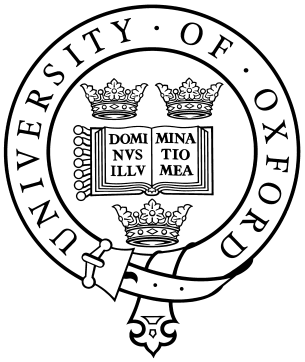


MODELLING TECHNOLOGY IN AGRICULTURE AND MANUFACTURING USING CROSS-COUNTRY PANEL DATA

Markus Eberhardt

St John's College

University of Oxford

A thesis submitted for the

degree of

doctor of philosophy

at the university of oxford

hilary term 2009

für meine Eltern

MODELLING TECHNOLOGY IN AGRICULTURE AND MANUFACTURING USING CROSS-COUNTRY PANEL DATA

Markus Eberhardt
St John's College, University of Oxford
Doctor of Philosophy in Economics
Hilary Term 2009
Supervisors: Dr. Francis Teal, Prof. Stephen Bond

Thesis Abstract

Why do we observe such dramatic differences in labour productivity across countries in the macro data? This thesis argues that the growth empirics literature oversimplifies the complexity of the production process across countries and neglects data cross-section and time-series properties, leading to bias in the empirical estimates.

Chapter 1 presents two general empirical frameworks for cross-country productivity analysis and demonstrates that they encompass the growth empirics literature of the past decades. We introduce our central argument of cross-country heterogeneity in the impact of observables and unobservables on output and develop this against the background of the pertinent time-series and cross-section properties of macro panel data.

Chapter 2 uses data from 48 countries to estimate manufacturing production functions. We discuss standard and novel estimators, focusing on their treatment of parameter heterogeneity and data time-series and cross-section properties. We develop the Augmented Mean Group (AMG) estimator and show its similarity to the Pesaran (2006) Common Correlated Effects (CCE) approach. Our results confirm parameter heterogeneity across countries in the impact of observable inputs on output. We check the robustness of this finding and highlight its implications for empirical measures of TFP.

Chapter 3 investigates the heterogeneity of agricultural production technology using data for 128 countries. We develop an extension to the CCE estimators which allows us to suggest that TFP is structured such that countries with similar agro-climatic environment are influenced by the same unobserved factors. This finding offers a possible explanation for the failure of technology-transfer from advanced countries of the temperate 'North' to developing countries of the arid/equatorial 'South'.

Our Monte Carlo simulations in Chapter 4 investigate the performance of the AMG, CCE and standard (micro-)panel estimators. Failure to account for cross-section dependence is shown to result in serious distortion of the empirical estimates. We highlight scenarios in which the AMG is biased and offer simple remedies.

Acknowledgements

I first and foremost want to thank my two thesis supervisors, Francis Teal and Steve Bond. Francis has been extremely supportive and helpful during the course of my time in the Economics Department, and Steve's comments and suggestions reassured me that the empirical strategy pursued in this thesis is both meaningful and sound. The opportunity to pursue postgraduate studies at Oxford did not arise out of nowhere, and I am particularly indebted to Vegard Iversen, John Thoburn and Andrew Millington for their tremendous support and advice during my previous work and studies. Financial support from the UK Economic and Social Research Council and St John's College is gratefully acknowledged.

I would also like to thank my fellow DPhil students and friends in the Department of Economics who have been a great collective resource for issues related and unrelated to economics, development and econometrics, in particular Simon Baptist, Jenny Castle, Mark Koyama, Courtney Monk, James Reade, Nicolas van de Sijpe and Bruno Versailles. Outside the Economics Department I want to thank a number of people who have endured my passion for 'economometrics' and despite this have become close friends: James Briaris, Betsy Cooper, Neli Demireva, Mark Faulkner, Victoria Gardner, Alan & Melinda Gamlen, Jenny Morrissey, Liam McNamara, Ashby Monk and Filiz Öztanriverdi.

Finally, I particularly wish to thank my family and my girlfriend Sarah for their love, support and encouragement.

Word Count

The main body of this thesis contains 210 pages, with a typical page (p.144) containing around 393 words, making the length of the document approximately 83,000 words (excluding references).

Contents

Table of Contents i
List of Tables iii
List of Figures iv

1 Econometrics for Grumblers: An Interpretation of the Literature on Cross-Country Growth Empirics 1
1.1 A general empirical framework for production analysis with cross-country panel data 6
1.1.1 An encompassing framework 6
1.1.2 A note on factor shares and production function parameters 13
1.1.3 Complete parameter homogeneity 14
1.1.4 Heterogeneous TFP levels 17
1.1.5 Heterogeneous TFP levels and growth rates 21
1.1.6 Full parameter heterogeneity in a stationary variable model 23
1.1.7 Full parameter heterogeneity in a model with nonstationary inputs . . . 24
1.2 Accounting for time-series properties in macro panel data 28
1.3 Cross-section dependence in macro panel data 33
1.4 Parameter heterogeneity in the face of nonstationarity 36
1.5 Investigating growth and development at the sectoral level 39
1.6 Concluding remarks 42
1.A 'Finding productivity': empirical extraction of TFP 44
1.B A preoccupation with value-added based models 48

