JOURNAL OF ECONOMIC SURVEYS

doi: 10.1111/j.1467-6419.2010.00624.x

ECONOMETRICS FOR GRUMBLERS: A NEW LOOK AT THE LITERATURE ON CROSS-COUNTRY GROWTH EMPIRICS

Markus Eberhardt

Centre for the Study of African Economies, University of Oxford

Francis Teal

Centre for the Study of African Economies, University of Oxford, and IZA

Abstract. Since the seminal contribution of N. Gregory Mankiw, David Romer and David N. Weil in 1992 the growth empirics literature has used increasingly sophisticated methods to select relevant growth determinants in estimating cross-section growth regressions. The vast majority of empirical approaches, however, limit cross-country heterogeneity in production technology to the specification of total factor productivity, the 'measure of our ignorance'. In this survey, we present two general empirical frameworks for cross-country growth and productivity analysis and demonstrate that they encompass the various approaches in the growth empirics literature of the past two decades. We then develop our central argument, that cross-country heterogeneity in the impact of observables and unobservables on output as well as the time-series properties of the data are important for reliable empirical analysis.

Keywords. Common Factor Model; Cross-country Empirical Analysis; Cross-section Dependence; Non-stationary Panel Econometrics; Parameter Heterogeneity

One model is supposed to apply everywhere and always. Each country is just a point on a cross-section regression, or one among several essentially identical regressions, leaving only grumblers to worry about what is exogenous and what is endogenous, and whether simple parameterizations do justice to real differences in the way the economic mechanism functions in one place or another. (Solow, 1986, p. S23)

As a careful reading of Solow (1956, 1970) makes clear, the stylized facts for which this model was developed were not interpreted as universal properties for every country in the world. In contrast, the current literature imposes very strong homogeneity assumptions on the cross-country growth process as each country is assumed to have an identical . . . aggregate production function. (Durlauf *et al.*, 2001, p. 929)

Journal of Economic Surveys (2011) Vol. 25, No. 1, pp. 109–155

Commenting on the state of empirical macroeconomics more generally, Robert Solow gave us the image of an 'econometrics for grumblers' which concerns itself with matters of empirical misspecification and variable endogeneity. Interpreted as a criticism of empirical cross-country analysis we believe this polemic assessment rings as true today as it did in 1986, despite the explosion of empirical papers on cross-country growth and development in the 1990s following the seminal contributions by Barro (1991) and Mankiw *et al.* (1992). In essence, as Durlauf *et al.* (2001) lament, one model of *common* technology parameters and (in the original representation) *common* total factor productivity (TFP) is supposed to apply everywhere and always; one aggregate model is assumed to correctly summarize a sample of economies at very different stages of industrial development and made up of diverse industrial sectors. Later variations on the theme have introduced limited forms of heterogeneity in TFP, using intercepts and trends for heterogeneous TFP levels and constant TFP growth rates, but at the same time cross-country regression models have *for the most part* held on to the common technology specification.[1]

We will argue that there are a number of important reasons why the standard cross-country growth regression framework and its panel cousins need to be reconsidered. Rather than to merely 'grumble' about parameter heterogeneity in the observables and unobservables, we show that this property has important implications for estimation and inference once the pertinent time-series properties of the data (non-stationarity, cointegration, cross-country correlation) are recognized. Intuitively, the heterogeneity in production technology could be taken to mean that countries can choose an 'appropriate' production technology from a menu of feasible options. Further, the cross-country heterogeneity in unobservables (TFP) relates to differences both in the underlying processes that make up TFP and in the impact of those processes on output.

The majority of regression approaches following the standard empirical models tend to yield implied technology coefficients considerably out of line with macro evidence on factor income shares. Following Mankiw *et al.* (1992) most empirical studies put this down to the failure to account for forms of intangible capital (human capital, 'social capital') in the regression model. This belief has led to a growth empirics literature that for the most part neglects technology-parameter heterogeneity across countries and simplifies dynamics. The *mainstream* literature favours ever more sophisticated statistical devices – most recently Bayesian model averaging (Sala-i-Martin *et al.*, 2004; Moral-Benito, 2009) and 'general-to-specific' automatic model selection algorithms (Hendry and Krolzig, 2004; Ciccone and Jarocinski, 2008) – to pick out the 'relevant' variables in an augmented Solow regression model with time-averaged variables, so-called 'Barro regressions'. At the last count no fewer than 145 variables have been investigated in their impact on growth (Durlauf *et al.*, 2005) *and most were found to matter* in at least some studies.

A number of papers, however, question this paradigm and have integrated considerations of parameter heterogeneity into their cross-country empirics, also considering the time-series properties of the data, an issue largely ignored in the standard cross-country growth regression framework. Their regression results and diagnostic tests for variable non-stationarity and parameter heterogeneity confirm their importance in the empirical analysis (Pedroni, 2007; Canning and Pedroni, 2008). Our approach in this paper further highlights the importance of cross-section dependence in macro productivity analysis (Costantini and Destefanis, 2009; Eberhardt and Teal, 2009a, b). We will make use of recent developments in the panel time-series literature which have relaxed the standard assumption of cross-section independence. This has led to the development of analytical methods robust to the impact of correlation across panel units (Bai and Ng, 2004; Pesaran, 2006, 2007; Bai, 2009; Bai *et al.*, 2009; Kapetanios *et al.*, 2009). In the context of cross-country growth and development analysis, the potential for this type of data dependency is particularly salient, given the interconnectedness of countries through history, geography and trade relations.

Time-series and cross-section correlation properties of macro panel data have not been considered in great detail in the *empirical* growth literature (Durlauf and Quah, 1999; Temple, 1999; Durlauf *et al.*, 2005), but will be shown to have solid foundations in the *theoretical* literatures on growth and econometrics. Existing empirical work instead has primarily concerned itself with the potential endogeneity of regressors in the empirical framework, an issue that is given considerably more attention in the literature than the data properties or the potential misspecification of the empirical regression model.

We believe that our review of the cross-country growth empirics literature is particularly timely, as over recent years empirical development economics has witnessed the emergence of two powerful analytical tools that explicitly or implicitly question the validity of cross-country regressions and signal nothing less than a paradigm shift in the field. The first of these new tools is the increased use of ‘randomized experiments’ in economics (see Banerjee and Duflo, 2008), which have been brought to great prominence by Abhijit Banerjee, Esther Duflo and collaborators at the Abdul Latif Jameel Poverty Action Lab (J-PAL), MIT. The work by these authors frequently refers to ‘hard evidence’ of causal links between economic and social processes and implies that cross-country growth regression results certainly do not attain this status. The second approach is that of ‘growth diagnostics’ (see Hausmann *et al.*, 2006), developed by a group of economists at the Kennedy School of Government, Harvard University, including Dani Rodrik, Ricardo Hausmann and Lant Pritchett. Their method calls for country-specific analysis by development experts to identify the most important binding constraints to growth and development. Both of these approaches are in agreement that cross-country growth regressions are uninformative as to the causes of growth and this empirical approach is now deeply unfashionable. The fact that we have not learnt the causes of growth from cross-country regressions does not mean that we have learnt nothing and – as this survey will seek to demonstrate – it does not mean that we cannot learn more by using appropriate methods.

The remainder of this paper is organized as follows. The next section presents and motivates two general empirical frameworks for cross-country production analysis, one for the production function approach, a second for the convergence regression equation. In Section 2, these frameworks are shown to encompass a wide variety of modelling approaches representing the full evolution of growth empirics over the past two decades. Our aim here is not to provide an exhaustive review of the empirical growth literature, but to highlight the gradual relaxation of assumptions over the course of this period. In Sections 3, 4 and 5, we discuss the central issues of parameter heterogeneity, variable non-stationarity and cross-section dependence in great detail and show how their interplay leads to the breakdown of standard assumptions in the empirical estimators commonly applied in the literature. A brief conclusion summarizes the arguments.

1. A General Empirical Framework for Production Analysis with Cross-country Panel Data

This section sets out two general empirical frameworks for the macro production function and for the convergence equation from the canonical Solow model, based on a Cobb–Douglas specification.[2] We specify the general production function and convergence equation models below as a means to build an encompassing framework for the literature that affords maximum flexibility with regard to specification. This flexibility is briefly motivated from an economic and econometric standpoint.

1.1 *An Encompassing Framework*

We assume panel data for N countries, with a substantial time-series dimension T which may vary across countries (unbalanced panel). For the empirical production function let

$$O_{it} = \alpha_i L_{it} + \beta_i K_{it} + \gamma_i M_{it} + u_{it} \qquad u_{it} = A_{0,i} + \boldsymbol{\lambda}_i' \boldsymbol{f}_t + \varepsilon_{it} \tag{1}$$

for $i = 1, \ldots, N$ and $t = 1, \ldots, T$. We can think of this framework as representing N country equations, or a single pooled equation. O represents gross output, L labour force, K capital stock and M material inputs (all in logarithms). These represent the observable variables of the model.[3] For TFP we employ the combination of a country-specific TFP level $A_{0,i}$ and a set of common factors $\boldsymbol{f}_t$ with country-specific factor loadings $\boldsymbol{\lambda}_i$ – TFP is unobserved. We introduce common factor models in detail in Section 4; for the time being we highlight this specification of unobservables as a flexible way to allow for each country to be influenced differentially.

Much of the micro-econometric literature on production functions adopts gross-output based models, but, at the macro level, a specification using value-added (Y in logs) as the dependent variable is more common:

$$Y_{it} = \alpha_i^{va} L_{it} + \beta_i^{va} K_{it} + u_{it} \qquad u_{it} = A_{0,i}^{va} + \boldsymbol{\lambda}_i' \boldsymbol{f}_t + \varepsilon_{it} \tag{2}$$

Our notation in (2) indicates that parameter values and interpretation will differ between a value-added based and gross-output based empirical specification, but under certain assumptions we can transform results to make them directly comparable.[4]

We maintain the following assumptions for the general output based production function model and analogously for the value-added variant:

A.1. The parameters α_i, β_i and γ_i are unknown random coefficients with fixed means and finite variances. The same applies for the unknown factor loadings, i.e. $\boldsymbol{\lambda}_i = \boldsymbol{\lambda} + \boldsymbol{\eta}_i$ where $\boldsymbol{\eta}_i \sim \text{iid}(0, \Omega_\eta)$.[5] The unobserved TFP evolution, captured by $\boldsymbol{\lambda}_i' \boldsymbol{f}_t$, can contain elements which are common across countries as well as elements which are country specific.
A.2. Error terms $\varepsilon_{it} \sim N(0, \sigma_i^2)$, where the variances σ_i^2 are finite.
A.3. Observable inputs $\boldsymbol{x}_{it} = \{L_{it}, K_{it}, M_{it}\}$ and output O_{it}, as well as the unobserved common factors $\boldsymbol{f}_t$, are *not a priori* assumed to be stationary variables/processes.
A.4. The observable inputs $\boldsymbol{x}_{it}$ may be functions of some of the unobserved common factors $\boldsymbol{f}_t$ driving output. This would lead to correlation between regressors and unobservables u_{it}, creating difficulties for the identification of the technology parameters $\alpha_i, \beta_i, \gamma_i$.

The majority of the growth empirics literature does not estimate production functions, but focuses on versions of the ‘convergence regression equation’ (Mankiw *et al.*, 1992), which is commonly estimated for aggregate economy data. In its simplest form – without human capital augmentation – this links country i’s per capita GDP (y_i in logs) to a proxy for its savings rate (s_i^k in logs). For the most general, ‘unrestricted’ form of this convergence equation let

$$y_{it} - y_{i,t-\tau} = -\left(1 - e^{-\xi_i \tau}\right) y_{i,t-\tau} - \left(1 - e^{-\xi_i \tau}\right) \left(\frac{\beta_i^{va}}{1 - \beta_i^{va}}\right) \ln(\delta + \bar{n}_i + \mu^\star)$$
$$+ \left(1 - e^{-\xi_i \tau}\right) \left(\frac{\beta_i^{va}}{1 - \beta_i^{va}}\right) \bar{s}_i^k + u_{it} \tag{3}$$

$$u_{it} = \left(1 - e^{-\xi_i \tau}\right) A_{0,i}^{va} + \left(t - e^{-\xi_i \tau}(t - \tau)\right) \boldsymbol{\lambda}_i' \boldsymbol{f}_t + \varepsilon_{it} \tag{4}$$

where τ is the time period between two observations on the income variable (more on this notation below), $\bar{n}_i$ is the average population growth rate in country i over the period between $t - \tau$ and t and $\bar{s}_i^k$ is the average savings rate over the same period. δ is the depreciation rate for physical capital and ξ_i indicates the speed of convergence (per annum) to either a common or country-specific steady-state equilibrium, depending on specification.

$\bar{n}_i$, $\bar{s}_i^k$ and y_i are the observable variables in this framework, whereas TFP ($A_{0,i}^{va}$ and $\boldsymbol{\lambda}_i' \boldsymbol{f}_t$) is unobserved. This setup allows for TFP evolution to differ across time and countries. Note that there is an additional parameter which also refers to TFP growth, namely $\mu^\star$: this is the outcome of Solow’s steady-state equation for income, where capital stock per effective worker k^* is a function of both the savings rate s^k and population growth, growth in labour-augmenting efficiency and capital depreciation ($n + \mu^\star + \delta$). We adopt this notation to highlight that empirical work using this type of framework has commonly neglected $\mu^\star$, which together with depreciation δ is simply assumed to equal a constant 5% per annum in all countries.

Following estimation, the coefficients on the observables can be transformed to yield the 'implied' capital coefficients β_i^{va}. As in the production function framework, β_i^{va} is the technology parameter in the Cobb–Douglas production function that frames the production analysis for country i. Although the literature refers to equations like (3) as 'growth regressions', this equation is in fact a levels regression where $y_{i,t-\tau}$ has been subtracted from both sides to investigate out-of-steady-state behaviour – this approximation is only valid *in proximity* to the steady state.[6] The convergence regression model can be rewritten as an equivalent dynamic panel equation with variables in levels *on both* sides (see Islam, 1995; Bond, 2002) to underline this point.

This framework provides us with a panel of T/τ equations for N countries. We briefly illustrate the implications of adopting different values for τ: at one extreme, if $\tau = T$, we have a single cross-section regression equation with N observations. Let $t = 1985$ and $t - \tau = 1960$, thus $\tau = 26$; then (3) transforms into

$$\begin{aligned} y_{i,1985} - y_{i,1960} = &- \left(1 - e^{-\xi_i \tau}\right) y_{i,1960} \\ &- \left(1 - e^{-\xi_i \tau}\right) \left(\frac{\beta_i^{va}}{1 - \beta_i^{va}}\right) \ln(\delta + \bar{n}_i + \mu^\star) \\ &+ \left(1 - e^{-\xi_i \tau}\right) \left(\frac{\beta_i^{va}}{1 - \beta_i^{va}}\right) \bar{s}_i^k + \left(t - e^{-\xi_i \tau}(t - \tau)\right) a^{va} + \varepsilon_i \end{aligned} \tag{5}$$

Equation (5) represents an approximation which works best in the vicinity of the steady state, and the intercept term a^{va} represents initial TFP level A_0^{va} *as well as* the constant for TFP evolution $\boldsymbol{\lambda}_i \boldsymbol{f}_t$. Our discussion in Section 2.1 will highlight the important assumptions implicit in this specification of TFP.

At the other extreme, if $\tau = 1$ we have a panel of dimension $T \times N$:

$$\begin{aligned} \Delta y_{it} = &- \left(1 - e^{-\xi_i}\right) y_{i,t-1} - \left(1 - e^{-\xi_i}\right) \left(\frac{\beta_i^{va}}{1 - \beta_i^{va}}\right) \ln(\delta + n_{it} + \mu^\star) \\ &+ \left(1 - e^{-\xi_i}\right) \left(\frac{\beta_i^{va}}{1 - \beta_i^{va}}\right) s_{it}^k + u_{it} \end{aligned} \tag{6}$$

$$u_{it} = \left(1 - e^{-\xi_i}\right) A_{i,0}^{va} + \left(t - e^{-\xi_i}(t - 1)\right) \boldsymbol{\lambda}_i' \boldsymbol{f}_t + \varepsilon_{it} \tag{7}$$

where $A_{i,0}^{va}$ and $\boldsymbol{\lambda}_i' \boldsymbol{f}_t$ capture the unobservables. Note that n_{it} and s_{it}^k are no longer average values.

The two examples indicate that the choice of τ has considerable impact on empirical implementation, as will become clearer when we discuss the applied literature below. We maintain the following assumptions for the general convergence model in (3) and (4) and the data:

B.1. The parameters β_i^{va} and ξ_i are unknown random coefficients with individual means and finite variances.[7] The same applies for the unknown factor loadings $\boldsymbol{\lambda}_i$.

