ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Fischer, Manfred M.; LeSage, James P.

Conference Paper

A Bayesian space-time approach to identifying and interpreting regional convergence clubs in Europe

53rd Congress of the European Regional Science Association: "Regional Integration: Europe, the Mediterranean and the World Economy", 27-31 August 2013, Palermo, Italy

Provided in Cooperation with:
European Regional Science Association (ERSA)

Suggested Citation: Fischer, Manfred M.; LeSage, James P. (2013) : A Bayesian space-time approach to identifying and interpreting regional convergence clubs in Europe, 53rd Congress of the European Regional Science Association: "Regional Integration: Europe, the Mediterranean and the World Economy", 27-31 August 2013, Palermo, Italy, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:
https://hdl.handle.net/10419/123835

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

A Bayesian space-time approach to identifying and interpreting regional convergence clubs in Europe

February 5, 2013

Manfred M. Fischer
Department for SocioEconomics
Vienna University of Economics and Business
Vienna, Austria
E-mail: Manfred.Fischer@wu.ac.at
and
James P. LeSage
Fields Endowed Chair
Texas State University-San Marcos
Department of Finance & Economics
San Marcos, TX 78666
E-mail: jlesage@spatial-econometrics.com

Abstract

This study suggests a two-step approach to identifying and interpreting regional convergence clubs in Europe. The first step calculates Bayesian probabilities for various assignments of regions to two clubs using a general stochastic space-time dynamic panel relationship between growth rates and initial levels of income as well as endowments of physical, knowledge and human capital. This approach produces club assignments that are unconditional on specific parameter estimates. The second step uses the club assignments in a dynamic space-time panel data model to assess long-run dynamic direct and spillover responses of regional income levels to changes in initial period endowments for clubs that were identified. Correctly determining the dynamic partial derivative impacts of changes in initial endowments on regional income levels is an important contribution of our study. The dynamic trajectories of regional income levels over time allow us to draw inferences regarding the timing and magnitude of regional income responses to changes in (physical capital, human capital and knowledge capital) endowments for the clubs that have been identified in the first step. We find different responses to endowments by regions in two clubs that appear consistent with low- and high-income regions as clubs.

KEYWORDS: Dynamic space-time panel data model, Bayesian Model Comparison, European regions.

JEL: C11, C23, O47, O52

Introduction

During the past two decades, the question of systematically identifying and interpreting convergence clubs has received increasing attention.[1] A convergence club is a group of economies whose initial conditions are near enough to converge towards the same long-term equilibrium. This notion of club convergence can be traced back to Baumol (1986), but owes its more rigorous formulation to Durlauf and Johnson (1995), and Galor (1996). The concept is based on new growth theories that yield multiple, locally stable steady state equilibria in per capita output.[2] In contrast to conventional wisdom Galor (1996) has demonstrated that if heterogeneity is permitted across individuals, multiplicity of stationary equilibria may also occur in Solow (1956) and Mankiw et al. (1992) worlds, and in these cases the distribution of initial income per capita determines the club to which a particular region will belong.[3] But neither neoclassical nor new growth theories offer explicit guidance in determining the number and composition of clubs within a given cross-section of regions.

The determination of such clubs is a difficult issue. Some authors use a priori (exogenous) criteria to define the clubs, such as belonging to the same geographic area or having similar initial per capita incomes. But endogenous club determination dominates in the literature. Endogeneity may involve either the number of clubs, the composition of clubs, or both. Endogenous methods of club determination are quite diverse and include classification and regression tree methods (Durlauf and Johnson 1995; Fagerberg and Verspagen 1996), projection pursuit methods (Desdoigts 1999), cluster-analytic methods combined with cointegration tests (Hobijn and Franses 2000; Corrado et al. 2005), cluster-analytic methods along with time series regression tests (Phillips and Sul 2007; Bartkowska and Riedl 2012), switching regression methods (Fève and Le Pen 2000), and Bayesian methods that identify a mixture distribution for the predictive density of per capita output (Canova 2004), among others. All these club selection methodologies, however, are non-spatial in nature, and hence neglect the importance of spatial dependence for club detection.

Some other methods of club determination explicitly take into account the spatial dimension of the data and use exploratory spatial data analysis tools to detect spatial regimes. Examples include Moran scatterplots, based on the initial per capita income of a sample of European regions, to determine spatial clubs (Ertur, LeGallo and Baumont 2006). Fischer and Stirböck (2006), LeGallo and Dall'erba (2006), and Dall'erba and LeGallo (2008)

alternatively take the Getis-Ord statistics, applied to initial per capita income. This generates a two-way partitioning of the sample: spatial clusters of high values of per capita income (corresponding to positive values of the statistic) and spatial clusters of low values of per capita income (corresponding to negative values of the statistic). These methods may be labelled semi-endogenous, since the number of clubs is fixed a priori (four in the case of Moran scatterplots and two in the case of the Getis-Ord statistic), but the regions are endogenously allocated to the clubs.

This paper suggests a novel methodological approach for (semi-endogenous) club determination that draws on Bayesian ideas to identify regional convergence clubs in Europe. A very general stochastic panel *relationship*[4] between growth rates and initial levels of income as well as endowments of physical, knowledge and human capital with regional random effects is used to derive posterior probabilities for various assignments of regions to two clubs.[5] These are based on splits of the sample of regional growth rates according to varying levels of initial period income. This stochastic space-time dynamic panel relationship serves not as a model with corresponding parameters to be estimated, but rather as a basis for integrating over all parameters of the relationship to find posterior probabilities for the various assignment of regions to the two clubs based on sample splits according to varying initial period income levels.

It is important to note that rather than estimating a single set of panel data parameters for the growth rates versus initial levels (of income and endowments), our approach is to integrate over all possible values that can be taken by the parameters of this relationship. Determining posterior probabilities for club assignment based on the varying initial levels of income using this approach means that our club assignments do not depend on particular parameter values taken by variables in the relationship, (including parameters for space, time and space-time covariance of the growth rates process, as well as the regional random effects parameters, and parameters that relate initial period income and endowments levels to the growth rates).

In other words, our club selection approach avoids dependence on specific parameter estimates and relies on a very general dynamic space-time panel relationship of growth rates and initial period income and endowments, and this is in contrast to non-spatial and spatial cross-section studies/relationships as well as studies based on both non-spatial and spatial panel data models. This is one important contribution of our methodology, a

procedure for club assignment of regions that is unconditional on any specific set of model estimates.

In addition to the club selection methodology, our paper distinguishes itself from previous work by using a Bayesian space-time panel data model to assess long-run dynamic direct and spillover responses of regional income levels to changes in initial period endowments that are not part of other studies.[6] Correctly determining the partial derivative impact of changes in initial endowments on regional income levels is another important contribution of our study. The dynamic trajectories of changes in regional income levels over time allow us to draw inferences regarding the timing and magnitude of regional income responses to changes in (physical capital, human capital and knowledge capital) endowments for the clubs that have been identified in the first step of the methodology. We find different responses to endowments by regions in the two clubs that appear to be considered with low- and high-income regions as clubs.

The next section outlines the general stochastic panel relationship, and the formal methodology for calculating posterior probabilities for club assignment of regions, as it applies to our work here.[7] Of course, the resulting club classification is conditional on the dynamic space-time panel relationship used in the comparison procedure.

The methodology for identifying clubs

The first step of our approach uses a formal Bayesian methodology to classify European regions into clubs. Each region must be classified into one of two clubs. The classification takes place conditional on a space-time (random effects) panel data relationship[8] of regional income growth given by

$$\begin{aligned} g_t &= \phi g_{t-1} + \rho W g_t + \theta W g_{t-1} + \psi \ln y_{t-1} + \alpha \iota_N + X_{t-1}\beta + \varphi_t, \qquad (1) \\ \varphi_t &= \mu + \varepsilon_t, \\ g_t &= \ln(y_t) - \ln(y_{t-1}), \qquad t = 2, \ldots, T. \end{aligned}$$

The growth regression relationship is between the $N \times 1$ vector of time t growth rates (g_t) and those from the previous time period (g_{t-1}), neighboring regions in the current time

period (Wg_t), and also to those of neighboring regions in the previous time period (Wg_{t-1}). $g_t = (g_{1t}, \ldots, g_{Nt})'$ is the $N \times 1$ vector of observed income growth rates for the tth time period, with y_t denoting income levels at time t, and ψ the parameter reflecting dependence on previous period levels. The intercept parameter is α and ι_N is an $N \times 1$ column vector of ones. Previous period endowments of physical capital, knowledge capital and human capital which are thought to exert an influence on regional income growth are contained in the $N \times K$ matrix X_{t-1} with K denoting the number of (conditioning) variables included to capture proximate determinants of economic growth and β representing the associated parameter vector.

The vector $\varphi_t = \mu + \varepsilon_t$ represents the summation of two unobserved normally distributed random components: μ an $N \times 1$ column vector of random effects with $\mu_i \sim \mathcal{N}(0, \sigma_\mu^2), i = 1, \ldots, N$, that are the same for all time periods, and the $N \times 1$ stochastic disturbance ε_t, assumed to be independent and identically distributed with zero mean and scalar variance $\sigma_\varepsilon^2 I_N, t = 1, \ldots, T$. We make the traditional assumption that μ is uncorrelated with ε_t for identification purposes. W is a known $N \times N$ spatial weight matrix whose diagonal elements are zero. This matrix defines the dependence between cross-sectional (spatial) observational units. We will also assume that W is row-normalized from a symmetric matrix, so that all eigenvalues are real and less than or equal to one. The strength of spatial dependence is measured by the parameter ρ, the first order time dependence is reflected in the scalar parameter ϕ, and θ represents the component mixing space and time dependence.

We note for future reference that the relationship in (1) can be viewed as applying a space-time filter to the regional growth rates. Specifically we have expressions (2) and (3), where L is a time lag operator: $Lg_t = g_{t-1}$.

$$
\begin{aligned}
(I_N - \phi L)(I_N - \rho W) g_t &= \psi \ln y_{t-1} + \alpha \iota_N + X_{t-1}\beta + \varphi_t, \qquad t = 2, \ldots, T \qquad (2) \\
(I_N - \phi L - \rho W + \phi \rho L W) g_t &= \psi \ln y_{t-1} + \alpha \iota_N + X_{t-1}\beta + \varphi_t \\
(I_N - \phi L - \rho W - \theta L W) g_t &= \psi \ln y_{t-1} + \alpha \iota_N + X_{t-1}\beta + \varphi_t. \qquad (3)
\end{aligned}
$$

This view of the relationship implies a restriction on the parameter $\phi\rho$ of the cross-product term (LW), but we introduce a parameter θ indicated in (3) that does not impose the restriction $\theta = -\phi\rho$.[9]

Most theoretical models of multiple steady states (see, for example, Azariadis and Drazen 1990; Galor 1996) predict that if (regional) economies are concentrated around several steady states, then their initial per capita output levels (here measured in terms of GVA per capita levels) will fall into distinct (i.e. non-overlapping) categories (Durlauf and Johnson 1995). This motivates our use of a panel data variant of the traditional cross-sectional growth regression relationship to identify/classify regions into low- and high-income clubs (which we label Club 1 and Club 2). The relationship of focus is that between *income growth rates* and the previous period (logged) level of income and (logged) previous period endowments of physical, knowledge and human capital.

Islam (1995) was one of the first studies to examine conventional cross-sectional growth regressions using a panel data setting. He proposed splitting the overall sample into several shorter time spans, with the motivation being that annual growth rates are "too short to be appropriate for studying growth convergence" since "short-term disturbances may loom large in such brief time spans". He relied on five-year time intervals for his panel data model estimation. Despite this brief and informal argument against using annual growth rates in a panel data setting, almost all panel data growth convergence studies have followed the approach of Islam (1995). This includes the space-time dynamic panel data relationship used in a growth convergence study by Yu and Lee (2012), where four and five year intervals were used.