2 Analysing Heterogeneity in Global Production Technology and TFP: The Case of Manufacturing 50
2.1 Empirical implementation of a production function for cross-country analysis . 53
2.1.1 Data time-series and cross-section correlation properties 55
2.1.2 Empirical estimators 57
2.1.3 Summary 70
2.2 Data and main empirical results 70
2.2.1 Data 70
2.2.2 Time-series properties of the data 71
2.2.3 Pooled regressions 72
2.2.4 Common TFP 76
2.2.5 Country regressions 77
2.2.6 The importance of constant returns to scale 83
2.2.7 Residual diagnostics 85
2.2.8 Discussion of main empirical results 89
2.3 Testing parameter heterogeneity 91
2.4 Estimating TFP 97

2.4.1 Deriving TFP via regression estimates . . . 98
2.4.2 Using accounting strategies for TFP determination . . . 100
2.5 Reverse causality . . . 101
2.6 Concluding Remarks . . . 103
2.A Data construction and descriptives . . . 106
2.B Time-series properties of the data . . . 109
2.C The Phillips & Moon (1999) result . . . 114
2.D Robustness check: dynamic specification . . . 115
2.E Additional tables, graphs and figures . . . 120

3 Modelling Spatial Heterogeneity in Agricultural Productivity Analysis 126
3.1 Literature review . . . 129
3.2 Technology Heterogeneity and the Analysis of Agricultural Production . . . 134
3.2.1 Heterogeneity in the Agro-Climatic Environment . . . 134
3.2.2 The Case for Technology Heterogeneity and Cross-section Dependence in the Data . . . 136
3.3 Empirical Model and Implementation . . . 139
3.3.1 Common factor model representation . . . 139
3.3.2 Empirical estimators . . . 140
3.3.3 Data . . . 142
3.4 Empirical Results . . . 142
3.4.1 Time-Series Properties and Cross-Section Dependence . . . 142
3.4.2 Pooled estimation results . . . 143
3.4.3 Averaged country regression estimates . . . 148
3.5 Interpretation and Conclusions . . . 154
3.A Data construction and descriptives . . . 157
3.B Time-series properties of the data . . . 162
3.C Cross-section dependence in the data . . . 167
3.D Additional tables and figures . . . 174

4 Estimation and Inference in Nonstationary Heterogeneous Panels with Multi-Factor Error Structure: New Insights 175
4.1 Two recent Monte Carlo studies on multifactor models . . . 178
4.1.1 Coakley, Fuertes and Smith (2006) . . . 178
4.1.2 Kapetanios, Pesaran and Yamagata (2008) . . . 182
4.2 New simulations . . . 184
4.2.1 Our Monte Carlo setup . . . 185
4.2.2 Results . . . 186
4.3 Interpretation and concluding remarks . . . 188

Summary and Conclusion 206

Bibliography 211

List of Tables

1.1 Share of Agriculture in GDP (in %) . . . 41

2.1 Pooled regressions (unrestricted returns to scale) . . . 74
2.2 Pooled regressions (CRS) . . . 75
2.3 Country regression averages (levels, CRS imposed) . . . 81
2.4 Country regression averages (FD, CRS imposed) . . . 82
2.5 Country regressions (unrestricted returns to scale) . . . 84
2.6 Testing Residuals for Normality and Homoskedasticity . . . 86
2.7 Testing Residuals for Stationarity (i) . . . 87
2.8 Testing Residuals for Stationarity (ii) . . . 88
2.9 Parameter Heterogeneity — Pooled Tests (levels) . . . 92
2.10 Parameter Heterogeneity — Pooled Tests (FD) . . . 92
2.11 Parameter Heterogeneity — Swamy (1970) Tests . . . 93
2.12 Parameter Heterogeneity — Wald Tests (levels and FD) . . . 95
2.13 Parameter Heterogeneity — F-Tests . . . 96
2.14 Country regressions using FMOLS . . . 102
2.15 Descriptive statistics . . . 107
2.16 Sample of countries and number of observations . . . 108
2.17 Naïve AR regressions . . . 110
2.18 Time-series unit root tests — rejection frequency . . . 111
2.19 First generation panel unit root tests . . . 112
2.20 Second generation panel unit root tests . . . 113
2.21 Pooled regressions (Error Correction Model) . . . 117
2.22 Country regression averages (dynamic model, CRS imposed) . . . 119
2.23 Country regressions — data in deviation from x-section mean . . . 120
2.24 Country regressions (FD, unrestricted returns to scale) . . . 121
2.25 Country rankings by TFP-level . . . 122

3.1 Selective Review of the Literature . . . 133
3.2 Pooled regressions (unrestricted returns to scale) . . . 146
3.3 Pooled regressions (CRS imposed) . . . 147
3.4 Mean Group type estimators (no restrictions on returns to scale) . . . 150
3.5 Mean Group type estimators (CRS imposed) . . . 151
3.6 Dynamic Specification — MG-type estimators (CRS imposed) . . . 153
3.7 Descriptive statistics . . . 159
3.8 Sample of countries and number of observations . . . 160
3.9 Country groupings . . . 161
3.10 Climate Zones following Köppen-Geiger . . . 161
3.11 Time-series unit root tests — rejection frequency . . . 162
3.12 First generation panel unit root tests: Fisher tests . . . 164

3.13 Second generation panel unit root tests . . . 166
3.14 'Global' Cross-section Dependence (i) . . . 168
3.15 'Global' Cross-section Dependence (ii) . . . 169
3.16 Regional Cross-section Dependence (i) . . . 171
3.17 Regional Cross-section Dependence (ii) . . . 172
3.18 Climate-zone Cross-section Dependence . . . 173
3.19 Dynamic specification — Pooled regressions (CRS imposed) . . . 174