B.2. Error terms $\varepsilon_{i\tau} \sim N(0, \sigma_i^2)$, where the variances σ_i^2 are finite.
B.3. The specification of TFP levels and their evolution paths is equivalent to an unobserved common factor model that allows for common and idiosyncratic elements of TFP.
B.4. We allow for the possibility that the observable inputs (s_{it}^k, n_{it}), per capita GDP (y_{it}) and the unobserved factors driving TFP evolution ($\boldsymbol{f}_t$) are non-stationary.[8] The consequences of this property will be investigated.
B.5. The observable inputs may contain some of the unobserved common factors $\boldsymbol{f}_t$ driving output. This would lead to correlation between regressors and unobservables u_{it}, making the technology parameters difficult to identify.

In *econometric* terms, the setup in the general empirical production function and convergence equation frameworks allows for parameter heterogeneity across countries in the impact of observables (inputs) and unobservables (TFP) on output. In line with the cross-country growth empirics literature (Durlauf *et al.*, 2005) our setup assumes parameter constancy over time.[9] The common factor model specification for the error terms u_{it} operationalizes cross-section dependence in the panel, whereby unobserved processes are correlated across countries. Note that this specification encompasses the case of serial correlation in the u_{it}, which can be thought of as arising from persistence in one or more of the common factors, i.e. $\boldsymbol{f}_t = \mu + \boldsymbol{\rho} \boldsymbol{f}_{t-1} + e_t$ where μ is a drift term and $0 < \boldsymbol{\rho} \leq 1$. The impact of processes driving TFP evolution can be heterogeneous ($\boldsymbol{\lambda}_i$) or homogeneous ($\boldsymbol{\lambda}_i \equiv \boldsymbol{\lambda}\, \forall i$) across countries, or a mixture of both. 'Weak cross-section dependence' (Chudik *et al.*, 2009) such as spatial correlation can also be accommodated in this general setup (see Pesaran and Tosetti, 2009) by assumption of additional factors.[10] Furthermore, since observable inputs may be driven by (some of) the same unobserved factors as output we encounter variable endogeneity, which makes it difficult to identify the technology parameters separately from the factor loadings $\boldsymbol{\lambda}_i$. Input variables and output as well as the unobserved factors may be non-stationary. This allows for a number of potential cases: first, factor inputs and output are non-stationary and cointegrated; second, factor inputs, output and (some) common factors are non-stationary and cointegrated; and third, neither of the above (non-cointegration). Whether the cointegrating vectors are homogeneous or heterogeneous across countries depends on the nature of the factor input parameters (α, β, γ). Thus the general empirical frameworks provide maximum flexibility with regard to the time-series and cross-section properties of the variable series and unobserved processes analysed.

Note that our discussion thus far has focused on 'Solow-type regression models': here the specification arises out of the steady-state solutions of the theory model, as exemplified by the human-capital-augmented Solow model of Mankiw *et al.* (1992). Due to the origin of the empirical equation, cross-coefficient restrictions apply and the technology parameter of the underlying (Cobb–Douglas) production function can be backed out. Many empirical papers have entered additional regressors into the convergence specification of Mankiw *et al.* (1992) without deriving their impact in the solution to an economic theory model and estimate these models without consideration for cross-coefficient restrictions. These approaches are also referred to as 'Barro regressions', following the seminal contribution by Robert Barro (1991). In either case it is difficult to argue that the variables entering the regression model can capture the underlying 'structural heterogeneity' across countries if they are potentially driven by initial conditions and thus endogenous (Durlauf *et al.*, 2005).[11] Typically a range of instrumentation strategies is employed to deal with potential variable endogeneity (see Clemens and Bazzi, 2009), while more recently Bayesian model averaging was adopted to argue for or against the robustness of specific growth determinants (Sala-i-Martin *et al.*, 2004). Note that our common factor model setup allows for variable endogeneity, which will need to be addressed in estimation.

In *economic* terms, the above frameworks in equations (1)–(4) are *as general as possible*, allowing for individual countries to possess idiosyncratic production technologies with regard to factor input parameters, TFP levels and TFP evolution. We briefly motivate this in the following paragraphs.

A theoretical justification for *heterogeneous technology parameters* can be found in the 'new growth' literature. This strand of the theoretical growth literature argues that production functions differ across countries and seeks to determine the sources of this heterogeneity (Durlauf *et al.*, 2001). As Brock and Durlauf (2001, pp. 8–9) put it:

> ... the assumption of parameter homogeneity seems particularly inappropriate when one is studying complex heterogeneous objects such as countries ...

The model by Azariadis and Drazen (1990) can be seen as the 'grandfather' for many of the theoretical attempts to allow countries to possess different technologies from each other (and/or at different points in time). Their model incorporates a qualitative change in the production function, whereby upon reaching a critical 'threshold' of human capital, economies will jump to a higher steady-state equilibrium growth path represented by a different production function. Further theoretical work leads to multiple equilibria interpretable as differential production technology across countries (e.g. Murphy *et al.*, 1989; Banerjee and Newman, 1993; Durlauf, 1993). A simpler justification for heterogeneous production functions is offered by Durlauf *et al.* (2001), as quoted at the beginning of this chapter: the Solow model was never intended to be valid in a homogeneous specification for *all* countries, but may still be a good way to investigate *each* country, i.e. if we allow for parameter differences *across* countries.

The heterogeneity of technology parameters has been considered extensively in the empirical convergence literature.[12] Originally technology parameters were assumed group-specific rather than fully heterogeneous and studies opted for *ad hoc* country groupings by income or geographic location (e.g. Durlauf and Johnson, 1995; Caselli *et al.*, 1996; Liu and Stengos, 1999). More recent studies have adopted data-driven 'endogenous clustering models' whereby the makeup of the convergence clubs is determined endogenously – for instance Hobijn and Franses (2000) and Paap *et al.* (2005) applied this approach to annual growth rates, whereas the analysis by Basturk *et al.* (2008) also takes the differential effect of standard 'growth determinants' into account. In a parallel development, threshold regression models (Hansen, 2000; Caner and Hansen, 2004; Kourtellos *et al.*, 2008) have been employed to split the sample of countries into different 'regimes' in Barro-type regressions (e.g. Papageorgiou and Chmelarova, 2005; Foster, 2006; Crespo Cuaresma and Doppelhofer, 2007; Aidt, 2009). Durlauf *et al.* (2001) then allowed for technology parameters to vary *by country* as a function of initial per capita income in a simple cross-section regression model, referred to as a 'local' Solow growth model.

Many empirical growth models in the applied literature assume common technology parameters in a framework allowing for *differential TFP* across countries – a specification derived from further extension of the empirical version of a neoclassical Solow–Swan model by Mankiw *et al.* (1992). Mankiw *et al.* (1992) include human capital in their theory model and derive an empirical equivalent using secondary education as a proxy for human capital. In alternative theoretical models determinants such as institutional environment, good governance or geographical features are all deemed important for the production process (e.g. Sachs, 2003; Glaeser *et al.*, 2004; Rodrik *et al.*, 2004). However, identifying good proxy variables in these cases proved more cumbersome, especially if the empirics were conducted allowing for dynamics, rather than in a single cross-section regression. In a black-box approach to the specification of an underlying production function, empirical models therefore can allow for heterogeneity in TFP levels and growth rates across countries: since including *all* non-standard growth determinants (i.e. those other than labour, capital and materials) in the model is difficult, it is assumed that a flexible TFP specification can empirically 'mop up' any excess variation in the data – this is essentially the approach in growth accounting exercises. At the same time, any notion of common TFP evolution as implicitly assumed by Solow (1956)[13] and implemented empirically by Mankiw *et al.* (1992) has been dropped. Another motivation for a flexible TFP specification is that an overly rigid homogeneous specification is likely to induce parameter non-constancy across countries; adding non-standard growth determinants alone may not be sufficient to 'remove' the impact of apparent parameter heterogeneity, but flexible TFP may go some way to achieve this.

In our general empirical models above we emphasize a view of TFP as a 'measure of ignorance' (Abramowitz, 1956), more broadly defined than a narrow notion of 'technology', incorporating a wider set of factors that can shift the production possibility frontier ('resource endowments, climate, institutions, and so on', Mankiw *et al.*, 1992, pp. 410–411). This is in contrast to the notion of TFP as a definitive efficiency index, as commonly adopted in the microeconometric literature on production analysis. Furthermore, we allow for the possibility that TFP is *in part* common to all countries, e.g. representing the global dissemination of non-rival scientific knowledge. Finally, we do not impose restrictions on the individual evolution paths of these unobservables.[14] These considerations lead to the adoption of a TFP structure that allows for common and/or country-specific evolution, which we implement using the common factor model approach.

The following section will show that our frameworks encompass the empirical specifications in the literature of the past two decades. In our presentation, we highlight technology parameter estimates reported in different studies and, to put these estimates into context, we briefly discuss parameter values and their consistency with macro data on factor input shares.

1.2 *A Note on Factor Shares and Production Function Parameters*

Conducting empirical growth analysis with *aggregate* economy data has an advantage over many other empirical exercises, in that we already know parts of the answers we are seeking: under the assumption of perfect competition the values for β^{va} in the above value-added based production function (2) and implied β^{va} in the convergence equation (3) should be equal to the physical capital shares in aggregate income. Macroeconomic data for labour are available through the aggregate data on wages and welfare payments to labour.[15] From this it can be deduced that the labour share in income is roughly two-thirds, with capital's share around one-third (Mankiw *et al.*, 1992, p. 415).[16] It has been pointed out that while country data show high persistence over time, there is considerable variation in the factor shares *across* countries, with the labour share ranging from 5% to 80% of aggregate value-added (UN, 2004, national accounts data). Deviation from perfect competition may account for some of this dispersion. Gollin (2002) attributes this phenomenon to the mismeasurement of labour income in small firms, which is particularly the case in less-developed countries, and concludes that adjusted labour shares are in a range of 65% to 80% in the majority of countries. As Islam (2003) and Pedroni (2007) point out, the majority of empirical approaches tend to produce capital coefficients far in excess of 0.3 and this paper points to empirical misspecification (neglected heterogeneity in the impact of observables and unobservables) and the neglect of time-series and cross-section dependence properties of the data as potential explanations for this 'empirical puzzle'.

2. A Review of Growth Empirics

In this section we demonstrate that the general models introduced above encompass the most influential models of the empirical growth literature in the past two decades as well as approaches from the emerging non-stationary panel literature. Our focus in this selective survey is on cross-country *macro* data. The question whether the empirical specification is most appropriately applied to data at the aggregate economy or at the sectoral level, although in our view of great importance, is not central to the discussion.[17] This aside, the question as to *how to empirically implement these general models* is a separate issue and we refer to recent contributions to the non-stationary panel econometric literature for suggestions (Bai and Ng, 2004; Coakley *et al.*, 2006; Pesaran, 2006; Bai, 2009; Kapetanios *et al.*, 2009). In our review of the growth empirics literature we discuss the restrictions made by each of the seminal contributions on the general model and set out their implications for estimation and inference.

2.1 *Complete Parameter Homogeneity*

The canonical Mankiw *et al.* (1992) single cross-section regression model is explicitly derived from the Solow–Swan growth model (Solow, 1956, 1957; Swan, 1956), where the steady-state solutions are given empirical equivalents. We simplify their notation to

$$y_{i,1985} = -\left(\frac{\beta^{va}}{1-\beta^{va}}\right)\ln(\delta + \bar{n}_i + \mu^{\star}) + \left(\frac{\beta^{va}}{1-\beta^{va}}\right)\bar{s}_i^k + a^{va} + \varepsilon_i \qquad (8)$$

The regression applies *end of period level* of per capita GDP and *total period averages* for population growth ($\bar{n}$) and the savings rates for physical capital ($\bar{s}^k$) – averages are employed to reduce measurement error and to boost the robustness of results in the presence of business cycles and/or changes in capacity utilization over time (Barro, 1991). The a^{va} term is said to include not only 'technology', but also resource endowment, geography, climate, institutions etc. – thus in the spirit of a 'measure of ignorance'.

In a first step, Mankiw *et al.* (1992) estimate the above equation by ordinary least squares (OLS) using Penn World Table (PWT) data – in the following we concentrate on results for the samples including both developed and developing economies in each paper discussed. The authors' first important empirical result derives from the theory-driven inclusion of human capital[18] into this model, which yields implied income elasticities for physical and human capital of 0.31 and 0.28, respectively. This 'augmented Solow model' explains around 80% of the variation in the data.

The second part of the Mankiw *et al.* (1992) analysis concerns the *out-of-steady-state* behaviour of the augmented Solow model, thus addressing the convergence argument, which has found huge interest in the empirical literature: convergence refers to 'the tendency of differences between countries [in income terms] to disappear over time' (Durlauf and Johnson, 2008). Via approximation around the steady state and minor transformation, the authors derive a convergence regression as represented in equation (5) above. They find that the conditional convergence rate in the model augmented with human capital is statistically significant and relatively slow (albeit higher than in the unaugmented case), around 1.4% per annum. The implied income elasticities for physical capital in the unaugmented and human-capital-augmented models are considerably above those in the steady-state model,[19] while the implied human capital coefficient is still broadly in line with the cross-section regression result.

The paper received most criticism for its assumption of random cross-country differences in TFP levels, which allow for the empirical representation $A_{i,0}^{va} = a^{va} + \epsilon_i$ by assuming error term exogeneity such that $\epsilon_i \perp n,\ s^k,\ s^h$ (Islam, 1995; Caselli *et al.*, 1996; Temple, 1999). In words, this assumption implies that the probability of Malawi having a higher initial TFP level than the USA is the same as the probability of the reverse (comment by Steven Durlauf). Furthermore, if the underlying unobserved factors drive both regressors and TFP growth contained in the error terms then regression parameters on the observables are unidentified, as will be shown in Section 4 below.

Later papers also pointed out that in using OLS the convergence specification *by construction* induces downward bias in the convergence rate (Knight *et al.*, 1993; Islam, 1995; Caselli *et al.*, 1996): the presence of the initial income variable on the right-hand side induces non-zero correlation between $y_{i,1960}$ and the regression error, as both contain the omitted country-specific effect ϵ_i from $A_{0,i} = a + \epsilon_i$. The parameter estimate on $y_{i,1960}$ is therefore biased towards zero, with the implication of a reduced convergence rate ξ. Nevertheless, this specification forms the basis for the popular 'Barro regressions' where the convergence equation is expanded by any variable deemed a relevant growth determinant. The inclusion of additional variables in the model can be viewed as relaxing the assumption of *random* TFP level differences (Durlauf *et al.*, 2005, p. 579), but the problems raised above remain.[20]

The Mankiw *et al.* (1992) convergence regression model in (5) is nested within our more general model, under the assumption of

(i) common technology parameters across countries ($\beta_i^{va} \equiv \beta^{va}$), or no bias from imposition of a common parameter on heterogeneous country coefficients, i.e. the resulting estimate is the unweighted mean of underlying 'micro-parameters';
(ii) constant savings rates s^k and population growth rates n for each country over the *full* period of observation;
(iii) common and constant TFP evolution across countries, equivalent to common factor loadings on unobserved common factors which themselves are linear in evolution;
(iv) random differences in TFP levels across countries; and
(v) cross-section independence

The essence of the Mankiw *et al.* (1992) empirics is that if we assume all countries are in *steady state* and can be represented by the same underlying production function, then the augmentation of an empirical Solow model with a proxy for human capital yields sensible values for the implied physical capital elasticity, and can account for a large share of the variation in the data. The same cannot be said for the *out-of-steady-state* version of their model (convergence equation), where the implied capital coefficient in the human-capital-augmented model changes considerably. This aside, the assumption of error term exogeneity required for the specification of $A_{i,0}^{va} = a^{va} + \varepsilon_i$ is econometrically useful but conceptually questionable, while the convergence equation setup induces downward bias in the estimate for ξ by construction.