Our limited sample size of 11 years does not allow us to fully explore this issue. In the context of Islam's non-dynamic panel data model, use of initial period endowments from four or five years ago may make sense. It is less clear how one should proceed for the case of a space-time dynamic relationship of the type in (1). For a dynamic model that takes into account space, time and space-time diffusion of regional growth rates, imposing multi-year time intervals should not be required when calculating growth rates. We also show how the partial derivatives from these models imply both short- and long-run responses of the dependent variable to changes in values of the explanatory variables. This issue requires further econometric study, perhaps in a Monte Carlo or simulation setting. To explore this issue, we consider relationship (4) that imposes a 3-year lag on the initial period endowment variables. This specification could be justified on an ad-hoc basis as in Islam (1995) because it allows a longer time lag before endowment levels influence income growth rates. We present classification results for regions into the two clubs based on relationships (1) and

(4), and these produced similar results.

$$
\begin{aligned}
g_t &= \phi g_{t-1} + \rho W g_t + \theta W g_{t-1} + \psi \ln y_{t-3} + \alpha \iota_N + X_{t-3}\beta + \varphi_t, \qquad (4)\\
\varphi_t &= \mu + \varepsilon_t,\\
g_t &= \ln(y_t) - \ln(y_{t-3}), \qquad t = 4, \ldots, T.
\end{aligned}
$$

The dynamic space-time panel data relationship in (1) expressed in matrix/vector form shown in (5) is used in conjunction with Bayesian methods to assign regions to clubs.

$$
\begin{aligned}
Pg &= H\psi + \iota_{N(T-1)}\alpha + X\beta + \varphi \qquad (5)\\
P &= \begin{pmatrix} B & 0_{N\times N} & 0_{N\times N} & \ldots & 0_{N\times N} \\ A & B & 0_{N\times N} & \ldots & 0_{N\times N} \\ 0_{N\times N} & A & B & & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0_{N\times N} \\ 0_{N\times N} & \ldots & 0_{N\times N} & A & B \end{pmatrix} \qquad (6)\\
H &= \begin{pmatrix} \ln(y_1) & \ldots & \ln(y_{T-1}) \end{pmatrix}'\\
X &= \begin{pmatrix} X_1 & \ldots & X_{T-1} \end{pmatrix}'\\
A &= -(\phi I_N + \theta W)\\
B &= I_N - \rho W\\
\varphi &\sim \mathcal{N}(0, \Omega)\\
\Omega &= [(T-1)\sigma_\mu^2 + \sigma_\varepsilon^2](\bar{J}_{T-1} \otimes I_N) + \sigma_\varepsilon^2 \left[(I_{T-1} - \bar{J}_{T-1}) \otimes I_N\right]. \qquad (7)
\end{aligned}
$$

We use $\otimes$ to denote the Kronecker product in the expression for Ω in (7), which represents a decomposition proposed by Wansbeek and Kapteyn (1982), that replaces $J_{T-1} = \iota_{T-1}\iota'_{T-1}$ by its idempotent counterpart $\bar{J}_{T-1} = J_{T-1}/(T-1)$ (see Parent and LeSage 2012).

The scalars σ_μ^2 and σ_ε^2 denote the variances of the random effects vector μ and noise vector ε, respectively. This specification uses the first time period to "feed the lag", leading to the $N(T-1) \times NT$ matrix P in (6). Treating the first period in this way simplifies work involved in analytically calculating the log-marginal likelihood needed to compute posterior

probabilities for model comparison purposes, and should have little impact in cases where T is reasonably large.

In our empirical application $N = 216$ European regions and $T = 11$ years covering the period from 1995 to 2005, with the initial period being 1995, so T is not excessively large here. To assign regions to candidate clubs we introduce a dummy variable that splits the sample according to initial year (1995) regional income levels above and below (or equal to) m during the initial year 1995. Regions with incomes below m are assigned to Club 1 and those with incomes above this level to Club 2. In (8), we express the dynamic panel relationship including the $N \times 1$ dummy vector D with zero values for regions where $y_1 \leq m$ and ones for $y_1 > m$, and an $N \times K$ dummy matrix $\tilde{D} = \begin{pmatrix} D & D & \dots & D \end{pmatrix}$. The Hadamard (element-by-element) product $\odot$ is used in conjunction with the dummy matrix $\tilde{D}$, where we use $\tilde{\alpha}, \tilde{\beta}, \tilde{\psi}$ for parameters associated with the club dummy.

$$\begin{aligned} g_t &= \phi g_{t-1} + \rho W g_t + \theta W g_{t-1} + \psi \ln(y_{t-1}) + \tilde{\psi} D \ln(y_{t-1}) \\ &+ \alpha \iota_N + \tilde{\alpha} D \iota_N + X_{t-1}\beta + (\tilde{D} \odot X_{t-1})\tilde{\beta} + \varphi_t, \qquad t = 2, \dots, T. \end{aligned} \tag{8}$$

Parent and LeSage (2011) show that the log-likelihood for this relationship (with the random effects vector μ integrated out) can be expressed as in (9). For simplicity we use Z to denote a matrix containing all explanatory variables for each time period, and we define: $\lambda = \sigma_\mu^2/\sigma_\varepsilon^2$.

$$\begin{aligned} \ln \mathcal{L}_{T-1}(\upsilon) &= -\frac{N(T-1)}{2}\ln(2\pi) - N(T-1)[\ln(\sigma_\mu^2) - \ln(\lambda)] - N\ln((T-1)\lambda + 1) \\ &+ T\ln|I_N - \rho W| - \frac{1}{2(\sigma_\mu^2/\lambda)} e'\Omega^{-1}e \\ \lambda &= \sigma_\mu^2/\sigma_\varepsilon^2 \\ e &= (Pg - Z\delta) \\ Z &= \begin{pmatrix} Z_1 & \dots & Z_{T-1} \end{pmatrix}' \\ Z_{t-1} &= \begin{pmatrix} \ln y_{t-1} & D\ln y_{t-1} & \iota_N & D\iota_N & X_{t-1} & (\tilde{D} \odot X_{t-1}) \end{pmatrix} \\ \delta &= \begin{pmatrix} \psi & \tilde{\psi} & \alpha & \tilde{\alpha} & \beta & \tilde{\beta} \end{pmatrix}' \\ \upsilon &= (\phi, \rho, \theta, \lambda, \delta'). \end{aligned} \tag{9}$$

To determine Bayesian probabilities for various assignments of regions to the two clubs, we wish to find an expression for the log-marginal likelihood. Zellner (1971) sets forth the basic Bayesian approach to this. This involves specifying prior probabilities for each split of the sample as well as prior distributions for the regression parameters. Posterior probabilities are calculated for each split of regions according to initial income levels and used for inferences regarding the "best split" of the sample into two clubs. The Bayesian theory involves specifying prior probabilities for each of the r alternative sample splits $\{R_1, R_2, \ldots, R_r\}$ under consideration, which we label $\pi(R_q), q = 1, \ldots, r$, as well as prior distributions for the parameters $\pi(\upsilon)$. If the sample data are to determine the posterior probabilities for the various splits, the prior probabilities should be set to equal values of $1/r$, making each split equally likely a priori. We treat the spatial weight matrix W as fixed and exogenous, relying on a weight structure consisting of the 10 nearest neighboring regions (measured in terms of great circle distances). The motivation for this is that use of the 10 nearest neighboring regions allows the island regions of Greece to be connected to mainland Greece.[10] We also treat the number of clubs as fixed at two, but hope to extend this in future research.

The prior distributions for the parameters are combined with the likelihood for (g, Z, W) conditional on υ as well as the set of models R, which we denote $p(g|\upsilon, R, Z, W)$. The joint probability for R_q, υ, and g takes the form in (10), for the qth sample split at initial period income level $m = m_q$.

$$p(R_q, \upsilon, g, Z, W, m = m_q) = \pi(R_q)\pi(\upsilon|R_q)p(g|\upsilon, R, Z, W). \tag{10}$$

Application of Bayes rule produces the joint posterior for both split levels and parameters as:

$$p(R_q, \upsilon|g, Z, W) = \frac{\pi(R_q)\pi(\upsilon|R_q)p(g|\upsilon, R, Z, W)}{p(g)}. \tag{11}$$

The posterior probabilities regarding the split levels take the form:

$$p(R_q|g, Z, W) = \int p(R_q, \upsilon|g, Z, W)d\upsilon \tag{12}$$

which requires integration over the parameter vector υ. We follow LeSage and Parent (2007)

who develop expressions for the log-marginal likelihood in the case of a cross-sectional model by analytically integrating out the parameters δ and σ_ε, leaving a simple univariate numerical integration over the spatial dependence parameter ρ. Things are more complicated here, but we are able to analytically integrate out the parameters δ (see the Appendix for technical details), and numerically integrate over the space and time dependence parameters ρ and ϕ. This requires that we fix $\lambda = \sigma_\mu^2/\sigma_\varepsilon^2$, the variance ratio of the random effects and noise.

We make the following observation regarding λ. For small values of λ the effects magnitudes are likely to be close to their mean values of zero and not of substantive importance. Large values for the effects magnitudes accompanied by large values of λ likely suggest the relationship is not consistent with the sample data. This leads us to posit that a well-specified relationship would exhibit probabilities for various splits of the sample that should not be sensitive to fixing the value of λ, based on say, estimates for the parameters $\sigma_\mu^2, \sigma_\varepsilon^2$ from a panel data model with no dummy variables. We examine the resulting posterior probabilities at values around the estimated value: $\hat{\lambda} = \hat{\sigma}_\mu^2/\hat{\sigma}_\varepsilon^2$, to check robustness of results with regard to this ratio of variances.

Another simplification can be achieved by fixing the parameter $\theta = -\rho\phi$ which is a restriction implied by the space-time filter view of the panel data relationship. Parent and LeSage (2012) discuss the role of this restriction which simplifies both estimation and interpretation in the context of using our relationship to actually estimate the parameters of a model. They also show that the restriction is often consistent with sample data sets, a finding in the empirical application undertaken here. The advantage of this restriction is that we have a bivariate numerical integration problem involving the parameters ϕ and ρ rather than trivariate numerical integration.

To conclude this discussion, we note that our methodological approach for club determination draws on Bayesian ideas to identify regional convergence clubs in Europe. Using a general stochastic panel relationship between growth rates and initial levels of income as well as endowments of physical, knowledge and human capital with regional random effects we derive posterior probabilities for various assignments of regions to two clubs. This stochastic space-time dynamic panel relationship serves not as a model with corresponding parameters to be estimated, but rather as a basis for integrating over all parameters of the relationship to find posterior probabilities for the various assignment of regions to the

two clubs based on sample splits according to varying initial period income levels. This procedure assigns regions to clubs in such a way that the resulting clubs do not depend on specific parameter estimates from a single model specification. This approach is in contrast to non-spatial and spatial cross-section and panel data studies found in the literature. This is one important contribution of our methodology, a procedure for club assignment of regions that is unconditional on any specific set of model estimates.

An alternative approach to assigning regions to clubs would be to attempt random assignment of a single region to each club and then examine the posterior probabilities for these two assignments. Postiglione et al. (2012) adopt this type of random assignment of individual regions strategy, but rely on the Potts model from image processing to impose a contiguity penalty on assignment of regions to clubs. The Potts smoothing penalty arises with image restoration because nearby pixels from images are likely to be the same as neighboring pixels. A measure of fit (normalized sum of squared errors) based on residuals from a spatial Durbin model is used in conjunction with the Potts penalty term in either a simulated annealing or iterated conditional modes optimization routine. They find evidence of four clubs using a sample of 187 EU regions over the 1981 to 2004 period.

We attempted a random assignment approach similar to that of Postiglione et al. (2012) using our log-marginal likelihood as a Metropolis-Hastings accept/reject step in a formal Bayesian Markov Chain Monte Carlo comparison approach. Our attempt to implement this strategy indicated that posterior probabilities changed very slightly when a single region was moved from one club assignment to another. A reason for this can be see by considering Figure 1 which shows a frequency distribution of 1995 GVA per capita levels for regions.[11] Large changes in the number of regions assigned to each club from splitting the sample of regions at some income levels would lead to more dramatic changes in the posterior probabilities for these splits of the sample. This should be clear by considering that adding or subtracting a single region from the set of Club 1 regions should lead to small changes in the log-marginal likelihood (and associated probabilities). In contrast, changing the subsamples through addition or subtraction of many regions would lead to larger changes in the posterior probabilities.

(Figure 1 about here)

The smaller changes in posterior probabilities meant that our formal approach to assigning single regions to one of the two clubs failed to produce convergence. The Postiglione et al. (2012) approach should be similar to the one we tried, with the important difference being the contiguity smoothing penalty. Since this penalty was not included in our more formal Bayesian comparison approach, it could mean that their results are driven by the Potts penalty term. A spatial contiguity penalty may not be consistent with economic theory that predicts regional economies concentrated around several steady states will have initial per capita output levels that fall into distinct (non-overlapping) categories (Durlauf and Johnson 1995).

This motivated our approach that calculates posterior probabilities for sample splits based on a number of different candidate initial period income levels. The empirically determined club assignments are reported in the next section.