4.1 Coakley, Fuertes and Smith (2006) . . . 191
4.2 Kapetanios, Pesaran and Yamagata (2008) . . . 192
4.3 New Simulations — (i) Baseline setup . . . 194
4.4 New Simulations — (ii) Additional country trend . . . 196
4.5 New Simulations — (iii) Feedback setup . . . 198
4.6 New Simulations — (iii)* Feedback and country trend . . . 200
4.7 New Simulations — (iv) Two 'clubs' for β . . . 202
4.8 New Simulations — (iv)* Two 'clubs', country trends . . . 204

List of Figures

2.1 Evolution of the 'Common dynamic process' $\hat{\mu}_t^{\bullet}$ . . . 77
2.2 Regression intercepts and TFP level estimates . . . 98
2.3 Kernel densities for technology parameter estimates (gross-output) . . . 123
2.4 Kernel densities for technology parameter estimates (value-added) . . . 124
2.5 TFP levels — adjusted from AMG(ii) and CCE(i) estimates* . . . 125

3.1 Geographical regions and climate zones — weighted means . . . 135
3.2 Agro-climatic 'distance' — the view from Kenya . . . 136

Chapter 1

Econometrics for Grumblers: An Interpretation of the Literature on Cross-Country Growth Empirics

Abstract:

Since the seminal contribution of Gregory Mankiw, David Romer and David Weil (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' (Abramowitz, 1956). The central theme of this thesis is an investigation of this choice of specification against the background of pertinent data properties when the units of observations are countries or regions and the time-series dimension of the data becomes substantial. In this first introductory chapter we present two general empirical frameworks for cross-country productivity analysis and demonstrate that they encompass the 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 is important for reliable empirical analysis. This idea is developed against the background of the pertinent time-series and cross-section properties of macro panel data.

> *"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, Kourtellos, and Minkin (2001, p.929)

Robert Solow's criticism of empirical cross-country growth analysis rings as true today as it did in 1986, despite the explosion of empirical papers on the topic in the 1990s following the seminal contributions by Barro (1991) and Mankiw, Romer, and Weil (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 summarise a sample of economies at very different stages of industrial development and made up of diverse industrial sectors.[1] 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.

We believe that there are a number of important reasons why one might want to side with Solow and reject the standard cross-country growth regression framework and its panel cousins. Rather than to merely 'grumble' about parameter heterogeneity in the observables and unobservables, we show that this property has very important implications for estimation and inference once the pertinent time-series properties of the data (nonstationarity, cointegration) are recognised. Intuitively, the heterogeneity in production technology could be taken to mean that countries can chose an 'appropriate' production technology from a menu of feasible options. Cross-country heterogeneity in unobservables (TFP) relates to both differences in the underlying processes that make up TFP and in the impact of these processes on output.

[1] Here and in the remainder of the thesis we refer to heterogeneity in 'technology parameters' to indicate differential production function parameters on observable factor-inputs across countries.

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 parts neglects technology-parameter heterogeneity across countries as well as dynamics. Instead, the *mainstream* literature favours ever more sophisticated statistical devices — most recently Bayesian Model Averaging and 'general-to-specific' automatic model selection algorithms — 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, Johnson, & Temple, 2005) and most were found to matter...

A small number of papers however question this paradigm and have integrated considerations of parameter heterogeneity and variable nonstationarity into their cross-country empirics. Their regression results are encouraging and further diagnostic tests for variable nonstationarity and parameter heterogeneity confirm both of these properties in a formal manner (Pedroni, 2007; Canning & Pedroni, 2008). Our critique in this initial chapter is in the same spirit but further highlights the importance of cross-section dependence in cross-country productivity analysis. We will make use of recent theoretical 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 & Ng, 2004; Pesaran, 2006, 2007; Kapetanios, Pesaran, & Yamagata, 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.

Nonstationarity, TFP heterogeneity (including the potential for cross-section dependence), and technology heterogeneity — these three themes thus form the theoretical bedrock of this thesis, which is laid out in the following sections of this chapter. The issues surrounding the time-series and cross-section correlation properties of macro panel data in their impact on model specification have previously not been considered in great detail in the *empirical* growth literature (Temple, 1999; Durlauf et al., 2005), but have solid foundations in the *theoretical* literatures

on growth and econometrics respectively. As our discussion will develop, we believe that they should be recognised as integral parts of the growth empirics canon. Existing empirical work instead has worried about the potential endogeneity of regressors in the empirical framework, an issue which is given considerably more attention in the literature than the potential misspecification of the empirical regression model.

Our focus in this chapter 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 in Chapter 1.

The two empirical implementations for manufacturing and agriculture in this thesis (Chapters 2 and 3) have considerable overlap in terms of econometric theory and implementation, but go on to focus on different issues in the empirical analysis. In part this reflects the data properties in terms of data availability, but also the outcome of diagnostic testing. For instance, our manufacturing dataset is highly unbalanced and furthermore has missing observations, thus hampering a formal analysis of the cross-section dependence properties. We begin Chapter 2 by discussing the properties of standard panel estimators in the light of our general empirical framework and introduce a novel two-stage regression procedure — the Augmented Mean Group (AMG) estimator — which accounts for parameter heterogeneity and cross-section dependence. Following the estimation of production functions and residual diagnostics for manufacturing production in 48 countries our focus is then on providing formal evidence for cross-country parameter heterogeneity. Our main results are vindicated when we adopt a non-stationary panel econometric approach allowing for reverse causality, which thus plays a minor role in the consistent estimation of cross-country production functions for manufacturing. We also develop and investigate the difficulties of obtaining meaningful TFP-level estimates in heterogeneous parameter models and suggest a solution to this problem.