2.2 *Heterogeneous TFP Levels*

Given the availability of longer time-series data, the application of panel data methods to investigate the growth and development process was a natural next step, notably the contributions by Knight *et al.* (1993), Loayza (1994) and Islam (1995). The latter tackles the critical assumption of error term exogeneity in the Mankiw *et al.* (1992) convergence framework by transforming the single convergence equation into a dynamic panel model in levels with panels of 5-year averages ($\bar{n}_i$, $\bar{s}_i^k$) and then introducing country (and period) fixed effects:

$$y_{it} = -\left(-e^{-\xi_i \tau}\right) y_{i,t-\tau} - \left(-e^{-\xi_i \tau}\right) \left(\frac{\beta^{va}}{1-\beta^{va}}\right) \ln(\delta + \bar{n}_i + \mu^\star) + \left(-e^{-\xi_i \tau}\right) \left(\frac{\beta^{va}}{1-\beta^{va}}\right) \bar{s}_i^k + \left(-e^{-\xi_i \tau}\right) A_{0,i}^{va} + \left(-e^{-\xi_i \tau}\right) \sum_{s=\tau}^{T} \mu_s D_s + u_{it} \tag{9}$$

Due to the change in dependent variable the interpretation of estimated coefficients changes slightly,[21] but the implied technology coefficient β^{va} is equivalent to that in our general model in equations (3) and (4) with $\tau = 5$.[22] The crucial innovation over Mankiw *et al.* (1992) is the introduction of country fixed effects (operationalized using the 'within-groups' transformation), which allow for differential TFP levels across countries. Islam further uses period dummies (the sum of *D*s with parameter coefficients μ in (9)) to account for common TFP evolution – this specification imposes no assumption about linearity (or stationarity) on the underlying common TFP evolution, although time-averaging arguably diminishes the model's ability to capture TFP dynamics (Lee *et al.*, 1997). Islam also uses the PWT data set with sample and variable construction close to that in Mankiw *et al.* (1992).

Following some OLS estimations to show that a pooled convergence regression of 5-year averages yields next to identical results to the single convergence regression, the above panel model is estimated *without* the human capital variable but accounting for country fixed effects. The implied convergence rate is much higher than in Mankiw *et al.* (1992), around 5% per annum, while the implied income elasticity for capital is 0.44 – it is thus suggested that individual country effects (TFP levels) play an important role in the development process.[23] Inclusion of the human capital variable slightly lowers the convergence rate but the implied capital coefficient at 0.52 is quite similar to the regression without augmentation.

The major criticism levelled at Islam (1995) and other panel approaches was related to the potential endogeneity issues in the empirical model. Although the within-groups transformation wipes out the country fixed effects, Nickell (1981) has shown that in the presence of a lagged dependent variable this approach requires sizeable T for consistency, which is not given in the averaged panel case. The estimate on $y_{i,t-\tau}$ is therefore likely to be biased *downward*. Islam (1995) remarks that his own Monte Carlo experiments indicate that this bias is likely to be small. Potential bias due to the dynamic setup aside, Caselli *et al.* (1996) argue that violation of the exogeneity assumption with regard to the input variables renders Islam's results, as well as those from similar panel fixed effects models, inconsistent. Lee *et al.* (1997, p. 321) highlight that the validity of Islam's estimates depend critically on the assumption of *common* TFP evolution for all countries[24] – if this is violated, country residuals will be serially correlated with a unit coefficient, causing *upward* bias in the convergence rate estimate.

The model in (9) is nested within our more general model in (3), under the assumption of

(i) common technology parameters across countries ($\beta_i^{va} \equiv \beta^{va}$), or no bias from imposition of a common parameter on heterogeneous country coefficients;
(ii) roughly constant population growth rates n and savings rates s^k over the 5-year intervals;
(iii) common TFP evolution across countries, equivalent to common factor loadings on the unobserved common factors;
(iv) heterogeneous TFP levels across countries;
(v) stationary input and output variable series;
(vi) cross-section independence; and
(vii) no dynamic misspecification through period averaging.

The last point merits further comment. The Islam (1995) framework of short period-averaged panels has become a standard alternative to the single convergence regression model and both are commonly implemented without any concern for the time-series properties of the data. If we were to apply the period-averaged panel approach to the most recent PWT data (Heston *et al.*, 2009) we would obtain nine time periods; given that averages constructed from non-stationary variables are non-stationary themselves (Granger, 1988; Granger and Siklos, 1995; Marcellino, 1999), the use of longer T eventually faces the same issues as annual data, outlined in detail in Section 3. As our discussion of Caselli *et al.* (1996) below indicates, high levels of persistence in the data (with unit roots the extreme case) in general can have a serious impact on estimation and inference. In defense of Islam (1995), his approach successfully challenges single convergence regressions by providing evidence of the significance of fixed effects in a panel setup.

A variant on the above panel estimation approach is presented by Caselli *et al.* (1996). Instead of fixed effects estimation they apply dynamic panel generalized method of moments (GMM) estimation techniques (Arellano and Bond, 1991), which eliminate heterogeneous TFP levels by differencing and overcome endogeneity issues by instrumentation making use of the panel structure (for a detailed discussion of the estimation approach see Bond, 2002). Their model setup and assumptions are identical to those in Islam (1995) described above – the difference between the approaches is in the empirical estimator, which is argued to deal with variable endogeneity in a short-T dynamic panel data model. Throughout their empirics the authors use variables in deviation from the cross-section mean to account for common TFP evolution, a practice which is valid if technology parameters are homogeneous across countries, but otherwise introduces misspecification bias (Pedroni, 2000). TFP evolution is thus assumed identical across countries.

In their estimation without human capital the authors find an implausibly low income elasticity for capital of 0.10, with convergence very high at 13% per annum. In the human-capital-augmented regression the income elasticities for physical and human capitals are 0.50 and −0.26, respectively, with convergence around 6.8% per annum. The implied physical capital parameter is relatively close to the result in Islam (1995), while the implied human capital elasticity is significantly negative and the convergence rate λ is almost doubled. Since the Islam results in turn have lower implied capital elasticities but roughly double the convergence rate of the Mankiw *et al.* (1992) convergence regression, Caselli *et al.* (1996) argue that the step-wise 'improvement' in the estimate for conditional convergence from Mankiw *et al.* (1992) to panel approaches like Islam (1995) and further to their own results is the outcome of appropriate accounting for country fixed effects and endogeneity, respectively. They interpret the low capital estimate in their unaugmented model as a rejection of the empirical Solow model, and embark on more general 'Barro-type' regressions to analyse convergence using the same difference GMM estimation method. Augmentation is carried out using variables such as government consumption, male and female education, and black market premium. Results suggest consistently high convergence rates of up to 10% per annum.

The difference GMM estimator used in the analysis by Caselli *et al.* (1996) was found to be liable to considerable small sample bias for highly persistent variable series (Blundell and Bond, 1998, 1999).[25] Furthermore, if technology parameters are heterogeneous across countries, 'there exists no consistent instrumental variables estimator' (Lee *et al.*, 1997, p. 367).[26] Implied capital coefficients estimated by Caselli *et al.* (1996) are vastly different from the expected 30% share in value-added – a finding which led them to reject the empirical Solow model, but which may be attributable to misspecification and weak instrument bias.

Note that both Islam (1995) and Caselli *et al.* (1996) favour a dynamic empirical specification but in the form of a 'partial adjustment model' (PAM) ($Y_t = f(Y_{t-1}, X_t)$), which in turn is a (Koyck) transformation of a distributed lag model with infinite lag structure ($Y_t = f(X_t, X_{t-1}, X_{t-2}, \ldots)$) and a geometric rate of decline in the impact of lagged values of X on Y. This structure emerges in the panel from the convergence equation by Mankiw *et al.* (1992), as is shown in Islam (1995, p. 1135). A more general dynamic specification is the autoregressive distributed lag (ARDL) model ($Y_t = f(Y_{t-1}, X_t, X_{t-1})$), for which 'virtually every type of single equation model in empirical time-series econometrics is a special case' (Hendry, 1995, p. 212), including the error correction model (ECM). Although the PAM has 'respectable pedigree in economic analysis' (Hendry, 1995, p. 256) it nevertheless somewhat awkwardly imposes a zero coefficient on X_{t-1} and a more general ARDL/ECM specification of dynamics may therefore be recommended.

2.3 *Heterogeneous TFP Levels and Growth Rates*

A further variant on the stationary panel estimation approach is presented by Martin and Mitra (2002), who estimate sectoral production functions for agriculture and manufacturing using Crego *et al.* (1998) data for 1967 to 1992 – to our knowledge this is the only paper in the literature specifying separate production functions at this level of aggregation in a sample including developed and developing countries. In a further contrast to the previous papers, the authors allow for differential TFP levels *and* growth rates across countries, modelled via country-specific intercepts and linear trend terms in a pooled panel estimation using *annual data* for around 50 countries. We adjust their notation for consistency:

$$\text{agriculture} \qquad y_{it} = \beta^{va} k_{it} + \theta^{va} n_{it} + u_{it} \qquad u_{it} = A^{va}_{0,i} + \mu_i t + \varepsilon_{it} \quad (10)$$

$$\text{manufacturing} \qquad y_{it} = \beta^{va} k_{it} + u_{it} \qquad u_{it} = A^{va}_{0,i} + \mu_i t + \varepsilon_{it} \quad (11)$$

where the additional factor input n in the agriculture equation refers to land per worker (in logarithms, as are sectoral value-added per worker, y, and capital stock per worker, k). TFP growth is captured by the country trends and thus assumed to be constant over time and heterogeneous across countries (and sectors).

As indicated Martin and Mitra (2002) impose constant returns to scale on this model and estimate it separately for agriculture and manufacturing data using least squares dummy variable. Their results indicate considerable variation in TFP growth rates between sectors and across countries, with TFP growth rates in agriculture commonly *in excess* of those in manufacturing.[27] The capital coefficient in the Cobb–Douglas estimation of the manufacturing data is estimated at 0.69; the authors highlight the magnitude of this coefficient in comparison to the macro data on factor shares and point to the omission of human capital from the model as a likely explanation. For agriculture, where arable land per worker is included as an additional regressor (coefficient 0.24), the estimated capital elasticity is 0.12. As in many other studies which obtain country-specific TFP levels or growth rates via production function regressions, the validity of the TFP estimates in the face of possibly *biased* factor parameters is not assessed.

The sector-level Martin and Mitra (2002) regression model is nested within the value-added version of our general production function model in (2), under the assumption of

(i) common technology parameters (within each sector) across countries ($\beta_i \equiv \beta$);
(ii) heterogeneous TFP growth across countries, constant over time, implying stationarity;
(iii) heterogeneous TFP levels across countries;
(iv) cross-section independence; and
(v) stationary input, output and TFP.

Martin and Mitra (2002) thus address the issue of heterogeneity in TFP levels and growth rates in a static pooled fixed effects model, which imposes common technology parameters across countries. Time-series and cross-section dependence properties of their data are not formally investigated. The estimation equations for agriculture and manufacturing are static and no investigation of error correlation is undertaken to justify this choice.

Keeping production technology constant across countries may be seen as a less restrictive assumption when investigating more homogeneous sets of economies, such as the group of OECD countries. Arnold *et al.* (2007) empirically compare two rival growth models, the human-capital-augmented Solow model and a two-sector *AK* model (Lucas–Uzawa), using annual panel data from 21 OECD countries over the 1971–2004 period. Their empirical specification allows for flexibility in the short-run dynamics across countries, while imposing common long-run production technology. The latter is 'consistent with the idea that the OECD countries have access to common technologies and have intensive intra-industry trade and foreign direct investment' (Arnold *et al.*, 2007, p. 6). Adopting an ECM variant of the Mankiw *et al.* (1992) human-capital-augmented convergence equation, the authors employ the Pesaran *et al.* (1999) 'pooled mean group' (PMG) estimator to estimate long-run technology parameters and average speed of adjustment. In comparison to Martin and Mitra (2002) this approach has the desirable feature of estimating a *dynamic* specification with country-idiosyncratic speed of convergence and deviations from the steady state, while still allowing for country-specific TFP levels and growth rates. Results suggest that the OECD data are more in line with the Lucas–Uzawa model, implying that human capital can have a persistent effect of income *growth*. Issues arising from variable stationarity and cross-section dependence are not considered in this study.

2.4 *Full Parameter Heterogeneity in a Stationary Variable Model*

Inspired by the 'new growth theory' literature, Durlauf *et al.* (2001) test a 'local' Solow growth model which specifies all parameters in the Mankiw *et al.* (1992) convergence regression as functions $\psi(\cdot)$ of some 'development index' z_i (here: initial period GDP per capita), thus modelling parameter heterogeneity explicitly:

$$y_{i,1985} - y_{i,1960} = \boldsymbol{\psi}(z_i)'\mathfrak{X}_i + u_i \tag{12}$$
$$\mathfrak{X}_i \equiv \left\{A^{va}_{0,i},\ \ln(\delta + \bar{n}_i + \mu),\ \bar{s}^k_i,\ \bar{s}^h_i,\ y_{i,1960}\right\}$$

where $\bar{s}^h_i$ is their measure for human capital and $\boldsymbol{\psi}(z_i)' \equiv \{\psi_0(z_i), \ldots, \psi_4(z_i)\}$. The notion underlying this specification is that a single Solow model was never intended to apply to *all* countries, but may still be a good representation for the production process in *each* country individually (Durlauf *et al.*, 2001, p. 929). In essence, all observables and the TFP levels are allowed to differ across countries but are constrained by the development index z_i; the empirical specification does not allow for TFP growth rates to differ from the constant 5% per annum for depreciation and TFP growth $(\delta + \mu)$.

The authors use the same sample and variable definitions as Mankiw *et al.* (1992) so as to obtain directly comparable results, employing a two-step semi-parametric method for estimation. They find strong evidence for parameter heterogeneity in the impact of all regressors, and note that there is no monotonic relationship between initial income (z_i) and the country-specific capital coefficient estimates.

The Durlauf *et al.* (2001) model is nested within our more general convergence equation model in (3), under the assumption of

(i) heterogeneous technology parameters across countries, which are functions of initial income $(\beta^{va}_i \equiv \beta^{va}(z_i))$;
(ii) common and constant TFP growth across countries, implying its stationarity;
(iii) heterogeneous TFP levels across countries; and
(iv) cross-section independence.

Their paper concludes that substantial parameter heterogeneity seems to exist across countries, and that 'empirical exercises which fail to incorporate parameter heterogeneity are likely to produce misleading results' (Durlauf *et al.*, 2001, p. 935).

2.5 *Full Parameter Heterogeneity in a Model with Non-stationary Inputs*

In the mainstream growth literature it is argued that attempting to identify and quantify 'deep' or 'structural' parameters of a model for production is overly ambitious if one were to accept parameter heterogeneity (Durlauf *et al.*, 2005, p. 616). There are, however, a small number of papers explicitly contemplating the implications of non-stationarity and cointegration in a cross-country empirical framework which allows for heterogeneity in technology parameters and TFP evolution, most notably Pedroni (2007) and Canning and Pedroni (2008).[28]

Of these, the most encompassing framework for a non-stationary panel analysis of growth is presented by Pedroni (2007). The author exploits the time-series properties of the data to argue for a simple heterogeneous, cointegrated panel specification

$$y_{it} = c_i + d_i t + \theta_i s_{it}^k + u_{it} \tag{13}$$

for $i = 1, \ldots, N$ and $t = 1, \ldots, T$, where s_{it}^k is the (log) investment share in GDP and y_{it} is (log) GDP per capita. The model setup and data transformations, however, allow for a much richer underlying framework, taking account of time-invariant and near time-invariant (unmeasured) variables such as social capital or other forms of 'intangible capital' H (with savings rate s_{it}^h). Building on an augmented production function framework of the form

$$Y_{it} = \left(A_{it}^{va} L_{it}\right)^{1-\beta_i^{va}-\varphi_i} K_{it}^{\beta_i^{va}} H_{it}^{\varphi_i} \tag{14}$$

the author argues that if H is accumulated using a share of income Y, a human-capital-augmented general specification can be written as

$$\begin{aligned} y_{it} = &-\left(\frac{\beta_i^{va} + \varphi_i}{1 - \beta_i^{va} - \varphi_i}\right) \ln(\delta + n_{it} + \mu^\star) + \left(\frac{\beta_i^{va} + \varphi_i}{1 - \beta_i^{va} - \varphi_i}\right) s_{it}^k \\ &+ \left(\frac{\varphi_i}{1 - \beta_i^{va} - \varphi_i}\right) s_{it}^h + u_{it} \end{aligned} \tag{15}$$

$$u_{it} = A_{i0}^{va} + \mu_i t + \varepsilon_{it} \tag{16}$$

where $A_{i,0}^{va}$ and μ_i are country-specific TFP levels (in logs) and constant growth rates, respectively. s_{it}^h is the country-specific savings rate for 'intangible investment' at time t. If H is accumulated *without* using a share of income Y, the term involving s_{it}^h in equation (15) is subsumed into the intercept term and/or the linear trend term in (13). The interpretation of θ_i in (13) as well as that of the deterministic fixed effect and trend terms thus varies depending on whether aggregate investment drives 'intangible investment' or not.[29] For the purpose of this exposition we limit discussion to the case where all forms of intangible capital H are accumulated *without* using a share of aggregate income. In this case, the s^h_{it} term can be thought of as time-invariant or time-variant in an approximately linear fashion and is therefore absorbed into the trend terms and fixed effects in the estimation equation (13).[30] The same argument applies for the population growth term $\ln(\delta + n_{it} + \mu^\star)$. Under these assumptions, φ_i drops out and the slope coefficient for measured physical capital θ_i in equation (13) is simply a function of the production function share parameter for 'tangible' capital, i.e. $\hat{\theta}_i = \hat{\beta}_i(1 - \hat{\beta}_i)^{-1}$ just like in the previous models presented.