Empirical club assignments

We present results from implementing the club assignment strategy that calculates Bayesian probabilities associated with splits according to varying 1995 income levels for a set of 216 European regions. This includes: a description of the sample data used to form the relationship in (1) between growth rates and initial period income as well as endowments of physical, knowledge and human capital; and a data generated example to demonstrate that the methodology works; and empirical results showing posterior probabilities associated with splits of the 216 regions at various initial period income levels.

The sample data

Our sample is a cross-section of 216 regions representing the 15 pre-2004 EU member states, Norway and Switzerland over the 1995-2005 period. The units of observation are the NUTS-2 regions[12] (NUTS revision 2003). These regions, though varying in size, are generally considered to be appropriate spatial units for modelling and analysis purposes. In most cases, they are sufficiently small to capture subnational variations. But we are aware that NUTS-2 regions are formal rather than functional regions, and their delineation does not represent the boundaries of regional growth processes very well. The sample regions include

regions located in Europe covering Austria (nine regions), Belgium (11 regions), Denmark (one region), Finland (five regions), France (22 regions), Germany (40 regions), Greece (13 regions), Ireland (two regions) Italy (20 regions), Luxembourg (one region), the Netherlands (12 regions), Norway (seven regions), Portugal (five regions), Spain (16 regions), Sweden (eight regions), Switzerland (seven regions) and United Kingdom (37 regions).

We use gross-value added, GVA, rather than gross regional product at market prices as a proxy for regional income. The proxy is measured in accordance with the European Systems of Accounts (ESA) 1995. The data for the EU regions come from Eurostat's Regio database, and those for Norway and Switzerland from Statistics Norway (Division for National Accounts) and the Swiss Office Féderal de la Statistique (Comptes Nationaux), respectively.

We use three variables in the dynamic space-time growth relationship to group regions based on initial levels: physical capital, knowledge capital and human capital. Physical capital stock data is not available in Cambridge econometrics database, but gross fixed capital formation in current prices is. Thus, the stocks of physical capital were derived for each region from investment flows, using the perpetual inventory method. We applied a constant rate of 10 percent depreciation, and the annual flows of fixed investments were deflated by national gross-fixed capital formation deflators. The mean annual rate of growth, which precedes the benchmark year 1995, covers the period 1990-1994 to estimate initial regional physical capital stocks.

Corporate patent applications are used to proxy knowledge capital. Corporate patents cover inventions of new and useful processes, machines, manufactures, and compositions of matter. To the extent that patents document inventions, an aggregation of patents is arguably more closely related to a stock of knowledge than is an aggregation of R&D expenditures. However, a well known problem of using patent data is that technological inventions are not all patented. This could be because applying for a patent, is a strategic decision and, thus, not all patentable inventions are actually patented. Even if this is not an issue, as long as a large part of knowledge is tacit, patent statistics will necessarily miss that part, because codification is necessary for patenting to occur.

Patent stocks were derived from European Patent Office (EPO) documents. Each EPO document provides information on the inventor(s), his or her name and address, the company or institution to which property rights have been assigned, citations to previous

patents, and a description of the device or process. To create the patent stocks for 1995-2005, the EPO patents with an application date 1990-2005 were transformed from individual patents into stocks by first sorting based on the year that a patent was applied for, and second the region where the inventor resides. In the case of cross-region inventor teams we used the procedure of fractional rather than full counting. Then for each region i, patent stocks were derived from patent data, using the perpetual inventory method. Because of evident complications in tracking obsolescence over time, we used a constant depreciation rate of 12 years that corresponds to the rate of knowledge obsolescence in the US over the past century, as found in Caballero and Jaffe (1993). Patent stocks were initialized the same way as physical capital (see Fischer et al. 2009 for details).

There is no clear-cut consensus of how human capital should be represented and measured. In this study we follow Fischer et al. (2009) and measure human capital in terms of educational attainment based on data for the active population aged 15 years and older that attained the level of tertiary education, as defined by the International Standard Classification of Education (ISCED) 1997 classes five and six. This variable is clearly imperfect: it completely ignores primary and secondary education, and on-the-job training, and does not account for the quality of education.

A test of the club assignments methodology

We test the Bayesian comparison procedure using a generated vector of growth rates constructed from our sample data for 216 European regions. A set of parameter estimates were used to produce predicted values that reflected two regimes with regions' split at an income level of 20,000. When generating predicted values, parameters $\hat{\rho} = 0.65, \hat{\phi} = -0.18$ were used, in conjunction with a value of $\theta = 0.025$, which does not obey the restriction on the parameter $\theta = -\phi\rho$. Specifically, $\theta = -\phi\rho = -(-0.18 \times 0.65) = 0.117$, rather than the value $\theta = 0.025$ used to produce a sample of growth rates. However, posterior probabilities were calculated based on the assumption that $\theta = -\phi\rho$, as a test of the impact on performance in this type of setting where the assumption is violated. For this experiment, a dummy variable vector was used to split the sample at initial period income levels of $m = 20,000$.

The distributions of generated growth rates for the two clubs that resulted from this approach are shown in Figure 2, where we see the high income club exhibiting a slightly lower mean growth rate than the lower income club. This is of course consistent with

the usual notion of β-convergence, where regions with lower initial levels of income exhibit higher growth rates than higher income regions.

(Figure 2 about here)

The estimated ratio of variances $\hat{\lambda} = \hat{\sigma}_{\mu}^2/\hat{\sigma}_{\varepsilon}^2$ equalled 0.2594. Table 1 shows posterior probabilities derived from a comparison of splits of the regions based on initial period income levels ranging from $10,000$ to $32,000$ in increments of $2,000$, for values of $\hat{\lambda}$ as well as $(1/2)\ \hat{\lambda}$ and $2\hat{\lambda}$. The resulting posterior probabilities point to the correct split of the regions at the $m = 20,000$ level for all three settings of λ. As we would expect, there is some degradation of performance for values based on $1/2\lambda$ and 2λ, but the correct inference would be drawn in these cases. We also note that this test of the methodology relied on estimates from the growth relationship in (8), but altered values of ϕ and ρ so they did not obey the restriction $\theta = -\phi\rho$. This did not appear to produce erroneous inferences regarding the correct split level.

(Table 1 about here)

Club assignments of the regions

As indicated, there is some question regarding whether growth rates should be calculated using a single year or multiple year interval. Elhorst et al. (2010) study a panel data relationship between EU regional income growth rates as the dependent variable and growth rates of: savings, population, technical progress, and the depreciation rate, as explanatory variables. They make a number of methodological points including an observed decrease in (small T) bias arising from decreasing the time interval over which the growth rates are measured. Yu and Lee (2012) report estimation results from use of a dynamic spatial panel model covering 48 years but having no explanatory variables. They explore four and five year time intervals for calculating growth rates and find no difference in estimates for the time (ϕ), space (ρ) and space-time dependence (θ) parameters. They also find the data to be consistent with the restriction $\theta = -\phi\rho$.

We report posterior probability results for sample splits of the regions by initial period income levels using growth rates calculated from one- and three-year intervals in Table 2. Given our small number of 11 time periods, we were limited in our ability to explore this

issue. However, we find identical inferences regarding the initial period income level at which to split the sample of regions into two clubs.

(Table 2 about here)

There appears to be support for a split of the sample around 16,000, for the case of relation (1), with splits based on this level for initial period (1995) income levels exhibiting the highest posterior probabilities. These results were relatively stable across values of the noise variance ratio parameter λ, always giving slightly more posterior probability support for a split at 16,000. It should be noted that we are forcing a choice of "the best split" from this finite set of initial period income levels ranging from 8,000 to 30,000. This means that the posterior probabilities sum to unity, with all mass being assigned to the finite set of splits. For the specification in relation (4), results point to a split of the regions into two clubs based on 14,000 initial period income levels, which is close to the results for relation (1). It should be noted that these results involve a smaller panel of only eight periods because of the imposition of a 3-year lag on endowments.

The conclusion we draw from relation (1) is that the preponderance of evidence points to the existence of two clubs based on splitting the sample at initial period (1995) per capita GVA levels of 16,000. If we used relation (4), the split would occur at 14,000, producing a similar set of club assignments of regions. Figure 3 shows a map of the European regions classified into the two clubs based on a split according to the 1995 GVA per capita levels for regions above and below 16,000.

The next section describes the second step of our approach, which uses a dynamic space-time panel data model to analyze the space-time dynamic relationship between regional levels of income over time and space.[13] The model includes spatial and temporal dependence as well as space-time covariance so that changes in the endowments of a single region (say i) at time t can impact own- and other-regions ($j \neq i$) in the current and future time periods. In particular, we focus on the partial derivative impact of changes in the regional endowment variables on regional income levels at various time horizons, an issue that has received little attention in the spatial panel data model literature.[14]

(Figure 3 about here)

Space-time dynamics for the two clubs

The second step of our approach involves estimating a space-time dynamic panel data model that uses (logged) *levels* of regional income as the dependent variable and (logged) *levels* of previous period endowments of physical, knowledge and human capital stocks, to examine the response of regional income levels over space and time to changes in initial period endowments, in each of the two clubs of regions. Our focus is on the partial derivative effects associated with changing the physical, knowledge and human capital stocks. The next section sets forth the fixed effects variant of our dynamic space-time panel data model used for calculating dynamic response elasticities for regional income levels over space and time, to changes in initial period endowments of physical, knowledge and human capital stocks.

The space-time levels relationship

We use a fixed effects variant of our dynamic space-time panel data model, and focus on the (logged) levels relationship between the dependent y_t and explanatory variables X_{t-1} as well as a linear combination of neighboring region explanatory variables WX_{t-1}, with associated parameters η. The log transformation allows us to calculate dynamic response elasticities for regional income levels over space and time, to changes in initial period endowments of physical, knowledge and human capital stocks.[15] The (fixed effects) dynamic space-time panel model takes the form:

$$\begin{aligned} y_t &= \phi y_{t-1} + \rho W y_t + \theta W y_{t-1} + X_{t-1}\beta + (\tilde{D} \odot X_{t-1})\tilde{\beta} \\ &+ WX_{t-1}\eta + (\tilde{D} \odot WX_{t-1})\tilde{\eta} + \mathcal{F}\gamma + \varepsilon_t, \qquad t = 2, \ldots, T. \end{aligned} \tag{13}$$

where y_t, X_{t-1} have been log-transformed, ε_t is *i.i.d.* across i and t with zero mean and variance $\sigma_\varepsilon^2 I_N$, and $\mathcal{F}$ represent fixed effects with γ the associated parameters.

We rely on a Bayesian Markov Chain Monte Carlo estimation scheme described in Parent and LeSage (2012) to produce estimates of the parameters in the model. Our focus here is on the partial derivative effects associated with changing the explanatory variables in this

model, reflecting human and physical capital stocks as well as knowledge capital stocks.

This model has own- and cross-partial derivatives that measure the impact on own- and other-regions income. We will use y_{it} to reference elements in the $N \times 1$ vector y_t pertaining to the ith element/region at time t, and we drop the explicit log (ln) symbols for notational simplicity. The own-partial derivative: $\partial y_{it}/\partial X_{it}^k$, represents the time t *direct effect* on region i's (logged) income level (at time t), arising from a change in the kth explanatory variable (say logged physical capital levels) in region i (at time t). There is also a cross-partial derivative $\partial y_{jt}/\partial X_{it}^k$ that measures the time t *indirect effect*, that falling on regions (j) other than i, where most of the spatial spillover impacts fall on regions j that are nearby or neighbors to region i.

We are most interested in partial derivatives that measure how region i's (logged) income level responds over time to changes in the initial period (logged) endowment levels of physical and human capital, as well as knowledge capital, since this is the essence of the debate concerning regional convergence in levels of income over time. The model allows us to calculate partial derivatives that can quantify the magnitude and timing of regional income responses at various time horizons to changes in the initial period levels of the explanatory variables. Expressions for these are presented and discussed in the sequel. We simply note here that we are referring to: $\partial y_{it+T}/\partial X_{it}^k$ which measures the $T-$horizon own-region i response to changes in its initial endowments, and $\partial y_{jt+T}/\partial X_{it}^k$, that reflects spillovers/diffusion effects over time that impact other regions j when region i's initial period human, physical or knowledge capital are changed.