The agricultural production functions in Chapter 3 are based on panel data for 128 countries. Our review of the literature indicates that despite the desirability of modelling agricultural production as a heterogeneous process across countries the existing empirical approaches have narrowly focused on pooled estimation of static models and largely neglected the agro-climatic dimension of production in this sector. We address these issues and introduce a simple ex-

tension to the Pesaran (2006) Common Correlated Effects estimators which imposes more structure on the unobservable heterogeneity across countries. Empirical results support our general framework and indicate that agricultural TFP is structured along agro-climatic lines, thus providing a possible explanation for the anecdotal failure of agricultural technology transfer from the 'North' to the 'South'.

Chapter 4 represents a formal econometric analysis of the properties of standard and novel panel estimators in nonstationary panel data with a multi-factor error structure. The focal point here is the performance of our AMG estimator in comparison to the conceptually similar Pesaran (2006) CCE estimators: since these two approaches seem to agree in the manufacturing but disagree in the agriculture data we carry out Monte Carlo simulations to establish the small sample properties of the AMG and to identify the scenarios under which its estimates are biased. This exercise further reveals that the Monte Carlo setups adopted in the recent literature on this topic are unable to create similar magnitudes of differences in the results between the various estimators tested as those witnessed in real world data: the most serious bias in pooled panel estimators highlighted in these studies is wiped out when these models are augmented with year dummies — a standard practice in the applied literature. Our own simulation setups will go some way to expand our understanding of this point.

The conclusion summarises and interprets the overall findings, highlights the contributions to the literature made by this thesis and points to avenues for future research.

The remainder of this first chapter is organised 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 ('Barro') regression equation. These are shown to encompass a wide variety of modelling approaches representing the full evolution of growth empirics over the past two decades. In Sections 1.2, 1.3 and 1.4 we discuss the central issues of parameter heterogeneity, variable nonstationarity 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. Section 1.5 briefly motivates our empirical focus on sectoral, rather than aggregate economy data, followed by a brief conclusion which summarises the arguments.

The analysis in this thesis focuses on *regression* approaches to cross-country productivity analysis, however we also briefly discuss the separate empirical field of growth and development accounting in Appendix 1.A. We furthermore examine how econometric analysis at the sectoral level, in particular within manufacturing, may be severely affected by the choice between value-added and gross-output based models (Appendix 1.B) — an issue which turned out to be of minor importance at the level of aggregation of our empirical analysis.

1.1 A general empirical framework for production analysis with cross-country panel data

This section sets out general empirical frameworks for the macro production function and the convergence equation from the canonical Solow model based on a Cobb-Douglas specification.[2] It motivates their specification and shows that they encompass the most important models of the empirical growth literature in the past two decades. We set out the general production function and convergence equation models below as a means to build an encompassing framework for the literature which affords maximum flexibility with regard to specification. The question as to *how to empirically implement these general models* is a separate issue and will be tackled in the empirical chapters of this thesis. In our review of the growth empirics literature we discuss the restrictions on the general models made in each of the seminal contributions and set out their implications for estimation and inference.

1.1.1 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

$$O_{it} = \alpha_i L_{it} + \beta_i K_{it} + \gamma_i M_{it} + u_{it} \qquad u_{it} = A_{0,i} + \boldsymbol{\lambda}_i' \boldsymbol{f}_t + \varepsilon_{it} \tag{1.1}$$

for $i = 1, \ldots, N$ and $t = 1, \ldots, T$. O represents gross output, L labour force, K capital stock and M material inputs (all in logarithms). These represent the observable variables of the

[2] A number of recent papers replace the Cobb-Douglas representation with a more general CES production representation or the Christensen, Jorgenson, and Lau (1973) translog production representation (Bernard & Jones, 1996a; Duffy & 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 emprical literature.

model.[3] For TFP we employ the combination of a country-specific TFP level $A_{0,i}$ (in logarithm) and a set of common factors $\boldsymbol{f}_t$ with country-specific factor loadings $\boldsymbol{\lambda}_i$ — TFP is unobserved. We can think of the above framework as representing a single pooled equation or N country equations.

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

$$Y_{it} = \alpha_i^{va} L_{it} + \beta_i^{va} K_{it} + u_{it} \qquad u_{it} = A_{0,i}^{va} + \boldsymbol{\lambda}_i' \boldsymbol{f}_t + \varepsilon_{it} \tag{1.2}$$

Our notation in (1.2) 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 production function model and the data it is applied to:

A.1 The parameters α_i, β_i and γ_i are unknown random coefficients with fixed means and finite variances. Similarly 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 $\varepsilon_{it} \sim N(0, \sigma^2)$, where σ^2 is finite.

A.3 Observable inputs $\mathbf{x}_{it} = \{L_{it}, K_{it}, M_{it}\}$ and output O_{it} (Y_{it}), as well as the unobserved common factors $\boldsymbol{f}_t$ are <u>not</u> *a priori* assumed to be stationary variables/processes.

[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 reason: firstly, commonly applied schooling data (Barro & Lee, 1993, 1996, 2001) is only available in 5-year intervals and at the aggregate economy level, whereas we concentrate on annual data and (in our empirics in the later chapters) on sectoral production. Secondly, and perhaps more importantly, the accumulation of human capital is widely regarded to represent a slow-moving process. However, even if human capital accumulation were to take place in a non-linear or nonstationary fashion, our common factor model setup allows us to account for this and other unobserved processes and their heterogeneous impacts across countries.