Given that the non-stationary panel estimation approach extracts the N long-run relationships $\hat{\beta}_i$ (via the cointegrating vector $\hat{\theta}_i$) *without* the requirement of proximity to the steady state, the author argues that there is no need to specify a lagged dependent variable term as in the standard convergence equation or any other dynamics: if savings rate and per capita GDP are non-stationary and cointegrated, then their long-run relationship will be extracted via the non-stationary panel econometric method applied, regardless of any short-run dynamics (Pedroni, 2007). It bears remembering that the averaged (group mean) estimates from this empirical approach converge to the true parameter values at speed $T\sqrt{N}$ – thus even faster than the superconsistent time-series parameter estimates.[31]

The model in (13) is estimated using the aggregate PWT data (version 5.6, 1950–1992). The non-stationary panel approach depends on a reasonably long time-series dimension, non-stationarity of the variables series and cointegration between savings rate and per capita GDP. Unit root and cointegration tests therefore form an *integral part* of the sample selection process in Pedroni (2007), reducing the sample from 152 to 51 (full time series, $T = 43$), and finally 29 countries (non-stationary series, cointegrated).

Estimation is carried out using the Phillips and Hansen (1990) time-series fully modified OLS (FMOLS) in each country and then averaging the results (this method is referred to as group mean FMOLS – Pedroni, 2000). The FMOLS procedure modifies OLS to account for serial correlation in the errors and the endogeneity in the regressors resulting from the existence of a cointegrating relationship (Phillips, 1995). The non-stationary panel econometric approach does away with a great deal of assumptions required in stationary empirics: provided we have cointegration the FMOLS estimates are 'superconsistent', implying that variable endogeneity does not affect the results and that the equation can be interpreted as a long-run equilibrium relationship. He does not test for the direction of causation in the cointegrating vector, which can be implemented using error correction model regressions developed in another paper by the same author (Canning and Pedroni, 2008).

The resulting cross-country mean for the implied capital coefficient is 0.27 – Pedroni (2007) interprets the proximity of this value to macro data on factor shares as a vindication of his estimation approach. This value is surprisingly robust to the inclusion of countries previously dropped in the sample selection process: 0.25 for the sample of 51 countries. The adoption of the Pesaran (2006) common correlated effects estimator (CCEMG) to allow for cross-section dependence leads to an average implied capital coefficient of around 0.2.

Essentially, Pedroni (2007) extracts structural parameters by 'blending out' unobservable features of the data through the use of country-specific trends and intercepts. The apparent heterogeneity in country parameter estimates $\hat{\theta}_i$ (and thus $\hat{\beta}_i$) is not due to sampling variation and/or the relatively short time series of the country regressions: a combination of formal tests strongly reject factor parameter homogeneity in the sample of 29 countries.

The Pedroni (2007) model can be nested within our general model in (3) assuming

(i) heterogeneous technology parameters across countries (β_i^{va});
(ii) heterogeneous TFP levels and growth rates across countries, the latter constant over time and thus stationary;
(iii) cross-section independence;
(iv) non-stationary input and output variable series within-sample; and
(v) cointegration between the savings rate and GDP per capita within-sample.

The empirical implementation by Pedroni (2007) requires sample countries to satisfy strict data requirements (*full* 43 years of data, non-stationarity, cointegration), which in effect excludes the vast majority of less-developed countries from the analysis. One might suggest that the failure to identify a larger set of countries which satisfy the test for cointegration between per capita GDP and the savings rate is not the result of limited time-series data but instead of a misspecified cointegrating vector. Furthermore, Pedroni's preferred regression approach precludes the possibility of a common, non-stationary TFP evolution; once this assumption is relaxed the implied mean capital coefficient drops to around 0.2. One explanation for this finding may be that following dual economy arguments (Temple, 2005) aggregate economy data represent the wrong basic unit of analysis for growth empirics, which should be based on sector-level analysis of agriculture, manufacturing and services instead.

2.6 *Common Technology Parameters in a Non-stationary Panel Model with Cross-section Dependence*

As will be detailed in Section 4 the study of cross-section correlation has only very recently begun to concern panel time-series econometricians, such that the number of empirical papers applying the insights of this new field of research to macro-production function estimation is still limited. Bai *et al.* (2009) raise the estimation and inference problems created by unobserved common factors in a production function framework as one of the motivations for their novel 'continuous updated' (CUP) estimators, which are applied by Costantini and Destefanis (2009) in a sectoral study of Italian regions (although the aggregate economy model is also estimated).

The latter's empirical model is a homogeneous technology, human-capital-augmented production function with a common factor structure. We adjust their notation for consistency:

$$Y_{it}^{j} = A_{i,0}^{j} + \alpha^{j,va}\left(L_{it}^{j} h_{it}^{j}\right) + \beta^{j,va} K_{it}^{j} + u_{it} \qquad u_{it} = \boldsymbol{\lambda}_i' \boldsymbol{f}_t + \varepsilon_{it} \tag{17}$$

for $i = 1, \ldots, N$ and $t = 1, \ldots, T$, where the superscript j indicates different industrial sectors (construction, services, 'industry', agriculture) and Y^j, L^j and K^j are (sector-specific) value-added, labour force and capital stock, respectively (all in logarithms). Following Hall and Jones (1999) h^j is a transformation of the educational attainment of each labour unit in sector j, based on the returns to education coefficients from Mincerian wage equations for Italy and the average years of education for the male and female population. The sectors investigated in this study are the construction, services, industry and agricultural sectors.

Prior to estimation the authors carry out three sets of testing procedures: first, they test all variables in their regression model for correlation across regions by applying the Pesaran (2004) cross-section dependence (CD) test, which firmly rejects cross-section independence in all sectors and the aggregate economy data. Second, they carry out panel unit root tests for all variables following Bai and Ng (2004) – the latter's PANIC (panel analysis of non-stationarity in idiosyncratic and common components) stationarity tests allow for the presence of common factors, the number of which is determined by a method detailed in Bai and Ng (2002). These tests cannot reject non-stationarity of input and output variables in all sectors and the aggregate economy data. Furthermore, the authors establish that the non-stationarity of the observable variables derives both from common factors and idiosyncratic components. Finally, they defactor the data (i.e. remove the impact of unobserved common factors from the input and output variables) and apply cointegration tests following Pedroni (1999, 2004a). This approach follows a suggestion by Gengenbach *et al.* (2006) and is appropriate since the non-stationarity of input and output variables was not merely due to the presence of common factors. The panel cointegration tests reject the null hypothesis of 'no cointegration' between value-added, capital stock and human-capital-augmented labour in all sectors and the aggregate economy data in favour of homogeneous cointegration – this alternative is imposed on the data by the estimation approach favoured by the authors.

Using annual data for 20 Italian regions from 1970 to 2003 the authors estimate the above empirical model by sector, comparing results for the heterogeneous parameter group mean FMOLS Pedroni (2000) and the Bai *et al.* (2009) continuous-updated fully modified (CUP-FM) pooled estimator. Their CUP-FM results indicate relatively similar physical capital elasticities across sectors (ranging from 0.24 in construction to 0.31 in services), although returns to scale differ substantially: agriculture, construction and in particular the services sector are characterized by decreasing returns, while the 'industrial' sector (manufacturing and utilities) displays increasing returns. Their CUP-FM result for aggregate economy data displays a capital coefficient of 0.40 and rejects constant returns to scale in favour of the decreasing returns alternative. While empirical results for the Pedroni (2000) group mean FMOLS estimator do not systematically bias capital coefficients upward or downward, they yield considerably higher labour coefficients in all sectors and the aggregate data, implying increasing returns across *all* models tested.

The Costantini and Destefanis (2009) model can be nested within our general production function model in (2) assuming

(i) homogeneous technology parameters (within sector j or the aggregate economy) across Italian regions ($\alpha_i^{j,va} \equiv \alpha^{j,va}$, $\beta_i^{j,va} \equiv \beta^{j,va}$);
(ii) heterogeneous TFP levels and growth rates across regions, with TFP evolution not required to be stationary;
(iii) cross-section dependence;
(iv) non-stationary input and output variable series within-sample; and
(v) cointegration between (sectoral) value-added, human-capital-augmented labour and capital stocks within-sample.

In comparing and contrasting the empirical estimates and returns to scale implications for non-stationary panel methods which account for or neglect cross-section dependence the authors highlight the importance of this variable property for empirical estimation.

2.7 *Parameter Heterogeneity in a Model with Non-stationary Inputs and Cross-section Dependence*

Given the relatively recent emergence of cross-section correlation issues in macro panels only a small number of empirical papers combine cross-section correlation in macro panel data with heterogeneous production technology, including work by Bhattacharjee *et al.* (2009) and Fleisher *et al.* (2010) on production in Danish regions and Chinese provinces, respectively,[32] as well as work by Cavalcanti *et al.* (2009) investigating the 'natural resource curse' in a panel of 53 countries. This aside, we noted above that Pedroni (2007) adopts the CCEMG estimator as a robustness check in his analysis of PWT data from 29 countries. In our own work (Eberhardt and Teal, 2009a, b) we analysed cross-country macro data for the manufacturing (48 countries, 1970–2002; UNIDO, 2004) and agricultural (128 countries, 1961–2002; FAO, 2007) sectors, respectively. In the following we present the results from the latter study, since it builds on the seminal contribution by Pesaran (2006) and extends this approach by aiming to identify the structure of the unobserved common correlation driving the data.

In Eberhardt and Teal (2009b) we adopt the following empirical framework to study cross-country production in agriculture: for $i = 1, \ldots, N,\ t = 1, \ldots, T$ and $m = 1, \ldots, k$

$$y_{it} = \boldsymbol{\beta}_i' \boldsymbol{x}_{it} + u_{it} \qquad u_{it} = \alpha_i + \boldsymbol{\lambda}_i' \boldsymbol{f}_t + \varepsilon_{it} \tag{18}$$

$$x_{mit} = \pi_{mi} + \boldsymbol{\delta}_{mi}' \boldsymbol{g}_{mt} + \rho_{1mi} f_{1mt} + \cdots + \rho_{nmi} f_{nmt} + v_{mit} \qquad \boldsymbol{f}_{\cdot mt} \subset \boldsymbol{f}_t \tag{19}$$

$$\boldsymbol{f}_t = \boldsymbol{\varrho}' \boldsymbol{f}_{t-1} + \boldsymbol{e}_t \qquad \text{and} \qquad \boldsymbol{g}_t = \boldsymbol{\kappa}' \boldsymbol{g}_{t-1} + \boldsymbol{\epsilon}_t \tag{20}$$

We assume a production function with observed net output (y_{it}) and observed inputs ($\boldsymbol{x}_{it}$) labour, agricultural capital stock, livestock, fertilizer and land under cultivation (all in logarithms). Unobserved agricultural TFP is represented by a combination of country-specific TFP levels α_i and a set of common factors $\boldsymbol{f}_t$ with heterogeneous factor loadings ($\boldsymbol{\lambda}_i'$). In equation (19) we introduce an empirical representation of the observed inputs in order to indicate the possibility for endogeneity: the input variables $\boldsymbol{x}_{it}$ are driven by a set of common factors $\boldsymbol{g}_{mt}$ as well as an additional set of factors $\boldsymbol{f}_{\cdot mt}$, which may also drive output y. Equation (20) indicates that the factors are persistent over time, which allows for the setup to accommodate non-stationarity in the factors ($\varrho = 1,\ \kappa = 1$) and thus the observables. It further allows for various combinations of cointegration: between output y and inputs $\boldsymbol{x}$, between output y, inputs $\boldsymbol{x}$ and (some of) the unobserved factors $\boldsymbol{f}_t$, or non-cointegration.

We estimate the above empirical model using a number of standard and novel panel estimators, including the Pesaran (2006) common correlated effects (CCE) estimators. Rather than obtaining explicit estimates for the unobserved common factors $\boldsymbol{f}_t$ (Bai, 2009; Bai *et al.*, 2009), the CCE estimators account for their presence implicitly by adding cross-section averages for the dependent and independent variables to the regression equation. In both the pooled and heterogeneous variants of the estimator the specification assures that coefficients on the implied common factors are allowed to differ across countries (equivalent to $\boldsymbol{\lambda}_i$ differing across i). The Pesaran (2006) estimators yield consistent and efficient estimates of the technology parameters and are robust to structural breaks in the data. Alternative estimation approaches (Bai, 2009; Bai *et al.*, 2009) involve estimation of first the number of 'relevant' factors $\boldsymbol{f}_t$ and then the factors themselves, which hinges crucially on the assumption that all factors in the data generating process (DGP) are of the 'strong' type, thus excluding 'weak factors', e.g. spatial correlations such as neighbourhood effects (Pesaran, 2009). The CCE estimators can account for the presence of strong factors as well as an *infinite* number of weak factors, while no prior knowledge of the cointegrating properties of the observables and/or the unobservables is required, since the method is robust to all scenarios studied (Kapetanios *et al.*, 2009).

Our extension to the CCE estimators investigates a number of alternative structures for the nature of cross-section correlation in the agriculture data, by applying different weight matrices before the cross-section averages are computed.[33] Prior to estimation our empirical tests are unable to reject non-stationarity and cross-section dependence in the data. Our empirical results support the specification of a common factor model in intercountry production analysis, highlight the rejection of constant returns to scale in pooled models as an artefact of empirical misspecification and suggest that agro-climatic environment, rather than neighbourhood, drives similarity in TFP evolution across countries. The latter finding provides a possible explanation for the observed failure of technology transfer from advanced countries of the temperate 'North' to arid and/or equatorial developing countries of the 'South'. Empirical tests for cointegration and direction of causation support our interpretation of the regressions as empirical production functions, rather than disguised investment or labour demand equations.

The Eberhardt and Teal (2009b) model can be nested within our general production function model in (2) assuming

(i) heterogeneous technology parameters in agricultural production across countries;
(ii) heterogeneous TFP levels and growth rates across countries, with TFP evolution not required to be stationary;
(iii) cross-section dependence; and
(iv) non-stationary input and output variable series within-sample.

This concludes our selective discussion of the growth empirics literature. We have shown that our general frameworks encompass the models employed in the seminal contributions to the growth empirics literature as well as those employed in the emerging non-stationary panel literature. The evolution we charted is one away from single cross-section regressions imposing common technology parameters and towards models allowing for parameter heterogeneity and explicit treatment of variable time-series and cross-section correlation properties. We indicated on a number of occasions under which circumstances model assumptions in the literature are likely to be violated. We now provide a more detailed discussion of these concerns, beginning with variable time-series properties in the next section and subsequently discussing cross-section dependence and parameter heterogeneity.

3. Accounting for Time-series Properties in Macro Panel Data

> In some panel data sets like the Penn-World Table, the time series components have strongly evident non-stationarity, a feature which received virtually no attention in traditional panel regression analysis. (Phillips and Moon, 2000, p. 264)

Since the seminal empirical papers on cross-country growth by Barro (1991) and Mankiw *et al.* (1992) the theory of panel data econometrics has progressed rapidly, first, in its analysis of the dynamic specification and estimation of stationary panel data (Arellano and Bond, 1991; Blundell and Bond, 1998), and more recently in its treatment of non-stationarity and cointegration in a panel setup (Pedroni, 1995, 1999, 2000; Kao, 1999; Phillips and Moon, 1999). The latter development has thus far found limited attention in the *mainstream* literature on empirical growth modelling[34] – for instance the chapter on growth econometrics in the recent *Handbook of Economic Growth* by Durlauf *et al.* (2005) contains only limited discussion of non-stationarity and cointegration with respect to panel data regression. In the following paragraphs, we discuss important time-series properties some macro data series are likely to possess and highlight the implications for estimation – our focus is on the production function framework.[35]

In the long run, variable series such as gross output or capital stock often display high levels of persistence, such that it is not unreasonable to suggest for these series to be 'non-stationary' processes in *some* countries (Nelson and Plosser, 1982; Granger, 1997; Lee *et al.*, 1997; Rapach, 2002; Bai and Ng, 2004; Pedroni, 2007; Canning and Pedroni, 2008). Non-stationarity is a property that can be viewed in simple terms as an extreme form of variable persistence over time. If we have a stationary variable, adding more observations (e.g. when more years of data become available) allows us to get a better understanding of the mean, variance, i.e. the distribution (probability density function) of the underlying process of which our data are a realization. In the case of a non-stationary variable, adding more observations does not help us to get an idea of what the distribution looks like as the mean, variance etc. do not settle at ('converge' to) constant values.