We follow Yu et al. (2008) and treat the dynamic space-time process as conditional on the initial cross-section. A careful analysis of issues related to treatment of the first period observation can be found in Parent and LeSage (2012), and we do not address this issue here. For simplicity of exposition, we assume that the first period is only subject to spatial dependence, which allows us to write the model as in (14), with accompanying definitions in (15), (16), (17) and (18).

$$\begin{aligned} QY &= \begin{pmatrix} X & (I_{T-1} \otimes W)X \end{pmatrix} \begin{pmatrix} \beta \\ \eta \end{pmatrix} + (I_{T-1} \otimes \tilde{D}) \odot \begin{pmatrix} X & (I_{T-1} \otimes W)X \end{pmatrix} \begin{pmatrix} \tilde{\beta} \\ \tilde{\eta} \end{pmatrix} \\ &+ \mathcal{F}\gamma + \varepsilon \end{aligned} \tag{14}$$

$$
Q = \begin{pmatrix} B & 0_{N\times N} & 0_{N\times N} & \dots & 0_{N\times N} \\ A & B & 0_{N\times N} & \dots & 0_{N\times N} \\ 0_{N\times N} & A & B & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0_{N\times N} \\ 0_{N\times N} & \dots & 0_{N\times N} & A & B \end{pmatrix} \tag{15}
$$

$$
A = -(\phi I_N + \theta W) \tag{16}
$$

$$
B = (I_N - \rho W) \tag{17}
$$

$$
\mathcal{F} = \iota_{T-1} \otimes I_N. \tag{18}
$$

The dependent variable vector $Y = (y_2', \dots, y_T')'$, consisting of $N \times 1$ vectors of cross-sectional observations for each time period $y_t = (y_{1t}, \dots, y_{Nt})'$. The matrix $X = (X_2', \dots, X_T')'$, so that X_t denotes the $N \times K$ matrix of (lagged) non-stochastic regressors at time t. We use X_{it}^k to reference elements associated with the kth variable for region i at time t. The matrix product $[(I_{t-1} \otimes \tilde{D}) \odot X]$ applies the club dummy variables to the explanatory variables X and WX for each time period, allowing for parameters β, η associated with Club 1, the low initial period income club, and parameters $\beta + \tilde{\beta}, \eta + \tilde{\eta}$ for Club 2, the high initial period income club.

The $N \times 1$ column vector γ represents fixed effects parameters, and the $N(T-1) \times N$ matrix $\mathcal{F}$ the associated regional indicator variables. The disturbance vector $\varepsilon = (\varepsilon_2', \dots, \varepsilon_T')'$, with $\varepsilon_t = (\varepsilon_{1t}, \dots, \varepsilon_{Nt})'$, $t = 2, \dots, T$, assumed to be *i.i.d.* across i and t, with zero mean and variance σ^2. Spatial dependence is measured by the parameter ρ and time dependence is reflected in the scalar parameter ϕ, while the covariance between space and time is captured by the parameter θ. The space filter matrix $B = (I_N - \rho W)$ is nonsingular, where the scalar spatial dependence parameter is ρ and the $N \times N$ matrix W is assumed to be a known *row stochastic* spatial weight matrix (exogenous with row-sums of unity and with zeros on the diagonal). This matrix defines the dependence between cross-sectional spatial units. We will also assume that W was created by row-normalizing our 10 nearest neighbors matrix, so that all eigenvalues are less than or equal to one. To address time-specific effects, we apply the time mean differencing matrix transformation $J = I_{T-1} \otimes (I_N - (1/N)\iota_N \iota_N'))$ to put each time period in deviations from the time mean form.[16]

The associated data generating process (DGP) is shown in (19).

$$
\begin{aligned}
Y &= Q^{-1}\{\begin{bmatrix} X & (I_{T-1}\otimes W)X \end{bmatrix}\begin{pmatrix} \beta \\ \eta \end{pmatrix} \\
&+ (I_T\otimes \tilde{D})\odot \begin{bmatrix} X & (I_{T-1}\otimes W)X \end{bmatrix}\begin{pmatrix} \tilde{\beta} \\ \tilde{\eta} \end{pmatrix} + \mathcal{F}\gamma + \varepsilon\}
\end{aligned} \tag{19}
$$

Of course, the values taken by the kth explanatory variable change with time periods so we need to further elaborate expression (19). For future reference we note that Debarsy et al. (2012) show that the matrix Q^{-1} takes the form of a lower-triangular block matrix, containing blocks with $N \times N$ matrices. [17]

$$
Q^{-1} = \begin{pmatrix}
B^{-1} & 0_{N\times N} & 0_{N\times N} & 0_{N\times N} & \dots & 0_{N\times N} \\
C_1 & B^{-1} & 0_{N\times N} & 0_{N\times N} & \dots & 0_{N\times N} \\
C_2 & C_1 & B^{-1} & 0_{N\times N} & \dots & 0_{N\times N} \\
C_3 & C_2 & C_1 & B^{-1} & \ddots & \vdots \\
\vdots & \vdots & \ddots & \ddots & \ddots & 0_{N\times N} \\
C_{T-1} & C_{T-2} & \dots & C_2 & C_1 & B^{-1}
\end{pmatrix} \tag{20}
$$

$$
C_s = (-1)^s (B^{-1}A)^s B^{-1}, \quad s = 1, \dots, T-1.
$$

One implication of this is that we need only calculate the two $N \times N$ matrices: A and B^{-1} to analyze the partial derivative impacts for any time horizon T. This means we can use a panel involving say ten years to analyze the cumulative impacts arising from a permanent change in endowments at any time t extending to future horizons $t + T$. Of course, long horizons where T represents 30, 50 or 100 years are of interest for regional growth and convergence issues.

The one-period-ahead impact of a (permanent) change the kth variable at time t for regions in Club 1 are shown in (21) and those for regions in Club 2 are in (22).

$$
\partial Y_{t+1}/\partial X_t^k = C_1[I_N\beta_k + W\eta_k] \tag{21}
$$

$$
\begin{aligned}
&= -B^{-1}(\phi I_N + \theta W)B^{-1}[I_N\beta_k + W\eta_k] \\
\partial Y_{t+1}/\partial X_t^k &= C_1[I_N(\beta_k + \tilde{\beta}_k) + W(\eta_k + \tilde{\eta}_k)] \qquad (22) \\
&= -B^{-1}(\phi I_N + \theta W)B^{-1}[I_N(\beta_k + \tilde{\beta}_k) + W(\eta_k + \tilde{\eta}_k)].
\end{aligned}
$$

More generally, the T-period-ahead (cumulative) impact arising from a permanent change at time t in X_t^k takes the form in (23) for regions in Club 1 and (24) for Club 2 regions. Note that we are cumulating down the columns (or rows) of the matrix in (20). For interpretative purposes we follow LeSage and Pace (2009) who note that the columns represent partial derivative changes arising from a change in a single region, whereas the rows reflect changes in all regions.

$$
\begin{aligned}
\partial Y_{t+T}/\partial X_t^k &= \sum_{s=1}^{T} C_s[I_N\beta_k + W\eta_k] \qquad (23) \\
\partial Y_{t+T}/\partial X_t^k &= \sum_{s=1}^{T} C_s[I_N(\beta_k + \tilde{\beta}_k) + W(\eta_k + \tilde{\eta}_k)] \qquad (24) \\
C_s &= (-1)^s(B^{-1}A)^s B^{-1}.
\end{aligned}
$$

By analogy to LeSage and Pace (2009), the main diagonal elements of the $N \times N$ matrix sums for time horizon T represent (cumulative) own-region impacts that arise from both time and spatial dependence. The off-diagonal elements of these matrix sums reflect diffusion over space and time. We note that it is not possible to separate out the time from space and space-time diffusion effects in this model.[18]

The next section reports parameter estimates for the model along with scalar summary measures of the dynamic elasticity responses of income levels to changes in initial endowments.

Dynamic elasticity responses for the two clubs

We first report parameter estimates for the model, although these are not directly interpretable in terms of the space-time dynamic impacts associated with changes in the explanatory variables on the dependent variable (regional income levels). Posterior means, medians and standard deviations as well as a ratio of the mean/standard deviation are reported for the space-time dependence parameters ϕ, ρ, θ and the noise variance parameter

σ_ε^2 in Table 3.

(Table 3 about here)

From the table we see significant time, space and space-time dependence, with the restriction that $\theta = -\rho\phi$ discussed in Parent and LeSage (2012) being quite consistent with this dataset, since $0.72352 \times -0.78199 = -0.5658$, which is very close to the unrestricted estimate for $\theta = -0.5685$. In fact, the difference of 0.0027 between these two estimates is much smaller than the estimated standard deviation for θ equal to 0.01330. We come to a similar conclusion regarding the restriction that $\theta = -\rho\phi$ using the posterior medians in place of the means.

The table also reports coefficient estimates for the three explanatory variables used: (logged) regional levels of physical capital, knowledge capital and human capital as β coefficients, along with those from an average of the 10 neighboring regions recorded as η coefficients on the WX variables. The coefficients for the Club 2 dummy variables associated with these two sets of explanatory variables are denoted using $\tilde{\beta}, \tilde{\eta}$, and we note that neighboring region η coefficients for knowledge and human capital appear to exert a significant influence, as do the neighboring region dummy variable coefficients $\tilde{\eta}$ for knowledge and human capital.[19] It should be noted that none of these coefficients $(\beta, \eta, \tilde{\beta}, \tilde{\eta})$ are directly interpretable as indicating how the dependent variable responds to changes in the explanatory variables, a point that has frequently been overlooked in the dynamic panel data model literature.

The dynamic elasticity responses are shown in Table 4 for the direct (own-region) and indirect (other-region) responses to changes in the physical capital stock variable for both clubs. The *cumulative* direct effects estimates reported show time horizon zero effects that reflect simultaneous own-region spatial effects, while time horizons one to 30 years include the future period own-region impacts that arise from time dependence as well as some spatiotemporal feedback effects. Note that in this model regional income depends on neighboring regions, implying that future period changes in neighboring regions' income will set in motion a feedback loop that produces second order benefits/costs to the own-region as a result of spatial spillover benefits/costs generated for neighbors in earlier time periods.

(Table 4 about here)

The first column shows the time horizon $(t+T)$. The next three columns show a lower 0.99 credible interval constructed using 10,000 MCMC draws retained from a set of 50,000 draws, the posterior mean estimate, and an upper 0.99 credible interval. A positive mean with positive lower and upper 0.99 intervals should be interpreted as a positive and significant effect. Effects whose credible intervals span zero are not significant. Cumulative effects are interpreted as the percentage changes in regional income levels that would arise over time in response to a permanent percentage increase in the level of capital stock in the representative/typical region. The same format was used to report (cumulative) direct effects for Club 2 alongside those of Club 1 for comparison purposes. The table also includes indirect (or spatial spillover) effects in the same format.

The dynamic cumulative elasticity responses reveal that a 10 percent increase in physical capital stocks in Club 1 (initial period low income) regions would lead to a long-run own-region (direct) increase in income (GVA per capita) of 2.4 percent, and a very similar 2.1 percent increase for Club 2 (initial period high income) regions. The empirically derived credible intervals calculated for the responses show that increases in physical capital have a long-lived impact on regional incomes, since the effects level off at 19- and 20-years, respectively.

These results suggest no difference in how low and high income regions (Clubs 1 and 2 respectively) are able to convert increased physical capital stocks into higher regional income levels. As a test for significant differences between the responses for the two clubs, the top panel of Figure 4 shows a plot of the posterior mean difference between Club 1 and 2 (Club 1 minus Club 2), along with 0.99 lower and upper credible intervals. The credible intervals were empirically determined using the 10,000 retained MCMC draws for the difference between the two clubs effects estimates. The plot makes it clear that the difference between the effects is very close to zero and the credible intervals span zero.

(Figure 4 about here)

Table 4 also shows the cumulative indirect (spatial spillover) effects associated with a change in physical capital stocks. Here we see positive but not significant spillovers for both Club 1 and Club 2 regions. The cumulative spillover magnitudes of 0.74 for Club 1 and 0.12

for Club 2 appear very different. However, the second panel of Figure 4 shows posterior mean differences and empirical credible intervals indicating no significant difference.[20] We conclude that changes in physical capital for the typical region in both Club 1 and Club 2 do not produce significant spatial spillover impacts on income levels of neighboring regions.

One explanation for positive and long-lived direct effects arising from changes in regional physical capital but zero spatial spillover effects might be the type of capital stock changes taking place. For example, increases in physical capital representing resources shared between regions such as public transportation infrastructure would be more likely to produce spatial spillovers than physical capital put in place by private firms.[21]

Table 5 shows the direct cumulative effect responses to changes in knowledge capital stocks for regions in Clubs 1 and 2. Here we see (cumulative long-run) direct responses for the Club 2 regions (0.0925) that are positive and significant, while those for Club 1 regions are not different from zero. This would indicate that high income regions benefit from increased knowledge stocks while low income regions do not.