[4] If we assume constancy of the material-output ratio, then results are directly comparable (Söderbom & 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, footnote 2, p.81) the coefficients could alternatively be fixed but differing across groups. See also Kapetanios et al. (2009, p.6).

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 equation let

$$\begin{aligned} y_{i,t_2} - y_{i,t_1} &= -\left(1 - e^{-\xi_i \tau}\right) \left(\frac{\beta_i^{va}}{1-\beta_i^{va}}\right) \ln(\delta + \bar{n}_{i\tau} + \mu^\star) \\ &\quad + \left(1 - e^{-\xi_i \tau}\right) \left(\frac{\beta_i^{va}}{1-\beta_i^{va}}\right) \bar{s}_{i\tau}^k \\ &\quad - \left(1 - e^{-\xi_i \tau}\right) y_{i,t_1} + u_{i\tau} \end{aligned} \tag{1.3}$$

$$u_{i\tau} = \left(1 - e^{-\xi_i \tau}\right) A_{0,i}^{va} + \left(t_2 - e^{-\xi_i \tau} t_1\right) \boldsymbol{\lambda}_i' \boldsymbol{f}_\tau + \varepsilon_{i\tau} \tag{1.4}$$

where $\tau = t_2 - t_1$ (more on this notation below). $\bar{n}$ is the average population growth rate over period τ and $\bar{s}^k$ is the average savings rate over the same period. δ is the depreciation rate for physical capital and ξ indicates the speed of convergence to either a common or country-specific steady-state equilibrium, depending on specification.[6]

$\bar{n}$, $\bar{s}^k$ and y are the observable variables in this framework, whereas TFP ($A_{0,i}$ and $\boldsymbol{\lambda}_i' \boldsymbol{f}_\tau$) is unobserved. This framework 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 as well as 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.

[6] In the growth literature the convergence parameter is commonly expressed as λ, however we follow the convention in the common factor model literature where the unobserved factor loadings are marked by this Greek letter and specify the speed of convergence parameter as ξ.

Following estimation, the coefficients on the observables can be transformed to yield the 'implied' capital coefficient β_i^{va}. As in the production function framework, this β_i^{va} is the technology parameter in a Cobb-Douglas production function that frames the production analysis. Although the literature refers to equations like (1.3) as 'growth regressions', this equation in essence is a levels regression where y_{i,t_1} has been subtracted from both sides to investigate out-of-steady-state behaviour, although this approximation is only valid *in proximity* to the steady-state.[7] The model can be easily rewritten as an equivalent dynamic panel equation with variables in levels on both sides (see Islam, 1995; Bond, 2002).

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 $t_2 = 1985$ and $t_1 = 1960$, thus $\tau = 26$, then (1.3) transforms into

$$\begin{aligned} y_{i,1985} - y_{i,1960} \quad = \quad & -\left(1 - e^{-\xi_i \tau}\right) \left(\frac{\beta_i^{va}}{1 - \beta_i^{va}}\right) \ln(\delta + \bar{n}_i + \mu^{\star}) \\ & + \left(1 - e^{-\xi_i \tau}\right) \left(\frac{\beta_i^{va}}{1 - \beta_i^{va}}\right) \bar{s}_i^k \\ & - \left(1 - e^{-\xi_i \tau}\right) y_{i,1960} + \left(t_2 - e^{-\xi_i \tau} t_1\right) a^{va} + \varepsilon_i \end{aligned} \tag{1.5}$$

where TFP levels $A_{i,0}$ and evolution $\boldsymbol{\lambda}_i' \boldsymbol{f}_\tau$ have been subsumed into a single intercept term a^{va}. Equation (1.5) 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 *as well as* constant TFP evolution $\boldsymbol{\lambda}_i' \boldsymbol{f}_t$.[8] Our discussion in section 1.1.3 will highlight the important assumptions implicit in this specification of TFP.

[7] 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) the authors substitute for the steady-state solution y^* and then subtract y_0 from both sides to yield equation 1.3.

[8] As noted above for the inclusion of TFP evolution in $(\delta + \bar{n}_i + \mu^{\star})$ a value of .05 for $(\delta + \mu^{\star})$ is commonly applied (Islam, 1995).

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

$$\begin{aligned} \Delta y_{it} &= -\left(1 - e^{-\xi_i}\right)\left(\frac{\beta_i^{va}}{1-\beta_i^{va}}\right) \ln(\delta + n_{it} + \mu^\star) \\ &\quad + \left(1 - e^{-\xi_i}\right)\left(\frac{\beta_i^{va}}{1-\beta_i^{va}}\right) s_{it}^k \\ &\quad - \left(1 - e^{-\xi_i}\right) y_{i,t-1} + u_{it} \end{aligned} \tag{1.6}$$

$$u_{it} = \left(1 - e^{-\xi_i}\right) A_{i,0}^{va} + \left(t_2 - e^{-\xi_i} t_1\right) \boldsymbol{\lambda}_i' \boldsymbol{f}_t + \varepsilon_{it} \tag{1.7}$$

where $A_{i,0}^{va}$ and $\boldsymbol{\lambda}_i' \boldsymbol{f}_t$ capture the unobservables.

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 (1.3) and the data:

B.1 The parameters β_i^{va} and xi_i are unknown random coefficients[9] with individual means and finite variances. Similarly for the unknown factor loadings $\boldsymbol{\lambda}_i$.

B.2 Error terms $\varepsilon_{i\tau} \sim N(0, \sigma^2)$, where σ^2 is finite.