If the first difference of a variable series transforms it into a stationary process, then the untransformed series is said to be 'integrated of order one' or I(1). The order of integration indicates how many times a variable series needs to be differenced to be stationary, which implies that a series which is stationary to start with is referred to as I(0). Although economic time series in practice are usually not precisely integrated of any given order, it is for our purposes sufficient to assume that nominal and real value series typically behave as I(2) and I(1), respectively (Hendry, 1995; Jones, 1995) – note that this is a judgement derived from analysing data in a *developed country* context, but that less-developed countries may not submit to this generalization. Further, Pedroni has suggested that variable (non)stationarity (in his case for the savings rate) should not be seen as a 'global' property, valid for all times, but as a 'feature which describes local behaviour of the series within sample' (Pedroni, 2007, p. 432).

Since time-averaging does not alter the order of integration of variables (Granger, 1988; Granger and Siklos, 1995; Marcellino, 1999), any macro production function is thus likely to contain at least some countries with non-stationary input and output variables, and these time-series properties will need to be taken into account in the empirical approach to avoid bias and/or inefficiency.

Unfortunately, empirical testing of variable (non)stationarity is fraught with difficulty in panels of moderate dimension[36] and for diverse countries such as the PWT or alternative annual macro data on production (e.g. UNIDO, 2004; FAO, 2007). Single time-series unit root tests, applied to each country variable series, suffer from low power, while their panel cousins are difficult to interpret (Maddala, 1999; Smith and Fuertes, 2007): a Fisher-type test of unit roots in a panel (Maddala and Wu, 1999), for instance, investigates the null of *all* country series containing unit roots against the alternative that *at least one* series is stationary.[37] Thus while panel unit root tests have improved power over time-series unit root tests (Im *et al.*, 2003; Baltagi, 2005; Choi, 2007), the matter of their interpretation hinders a straightforward application to the data. A second generation of panel unit root tests has been introduced to account for cross-section dependence in the data (e.g. Bai and Ng, 2004; Pesaran, 2007) – an additional data property which was shown to matter greatly in the macro panel context. We will introduce this topic in detail below.

In a single time-series setup, regressing non-stationary output on non-stationary input variables in a linear model is an appropriate estimation strategy *if and only if* the regression error terms turn out to be stationary I(0), i.e. in the presence of what is termed a 'cointegrating relationship' between inputs and output. If this is not the case, we encounter a 'spurious regression': even if the variables in question were entirely unrelated, our regression results may indicate a highly significant relationship, since the standard tests for significance and goodness of fit in this case are invalid.[38] In contrast, when processes are indeed cointegrated, they define a 'long-run equilibrium trajectory' for the economy. At times the observed evolution will deviate from this path, but short-run 'error corrections' in the system will assure a return to the long-run equilibrium path (Hendry, 1995).

One could suggest that the macro production process is representative of a cointegrating relationship between output and 'some set of inputs' in the context of non-stationary variable series (Pedroni, 2007; Canning and Pedroni, 2008). This relationship could apply to all countries in the world *in the same way*, implying that all countries had the same long-run equilibrium trajectory (homogeneous cointegration). This means that the 'true' coefficients determining how inputs affect output (referred to as the DGP) would be the same for all countries. Alternatively, each country could follow a *different* long-run trajectory (heterogeneous cointegration), such that the DGP differed across countries. Naturally, if (some) country variable series are stationary the problem of non-cointegration and potential for spurious regression does not arise (for these countries).

Formulating the correct hypotheses to empirically test for cointegration in a panel data set involves severe conceptual difficulties and test results are often inconclusive, especially so in moderate N, T samples (Baltagi and Kao, 2000; Coakley *et al.*, 2001; Pedroni, 2007). In addition, it is unclear how the cointegration tests perform if variable series are non-stationary in some countries but stationary in others. Further, the theoretical literature on panel cointegration testing has only recently entered the second generation of tests (Westerlund and Edgerton, 2008) which account for cross-section dependence. An alternative testing strategy introduced by Gengenbach *et al.* (2006) defactors the data and then applies first generation cointegration tests (Pedroni, 1999, 2004a). All of these tests, however, still show comparatively low power for moderate T dimension and formally testing input and output series for cointegration in a diverse sample may leave many questions unanswered.

Thus far we have focused on time-series properties of observables (inputs, output), but the issue is also prevalent with regard to TFP, our 'measure of ignorance'. A number of empirical papers report (or assume without conceptual justification) that their measures of TFP display non-stationarity, whether analysed at the economy level (Coe and Helpman, 1995; Coe *et al.*, 1997; Kao *et al.*, 1999; Engelbrecht, 2002; Bond *et al.*, 2007) or at the sectoral level (Bernard and Jones, 1996b; Funk and Strauss, 2003). Further, Coakley *et al.* (2006) state explicitly with reference to cross-country production function estimation that technology shocks are plausibly non-stationary. A recent paper by Abdih and Joutz (2006) forms a conceptual link between TFP evolution and a 'knowledge production function', in the spirit of R&D-based endogenous growth models (Romer, 1990; Grossman and Helpman, 1991; Aghion and Howitt, 1998). They model knowledge (proxied as in Jones (1995) by 'measured' TFP[39]) as a function of patent stock, and the flow of new patents as a function of patent stock and R&D activity. Their analysis of aggregate US data from 1948 to 1997 confirms that 'inputs and output of the knowledge production function can be plausibly characterized as non-stationary and integrated of order one' (Abdih and Joutz, 2006, p. 243). Thus the non-stationarity of TFP/non-factor inputs is given a theoretical explanation by highlighting the link to R&D as postulated by parts of the endogenous growth literature. Palm and Pfann (1995, p. 691) explain the non-stationarity of TFP as 'reflecting the impact of common persistent innovations on TFP . . . [or] the impact of frequent but radical shocks'.

Misspecification of the dynamic evolution of TFP in the estimation equation leads to non-stationary errors, just as in the case of misspecification of factor parameter heterogeneity – see Section 5. A single time-series regression limited to T observations does not allow for the dynamic TFP process to be specified using an unrestricted set of $T - 1$ year dummies (Bond *et al.*, 2007), and a linear trend may not be flexible enough to capture the idiosyncrasies of non-stationary TFP evolution. The evolution of TFP in a non-stationary fashion implies that TFP becomes part of the cointegrating relationship.

A conclusion to be drawn from this discussion is that while formal testing for non-stationarity and cointegration may be a frustrating endeavour in a relatively short macro panel of diverse countries, we should nevertheless consider variable series in levels for *some* countries as non-stationary and adjust our estimation strategy accordingly. There are a number of implications: first, running separate country regressions in levels for the 'non-stationary countries' will only yield valid estimates if all elements of the underlying cointegrating relationship have been included in the empirical model; any non-stationary elements left out will enter the error terms and thus lead to the breakdown of the cointegrating relationship and potentially spurious regression results. Second, variable cointegration does not guarantee that the direction of causation is the same as that implied by our empirical framework: instead of a macro production function the data could represent an investment or labour demand equation. Given that systems approaches (panel vector autoregressions or vector error correction models) are difficult to implement in the moderate panel dimensions available for macro production data, one strategy could be to follow Canning and Pedroni (2008) in testing for the direction of causation among the variables of the cointegrating vector (see also Eberhardt and Teal, 2009b). Third, some countries will have stationary variable series, in which case the problem of non-cointegration does not arise but variable endogeneity may lead to bias in the parameter estimates.

4. Cross-section Dependence in Macro Panel Data

> When studying macroeconomic and financial data . . . , cross-sectional dependencies are likely to be the rule rather than the exception, because of strong inter-economy linkages. (Westerlund and Edgerton, 2008, p. 666)

Panel data econometrics over recent years has seen an increasing interest in models with unobserved time-varying heterogeneity induced by unobserved common shocks that affect all units (in our present interest: countries), but perhaps to a different degree (Coakley *et al.*, 2006). This type of heterogeneity introduces cross-section correlation or dependence between the regression error terms, which can lead to inconsistency and incorrect inference in standard panel econometric approaches (Phillips and Sul, 2003; Pesaran, 2006; Bai, 2009; Pesaran and Tosetti, 2009). The latter typically assume 'cross-section independence', not necessarily because this feature is particularly intuitive given real-world circumstances, but 'in part because of the difficulties characterizing and modelling cross-section dependence' (Phillips and Moon, 1999, p. 1092). In the context of cross-country productivity analysis, the presence of correlation between macro variable series across countries seems particularly salient. As an empirical illustration for his summary statistic of cross-section dependence Pesaran (2004, p. 23) uses GDP series from PWT and concludes that 'results clearly show significant evidence of cross section dependence in output innovations, that ought to be taken into account in cross country growth analysis'. Further, Durlauf and Quah (1999) discuss the possibility of cross-section dependence in a Lucas (1993) growth model with human capital spillovers. These spillovers markedly change the dynamics of convergence and the authors call for the modelling of cross-country interactions in empirical convergence analysis (Shepotylo, 2008). DeLong and Summers (1991) present a formal investigation of spatial correlation in their cross-country regression residuals and express surprise at its apparent absence in their sample.

Cross-section dependence can be addressed in the empirical specification in at least three ways (Pesaran and Tosetti, 2009): first, if particular drivers of the correlation are known, the dependence can be modelled explicitly. This is standard practice in spatial econometric models, where the strength of the correlation is determined by the location and distance of units in relation to each other. For instance models in empirical regional science and analytical geography often define spatial distance in terms of geographical proximity, but both in theory and practice distance can also be defined in terms of other variables, such as common colonial history, cultural affinity or bilateral trade volumes. To illustrate this approach, we can think of a 'spatial weight matrix', a pre-specified set of rules governing the spatial correlation, which for instance assigns a unity value to a country pair which share a common border, and a zero otherwise (Pesaran, 2004). Following standardization this weight matrix is then applied to a spatial error component or spatially lagged dependent variable in the empirical model. It is important to distinguish this impact of 'geography' on growth from that commonly discussed in the literature, most prominently in the work by Jeffrey Sachs (Sachs and Warner, 1995, 2001; Sachs, 2003). In these papers, geographical features such as proximity to the equator or the 'disease environment' have considerable power in explaining cross-country income patterns. Spatial correlation, in contrast, suggests that neighbourhood (or distance) introduces cross-section correlation which induces bias in standard OLS approaches. Its aim is to account for this dependence and provide country-specific 'counterfactual estimates' of the impact of a change in country i's x-variable on all other countries. This method incorporates assessing the impact on country i itself after accounting for spatial feedbacks in the system.

Naturally, there is a link between these two ideas since countries in close proximity will have similar disease environment and proximity to the equator; however, the empirical implementation (and interpretation) is clearly different. In practice, application of this approach is restricted by the enforced time-invariance of the spatial correlation (distance does not change over time) with much of the work thus limited to the single cross-section framework (e.g. Conley and Ligon, 2002; Ertur and Koch, 2007). Although we can point to recent developments in the theoretical literature (e.g. Kapoor *et al.*, 2007; Baltagi *et al.*, 2009; Hays *et al.*, 2009) which allow for a more flexible empirical specification, further discussion would go beyond the scope of this paper.[40]

Second, we can use fixed effects and year dummies to account for time-invariant and time-variant correlation across units. However, this assumes that the impact of the cross-section dependence is identical across all countries, which in our general model equates to identical $\boldsymbol{\lambda}_i$ across countries. The same applies to using data in deviation from the cross-section mean (Pedroni, 2000; Coakley *et al.*, 2006). In either case, violation of the homogeneity assumption leads to cross-country dependence in the error terms and thus cannot solve the problem.

A third alternative and altogether more promising approach is the adoption of a multi-factor error structure, where cross-section dependence is modelled to arise from ‘unobserved common factors’ which need to be appropriately accounted for in order to obtain unbiased estimates of the parameters on the observed regressors (Andrews, 2005; Moscone and Tosetti, 2009). We replace our general specification in (1) with a simple factor model for illustrative purposes:

$$y_{it} = \beta_i x_{it} + u_{it} \qquad u_{it} = \alpha_i + \lambda_i f_t + \varepsilon_{it} \tag{21}$$

$$x_{it} = \phi_i f_t + \psi_i g_t + v_{it} \tag{22}$$

where x is a single observed input, y is output, f_t and g_t are unobserved common factors with heterogeneous factor loadings ϕ_i and ψ_i, respectively, and ε_{it}, v_{it} are white noise. Note that we introduce g_t such that x is further driven by factors other than those driving y. The definition of x as being driven by the same factor f_t as y, albeit with a different factor loading, introduces endogeneity into the y-equation.

If we assume stationary factors f_t and g_t, the consistency of standard panel estimators such as a pooled fixed effect regression or a Pesaran and Smith (1995) mean group regression with country-specific intercepts rests on the parameter values (factor loadings) of the unobserved common factors: if their averages are jointly non-zero ($\bar{\lambda} \neq 0$ and $\bar{\phi} \neq 0$) a regression of y on x and N intercepts (in the pooled fixed effects regression case) will be subject to the omitted variable problem and hence misspecified. Regression error terms will be correlated with the regressor, leading to biased estimates and incorrect inference (Coakley *et al.*, 2006; Pesaran, 2006). In the case of non-stationary factors the consistency issues in the same setup are altogether more complex and will depend on the exact overall specification of the model (Kapetanios *et al.*, 2009). However, the latter scenario in any case will not yield an estimate of β or the mean of the β_i, as is easily shown by solving equation (22) for f_t and plugging this into equation (21):

$$
\begin{aligned}
y_{it} &= \alpha_i + \beta_i x_{it} + \lambda_i f_t + \varepsilon_{it} \\
&= \alpha_i + \beta_i x_{it} + \lambda_i \phi_i^{-1}(x_{it} - \psi_i g_t - e_{it}) + \varepsilon_{it} \\
&= \alpha_i + \underbrace{(\beta_i + \lambda_i \phi_i^{-1})}_{\varrho_i} x_{it} \underbrace{- \lambda_i \phi_i^{-1} \psi_i g_t - \lambda_i \phi_i^{-1} e_{it} + \varepsilon_{it}}_{\varsigma_{it}} \qquad (23)
\end{aligned}
$$

$$= \alpha_i + \varrho_i x_{it} + \varsigma_{it} \qquad (24)$$

From equation (24), it is clear that with standard panel estimators we can only obtain a consistent estimator of $\varrho_i = \beta_i + \lambda_i \phi_i^{-1}$ or its mean and not of β_i or its mean – β_i is unidentified (Kapetanios *et al.*, 2009). Under the specification described, a standard pooled fixed effects or Pesaran and Smith (1995) mean group estimator will therefore likely yield an inconsistent estimator of a parameter we are not interested in.

The problem laid out above in a common factor framework resembles that developed in a much earlier literature on production function estimation concerned with the endogeneity of input choices to the level of technology. This literature begins with Marschak and Andrews (1944), though in spirit Mendershausen (1938) seems to have been the first to worry about the question of identification in the production function context (see discussion in Griliches and Mairesse, 1998).

Finally, we noted above that temporal aggregation of time-series data is common practice in this literature with only a relatively small number of cross-country growth studies using annual data (Lee *et al.*, 1997; Martin and Mitra, 2002; Arnold *et al.*, 2007; Pedroni, 2007; Canning and Pedroni, 2008; Costantini and Destefanis, 2009). Lee *et al.* (1997, p. 359) suggest that both the cross-section regressions exemplified by Mankiw *et al.* (1992) and the panel regressions using period-averaged data (Islam, 1995; Caselli *et al.*, 1996) make it 'impossible to consider either the complex dynamic adjustments involved in the countries' [income] processes or the heterogeneity of [TFP] growth rates across countries'. Nevertheless, all issues arising from heterogeneity and variable time-series and cross-section properties aside, one might be concerned about the distorting impact of business cycle effects on regressions employing annual data. We can suggest that the common factor model approach is able to deal appropriately with *any* business cycle effects, whether they represent idiosyncrasies of a small number of economies, or global shocks with heterogeneous impacts: in the former case we can appeal to the Chudik *et al.* (2009) result whereby an infinite number of 'weak factors' (e.g. capturing neighbourhood effects) may be introduced to the model to account for highly idiosyncratic business cycle effects across the world. In the latter case, a 'strong factor' (i.e. of the nature we have assumed throughout) can be used to model the heterogeneous impact of a global shock.[41] Our general models are therefore intended for implementation using annual data, forcing the econometrician to deal with time-series and cross-section correlation properties explicitly.

