln(\widetilde{n}_{i,t}+g+\delta)$
CD test	2.95	-0.23	14.20	-0.71	1.37	21.46
p-value	0.003	0.815	0.000	0.475	0.170	0.000
Avg. ρ	0.004	0.000	0.021	-0.001	0.002	0.032
Avg. $\lvert\rho\rvert$	0.449	0.449	0.281	0.210	0.209	0.214

Notes: CD test statistic is asymptotically standard normal under the null hypothesis of no cross-section dependence; the test is based on the simple average of the $N(N-1)$ pairwise correlation coefficients between the regional series; avg. ρ and avg. $|\rho|$ are average and average absolute correlation coefficients.