B.3 We allow for the possibility that the observable inputs ($s_{i\tau}^k$, $n_{i\tau}$), per capita GDP (y_{i,t_1}) and the unobserved factors driving TFP evolution ($\boldsymbol{f}_t$) are nonstationary.[10] The consequences of this property will be investigated.

B.4 The specification of TFP levels and their evolution paths is equivalent to an unobserved common factor model which allows for common and idiosyncratic elements of TFP.

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 (factor-inputs) and unobservables (TFP) on output. The common factor model specification for the error terms u_{it}, equation (1.1) and (1.4), operationalises 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

[9] In both equations (1.5) and (1.6) we allow for parameter heterogeneity across countries for generality.
[10] Obviously, the time-series properties of y_{i,t_1} will have a bearing on those of $\Delta y_{i\tau} \equiv y_{i,t_2} - y_{i,t_1}$.

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. 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 $\boldsymbol{\lambda}_i$. Factor-input variables and output, as well as unobserved factors may be nonstationary. This allows for a number of potential cases: *firstly*, factor-inputs and output are nonstationary and cointegrated; *secondly* factor-inputs, output and (some) common factors are nonstationary and cointegrated; and thirdly, neither of the above (noncointegration). Whether cointegrating vectors are homogeneous or heterogeneous across countries depends on the nature of the factor-input parameters (α, β, γ). Thus our empirical frameworks provide maximum flexibility with regards 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 'structural regression models': in a structural regression equation the specification arises out of the steady-state solutions of a theory model, as exemplified by the human capital-augmented Solow model of Mankiw et al. (1992). In a structural model the regressors are assumed exogeneous by theoretical construction. Many empirical papers have however entered additional covariates into the convergence specification of Mankiw et al. (1992) without deriving their impact in the solution to an economic theory model. These 'reduced form' regressions are also referred to as 'Barro regressions', following the seminal contribution by Robert Barro (1991). Typically a range of instrumentation strategies are employed for Barro regressions to deal with potential variable endogeneity. Note that our common factor model setup allows for variable endogeneity which will need to be addressed in the estimation approach.

In *economic* terms, the above frameworks in equations (1.1)-(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, p.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 for 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 papers lead to multiple equilibria interpretable as factor parameter heterogeneity in the production function (e.g. Murphy, Shleifer, & Vishny, 1989; Durlauf, 1993; Banerjee & Newman, 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 empirical implementation of parameter heterogeneity is largely dominated by the empirical convergence literature.[11] Originally technology parameters were assumed group-specific rather than fully heterogeneous, e.g. Durlauf and Johnson (1995), Caselli, Esquivel, and Lefort (1996) or Liu and Stengos (1999). Durlauf et al. (2001) then allowed for 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. Thus the 'new growth theory' literature provides the foundations for the assumption of heterogeneous technology parameters in production. Existing empirical implementations have provided some indicative evidence for this and we therefore adopt this flexible approach for our general empirical frameworks.

Many empirical growth models in the applied literature assume common technology parameters in a framework allowing for *differential TFP* across countries — a specification derived

[11] 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, Pesaran, & Smith, 1998).

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 deemed equally important for the production process. 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 allowed 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' the excess variation in the data as displayed in standard neoclassical growth models with homogeneous technology parameters. At the same time, any notion of common TFP evolution as implicitly assumed by Solow (1956)[12] and implemented empirically by Mankiw et al. (1992) has been dropped.

In our general empirical models above we emphasise a view of TFP as a 'measure of ignorance' (Abramowitz, 1956), which is in contrast to the the notion of TFP as an efficiency index in the microeconometric literature on production analysis. Furthermore, we allow for the possibility that TFP is in parts 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. These considerations lead to the adoption of a TFP structure which allows for common and/or country-specific evolution as implemented via the common factor model approach.

The following sections will show that our framework encompasses the most common empirical specifications in the literature. In our presentation we highlight technology parameter estimates and to put these estimates into context, we briefly discuss parameter values and their consistency with macro data on factor-input shares.

[12] One should note that Solow never claimed for his model to be applied to cross-country analysis: his focus was on explaining the growth experience of the United States (Durlauf et al., 2001).

1.1.2 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: the values for α^{va} and β^{va} in the above value-added based production function (1.2), and implied β^{va} in the convergence equation (1.3) should be equal to the labour and capital shares in aggregate income. Macroeconomic data for labour are available through the aggregate data on wages and welfare payments to labour.[13] 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).[14] It has been pointed out that whilst country data shows high persistence over time, there is considerable variation in the factor shares *across* countries, with labour share ranging from 5% to 80% of aggregate value-added (from UN (2004) national accounts data). Gollin (2002) attributes this to the mismeasurement of labour income in small firms, which is particularly the case in Less Developed Countries (LDCs), 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 .3 and this chapter points to empirical misspecification (heterogeneity 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'.

1.1.3 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), whereby the steady-state solutions are given empirical equivalents. We simplify their notation to

$$y_{i,1985} = -\left(\frac{\beta^{va}}{1-\beta^{va}}\right)\ \ln(\delta+\mu^{\star}+\bar{n}_i)+\left(\frac{\beta^{va}}{1-\beta^{va}}\right)\bar{s}_i^{k}+a^{va}+\varepsilon_i \tag{1.8}$$

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 ($\bar{s}^k$) — this is never motivated, but reduction of measurement error and boosting the robustness of results (in the presence of

[13] A concern in this context is expressed by Hall (1990), who emphasises 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.