(Table 5 about here)

The top panel of Figure 5 shows that the difference between the mean cumulative direct effects (e.g., 0.0592 long-run mean response for Club 1 versus 0.0925 long-run mean response for Club 2) for the two clubs is not significant. However, the evidence from Table 5 suggests that we interpret the mean response for Club 1 as truly zero, and the mean response for Club 2 is significantly different from zero using the lower 0.01 and upper 0.99 credible intervals reported in the table.

(Figure 5 about here)

Indirect effects responses in Table 5 indicate positive and significant spatial spillovers from knowledge capital stocks for Club 1, but not for Club 2 regions. The magnitude is such that a 10 percent increase in knowledge stocks of neighboring regions would lead to a four percent (long-run) increase in income levels of the low income Club 1 regions. An implication of these results is that Club 1 (low income) regions that are close neighbors to Club 2 (high income) regions may benefit greatly from spatial spillovers and diffusion effects arising from increases in knowledge stocks in Club 2 regions. In contrast, Club 2 regions would not benefit from spillover and diffusion effects as a result of being neighbors to Club

2 regions where knowledge stocks are increasing. The bottom panel of Figure 5 shows that the difference between the mean indirect effects for Clubs 1 and 2 are significantly positive.

The magnitude of (cumulative) spillovers (e.g., 0.429 long-run response) may seem large, especially when compared to the long-run (cumulative) direct effects of 0.0925. It must be noted that these are cumulative spillovers, where the cumulation takes place over all neighboring regions, neighbors to the neighboring regions and so on. Effects falling on any individual region are smaller than the direct effects, consistent with spillovers being a "second order effect". This can be seen by considering that there are 10 first order neighbors alone, so if we divide the spillover/indirect effects estimates by a factor of 10, the marginal impacts of 0.043 associated with a single region are much smaller than the direct effects. Further note that we should in reality divide by a number much greater than the 10 first order neighbors, since these effects emanate out to more distant neighbors as time passes, a phenomenon representing spatial diffusion impacts. See Parent and LeSage (2010) for a decomposition of the effects into time-specific and space-specific as well as diffusion-specific impacts.

The direct effects from changes in human capital reported in Table 6 are positive and significant for Club 1 low income regions but not significant for Club 2 regions. The top panel of Figure 6 indicates that the difference between direct effects for the two clubs is significant. Club 1 regions benefit from increases in human capital whereas Club 2 regions do not.

(Table 6 about here)

Indirect effects from changes in human capital reported in Table 6 are negative for Club 1 regions and positive for Club 2 regions, but not significantly different from zero for either club. This indicates that changes in human capital do not produce spatial spillovers to neighboring regions. The bottom panel of Figure 6 indicates that the difference between the negative and positive mean indirect effects is significant, but this is irrelevant since both sets of mean effects are not significantly different from zero.

(Figure 6 about here)

Concluding remarks

This paper describes a two-step approach to identifying and interpreting regional convergence clubs in Europe. The first step uses a formal Bayesian methodology to produce posterior probabilities for alternative classifications of European regions into convergence clubs. Each region must be classified into one of two clubs. The classification takes place conditional on a space-time panel data relationship for regional income growth. Since observations are regions in our relationship, the comparison problem is one of comparing classifications based on different assignments of each observation (region) to one of the two club categories based on initial period income levels. A key contribution is that we use a dynamic (panel) growth regression relationship and integrate over the space of all parameters in the relation rather than produce classifications that are conditional on a single set of parameter estimates from a model.

Even for the case of two clubs, the classification problem leads to a high dimensional model space consisting of 2^N possible models where N is the number of regions in the sample that need to be compared for classification purposes. To overcome this, we use a procedure that splits the sample into clubs based on the initial period (per capita) income levels of the regions, and log-marginal likelihood expressions to calculate posterior probabilities for sample splits based on different initial period income levels. Deriving the log-marginal likelihood used for classification comparison purposes here involved a combined strategy that relied on: (i) analytical integration for some parameters of the model, (ii) numerical integration over the space and time dependence parameters, and (iii) fixing the variance ratio for the random effects versus noise vector.

Results from applying the classification procedure to split a sample of 216 European regions according to initial period income levels were reported. They suggest strong evidence of two clubs based on a split of regions at the 16,000 level of 1995 level of per capita GVA. A limitation of our approach is that we assumed only two clubs.

The second step of the approach involved estimating a space-time dynamic panel data model that used (logged) levels of regional income as the dependent variable and (logged) levels of previous period endowments of physical, knowledge and human capital stocks. Analytical expressions from Debarsy et al. (2012) for the partial derivatives showing dynamic

response elasticities were used to examine the response of regional income levels over space and time to changes in initial period endowments. These dynamic responses provide clear evidence of the distinct long-term behavior of the two clubs of regions.

Acknowledgements

The authors would like to acknowledge the Austrian Science Fund (FWF) for financial support (grant no. P19025-G11). We would like to thank Olivier Parent for his helpful comments on this manuscript, especially regarding the log-marginal likelihood derivation.

Notes

[1]See Durlauf et al. (2005), Magrini (2004), as well as LeGallo and Fingleton (2013) for useful growth empirics literature reviews.

[2]Modern growth theory has suggested that the distribution of per capita income of regions may display a tendency for the steady state distribution to cluster around a small number of poles of attraction, and hence lead to convergence clubs (Canova 2004). This tendency may be due to several factors: capital market imperfections, non-convexities, imperfectly competitive market structures, and spillovers due to physical and human capital accumulation (Galor 1996).

[3]Regions that are similar in their structural characteristics, but differ in their initial distribution of income, may cluster around different steady state equilibria (see Durlauf 1996; Quah 1996). It should be noted that if multiple equlibria depend on initial income cut-offs, the relationship between subsequent growth and initial income will not be linear.

[4]We use the term *relationship* to distinguish this from a model. Models have an associated set of parameter estimates and inferences about club assignments would be *conditional* on this set of estimates.

[5]For matters of simplicity, we concentrate on the determination of two clubs here. The existence of two clubs is supported by empirical regularities in the dynamics of cross-region income distributions in Europe as evidenced in Pittau and Zelli (2006), and Fischer and Stumpner (2008).

[6]Most approaches to estimating and testing the club convergence hypothesis have focused on calculating scalar measures of the speed of convergence taken from the cross-sectional models literature.

[7]Of course, there is a relationship between growth rates and level values taken by variables (such as income, physical and human capital) over time which is explored in detail for the case of spatially dependent sample data in LeSage and Fischer (2008).

[8]This type of space-time panel data relationship has been originally proposed by Anselin (2001), and explored by Yu et al. (2008) as well as Parent and LeSage (2012). Examples of empirical studies using this type of specification include Parent and LeSage (2010), and Debarsy et al. (2012).

[9] Parent and LeSage (2012) point to several computational advantages that arise when $\theta = -\phi\rho$, and there are also interpretative advantages regarding the partial derivatives of the model discussed in Debarsy

et al. (2012).

[10]Posterior probabilities of the type reported in Table 1 were constructed for varying weight matrices as a test of robustness to this assumption, and they showed no change regarding the level of income at which we split our sample.

[11]A restriction to regions with 1995 GVA per capita below 50,000 was implemented to improve scaling of the figure.

[12]We exclude the Spanish North African territories of Ceuta and Melilla, the Portuguese non-continental territories Azores and Madeira, the French Départements d'Outre-Mer Guadaloupe, Martinique, French Guayana and Réunion.

[13]The motivation for the use of this model type is that it can provide us with useful information about the clubs of regions not available from cross-section (spatial) regressions.

[14]Parent and LeSage (2010) as well as Debarsy et al. (2012) are exceptions.

[15]Ertur and Koch (2007) derive this type of expression where neighboring region explanatory variables arise in a growth regression framework from neoclassical growth theory.

[16]This transformation is applied to Y and X as well as $\mathcal{F}$ and it obliterates the intercept term from the model. For clarity we do not include this in the notation regarding our discussion of the partial derivative impacts on y_{t+T} arising from changes in X_{it}, since it does not influence these.

[17]See Parent and LeSage (2010) for the special case that arises when the restriction $\theta = -\rho\phi$ is imposed.

[18]See Parent and LeSage (2010) for the special case where space and time are separable.

[19] The role of neighboring region endowments was ignored by Yu and Lee (2012) in their implementation of this model for US data.

[20]There is evidence that the differences are significant using a 0.95 credible interval. This difference is positive indicating that low income (Club 1) regions enjoy larger spatial spillovers from changes in physical capital stock.

[21]The evidence of greater spillovers for Club 1 regions using the 0.95 intervals seems consistent with more public infrastructure investment in low income Club 1 regions over the period we examine.

References

Anselin, L. (2001). Spatial econometrics. In: Baltagi, Badi H. (Ed.), *A companion to theoretical econometrics.* Blackwell Publishers Ltd, Massachusetts, pp. 310-330.

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

Bartkowska, M., and A. Riedl. (2012). Regional convergence clubs in Europe: Identification and conditioning factors. *Economic Modelling* 29(1): 22-31.

Baumol, W. J. (1986) Productivity growth, convergence and welfare: What the long run data show? *American Economic Review* 76(5), 1072-1084.

Brock, W. A., and S. N. Durlauf. (2001). Growth empirics and reality. *The World Bank Economic Review* 15(2), 229-272.

Caballero, J. R., and J. B. Jaffe. (1993). How high are the giants shoulders: An empirical assessment of knowledge spillovers and creative destruction in a model of economic growth. In: U. Blanchard, and S. Fischer (eds.), *NBER Macroeconomics Annual 1993*, Vol. 8. Cambridge [MA]: The MIT Press, pp. 15-74.

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

Corrado, L., R. Martin, and M. Weeks. (2005). Identifying and interpreting regional convergence clusters across Europe. *The Economic Journal* 115(502), C133-C160.

Dall'erba, S., and J. LeGallo. (2008). Regional convergence and the impact of European structural funds over 1989-1999: a spatial econometric analysis. *Papers in Regional Science* 87(2), 219-244.

Debarsy, N., C. Ertur, and J. P. LeSage. (2012). Interpreting dynamic space-time panel data models, *Statistical Methodology* 9(1-2), 158-171.

Desdoigts, A. (1999). Patterns of economic development and the formation of clubs. *Journal of Economic Growth* 4(3), 305-330.

Durlauf, S. (1996). A theory of persistent income inequality. *Journal of Economic Growth* 1(1), 75-94.

Durlauf, S., and P. A. Johnson. (1995). Multiple regimes and cross-country growth behaviour. *Journal of Applied Econometrics* 10(4), 365-384.

Durlauf, S., P. A. Johnson, and J. R. W. Temple. (2005). *Growth economics*. In: P. Aghion and S. N. Durlauf (eds) *Handbook of Economic Growth, Volume IA*. Amsterdam: Elsevier, pp. 555-677.

Elhorst, J. P. (2012). Dynamic spatial panels: Models, methods, and inferences. *Journal of Geographical Systems* 14(1), 5-28.

Elhorst, J. P., G. Piras, and G. Arbia. (2010). Growth and convergence in a multiregional model with space-time dynamics. *Geographical Analysis* 42(3), 338-355.

Ertur, C., and W. Koch. (2007). Growth, technological interdependence and spatial externalities: Theory and evidence. *Journal of Applied Econometrics* 22(6), 1033-1062.

Ertur, C., J. LeGallo, and C. Baumont. (2006). The European regional convergence process, 1980-1995: do spatial dependence and spatial heterogeneity matter? *International Regional Science Review* 29(1), 2-34.

Fagerberg J., and B. Verspagen. (1996). Heading for divergence? Regional growth in Europe reconsidered. *Journal of Common Market Studies* 34(3), 431-448.

Fève P., and Y. Le Pen. (2000). On modelling convergence clubs. *Applied Economics Letters* 7(5), 311-314.

Fischer M. M., and C. Stirböck. (2006). Pan-European regional income growth and club-convergence. Insights from a spatial econometric perspective. *Annals of Regional Science* 40(4), 693-721.

Fischer M. M., and P. Stumpner. (2008). Income distribution dynamics and cross-region convergence in Europe. Spatial filtering and novel stochastic kernel representations. *Journal of Geographical Systems* 10 (2), 109-140.

Fischer, M.M., T. Scherngell, and M. Reismann. (2009). Knowledge spillovers and total factor productivity. *Geographical Analysis* 4(12), 204-220.

Galor, O. (1996). Convergence? Inference from theoretical models. *The Economic Journal* 106(37), 1056-1069.

Hobijn, B., and P. H. Franses. (2000). Asymptotically perfect and relative convergence of productivity. *Journal of Applied Econometrics* 15(1), 59-81.