5. Parameter Heterogeneity in the Face of Variable Non-stationarity

> What do Thailand, the Dominican Republic, Zimbabwe, Greece, and Bolivia have in common that merits their being put in the same regression? (Harberger, 1987, p. 256)

> The lesson for applied work is that if large T panels are available, the individual micro-relations should be estimated separately and the averages of the estimated micro-parameters and their standard errors calculated explicitly .... The hypothesis of homogeneity, common slope coefficients, can then be tested. Our experience is that it is almost always rejected.... (Pesaran and Smith, 1995, p. 102)

It is a common practice in the literature to pursue 'TFP extraction' adopting an empirical specification characterized by homogeneous factor parameters, coupled with heterogeneous TFP levels and evolution (Parente and Prescott, 1994, 1999; Klenow and Rodriguez-Clare, 1997; Caselli and Coleman II, 2006; Acemoglu *et al.*, 2007; Aiyar and Dalgaard, 2008). As the comments by Harberger (1987) above and Brock and Durlauf (2001) quoted earlier indicate, there is considerable unease about the standard empirical specification imposing *common* technology parameters and we have already provided a number of conceptual arguments for the heterogeneity of production technology and TFP across countries. Our general empirical specifications have gone to great lengths to be as flexible as possible regarding heterogeneity in the impact of observables (technology parameters) and unobservables (factor loadings) on output. Why this emphasis on parameter heterogeneity? Misspecification of technology parameter heterogeneity *in itself* may not be regarded as a serious problem for estimation: if slope parameters vary randomly across countries and are orthogonal to included regressors and the error terms, the pooled regression coefficient represents an unbiased estimate of the mean of the parameter across countries (Durlauf *et al.*, 2005, p. 617). Note that these are very strong assumptions, which are likely to be violated in the data (Caselli *et al.*, 1996). Nevertheless, it may be argued that variable omission and parameter heterogeneity may be interpreted as examples of the deviation of empirical growth models from a statistical 'ideal', allowing for the kind of inferences a researcher would wish to make in the growth context (Brock and Durlauf, 2001). Crucially, this viewpoint derives much of its justification from the assumption that variable series entering the empirical model are *stationary*.

Neglecting potential technology parameter heterogeneity and TFP heterogeneity in the empirical analysis, however, has more serious implications if observable and/or unobservable variables are *non-stationary*, namely, the breakdown of the cointegrating relationship between inputs and output. We can illustrate this point quite easily: a pooled estimation equation in levels imposes *common* technology parameters on all countries and thus creates non-stationary errors if 'true' technology parameters are heterogeneous and input variables are non-stationary. With reference to our general model in equation (1) pooled error terms may contain one or more of

$$(\alpha_i - \pi_L)\, L_{it} \qquad (\beta_i - \pi_K)\, K_{it} \qquad (\gamma_i - \pi_M)\, M_{it} \tag{25}$$

where π_L, π_K, π_M are the *common* regression coefficients for labour (*L*), capital stock (*K*) and material inputs (*M*), respectively (all in logarithms), while α_i, β_i, γ_i are the 'true' country-specific technology parameters. Each of the three terms in equation (25) is a linear combination of a non-stationary variable/process and thus will be non-stationary itself. The failure to account for technology parameter heterogeneity leads to the breakdown of the cointegrating relationship between inputs and output and thus produces potentially spurious results (Smith and Fuertes, 2007). Even if observed inputs and output cointegrate in each country equation (heterogeneous cointegration), the pooled equation does not and pooled estimation will not yield the mean of the cointegrating parameters across countries.

Similarly, a pooled estimation equation in levels augmented with $T - 1$ year dummies imposes *common* TFP evolution on all countries and thus creates non-stationary errors if 'true' TFP is heterogeneous and non-stationary; in our general notation

$$(\boldsymbol{\lambda}_i' \boldsymbol{f}_t - \pi_{TFP}) \tag{26}$$

where $\boldsymbol{\lambda}_i \boldsymbol{f}_t$ represents the 'true' country-specific TFP process (non-stationary) and π_{TFP} the estimated TFP evolution specified via a set of common year dummies. Note that if the 'true' TFP process is non-stationary, country-by-country regressions with linear trend terms can capture technology parameter heterogeneity but lead to a misspecified TFP evolution (stationary linear trend instead of a non-stationary evolution path) and thus result in non-stationary country-regression errors (Bai *et al.*, 2009). The same applies in a pooled regression equation with linear country trends.

A recent development in econometric theory implies that the spurious regression conclusion may need qualification. Phillips and Moon (1999, p. 1091) suggest that *pooled* regressions of level equations with I(1) errors will yield *consistent* estimates of 'interesting long-run relations' between observable input variables and output *provided N and T are large enough* (see also Kao, 1999; Phillips and Moon, 2000; Smith and Fuertes, 2007). This implies that regardless of whether we have homogeneous cointegration, heterogeneous cointegration or non-cointegration, the pooled fixed effects estimator in levels can provide a consistent point estimate of some average long-run correlation between inputs and output. Even if country regressions are spurious, the pooled panel regression is not! The result extends to the cross-section estimator investigated by Pesaran and Smith (1995) – this 'between estimator' runs a cross-section regression of time-series averages and is thus conceptually very close to the Mankiw *et al.* (1992) growth regression in equation (8). There are, however, a number of important qualifications for these striking results against the background of our general model setup. First, the asymptotic results build on the assumption of cross-section independence, which as was discussed above is seemingly unjustified in the macro panel context.

Monte Carlo simulations by Coakley *et al.* (2006) show that the mere presence of heterogeneous cross-section dependence (unobserved common factors with different factor loadings across countries) does not change the Phillips and Moon (1999) result: the pooled fixed effects estimator is unbiased even in small samples, but it is rather inefficient.[42] However, if unobserved factors drive both the input(s) and output variables, the Phillips and Moon (1999) results break down. The results continue to hold in the case of 'weak'/spatial cross-section correlation. Second, the cross-section estimator does not allow for any dynamic specification since it relies on strictly exogenous regressors (Pesaran and Smith, 1995). Abstracting from dynamics, the latter assumption on its own is already very restrictive in the cross-country growth analysis case. Third, as pointed out by Kao (1999), the conventional standard errors on estimates will be invalid, and the associated t-statistics will exaggerate the precision of the estimates. Any attempt to construct valid t-statistics requires knowledge of the true parameter values and thus cannot be implemented in empirical studies.

In Section 3 above, we have shown that there is a general consensus that macro data series such as output, GDP or capital stock should not *a priori* be considered as stationary processes for all countries analysed. We also provided arguments to suggest that TFP evolution may be best represented as a non-stationary process. In Section 4, we then argued that unobserved TFP processes can be conceptualized using common factor models, which allow for cross-country heterogeneity in the factor loadings. In the light of the discussion in this section, the assumption of parameter homogeneity, commonly adopted in the mainstream literature on growth empirics, is shown to have much more serious implications in the non-stationary than in the stationary context: any deviation from the homogeneity assumption no longer simply affects the precision of our estimate of the parameter 'mean', but will lead to the breakdown of cointegration and thus potentially spurious results. With regard to the specification of TFP evolution, we can state that the use of a linear trend may not capture the (potentially) non-stationary evolution and may equally lead to a breakdown in the cointegrating relationship, even if factor parameters were modelled correctly (Bai *et al.*, 2009).

In contrast to our suggestions, recent work by Hauk and Wacziarg (2009) investigating the performance of various panel estimators with Monte Carlo simulations seems to give support to the application of 'simple' pooled OLS estimators in cross-country growth regressions. It is necessary to put this result into perspective: the authors' simulation setups *assume away the three crucial features of modern panel time-series econometrics highlighted in this paper*, namely technology heterogeneity, variable non-stationarity and cross-section correlation.[43] The cautious conclusion from their work should therefore be that if the true DGP for cross-country growth had *common* technology parameters and *common* TFP growth, with all variables evolving in a *stationary* fashion and countries *cross-sectionally independent* from each other, then a simple pooled estimator may be best placed to tease out the correct parameter values from the data.[44]

The challenge for the econometrician is thus to accommodate the difficulties arising from parameter heterogeneity in the impact of observables (inputs) and unobservables (common factors with heterogeneous factor loadings), as well as the time-series properties of these variables and processes. Depending on the question of interest, there are a number of alternative routes available: if the primary interest lies in obtaining unbiased estimates for the input elasticities, the simple but powerful augmentation approach developed by Hashem Pesaran and co-authors (Pesaran, 2006, 2007; Chudik *et al.*, 2009; Kapetanios *et al.*, 2009; Pesaran and Tosetti, 2009) particularly lends itself to application – analysis of TFP processes can then be conducted by applying principal component analysis as an additional step. Alternatively, one may adopt an approach whereby the elements of TFP (common factors, factor loadings) are estimated 'jointly' with the input elasticities, following the methods introduced by Jushan Bai and co-authors (Bai and Ng, 2002, 2004; Bai and Kao, 2006; Bai, 2009; Bai *et al.*, 2009). It is suggested that the latter approach is dependent on the presence of strong factors, whereas the former can also accommodate weak factors as characteristic of spatial correlation (Pesaran, 2009). This field of research is developing very rapidly and it is hoped that these empirical methods will soon make their way into the mainstream growth empirics canon.

These econometric methods allow the econometrician to investigate an average relationship *across* panel members which allows for a *separate* equilibrium relationship in each cross-sectional unit but does not assume away any *common* influences and correlation between them. In the cross-country growth empirics case, as we have argued, these new methods can account for some of the puzzles introduced by more rigid empirical frameworks since the seminal work by Barro (1991) and Mankiw *et al.* (1992). In this context, it is important to stress that while the empirical models have changed the underlying research question has not; as Temple (1999, p. 126) put it: 'given that the purpose of cross-country empirical work is often to arrive at generalizations about growth, the averages [across countries] are important'. The heterogeneity introduced does not imply that each country result can be seen as *a reliable estimate or test statistic*: as Pedroni (2007, p. 440) explains, this interpretation is hazardous, since the 'long-run signals contained in [limited] years of data may be relatively weak', such that one should refrain from country-specific policy implications unless the single time-series analysis for this specific country is deemed reliable. However, previous empirical analysis averaging over individual country regressions has frequently found that while country estimates or tests were widely dispersed and at times economically implausible, averages represented very plausible results (Boyd and Smith, 2002; Baltagi *et al.*, 2003). Accounting for cross-section dependences aside, the panel dimension of empirical growth studies enables the researcher to boost the 'relatively weak' signal by looking at averages across countries. In our mind, adopting an empirical specification that allows for parameter heterogeneity across countries represents an important step towards an integrated treatment of the production technology *in its entirety*, which is able 'to explain *why* this parameter heterogeneity exists' (Durlauf *et al.*, 2001, p. 935, emphasis added).

6. Concluding Remarks

In his stimulating book *The Elusive Quest for Growth* Bill Easterly (2002) argues that after many decades of empirical work economists are still none the wiser as to what causes growth. In seeking a source for this failure many would see empirical growth regressions as the most likely guilty party. For the defence we would argue that, while not blameless, much has been learnt from cross-country regressions and the lesson of incomplete success is not to abandon the 'quest' but to seek to understand why success has been so limited. We noted in the introduction that the 'growth regression approach' has been under attack from those who think randomized experiments or country-based studies are the way forward. One of the themes of Easterly's book is that development economics has been dominated by fads and each new fad has been hailed as the solution to past problems. This survey has sought to show that one past fad, 'growth regressions', may still have life left in it. That is not to say it is the only approach; it may well in the long term turn out not to be the best one, but as we have argued, it is more informative and more flexible in the problems that it can address than its critics have allowed.

In this paper, we have provided two general empirical frameworks for cross-country growth and production analysis which encompass the various regression specifications favoured both in the mainstream and the emerging non-stationary panel literature over the past two decades. A selective review of this literature highlighted how parameter heterogeneity of the impact of observables and unobservables can be accommodated within our general modelling framework.

By placing the literature in this general framework we have been able to highlight two factors that may hold the key to some of the failures of growth regressions. The first is how central the assumption of homogeneity, of both technology and TFP, is to the results that have been derived. New econometric techniques allow us to relax these dimensions of homogeneity and the new evidence from these techniques certainly suggests that heterogeneity in both technology and TFP matter a lot. The second factor to which we have drawn attention is the assumptions empirical models make about the time-series and cross-section correlation properties of the data. The standard empirical estimators (e.g. fixed effects, difference and system GMM) not only impose homogeneous production technology, but they also implicitly assume stationary, cross-sectionally independent, variables. If we recognize the implications of these data properties for the estimators then we may well be able to reduce the gap between our models and the data they seek to explain – this, rather than banishment from the development discourse, is in our view the right verdict on growth regressions.

Acknowledgements

We are grateful to Stephen Bond, John Muellbauer, Hashem Pesaran, Jonathan Temple, Joakim Westerlund and two anonymous referees for helpful comments and suggestions. We also thank seminar/session participants at the Gorman Student Research Workshop and Productivity Workshop, University of Oxford (February 2008), the Nordic Conference for Development Economics, Stockholm

(June 2008), the Economic and Social Research Council (ESRC) Development Economics Conference, Brighton (September 2008), the International Conference for Factor Structures for Panel and Multivariate Time Series Data, Maastricht (September 2008), the CSAE Annual Conference, Oxford (March 2009), the 15th International Panel Data Conference, Bonn (July 2009), the 3rd World Conference of the Spatial Econometrics Association, Barcelona (July 2009), the Department of Economics, University of Gothenburg (September 2009), the Northeastern Universities Development Conference, Boston, and the Department of Economics, Brown University (November 2009), for their input. All remaining errors are our own. The first author gratefully acknowledges the financial support from the ESRC, Awards PTA-031-2004-00345 and PTA-026-27-2048.

Notes

1. Here and in the remainder of the paper we refer to heterogeneity in 'technology parameters' to indicate differential production function parameters on observable factor inputs across countries.
2. A number of recent papers replace the Cobb–Douglas representation with a more general constant elasticity of substitution production representation or the Christensen *et al.* (1973) translog production representation (Bernard and Jones, 1996a; Duffy and Papageorgiou, 2000). We confine the discussion to the Cobb–Douglas form as this allows us to show the importance of the econometric issues we highlight in the context of the most influential results in the empirical literature.
3. Many empirical studies of growth and development include some measure of human capital as additional input in the production function or convergence equation, following the example of Mankiw *et al.* (1992). In the present discussion, we abstract from this for two reasons. First, commonly applied schooling data (Barro and Lee, 1993, 1996, 2001) are only available in 5-year intervals, whereas we concentrate on annual data. Second, and perhaps more importantly, the accumulation of human capital is widely regarded to represent a slow-moving process, which can be captured via intercept and trend terms. However, even if human capital accumulation were to take place in a non-linear or non-stationary fashion, our common factor model setup allows us to account for this and other unobserved processes and their heterogeneous effects across countries.
4. If we assume constancy of the material–output ratio, then results are directly comparable (Söderbom and Teal, 2004). In our notation: $\beta_i^{va} = \beta_i/(1-\gamma_i)$ and in analogy for α_i^{va}.
5. The assumption of random coefficients is for convenience. Based on the findings by Pesaran and Smith (1995, p. 81, footnote 2) the coefficients could alternatively be fixed but differing across groups. See also Kapetanios *et al.* (2009, p. 6).
6. The convergence argument suggests that due to diminishing returns to capital in the Solow model, countries with low levels of per capita capital stock will have higher marginal product of capital, and thus (for similar savings rates) will grow faster than countries with an already high stock of per capita capital (Islam, 1995). The steady-state income equation in levels is defined as

$$y_t = \left(1 - e^{-\xi_i \tau}\right) y^* + e^{-\xi_i \tau} y_0$$

where y_t is GDP at time t, y_0 is the same at some original point in time 0 and y^* is steady-state level of income (all in logarithms of 'effective workers'). In the simple cross-section regressions of Mankiw *et al.* (1992) they substitute for the steady-state solution y^* and then subtract y_0 from both sides to yield equation (3).