[14] Data from the Federal Reserve Bank of Cleveland, for instance, shows an average labour share of 71.7% of value-added from 1970 to 2002 for the United States (Gomme & Rupert, 2004).

business cycles and/or changes in capacity utilization over time) might be plausible reasons for this choice. The a term is said to include not only 'technology', but also resource endowment, geography, climate, institutions, etc. thus is viewed as a 'measure of ignorance'.

In a first step, Mankiw et al. (1992) estimate the above equation by OLS using Penn World Table (PWT) data for 98 countries — in the following we concentrate on results for the samples including developing and developed economies in each paper discussed. Imposing parameter homogeneity Mankiw et al. (1992) find coefficients on the two input terms considerably above the magnitudes suggested by theory (.5), although of correct (opposite) signs. The implied income elasticity for capital ($\hat{\beta}^{va}$) is .60, s.t. 60% of income are attributed to capital — double the value observed in macro data for the capital share in income. The authors' first important empirical result derives from the theory-driven inclusion of human capital[15] into this model, which yields implied income elasticities for physical and human capital of .31 and .28 respectively. This 'augmented Solow model' can explain 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 & Johnson, 2008). Via approximation around the steady state and minor transformation, they derive a convergence regression as represented in equation (1.5) above, with 1985 as end of period and 1960 as base year, imposing parameter homogeneity ($\beta_i^{va} = \beta^{va} \, \forall \, i$, dto. for the convergence term ξ). Again using OLS they find that the conditional convergence rate in the model augmented with human capital is statistically significant and higher than in the unaugmented case: around 1.4% per annum — this is the second result in this paper which had a strong impact on the literature. The implied income elasticities for physical capital are now .82 and .48 in the unaugmented and human capital-augmented case respectively[16] — thus considerably different from the steady state model results. The implied human capital coefficient in the augmented model is .23, thus very similar to the estimate of .28 in the cross-section regression.

[15] Mankiw et al. (1992) 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, readily available from Barro and Lee (1993, 1996, 2001)

[16] Mankiw et al. (1992) do not report results for the restricted case of the convergence regression without human capital. We therefore report implied $\hat{\beta}^{va}$ for the unrestricted case computed from result in Table IV of their paper. The imposition of the restriction has a second order effect on this estimate in all other regressions they present. Further, Islam (1995) reports the implied $\hat{\beta}^{va}$ for a slightly different sample, obtaining .83.

The paper received most criticism for its assumption of random cross-country differences in TFP levels, which allow for the empirical representation $A_{i,0} = a + \epsilon_i$ 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 US 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 1.3 below.

Later papers also pointed out that in using OLS the convergence specification *by construction* induces downward bias in the convergence rate (Knight, Loayza, & Villanueva, 1993; Islam, 1995; Caselli et al., 1996): the presence of the initial income variable on the RHS 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 biased toward zero, with the implication of a reduced convergence rate ξ. Nevertheless, this specification forms the basis for the popular reduced form or '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.

Given the static, cross-section nature of this empirical framework, the time-series properties of the underlying data (potential nonstationarity) 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 is characterised as nonstationary (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.

The Mankiw et al. (1992) convergence regression model in (1.5) 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) roughly 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 implied physical capital, 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 $A_{i,0} = a + \varepsilon_i$ is econometrically useful but conceptually questionable, while the convergence equation setup induces downward bias in the estimate for ξ by construction.

1.1.4 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 non-overlapping panels of 5-year averages and then introducing country fixed effects

$$\begin{aligned} y_{i,\tau} \;=\; & -\left(-e^{-\xi_i \tau}\right)\left(\frac{\beta^{va}}{1-\beta^{va}}\right)\,\ln(\delta + \bar{n}_{i\tau} + \mu^{\star}) \\ & +\left(-e^{-\xi_i \tau}\right)\left(\frac{\beta^{va}}{1-\beta^{va}}\right)\bar{s}^{k}_{i\tau} \\ & -\left(-e^{-\xi_i \tau}\right)\,y_{i,\tau-1} + \left(-e^{-\xi_i \tau}\right)\,A^{va}_{0,i} + \left(-e^{-\xi_i \tau}\right)\sum_{\tau=2}^{T}\mu_\tau D_\tau + u_{i\tau} \end{aligned} \tag{1.9}$$

where due to the change in dependent variable the interpretation of estimated coefficients changes slightly,[17] but the implied technology coefficient β^{va} is equivalent to that in our general model in equation (1.3) with $\tau = 5$.[18] The crucial innovation over Mankiw et al. (1992) is the introduction of country fixed effects (operationalised using the 'within-groups' transformation), which allow for differential TFP levels across countries. Islam further uses period dummies (the sum of Ds with parameter coefficients μ in (1.9)) 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, Pesaran, & Smith, 1997). Islam also uses the PWT dataset 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 five-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 .44 — it is thus suggested that individual country effects (TFP levels) play an important role in the development process.[19] Inclusion of the human capital variable slightly lowers the convergence rate but the implied capital coefficient at .52 is quite similar to the regression without augmentation.

The major criticism leveled at Islam (1995) and other panel approaches was to highlight 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_1} 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

[17] 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 $\hat{\beta}^{va}$.

[18] 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.

[19] 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. As a consequence he does not carry out any formal test of statistically significant differences across the $A^{va}_{0,i}$.

that the validity of Islam's estimates depend critically on the assumption of *common* TFP evolution for all countries[20] — if this is violated, country-errors will contain a term which is serially correlated with unit coefficient, causing *upward* bias in the convergence rate estimate. Variable endogeneity due to the presence of common factors (see Section 1.3) is also a matter for concern.