Islam, N. (1995). Growth empirics: A panel data approach. *The Quarterly Journal of Economics* 110(4), 1127-1170.

LeGallo, J., and S. Dall'erba. (2006). Evaluating the temporal and spatial heterogeneity of the European convergence process, 1980-1999. *Journal of Regional Science* 46(2), 269-288.

LeGallo, J., and B. Fingleton. (2013). Regional growth and convergence empirics. In M. M. Fischer, and P. Nijkamp (eds) Handbook of Regional Science. Springer, Berlin, Heidelberg (in press)

LeSage, J. P., and M. M. Fischer. (2008). Spatial growth regressions: model specification, estimation and interpretation. *Spatial Economic Analysis* 3(3), 275-304.

LeSage, J. P., and R. K. Pace. (2009). *Introduction to Spatial Econometrics* London: Taylor & Francis/CRC Press.

LeSage, J. P., and O. Parent. (2007). Bayesian model averaging for spatial econometric models. *Geographical Analysis* 39(3), 241-267.

Magrini, S. (2004). Regional (di)convergence. In V. Henderson and J. F. Thisse (eds.), *Handbook of Regional and Urban Economics*. Amsterdam: Elsevier, pp. 2741-2796.

Mankiw, N. G., D. Romer, and D. Weil. (1992). A contribution to the empirics of economic growth. *The Quarterly Journal of Economics* 107(2), 407-437.

Mur, J., F. López, and A. Angulo. (2010). Instability in spatial error models: an application to the hypothesis of convergence in the European case. *Journal of Geographical Systems* 12(3), 259-280.

Parent, O., and J. P. LeSage. (2010). A spatial dynamic panel model with random effects applied to commuting times. *Transportation Research Part B: Methodological* 44B(5), 633-645.

Parent, O., and J. P. LeSage. (2011). A space-time filter for panel data models containing random effects. *Computational Statistics & Data Analysis* 55(1), 475-490.

Parent, O., and J. P. LeSage. (2012). Spatial dynamic panel data models with random effects, *Regional Science & Urban Economics* 42(4), 727-738.

Phillips, P. C. B., and D. Sul. (2007). Transition modelling and econometric convergence tests. Econometrica 75(6), 1771-1855.

Pittau, M.G., and R. Zelli. (2006). Empirical evidence of income dynamics across EU regions. *Journal of Applied Econometrics* 21(5), 605-628.

Postiglione, P., S. Andreano, and R. Benedetti. (2012). Using constrained optimization for the identification of convergence clubs, *Computational Economics*, (in press).

Quah, D. (1996). Convergence empirics across countries with (some) capital mobility. *Journal of Economic Growth* 1(1), 95-124.

Solow, R. M. (1956). A contribution to the theory of economic growth. *Quarterly Journal of Economics* 70(1), 6594.

Wansbeek, T.J., and A. Kapteyn. (1982). A simple way to obtain the spectral decomposition of variance components models for balanced data. *Communications in Statistics Part A - Theory Methods* 11(18), 2105-2112.

Yu, J., and L. F. Lee. (2012). Convergence: A spatial dynamic panel data approach. *Global Journal of Economics* 1(1), 1250006-12500042.

Yu, J., R. de Jong, and L. F. Lee. (2008). Quasi-maximum likelihood estimators for spatial dynamic panel data with fixed effects when both n and T are large. *Journal of Econometrics* 146(1), 118-134.

Zellner, A. (1971). *An Introduction to Bayesian Inference in Econometrics*. New York: John Wiley & Sons.

Zellner, A. (1986). On assessing prior distributions and Bayesian regression analysis with g-prior distributions. In P. Goel, and A. Zellner (eds.), *Bayesian Inference and Decision Techniques: Essays in Honor of Bruno de Finetti*. North-Holland/Elsevier: Amsterdam, pp. 233-243.

Appendix

Deriving the log-marginal likelihood used for model comparison purposes in our study involves a combined strategy that relies on analytical integration for some parameters of the model, numerical integration over the space and time dependence parameters, and fixing the variance ratio for the random effects versus noise vector. We will develop the log-marginal likelihood expressions to calculate posterior probabilities for models involving splits based on different initial period income levels of the sample of regions. Let us start with the task of analytically integrating out the parameters $\delta = (\ \psi \quad \tilde{\psi} \quad \alpha \quad \tilde{\alpha} \quad \beta \quad \tilde{\beta}\ )'$.

Proceeding to the task of analytically integrating out the parameters δ, we can concentrate out the parameters δ using:

$$\hat{\delta} \;=\; (Z'Z)^{-1}Z'Pg$$

which can be strategically written using the following expressions:

$$\begin{aligned}
\hat{\delta} &= (\delta_0 - \phi\delta_\phi - \rho\delta_\rho - \theta\delta_\theta) \\
\delta_0 &= (Z'Z)^{-1}Z'(F \otimes I_N)g \\
\delta_\phi &= (Z'Z)^{-1}Z'(L \otimes I_N)g \\
\delta_\rho &= (Z'Z)^{-1}Z'(F \otimes W)g \\
\delta_\theta &= (Z'Z)^{-1}Z'(L \otimes W)g
\end{aligned}$$

where

$$L \;=\; \begin{pmatrix} -1 & 0 & 0 & \dots & 0 \\ 0 & -1 & 0 & \dots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & -1 & 0 \end{pmatrix}$$

$$F = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & & 0 & 1 \end{pmatrix}$$

with L and F being $(T-1) \times T$ matrices.

Now consider the errors: $e = Pg - Z\delta$, which can be written using:

$$\begin{aligned}
e &= \begin{pmatrix} 1 & -\phi & -\rho & -\theta \end{pmatrix} \begin{pmatrix} E^{(1)} \\ E^{(2)} \\ E^{(3)} \\ E^{(4)} \end{pmatrix} \\
E^{(1)} &= (F \otimes I_N)g - Z(Z'Z)^{-1}Z'(F \otimes I_N)g \\
E^{(2)} &= (L \otimes I_N)g - Z(Z'Z)^{-1}Z'(L \otimes I_N)g \\
E^{(3)} &= (F \otimes W)g - Z(Z'Z)^{-1}Z'(F \otimes W)g \\
E^{(4)} &= (L \otimes W)g - Z(Z'Z)^{-1}Z'(L \otimes W)g \\
e'\Omega^{-1}e &= \tau' Q \tau \\
\tau &= \begin{pmatrix} 1 & -\phi & -\rho & -\theta \end{pmatrix} \\
Q_{ij} &= \mathrm{tr}(E^{(i)\prime}\Omega^{-1}(\lambda)E^{(j)}), \quad i = 1, \dots, 4 \quad j = 1, \dots, 4.
\end{aligned}$$

The advantage of this specification is that the likelihood can be written expressing the sum of squared residuals $e'\Omega^{-1}e$ as a function of only the parameters ϕ, ρ, θ in the vector τ and the parameter λ, plus sample data information g, Z, W.

We assign an inverse gamma prior $IG(a, b)$ for σ_μ^2/λ:

$$\pi_s(\sigma_\mu^2/\lambda) \sim \frac{(ab/2)^{a/2}}{\Gamma(a/2)}(\sigma_\mu^2/\lambda)^{-(\frac{a+2}{2})}\exp\left(-\frac{ab}{2\sigma_\mu^2/\lambda}\right),$$

where a, b are parameters of the inverse gamma prior. We follow LeSage and Parent (2007) and assign Zellner's g-prior (Zellner 1986) to the parameters δ:

$$
\begin{aligned}
\pi_d(\delta|\sigma_\mu^2/\lambda) &\sim \mathcal{N}(0,(\sigma_\mu^2/\lambda)V^{-1}) \\
V &= GZ'Z.
\end{aligned}
$$

Using Bayes theorem the marginal likelihood for the model can be written as the integral below, where we use $\mathcal{D}$ to denote the data g, Z, W.

$$
\begin{aligned}
& \int \pi_d(\delta|\sigma_\mu^2/\lambda)\pi_s(\sigma_\mu^2/\lambda)p(\mathcal{D}|\alpha,\delta,\rho,\phi,\theta,\sigma_\mu^2/\lambda)\ d\delta\ d(\sigma_\mu^2/\lambda)\ d\rho\ d\phi\ d\theta \\
= \ & \kappa(2\pi)^{-(N(T-1)+K)/2}((T-1)\lambda+1)^{-N}|V|^{1/2} \\
\times \ & \int |I_N-\rho W|^T \frac{\lambda^{[N(T-1)]+a+2K+1}}{\sigma_\mu^2} \\
\times \ & \exp\left(-\frac{1}{2\sigma_\mu^2/\lambda}[ab+e'\Omega^{-1}e+\delta'V\delta+(\delta-\hat{\delta}(\phi,\rho,\theta))'(Z'Z)(\delta-\hat{\delta}(\phi,\rho,\theta))]\right) \\
\times \ & \pi_\delta\pi_\phi\pi_\rho\pi_\theta d\delta\ d\phi\ d\rho\ d\theta \\
\kappa = \ & \Gamma\left(\frac{a}{2}\right)^{-1}\left(\frac{ab}{2}\right)^{a/2}.
\end{aligned}
$$

We can use the properties of the multivariate normal pdf and the inverted gamma pdf to analytically integrate out the parameters δ and σ_μ^2/λ which produces an expression for the marginal likelihood as a function of the three parameters $\zeta = (\phi, \rho, \theta)$ only.

An expression that is analogous to that from LeSage and Parent (2007) arises:

$$
\begin{aligned}
p(\zeta|\mathcal{D}) &= \tilde{\kappa}(\frac{G}{1+G})^{K/2}(T\lambda+1)^{-N} \\
&\times \int |I_N-\rho W|^T[ab+\mathcal{R}(\zeta)+\mathcal{S}(\zeta)]^{-[N(T-1)+a-1]/2}\pi_\phi\pi_\rho\pi_\theta\ d\phi\ d\rho\ d\theta
\end{aligned}
$$

where

$$
\begin{aligned}
\tilde{\kappa} &= \frac{\Gamma[(N(T-1)+a-1)/2]}{\Gamma(a/2)}(ab)^{a/2}\pi^{-[N(T-1)-1]/2} \\
\mathcal{R}(\zeta)+\mathcal{S}(\zeta) &= \frac{1}{G+1}\tau'Q\tau \\
&+ \frac{G}{G+1}\left(Pg-\hat{\alpha}\iota_{NT}\right)'\left(Pg-\hat{\alpha}\iota_{NT}\right) \\
\hat{\alpha} &= U^{(1)}-\phi U^{(2)}-\rho U^{(3)}-\theta U^{(4)}
\end{aligned}
$$

$$
\begin{aligned}
U^{(1)} &= (F \otimes I_N)g \\
U^{(2)} &= (L \otimes I_N)g \\
U^{(3)} &= (F \otimes W)g \\
U^{(4)} &= (L \otimes W)g
\end{aligned}
$$

with Γ denoting the gamma function. Recall that $e'\Omega^{-1}e = \tau' Q \tau$ and Ω is a function of λ which we are treating as a fixed scalar, so Ω is presumed known. Without loss of generality we can view λ as equal to any fixed value here, but in practice we should test for robustness across various values of this parameter reflecting the variance ratio of the random effects to noise.

While we developed these expressions for the case of unrestricted θ, we can reduce the trivariate numerical integration to a bivariate problem by imposing the restriction $\theta = -\rho\phi$, which is the approach we take in our application.