7. In both equations (5) and (6), we allow for parameter heterogeneity across countries for generality.
8. Obviously, the time-series properties of y_{i,t_1} will have a bearing on those of $\Delta y_{i\tau} \equiv y_{i,t_2} - y_{i,t_1}$.
9. While this is clearly restrictive it is difficult to relax this assumption in a parametric framework with parameter heterogeneity across panel units. In practice, provided the time-series dimension allows for this, empirical results could be tested for parameter constancy over time by running separate regressions for an earlier and later period.
10. As the work of these authors shows the Pesaran (2006) common correlated effects estimators are robust to an *infinite* number of 'weak' factors representing spatial correlation.
11. We employ the notion of a structural model as one in which 'parameters are either "deep" or correspond to precisely defined causal effects within a coherent theoretical framework' (Durlauf *et al.*, 2005, p. 622). The important distinction is between 'structural characteristics' and initial conditions, with only steady-state effects of initial conditions implying convergence clubs.
12. This is despite the problem that with full parameter heterogeneity the conventional notion of 'β-convergence' across countries changes its meaning fundamentally as countries converge to *their own* steady-state equilibrium paths (Lee *et al.*, 1998).
13. One should note that the focus of Solow's work was to explain the growth experience *of the USA*. He thus did not postulate the application of his model to cross-country analysis (Durlauf *et al.*, 2001).
14. It is worth emphasizing that the common factor structure encompasses a specification with heterogeneous linear trends, whereby $f_t \equiv t$ and the λ_i are equivalent to the country coefficients on the trends.
15. A concern in this context is expressed by Hall (1990), who emphasizes that under imperfect competition we cannot assume that required payments to capital are simply a residual after all other factors are paid. This calls for the use of cost-based instead of revenue-based factor shares.
16. Data from the Federal Reserve Bank of Cleveland, for instance, show an average labour share of 71.7% of value-added from 1970 to 2002 for the USA (Gomme and Rupert, 2004).
17. See Temple (2005) for a discussion of dual economy models and Felipe and Fisher (2003) for more general concerns about aggregation.
18. They use the share of adult population in secondary education, thus a 'flow' variable, to proxy for human capital. Later implementations shifted to 'stock' variables such as the population average of total years of education.
19. Mankiw *et al.* (1992) do not report results for the restricted case of the convergence regression without human capital. The imposition of the restriction has a second order effect on this estimate in all other regressions they present (see Islam, 1995).
20. Given the static, cross-section nature of this empirical framework, the time-series properties of the underlying data (potential non-stationarity) would seem not to have any bearing on consistent estimation. There is a limited theoretical literature on the identification of cointegrating vectors in cross-sections when the underlying data are characterized as non-stationary (Madsen, 2005). For our present purposes we

assert that single convergence regressions are free from the impact of time-series properties of the underlying data series. Cross-section dependence induced by the presence of unobserved common factors leads to non-randomness in the regression errors if TFP growth differs across countries; the absence of endogeneity between the regressors and the error terms for identification is also questionable.

21. The coefficient on y_{i,t_1} (the lagged dependent variable) is now $-e^{-\xi_i \tau}$, rather than $1 - e^{-\xi_i \tau}$, which needs to be accounted for in the computation of $\hat{\beta}^{va}$.
22. The factor variables n and s^k are 5-year averages, starting with 1960–1964; if the output variable is $y_{i,1965}$ for y_{i,t_2}, then y_{i,t_1} is $y_{i,1960}$ and so on.
23. Islam does not directly estimate the TFP levels which are subjected to a great deal of comparison and analysis in the literature. Instead he 'backs them out' using predictions from his 'within-groups' regression. Consequently, he does not carry out any formal test of statistically significant differences across the $A^{va}_{0,i}$.
24. This can be interpreted as identical factor parameters $\boldsymbol{\lambda}_i = \boldsymbol{\lambda}$ on the unobserved common factors.
25. The parameter on the lagged dependent variable is biased downward (in absolute terms), implying a higher rate of convergence (Bond *et al.*, 2001; Durlauf *et al.*, 2005).
26. This result is due to Pesaran and Smith (1995, p. 84). The reason for this is that $E[y_{i,t_1} u_{i\tau}] \neq 0$ as the error contains $[(-e^{-\xi \tau}) - (-e^{-\xi_i \tau})] y_{i,t_1}$. All variables that are correlated with the lagged dependent variable will also be correlated with $u_{i\tau}$. If variables exist that are not correlated with the error, they are not informative and thus cannot act as instruments.
27. While this result is not investigated for statistical significance at the individual country level, they construct a number of test statistics which reject the null that agricultural TFP growth is the same as manufacturing TFP growth in favour of the alternative that the former is larger.
28. Our discussion here focuses on empirical papers that make use of the panel element of data in estimation and inference. We can point to a separate empirical growth literature employing diagnostic tests and/or estimation methods from the time-series econometric literature to individual country data series (e.g. Jones, 1995; Cellini, 1997) – see also Durlauf *et al.* (2005, Section 6.1).
29. See Pedroni (2007, pp. 435–436) for details. The working paper (Pedroni, 2004b) expands on this.
30. Deviations from linearity will result in additional terms in the residuals. Crucially, these are I(0) and merely affect the precision of the estimation approach, although if inputs and output are cointegrated as is suggested by Pedroni (2007) this effect becomes unimportant.
31. Nevertheless, in analogy to the Engle–Granger case in single time-series regressions, it may in practice be preferable to adopt a *dynamic* specification so as to avoid the distorting influence of short-run equilibrium corrections.
32. Our focus in this paper is on empirical studies of cross-country growth and production. A separate but related literature on the more narrowly defined question of income convergence across countries uses panel unit root and cointegration tests to evaluate 'the presence of statistically verifiable stylized facts [Do economies converge? Which type of convergence, if any, is observed?], without being too concerned with "deep parameters questions" (Costantini and Lupi, 2005, p. 1). Important representatives of this literature include Bernard and Durlauf (1996), Lee *et al.* (1997) and Phillips and Sul (2003, 2007), with the more recent work

accounting for heterogeneity and cross-section correlation – see also Islam (2003) and the relevant sections in Durlauf *et al.* (2005) for a detailed overview.

33. These alternative structures represent a neighbourhood effect (only countries sharing a common border affect each other), a distance effect (distance between countries is inversely related to strength of correlation) and an agro-climatic distance effect (similarity of countries in terms of their agro-climatic makeup, measured using data on the share of each country's arable land in different climatic zones (Matthews, 1983), determines the strength of correlation).
34. This is surprising given that in their important contribution to the non-stationary panel literature, Phillips and Moon (1999, p. 1057) mention the Penn World Tables in their opening paragraph and the same data set is also highlighted in the opening paragraphs of review chapters on non-stationary panel econometrics by Kao and Chiang (2000) and Baltagi (2005).
35. The same arguments apply for the static Mankiw *et al.* (1992) specification in equation (8) (see Pedroni, 2007).
36. The term 'moderate' here is to be interpreted from a time-series econometrics standpoint. Data for 50 years may not be sufficient to detect the time-series properties appropriately, and many macro data sets including observations from less-developed countries are considerably shorter.
37. Panel-based tests that provide explicit information on which country series are I(1) and which I(0) (Breuer *et al.*, 2001) were found to be highly sensitive to sample selection (Ford *et al.*, 2006).
38. Granger and Newbold (1974) illustrated this in a simulation using two independent random walks.
39. This is obtained from the US Bureau of Labor Statistics. From the parsimonious description it has to be assumed that TFP is constructed via an accounting exercise.
40. Baltagi (2009) recently suggested that macro panels seem to be more commonly subject to 'strong' rather than 'weak cross-sectional dependence', i.e. common factors rather than spatial correlation. Chudik *et al.* (2009) have pointed out that common factor model approaches such as the Pesaran (2006) CCE estimators can accommodate an infinite number of weak common factors.
41. In terms of empirical implementation the Pesaran (2006) CCE estimators were already shown to be robust to these types of data dependences.
42. Bond and Eberhardt (2009) develop and interpret the practical implications of various simulation results from the panel time-series literature for the applied economist.
43. In their study 'heterogeneity' across countries is limited to the fixed effects (TFP levels).
44. It is important to point out that Hauk and Wacziarg (2009, Sections 4.4, 5) acknowledge some of these assumptions and their implications on the findings.

References

Abdih, Y. and Joutz, F. (2006) Relating the knowledge production function to total factor productivity: an endogenous growth puzzle. *IMF Staff Papers* 53(2): 242–271.

Abramowitz, M. (1956) Resource and output trends in the United States since 1870. *American Economic Review* 46(2): 5–23.

Acemoglu, D., Antras, P. and Helpman, E. (2007) Contracts and technology adoption. *American Economic Review* 97(3): 916–943.

Aghion, P. and Howitt, P. (1998) *Endogenous Growth Theory*. Cambridge, MA: MIT Press.

Aidt, T.S. (2009) Corruption, institutions, and economic development. *Oxford Review of Economic Policy* 25(2): 271–291.

Aiyar, S. and Dalgaard, C.-J. (2008) Dual economies and international total factor productivity differences: channeling the impact from institutions, trade, and geography. *Economica* 75(300): 629–661.

Andrews, D.W.K. (2005) Cross-section regression with common shocks. *Econometrica* 73(5): 1551–1585.

Arellano, M. and Bond, S. (1991) Some tests of specification for panel data. *Review of Economic Studies* 58(2): 277–297.

Arnold, J., Bassanini, A. and Scarpetta, S. (2007) Solow or Lucas? Testing growth models using panel data from OECD countries. OECD Economics Department Working Paper 592.

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

Bai, J. (2009) Panel data models with interactive fixed effects. *Econometrica* 77(4): 1229–1279.

Bai, J. and Kao, C. (2006) On the estimation and inference of a panel cointegration model with cross-sectional dependence. In B.H. Baltagi (ed.), *Panel Data Econometrics: Theoretical Contributions and Empirical Applications*. Amsterdam: Elsevier Science.

Bai, J. and Ng, S. (2002) Determining the number of factors in approximate factor models. *Econometrica* 70(1): 191–221.

Bai, J. and Ng, S. (2004) A PANIC attack on unit roots and cointegration. *Econometrica* 72: 191–221.

Bai, J., Kao, C. and Ng, S. (2009) Panel cointegration with global stochastic trends. *Journal of Econometrics* 149(1): 82–99.

Baltagi, B.H. (2005) *Econometric Analysis of Panel Data*. New York: Wiley.

Baltagi, B.H. (2009) *Spatial Panels*. Keynote speech, Invited Session in Honour of Cheng Hsiao, 15th Conference on Panel Data, Bonn, 4 July.

Baltagi, B.H. and Kao, C. (2000) Nonstationary panels, cointegration in panels and dynamic panels. In B.H. Baltagi (ed.), *Nonstationary Panels, Cointegration in Panels and Dynamic Panels: A Survey*. Amsterdam: Elsevier.

Baltagi, B.H., Bresson, G., Griffin, J.M. and Pirotte, A. (2003) Homogeneous, heterogeneous or shrinkage estimators? Some empirical evidence from French regional gasoline consumption. *Empirical Economics* 28(4): 795–811.

Baltagi, B.H., Song, S.H. and Kwon, J.H. (2009) Testing for heteroskedasticity and spatial correlation in a random effects panel data model. *Computational Statistics and Data Analysis* 53(8): 2897–2922.

Banerjee, A.V. and Duflo, E. (2008) The experimental approach to development economics. NBER Working Papers 14467, National Bureau of Economic Research.

Banerjee, A.V. and Newman, A.F. (1993) Occupational choice and the process of development. *Journal of Political Economy* 101(2): 274–298.

Barro, R.J. (1991) Economic growth in a cross-section of countries. *Quarterly Journal of Economics* 106(2): 407–443.

Barro, R.J. and Lee, J.-W. (1993) International comparisons of educational attainment. *Journal of Monetary Economics* 32(3): 363–394.

Barro, R.J. and Lee, J.W. (1996) International measures of schooling years and schooling quality. *American Economic Review* 86(2): 218–223.

Barro, R.J. and Lee, J.-W. (2001) International data on educational attainment: updates and implications. *Oxford Economic Papers* 53(3): 541–563.

Basturk, N., Paap, R. and van Dijk, D. (2008) Structural differences in economic growth. Tinbergen Institute Discussion Paper 2008-085/4.

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

Bernard, A.B. and Jones, C.I. (1996a) Comparing apples to oranges: productivity convergence and measurement across industries and countries. *American Economic Review* 86(5): 1216–1238.

Bernard, A.B. and Jones, C.I. (1996b) Productivity across industries and countries: time series theory and evidence. *Review of Economics and Statistics* 78(1): 135–146.

Bhattacharjee, A., Castro, E. and Jensen-Butler, C. (2009) Regional variation in productivity: a study of the Danish economy. *Journal of Productivity Analysis* 31(3): 195–212.

Blundell, R. and Bond, S. (1998) Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics* 87(1): 115–143.

Blundell, R. and Bond, S. (1999) GMM estimation with persistent panel data: an application to production functions. IFS Working Papers W99/04.

Bond, S. (2002) Dynamic panel data models: a guide to micro data methods and practice. *Portuguese Economic Journal* 1(2): 141–162.

Bond, S. and Eberhardt, M. (2009) Cross-section dependence in nonstationary panel models: a novel estimator. Paper prepared for the Nordic Econometrics Meeting in Lund, Sweden, 29–31 October.

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

Bond, S., Leblebicioglu, A. and Schiantarelli, F. (2007) Capital accumulation and growth: a new look at the empirical evidence. Unpublished Working Paper.

Boyd, D. and Smith, R.P. (2002) Some econometric issues in measuring the monetary transmission mechanism, with an application to developing countries. In L. Mahadeva and P. Sinclair (eds), *Monetary Transmission in Diverse Economies* (pp. 68–99). Cambridge: Cambridge University Press.

Breuer, J.B., McNown, R. and Wallace, M.S. (2001) Misleading inferences from panel unit-root tests with an illustration from purchasing power parity. *Review of International Economics* 9(3): 482–493.

Brock, W. and Durlauf, S. (2001) Growth economics and reality. *World Bank Economic Review* 15(2): 229–272.

Caner, M. and Hansen, B.E. (2004) Instrumental variable estimation of a threshold model. *Econometric Theory* 20(5): 813–843.

Canning, D. and Pedroni, P. (2008) Infrastructure, long-run economic growth and causality tests for cointegrated panels. *Manchester School* 76(5): 504–527.

Caselli, F. and Coleman II, W.J. (2006) The world technology frontier. *American Economic Review* 96(3): 499–522.

Caselli, F., Esquivel, G. and Lefort, F. (1996) Reopening the convergence debate: a new look at cross-country growth empirics. *Journal of Economic Growth* 1(3): 363–389.

Cavalcanti, T., Mohaddes, K. and Raissi, M. (2009) Growth, development and natural resources: new evidence using a heterogeneous panel analysis. Cambridge Working Papers in Economics (CWPE) 0946.

Cellini, R. (1997) Implications of Solow's growth model in the presence of a stochastic steady state. *Journal of Macroeconomics* 19(1): 135–153.

Choi, I. (2007) Nonstationary panels. In T.C. Mills and K. Patterson (eds), *Palgrave Handbook of Econometrics* (Vol. 1). Basingstoke: Palgrave Macmillan.

Christensen, L.R., Jorgenson, D.W. and Lau, L.J. (1973) Transcendental logarithmic production frontiers. *Review of Economics and Statistics* 55(1): 28–45.

Chudik, A., Pesaran, M.H. and Tosetti, E. (2009) Weak and strong cross section dependence and estimation of large panels. Cambridge Working Papers in Economics (CWPE) 0924.

Ciccone, A. and Jarocinski, M. (2008) Determinants of economic growth: will data tell? European Central Bank Working Paper Series 852.

Clemens, M. and Bazzi, S. (2009) Blunt instruments: on establishing the causes of economic growth. Center for Global Development Working Paper 171.

Coakley, J., Fuertes, A. and Smith, R. (2001) Small sample properties of panel time-series estimators with I(1) errors. Unpublished Working Paper.

Coakley, J., Fuertes, A.M. and Smith, R. (2006) Unobserved heterogeneity in panel time series models. *Computational Statistics and Data Analysis* 50(9): 2361–2380.

Coe, D.T. and Helpman, E. (1995) International R&D spillovers. *European Economic Review* 39(5): 859–887.

Coe, D.T., Helpman, E. and Hoffmaister, A.W. (1997) North–south R&D spillovers. *Economic Journal* 107(440): 134–149.

Conley, T. and Ligon, E. (2002) Economic distance and long-run growth. *Journal of Economic Growth* 7(2): 157–187.

Costantini, M. and Destefanis, S. (2009) Cointegration analysis for cross-sectionally dependent panels: the case of regional production functions. *Economic Modelling* 26(2): 320–327.

Costantini, M. and Lupi, C. (2005) Stochastic convergence among European economies. *Economics Bulletin* 3(38): 1–17.

Crego, A., Larson, D., Butzer, R. and Mundlak, Y. (1998) A new database on investment and capital for agriculture and manufacturing. World Bank Policy Research Working Paper Series 2013.

Crespo Cuaresma, J. and Doppelhofer, G. (2007) Nonlinearities in cross-country growth regressions: a Bayesian averaging of thresholds (BAT) approach. *Journal of Macroeconomics* 29(3): 541–554.

DeLong, J.B. and Summers, L.H. (1991) Equipment investment and economic growth. *Quarterly Journal of Economics* 106(2): 445–502.

Duffy, J. and Papageorgiou, C. (2000) A cross-country empirical investigation of the aggregate production function specification. *Journal of Economic Growth* 5(1): 87–120.