Tables

Table 1: Posterior probabilities for splits of the generated data example

Sample split y_0 levels	Prob(split q) $\lambda = (1/2)\hat{\lambda}$	Prob(split q) $\lambda = \hat{\lambda}$	Prob(split q) $\lambda = 2\hat{\lambda}$
10,000	0.0000	0.0000	0.0000
12,000	0.0000	0.0000	0.0000
14,000	0.0001	0.0000	0.0000
16,000	0.0000	0.0000	0.0000
18,000	0.2744	0.1157	0.2495
20,000*	0.6723	0.8658	0.6689
22,000	0.0009	0.0010	0.0005
24,000	0.0523	0.0000	0.0000
26,000	0.0000	0.0000	0.0000
28,000	0.0000	0.0000	0.0811
30,000	0.0000	0.0175	0.0000
32,000	0.0000	0.0000	0.0000
* indicates split that generated the growth rates			

Table 2: Posterior probabilities for various splits of the sample based on 1995 levels of income

Sample split y_0 levels	Relation (1) Probs ($\lambda = 0.3$)	Relation (1) Probs ($\lambda = 0.4$)	Relation (1) Probs ($\lambda = 0.5$)	Relation (1) Probs ($\lambda = 1$)
8,000	0.0206	0.0102	0.0739	0.0064
10,000	0.0251	0.0118	0.0214	0.0534
12,000	0.0063	0.0012	0.0011	0.0039
14,000	0.0260	0.0099	0.0089	0.0162
16,000	0.7165	0.5206	0.7062	0.4670
18,000	0.0013	0.0058	0.0109	0.0007
20,000	0.0070	0.0143	0.0039	0.0063
22,000	0.1777	0.4147	0.1634	0.1809
24,000	0.0150	0.0022	0.0065	0.0080
26,000	0.0009	0.0059	0.0004	0.0003
28,000	0.0031	0.0004	0.0008	0.2532
30,000	0.0006	0.0030	0.0026	0.0037
Sample split y_0 levels	Relation (4) Probs ($\lambda = 0.3$)	Relation (4) Probs ($\lambda = 0.4$)	Relation (4) Probs ($\lambda = 0.5$)	Relation (4) Probs ($\lambda = 1$)
8,000	0.0591	0.4197	0.1664	0.0442
10,000	0.0451	0.0535	0.0667	0.2515
12,000	0.0064	0.0089	0.0084	0.0046
14,000	0.4476	0.3962	0.3797	0.4338
16,000	0.3201	0.1038	0.1020	0.0719
18,000	0.0135	0.0036	0.0074	0.0030
20,000	0.0089	0.0022	0.0040	0.0063
22,000	0.0377	0.0071	0.0890	0.1828
24,000	0.0125	0.0017	0.0056	0.0002
26,000	0.0014	0.0010	0.1694	0.0008
28,000	0.0468	0.0004	0.0005	0.0003
30,000	0.0009	0.0021	0.0008	0.0006

Table 3: Dynamic space-time panel data model estimates

Posterior statistics	ϕ	ρ	θ	σ_ε^2
Mean	0.72352	0.78199	-0.56850	0.000660
Median	0.72382	0.78130	-0.57011	0.000659
Std	0.00812	0.01301	0.01330	0.000021
Mean/Std	89.08836	60.07254	-42.72632	30.746086
Variables	Mean	Std	Mean/Std	$t-$probability
β physical capital	0.05151	0.0137	3.73	0.00019
β knowledge capital	0.01048	0.0064	1.63	0.10247
β human capital	0.03411	0.0091	3.71	0.00021
η W (physical capital)	0.00822	0.0248	0.33	0.74036
η W (knowledge capital)	0.01919	0.0096	1.98	0.04780
η W (human capital)	-0.07255	0.0192	-3.77	0.00016
$\tilde{\beta}$ physical capital	0.01290	0.0206	0.62	0.53225
$\tilde{\beta}$ knowledge capital	0.01564	0.0089	1.74	0.08136
$\tilde{\beta}$ human capital	-0.04841	0.0147	-3.28	0.00105
$\tilde{\eta}$ W (physical capital)	-0.04952	0.0293	-1.68	0.09138
$\tilde{\eta}$ W (knowledge capital)	-0.04293	0.0109	-3.92	0.00009
$\tilde{\eta}$ W (human capital)	0.10238	0.0241	4.23	0.00002

Table 4: Dynamic elasticity (cumulative) direct and indirect responses for changes in physical capital

	Direct effects						Indirect effects					
	Club 1			Club 2			Club 1			Club 2		
Years	Lower 0.99	Mean	Upper 0.99	Lower 0.99	Mean	Upper 0.99	Lower 0.99	Mean	Upper 0.99	Lower 0.99	Mean	Upper 0.99
0	0.0296	0.0658	0.1039	0.0241	0.0590	0.0955	-0.0173	0.2137	0.4412	-0.1758	0.0398	0.2325
1	0.0509	0.1133	0.1790	0.0416	0.1017	0.1641	-0.0297	0.3643	0.7466	-0.3051	0.0664	0.3870
2	0.0663	0.1477	0.2330	0.0542	0.1324	0.2132	-0.0393	0.4710	0.9700	-0.4014	0.0842	0.4912
3	0.0774	0.1725	0.2723	0.0634	0.1546	0.2488	-0.0469	0.5467	1.1366	-0.4725	0.0961	0.5654
4	0.0854	0.1905	0.3010	0.0700	0.1707	0.2746	-0.0524	0.6008	1.2665	-0.5257	0.1042	0.6215
5	0.0912	0.2034	0.3219	0.0748	0.1823	0.2933	-0.0568	0.6395	1.3694	-0.5653	0.1097	0.6651
6	0.0953	0.2128	0.3371	0.0783	0.1907	0.3070	-0.0602	0.6674	1.4513	-0.5950	0.1135	0.6990
7	0.0983	0.2196	0.3482	0.0808	0.1967	0.3171	-0.0630	0.6875	1.5171	-0.6172	0.1161	0.7258
8	0.1004	0.2245	0.3564	0.0825	0.2011	0.3245	-0.0651	0.7022	1.5697	-0.6339	0.1180	0.7475
9	0.1019	0.2281	0.3623	0.0838	0.2043	0.3299	-0.0668	0.7128	1.6119	-0.6464	0.1193	0.7650
10	0.1030	0.2307	0.3666	0.0847	0.2066	0.3339	-0.0680	0.7206	1.6462	-0.6559	0.1202	0.7793
15	0.1053	0.2361	0.3760	0.0866	0.2115	0.3428	-0.0710	0.7377	1.7418	-0.6782	0.1223	0.8187
20	0.1057	0.2372	0.3780	0.0870	0.2124	0.3449	-0.0719	0.7417	1.7749	-0.6838	0.1228	0.8327
25	0.1057	0.2374	0.3785	0.0871	0.2126	0.3454	-0.0721	0.7427	1.7862	-0.6853	0.1230	0.8374
30	0.1057	0.2374	0.3786	0.0871	0.2127	0.3455	-0.0721	0.7430	1.7902	-0.6856	0.1230	0.8390

Table 5: Dynamic elasticity (cumulative) direct and indirect responses for changes in knowledge stocks

	Direct effects						Indirect effects					
	Club 1			Club 2			Club 1			Club 2		
Years	Lower 0.99	Mean	Upper 0.99	Lower 0.99	Mean	Upper 0.99	Lower 0.99	Mean	Upper 0.99	Lower 0.99	Mean	Upper 0.99
0	-0.0005	0.0147	0.0313	0.0095	0.0256	0.0426	0.0413	0.1212	0.2343	-0.0790	-0.0149	0.0695
1	-0.0011	0.0254	0.0538	0.0164	0.0441	0.0730	0.0699	0.2073	0.4048	-0.1374	-0.0257	0.1199
2	-0.0014	0.0330	0.0700	0.0214	0.0575	0.0947	0.0890	0.2686	0.5396	-0.1802	-0.0336	0.1563
3	-0.0018	0.0385	0.0818	0.0251	0.0671	0.1101	0.1016	0.3126	0.6475	-0.2118	-0.0393	0.1833
4	-0.0020	0.0425	0.0904	0.0277	0.0741	0.1212	0.1097	0.3441	0.7351	-0.2351	-0.0434	0.2033
5	-0.0023	0.0454	0.0968	0.0297	0.0792	0.1291	0.1149	0.3669	0.8063	-0.2527	-0.0464	0.2195
6	-0.0025	0.0475	0.1014	0.0311	0.0829	0.1349	0.1182	0.3834	0.8638	-0.2659	-0.0485	0.2326
7	-0.0026	0.0490	0.1049	0.0321	0.0855	0.1391	0.1203	0.3953	0.9103	-0.2758	-0.0500	0.2432
8	-0.0028	0.0500	0.1074	0.0329	0.0874	0.1421	0.1216	0.4041	0.9482	-0.2835	-0.0511	0.2518
9	-0.0029	0.0508	0.1093	0.0334	0.0888	0.1442	0.1224	0.4104	0.9788	-0.2893	-0.0519	0.2590
10	-0.0029	0.0514	0.1106	0.0338	0.0898	0.1458	0.1228	0.4151	1.0038	-0.2937	-0.0525	0.2648
15	-0.0031	0.0526	0.1138	0.0346	0.0920	0.1491	0.1233	0.4255	1.0731	-0.3046	-0.0536	0.2813
20	-0.0032	0.0528	0.1147	0.0348	0.0924	0.1498	0.1233	0.4280	1.0968	-0.3075	-0.0538	0.2873
25	-0.0032	0.0529	0.1149	0.0348	0.0925	0.1499	0.1233	0.4286	1.1049	-0.3084	-0.0538	0.2893
30	-0.0032	0.0529	0.1150	0.0348	0.0925	0.1500	0.1232	0.4288	1.1076	-0.3087	-0.0538	0.2901

Table 6: Dynamic elasticity (cumulative) direct and indirect responses for changes in human capital

	Direct effects						Indirect effects					
	Club 1			Club 2			Club 1			Club 2		
Years	Lower 0.99	Mean	Upper 0.99	Lower 0.99	Mean	Upper 0.99	Lower 0.99	Mean	Upper 0.99	Lower 0.99	Mean	Upper 0.99
0	0.0054	0.0269	0.0498	-0.0385	-0.0113	0.0153	-0.3671	-0.2021	0.0068	-0.0501	0.0830	0.2365
1	0.0094	0.0465	0.0860	-0.0663	-0.0195	0.0265	-0.6318	-0.3465	0.0099	-0.0848	0.1424	0.4040
2	0.0123	0.0607	0.1125	-0.0864	-0.0255	0.0347	-0.8317	-0.4499	0.0107	-0.1100	0.1849	0.5281
3	0.0144	0.0710	0.1319	-0.1009	-0.0299	0.0407	-0.9870	-0.5241	0.0105	-0.1288	0.2155	0.6206
4	0.0158	0.0784	0.1461	-0.1115	-0.0330	0.0452	-1.1076	-0.5775	0.0099	-0.1428	0.2376	0.6902
5	0.0169	0.0839	0.1566	-0.1191	-0.0353	0.0485	-1.2040	-0.6161	0.0093	-0.1531	0.2536	0.7441
6	0.0175	0.0878	0.1644	-0.1247	-0.0369	0.0509	-1.2800	-0.6441	0.0086	-0.1612	0.2652	0.7860
7	0.0179	0.0906	0.1701	-0.1288	-0.0381	0.0527	-1.3404	-0.6644	0.0080	-0.1673	0.2737	0.8181
8	0.0181	0.0927	0.1743	-0.1318	-0.0390	0.0541	-1.3892	-0.6791	0.0074	-0.1719	0.2798	0.8429
9	0.0182	0.0942	0.1774	-0.1340	-0.0396	0.0551	-1.4285	-0.6899	0.0070	-0.1753	0.2843	0.8625
10	0.0181	0.0953	0.1797	-0.1356	-0.0401	0.0558	-1.4604	-0.6978	0.0067	-0.1778	0.2876	0.8777
15	0.0177	0.0976	0.1848	-0.1391	-0.0410	0.0575	-1.5475	-0.7151	0.0059	-0.1839	0.2949	0.9170
20	0.0174	0.0980	0.1859	-0.1399	-0.0412	0.0579	-1.5760	-0.7190	0.0057	-0.1855	0.2966	0.9294
25	0.0173	0.0981	0.1862	-0.1401	-0.0413	0.0580	-1.5859	-0.7199	0.0057	-0.1859	0.2970	0.9335
30	0.0173	0.0981	0.1862	-0.1401	-0.0413	0.0580	-1.5894	-0.7201	0.0057	-0.1861	0.2971	0.9350

Figures

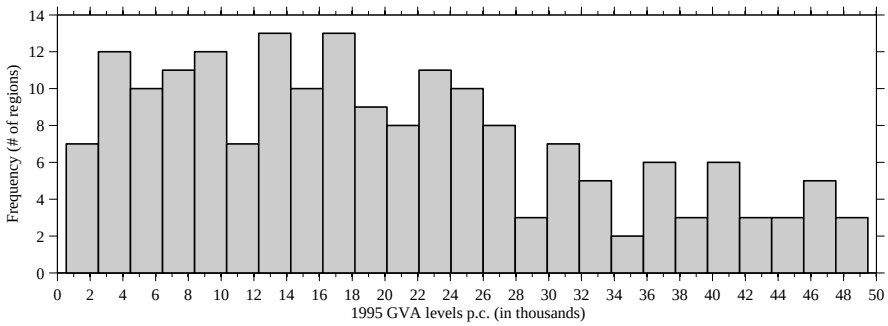

Figure 1: Frequency distribution of 1996 GVA per capita levels

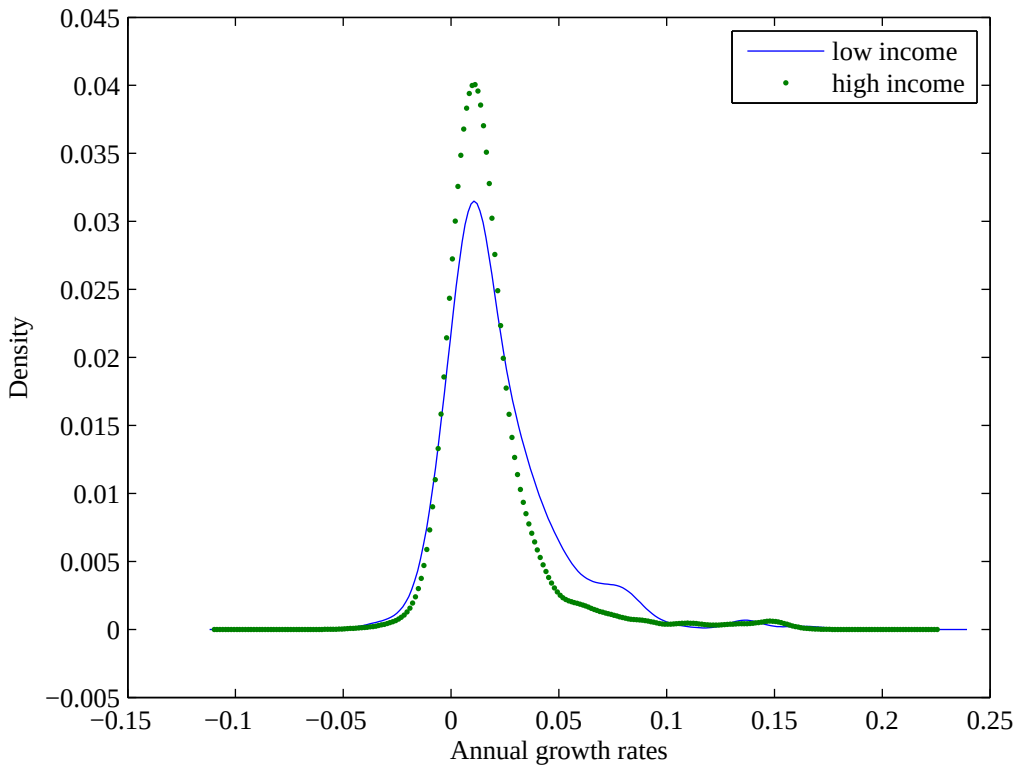

Figure 2: Distributions of low- and high-income regional growth rates generated using $m = 20,000$

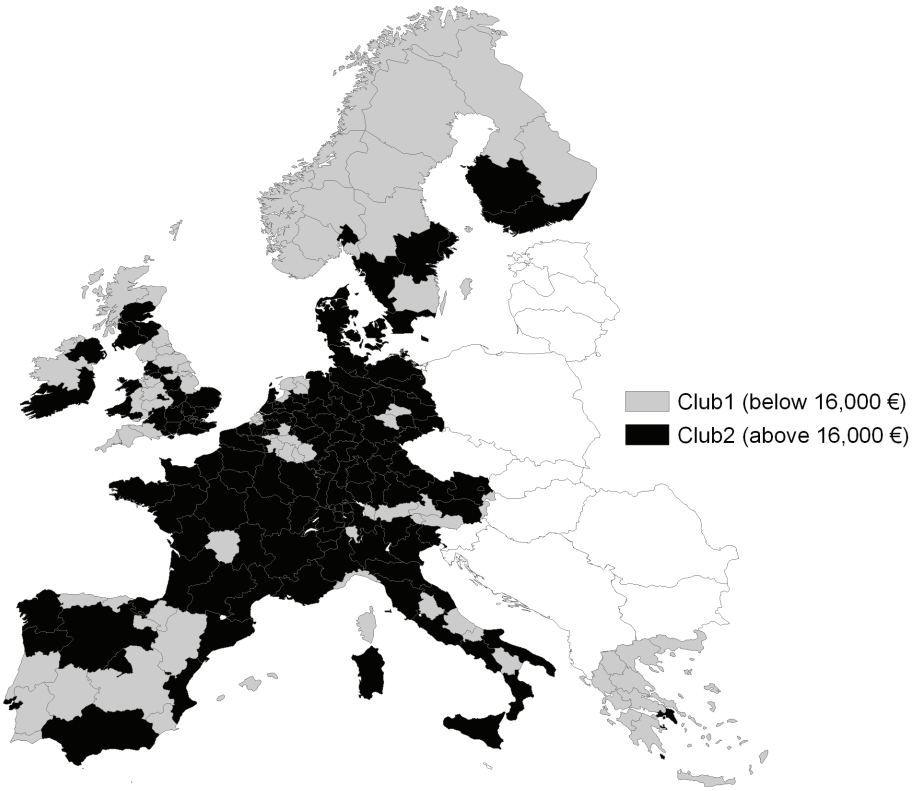

Figure 3: A map of regions classified into two clubs

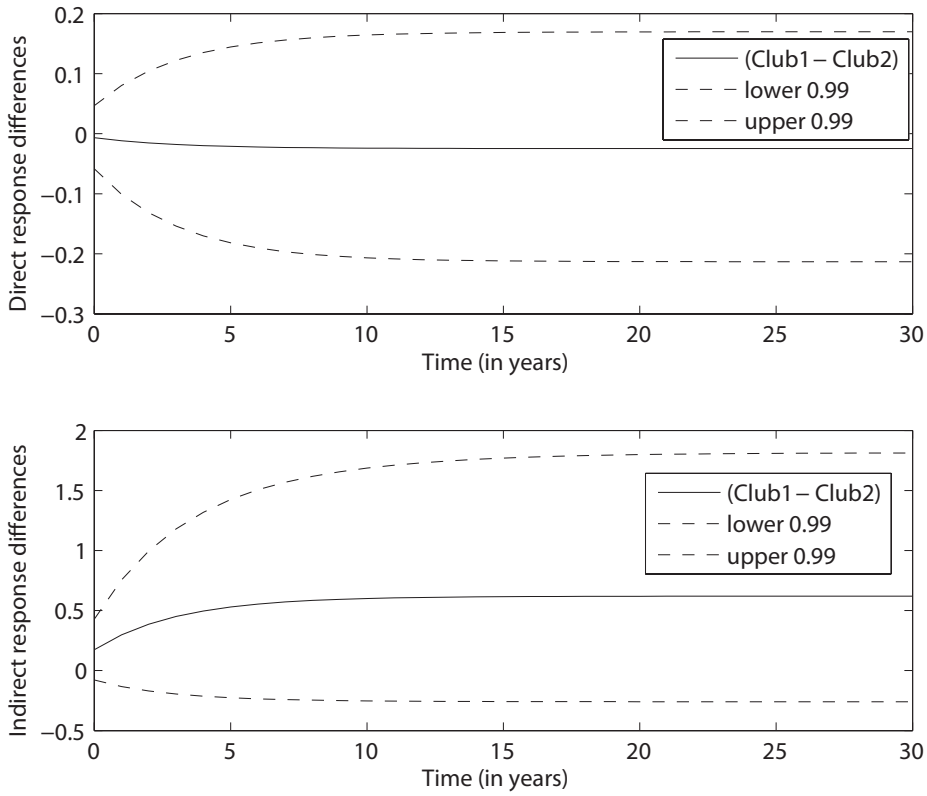

Figure 4: (Club 1 - Club 2) physical capital direct and indirect effect differences

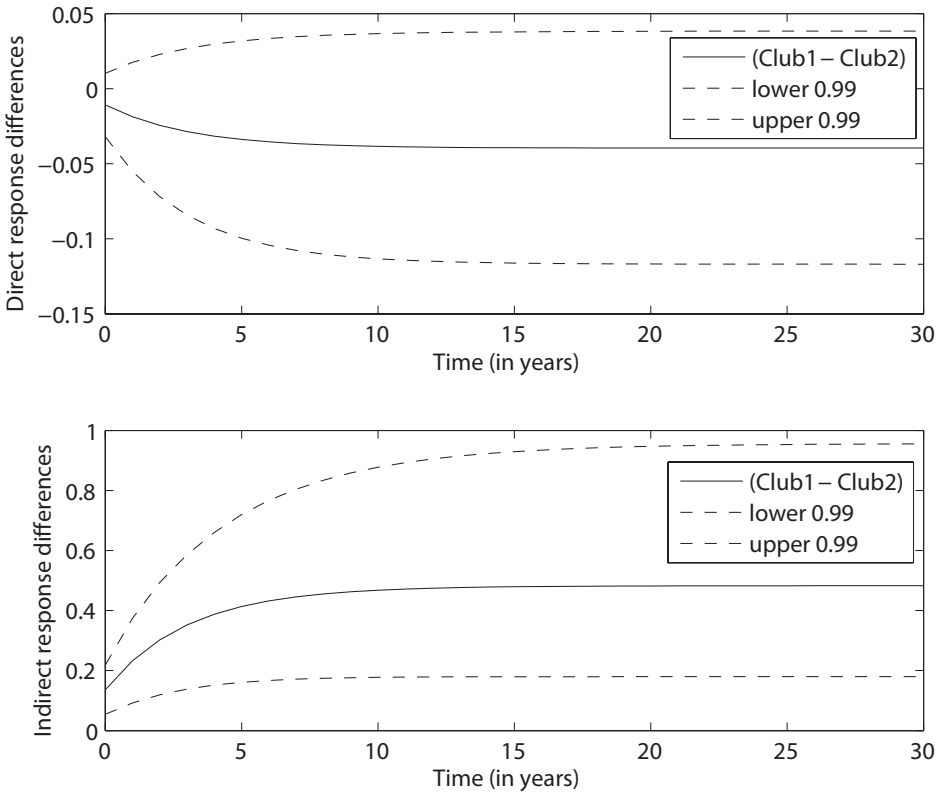

Figure 5: (Club 1 - Club 2) knowledge capital direct and indirect effect differences

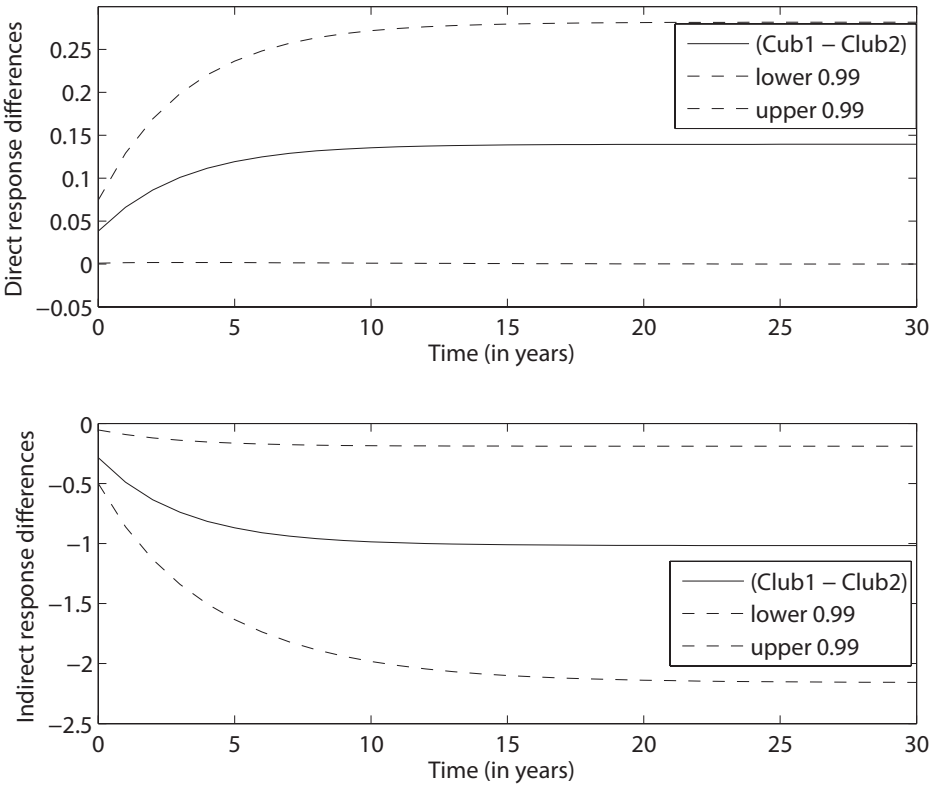

Figure 6: (Club 1 - Club 2) human capital direct and indirect effect differences