Durlauf, S.N. (1993) Nonergodic economic growth. *Review of Economic Studies* 60(2): 349–366.

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

Durlauf, S.N. and Johnson, P.A. (2008) Convergence. In S.N. Durlauf and L.E. Blume (eds), *The New Palgrave Dictionary of Economics Online*, 2nd edn. Palgrave Macmillan. (Accessed March 2009.)

Durlauf, S.N. and Quah, D.T. (1999) The new empirics of economic growth. In J.B. Taylor and M. Woodford (eds), *Handbook of Macroeconomics* (Vol. 1, pp. 235–308). Amsterdam: Elsevier.

Durlauf, S.N., Kourtellos, A. and Minkin, A. (2001) The local Solow growth model. *European Economic Review* 45(4–6): 928–940.

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

Easterly, W. (2002) *The Elusive Quest for Growth – Economists' Adventures and Misadventures in the Tropics*. Cambridge, MA: MIT Press.

Eberhardt, M. and Teal, F. (2009a) Analysing heterogeneity in global production technology and TFP: the case of manufacturing. MPRA Paper 10690, University Library of Munich.

Eberhardt, M. and Teal, F. (2009b) A common factor approach to spatial heterogeneity in agricultural productivity analysis. Paper presented at the North Eastern Universities Development Conference, Tufts University, Medford, MA, 7–8 November.

Engelbrecht, H.-J. (2002) Human capital and international knowledge spillovers in TFP growth of a sample of developing countries: an exploration of alternative approaches. *Applied Economics* 34(7): 831–841.

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

FAO (2007) FAOSTAT. Online database, Rome: FAO. (Accessed July 2007.)

Felipe, J. and Fisher, F.M. (2003) Aggregation in production functions: what applied economists should know. *Metroeconomica* 54(2–3): 208–262.

Fleisher, B., Li, H. and Zhao, M.Q. (2010) Human capital, economic growth, and regional inequality in China. *Journal of Development Economics*, forthcoming

Ford, G.S., Jackson, J.D. and Kline, A.D. (2006) Misleading inferences from panel unit root tests: a comment. *Review of International Economics* 14(3): 508–511.

Foster, N. (2006) Exports, growth and threshold effects in Africa. *Journal of Development Studies* 42(6): 1056–1074.

Funk, M. and Strauss, J. (2003) Panel tests of stochastic convergence: TFP transmission within manufacturing industries. *Economics Letters* 78(3): 365–371.

Gengenbach, C., Palm, F.C. and Urbain, J.-P. (2006) Cointegration testing in panels with common factors. *Oxford Bulletin of Economics and Statistics* 68(s1): 683–719.

Glaeser, E.L., Porta, R.L., Lopez-de-Silanes, F. and Shleifer, A. (2004) Do institutions cause growth? *Journal of Economic Growth* 9(3): 271–303.

Gollin, D. (2002) Getting income shares right. *Journal of Political Economy* 110(2): 458–474.

Gomme, P. and Rupert, P. (2004) Measuring labor's share of income. Federal Reserve Bank of Cleveland, Policy Discussion Paper.

Granger, C.W.J. (1988) Aggregation of time series variables: a survey. Federal Reserve Bank of Minneapolis Discussion Paper/Institute for Empirical Macroeconomics #1.

Granger, C.W.J. (1997) On modelling the long run in applied economics. *Economic Journal* 107(440): 169–177.

Granger, C.W.J. and Newbold, P. (1974) Spurious regressions in econometrics. *Journal of Econometrics* 2(2): 111–120.

Granger, C.W.J. and Siklos, P.L. (1995) Systematic sampling, temporal aggregation, seasonal adjustment, and cointegration theory and evidence. *Journal of Econometrics* 66(1–2): 357–369.

Griliches, Z. and Mairesse, J. (1998) Production functions: the search for identification. In S. Strøm (ed.), *The Ragnar Frisch Centennial Symposium* (pp. 169–203). Cambridge: Cambridge University Press.

Grossman, G.M. and Helpman, E. (1991) *Innovation and Growth in the Global Economy*. Cambridge, MA: MIT Press.

Hall, R.E. (1990) Growth/productivity/employment. In P. Diamond (ed.), *Invariance Properties of Solow's Productivity Residual* (pp. 71–112). Cambridge, MA: MIT Press.

Hall, R.E. and Jones, C.I. (1999) Why do some countries produce so much more output per worker than others? *Quarterly Journal of Economics* 114(1): 83–116.

Hansen, B.E. (2000) Sample splitting and threshold estimation. *Econometrica* 68(3): 575–604.

Harberger, A. (1987) Comment. In S. Fischer (ed.), *NBER Macroeconomics Annual* (pp. 255–258). Cambridge, MA: MIT Press.

Hauk, W. and Wacziarg, R. (2009) A Monte Carlo study of growth regressions. *Journal of Economic Growth* 14(2): 103–147.

Hausmann, R., Rodrik, D. and Velasco, A. (2006) Growth diagnostics. In J. Stiglitz and N. Serra (eds), *The Washington Consensus Reconsidered: Towards a New Global Governance*. New York: Oxford University Press.

Hays, J.C., Kachi, A. and Franzese, R.J., Jr (2009) The m-STAR model of dynamic, endogenous interdependence and network-behavior coevolution in comparative and international political economy. Paper presented at the 80th Annual Conference of the Southern Political Science Association, New Orleans, 8–10 January.

Hendry, D. (1995) *Dynamic Econometrics*. Oxford: Oxford University Press.

Hendry, D. and Krolzig, H.-M. (2004) We ran one regression. *Oxford Bulletin of Economics and Statistics* 66(5): 799–810.

Heston, A., Summers, R. and Aten, B. (2009) Penn World Table version 6.3, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.

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

Im, K., Pesaran, M.H. and Shin, Y. (2003) Testing for unit roots in heterogeneous panels. *Journal of Econometrics* 115(1): 53–74.

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

Islam, N. (2003) What have we learnt from the convergence debate? *Journal of Economic Surveys* 17(3): 309–362.

Jones, C.I. (1995) Time series tests of endogenous growth models. *Quarterly Journal of Economics* 110(2): 495–525.

Kao, C. (1999) Spurious regression and residual-based tests for cointegration in panel data. *Journal of Econometrics* 65(1): 9–15.

Kao, C. and Chiang, M. (2000) Nonstationary panels, cointegration in panels and dynamic panels. In B.H. Baltagi (ed.), *On the Estimation and Inference of a Cointegrated Regression in Panel Data*. Amsterdam: Elsevier.

Kao, C., Chiang, M. and Chen, B. (1999) International R&D spillovers: an application of estimation and inference in panel cointegration. *Oxford Bulletin of Economics and Statistics* 61(Special Issue): 691–709.

Kapetanios, G., Pesaran, M.H. and Yamagata, T. (2009) Panels with nonstationary multifactor error structures. Unpublished Working Paper, updated version of IZA Discussion Paper 2243.

Kapoor, M., Kelejian, H.H. and Prucha, I.R. (2007) Panel data models with spatially correlated error components. *Journal of Econometrics* 140(1): 97–130.

Klenow, P.J. and Rodriguez-Clare, A. (1997) The neoclassical revival in growth economics: has it gone too far? *NBER Macroeconomics Annual* 12: 73–103.

Knight, M., Loayza, N. and Villanueva, D. (1993) Testing the neoclassical theory of economic growth: a panel data approach. *IMF Staff Papers* 40(3): 512–541.

Kourtellos, A., Stengos, T. and Tan, C.M. (2008) THRET: threshold regression with endogenous threshold variables. University of Cyprus Working Papers in Economics 3-2008.

Lee, K., Pesaran, M.H. and Smith, R. (1997) Growth and convergence in a multi-country empirical stochastic Solow model. *Journal of Applied Econometrics* 12(4): 357–392.

Lee, K., Pesaran, M.H. and Smith, R. (1998) Growth empirics: a panel data approach – a comment. *Quarterly Journal of Economics* 113(1): 319–323.

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

Loayza, N. (1994) A test of the international convergence hypothesis using panel data. World Bank Policy Research Working Paper Series 1333.

Lucas Jr, R. E. (1993) Making a miracle. *Econometrica* 61(2): 251–272.

Maddala, G.S. (1999) On the use of panel data methods with cross-country data. *Annales d'Economie et de Statistique* 55/56: 429–449.

Maddala, G.S. and Wu, S. (1999) A comparative study of unit root tests with panel data and a new simple test. *Oxford Bulletin of Economics and Statistics* 61(Special Issue): 631–652.

Madsen, E. (2005) Estimating cointegrating relations from a cross section. *Econometrics Journal* 8(3): 380–405.

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

Marcellino, M. (1999) Some consequences of temporal aggregation in empirical analysis. *Journal of Business and Economic Statistics* 17(1): 129–136.

Marschak, J., Andrews, J. and William H. (1944) Random simultaneous equations and the theory of production. *Econometrica* 12(3/4): 143–205.

Martin, W. and Mitra, D. (2002) Productivity growth and convergence in agriculture versus manufacturing. *Economic Development and Cultural Change* 49(2): 403–422.

Matthews, E. (1983) Global vegetation and land use: new high resolution databases for climate studies. *Journal of Climate and Applied Meteorology* 22: 474–487.

Mendershausen, H. (1938) On the significance of Professor Douglas' production function. *Econometrica* 6(2): 143–153.

Moral-Benito, E. (2009) Determinants of economic growth: a Bayesian panel data approach. World Bank Policy Research Working Paper, WPS 4830.

Moscone, F. and Tosetti, E. (2009) A review and comparison of tests of cross-section independence in panels. *Journal of Economic Surveys* 23(3): 528–561.

Murphy, K.M., Shleifer, A. and Vishny, R.W. (1989) Industrialization and the big push. *Journal of Political Economy* 97(5): 1003–1026.

Nelson, C.R. and Plosser, C. (1982) Trends and random walks in macroeconomic time series: some evidence and implications. *Journal of Monetary Economics* 10(2): 139–162.

Nickell, S. (1981) Biases in dynamic models with fixed effects. *Econometrica* 49(6): 1417–1426.

Paap, R., Franses, P.H. and van Dijk, D. (2005) Does Africa grow slower than Asia, Latin America and the Middle East? Evidence from a new data-based classification method. *Journal of Development Economics* 77(2): 553–570.

Palm, F.C. and Pfann, G.A. (1995) Unraveling trend and stationary components of total factor productivity. *Annales d'Economie et de Statistique* 39: 67–92.

Papageorgiou, C. and Chmelarova, V. (2005) Nonlinearities in capital-skill complementarity. *Journal of Economic Growth* 10(1): 55–86.

Parente, S.L. and Prescott, E.C. (1994) Barriers to technology adoption and development. *Journal of Political Economy* 102(2): 298–321.

Parente, S.L. and Prescott, E.C. (1999) Monopoly rights: a barrier to riches. *American Economic Review* 89(5): 1216–1233.

Pedroni, P. (1995) Panel cointegration: asymptotic and finite sample properties of pooled time-series tests with application to the PPP hypothesis. Indiana Working Papers in Economics 95-013.

Pedroni, P. (1999) Critical values for cointegration tests in heterogeneous panels with multiple regressors. *Oxford Bulletin of Economics and Statistics* 61(Special Issue): 653–670.

Pedroni, P. (2000) Fully modified OLS for heterogeneous cointegrated panels. In B.H. Baltagi (ed.), *Nonstationary Panels, Cointegration in Panels and Dynamic Panels*. Amsterdam: Elsevier.

Pedroni, P. (2004a) Panel cointegration: asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis. *Econometric Theory* 20(3): 597–625.

Pedroni, P. (2004b) Social capital, barriers to production and capital shares: implications for the importance of parameter heterogeneity from a nonstationary panel approach. Working Paper, Williams College, June.

Pedroni, P. (2007) Social capital, barriers to production and capital shares: implications for the importance of parameter heterogeneity from a nonstationary panel approach. *Journal of Applied Econometrics* 22(2): 429–451.

Pesaran, M.H. (2004) General diagnostic tests for cross section dependence in panels. IZA Discussion Paper 1240 and CESifo Working Paper 1229.

Pesaran, M.H. (2006) Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica* 74(4): 967–1012.

Pesaran, M.H. (2007) A simple panel unit root test in the presence of cross-section dependence. *Journal of Applied Econometrics* 22(2): 265–312.

Pesaran, M.H. (2009) Weak and strong cross section dependence and estimation of large panels. Keynote speech, 5th Nordic Econometric Meeting, Lund, 29th October.

Pesaran, M.H. and Smith, R. (1995) Estimating long-run relationships from dynamic heterogeneous panels. *Journal of Econometrics* 68(1): 79–113.

Pesaran, M.H. and Tosetti, E. (2009) Large panels with common factors and spatial correlations. Unpublished Working Paper, June; revised version of IZA Discussion Paper 3032.

Pesaran, M.H., Shin, Y. and Smith, R. (1999) Pooled mean group estimation of dynamic heterogeneous panels. *Journal of the American Statistical Association* 94(446): 289–326.

Phillips, P.C.B. (1995) Fully modified least squares and vector autoregression. *Econometrica* 63(5): 1023–1078.

Phillips, P.C.B. and Hansen, B.E. (1990) Statistical inference in instrumental variables regression with I(1) processes. *Review of Economic Studies* 57(1): 99–125.

Phillips, P.C.B. and Moon, H.R. (1999) Linear regression limit theory for nonstationary panel data. *Econometrica* 67(5): 1057–1112.

Phillips, P.C.B. and Moon, H.R. (2000) Nonstationary panel data analysis: an overview of some recent developments. *Econometric Reviews* 19(3): 263–286.

Phillips, P.C.B. and Sul, D. (2003) Dynamic panel estimation and homogeneity testing under cross section dependence. *Econometrics Journal* 6(1): 217–259.

Phillips, P.C.B. and Sul, D. (2007) Some empirics on economic growth under heterogeneous technology. *Journal of Macroeconomics* 29(3): 455–469.

Rapach, D.E. (2002) Are real GDP levels nonstationary? Evidence from panel data tests. *Southern Economic Journal* 68(3): 473–495.

Rodrik, D., Subramanian, A. and Trebbi, F. (2004, 06) Institutions rule: the primacy of institutions over geography and integration in economic development. *Journal of Economic Growth* 9(2): 131–165.

Romer, P.M. (1990) Endogenous technological change. *Journal of Political Economy* 98(5): S71–S102.

Sachs, J.D. (2003) Institutions don't rule: direct effects of geography on per capita income. NBER Working Paper 9490.

Sachs, J.D. and Warner, A. (1995) Economic reform and the process of global integration. *Brookings Papers on Economic Activity* 1: 1–118.

Sachs, J.D. and Warner, A.M. (2001) The curse of natural resources. *European Economic Review* 45(4–6): 827–838.

Sala-i-Martin, X., Doppelhofer, G. and Miller, R.I. (2004) Determinants of long-term growth: a Bayesian averaging of classical estimates (BACE) approach. *American Economic Review* 94(4): 813–835.

Shepotylo, O. (2008) Spatial HAC estimator: analysis of convergence of European regions. National University of Kyiv-Mohyla Academy, Kyiv School of Economics (KSE), December.

Smith, R.P. and Fuertes, A.-M. (2007) Panel time series. Mimeo, April, Centre for Microdata Methods and Practice (cemmap).
Söderbom, M. and Teal, F. (2004) Size and efficiency in African manufacturing firms: evidence from firm-level panel data. *Journal of Development Economics* 73(1): 369–394.
Solow, R.M. (1956) A contribution to the theory of economic growth. *Quarterly Journal of Economics* 70(1): 65–94.
Solow, R.M. (1957) Technical change and the aggregate production function. *Review of Economics and Statistics* 39(3): 312–320.
Solow, R.M. (1970) *Growth Theory: An Exposition*. Oxford: Oxford University Press.
Solow, R.M. (1986) Unemployment: getting the questions right. *Economica* 53(210): S23–S34.
Swan, T.W. (1956) Economic growth and capital accumulation. *Economic Record* 32(2): 334–361.
Temple, J. (1999) The new growth evidence. *Journal of Economic Literature* 37(1): 112–156.
Temple, J. (2005) Dual economy models: a primer for growth economists. *The Manchester School* 73(4): 435–478.
UN (2004) *National Accounts Statistics: Main Aggregates and Detailed Tables, 1992*. New York: UN Publications Division.
UNIDO (2004) *UNIDO Industrial Statistics 2004*. Online database, Vienna: UNIDO.
Westerlund, J. and Edgerton, D.L. (2008) A simple test for cointegration in dependent panels with structural breaks. *Oxford Bulletin of Economics and Statistics* 70(5): 665–704.