The panel fixed effects model in (1.9) is nested within our more general model in (1.3), 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 $\bar{n}$ and savings rates $\bar{s}^k$ 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, Summers, & Aten, 2006) we would obtain $T = 9$; given that averages constructed from non-stationary variables are nonstationary themselves (Granger, 1988; Granger & Siklos, 1995; Marcellino, 1999), the use of longer T eventually faces the same issues as annual data, outlined in detail in section 1.2. 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.

[20] This can be interpreted as identical factor parameters $\boldsymbol{\lambda}_i = \boldsymbol{\lambda}$ on the unobserved common factors.

A variant on the above panel estimation approach is presented by Caselli et al. (1996). Instead of fixed effects estimation they apply dynamic panel GMM estimation techniques (Arellano & 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 (as is assumed here), 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 .10, with convergence very high at 13% per annum. In the human capital-augmented regression the income elasticities for physical and human capitals are .50 and -.26 respectively, with convergence around 6.8% per annum. The implied physical capital parameter is relatively close to the result in Islam (1995), while implied human capital parameter is significantly negative and the convergence rate λ is almost doubled. Since the Islam results in turn have lower implied capital coefficients 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. Their model is thus no longer 'structural', but instead 'reduced form'. 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% pa.

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 &

Bond, 1998, 1999).[21] Furthermore, if technology parameters are heterogeneous across countries, "there exists no consistent instrumental variables estimator" (Lee et al., 1997, p.367).[22] 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 specification of dynamics may therefore be commendable.

1.1.5 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, Larson, Butzer, and Mundlak (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, modeled 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

[21] The parameter on the lagged dependent variable is biased downward (in absolute terms), implying higher rate of convergence (Bond, Hoeffler, & Temple, 2001; Durlauf et al., 2005).

[22] 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 $[(-e^{-\xi\tau}) - (-e^{-\xi_i\tau})]y_{i,t_1}$. For more details refer to Section 1.4. 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.

consistency

$$\text{agriculture} \qquad y_{it} = \beta^{va} k_{it} + \theta^{va} n_{it} + u_{it} \qquad u_{it} = A_{0,i}^{va} + \mu_i t + \varepsilon_{it} \tag{1.10}$$

$$\text{manufacturing} \qquad y_{it} = \beta^{va} k_{it} + u_{it} \qquad u_{it} = A_{0,i}^{va} + \mu_i t + \varepsilon_{it} \tag{1.11}$$

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 (CRS) on this model and estimate it separately for agriculture and manufacturing data using LSDV.[23] 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. The capital coefficient in the Cobb-Douglas estimation of the manufacturing data is estimated at .69; the authors highlight the magnitude of this coefficient in comparison to 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 additional regressor (coefficient .24), the estimated capital elasticity is .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 questioned. The empirical approach furthermore assumes cross-section independence.

The sector-level Martin and Mitra (2002) regression model is nested within the value-added version of our general production function model in (1.2), under assumption of

(i) common technology parameters (within sector j) across countries ($\beta_i^j \equiv \beta^j$);
(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.

[23] They also estimate a more general translog model but results are qualitatively the same.

Martin and Mitra (2002) thus address the issue of heterogeneity in TFP levels and growth rates in a static pooled fixed effects model using annual data which imposes common technology parameters across countries. TFP growth rates (constant by construction) are found to be higher in agriculture than manufacturing.[24] Time-series ($T = 26$) 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.

1.1.6 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 modeling parameter heterogeneity explicitly:

$$\begin{aligned} y_{i,1985} - y_{i,1960} &= \boldsymbol{\psi}(z_i)'\mathfrak{X}_i + u_i \\ \mathfrak{X}_i &\equiv \left\{a,\ \ln(\delta + \bar{n}_i + \mu),\ \bar{s}_i^k,\ \bar{s}_i^h,\ y_{i,1960}\right\} \end{aligned} \tag{1.12}$$

where $\bar{s}^h$ 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% pa for depreciation plus 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. They introduce a local goodness-of-fit measure, a generalisation of the conventional R^2, which indicates that vis-à-vis a homogeneous parameter model the fit

[24] 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.

is improved for countries with higher base-year income $y_{i,1960}$.

The Durlauf et al. (2001) model is nested within our more general convergence equation model in (1.3), under the assumption of

(i) heterogeneous technology parameters across countries, which are functions of initial income ($\beta_i \equiv \beta(z_i)$);
(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 exists across countries, and that "empirical exercises which fail to incorporate parameter heterogeneity are likely to produce misleading results." (Durlauf et al., 2001, p.935).

1.1.7 Full parameter heterogeneity in a model with nonstationary inputs

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

Of these, the most encompassing framework for a nonstationary 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

$$y_{it} = c_i + \mu_i t + \theta_i \, s_{it}^k + u_{it} \tag{1.13}$$

for $i = 1, \ldots, N$ and $t = 1, \ldots, T$, where s^k is the (log) investment share in GDP and y 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 types of 'intangible capital' H, with savings rate s^h. If

H is accumulated using a share of income Y, the general specification can be written as

$$y_{it} = -\left(\frac{\beta_i^{va}\{+\varphi_i\}}{1-\beta_i^{va}\{-\varphi_i\}}\right)\ln(\delta+n_{it}+\mu^{\star})+\left(\frac{\beta_i^{va}\{+\varphi_i\}}{1-\beta_i^{va}\{-\varphi_i\}}\right)s_{it}^{k} + \left(\frac{\varphi_i}{(1-\beta_i^{va}-\varphi_i)}\right)s_{it}^{h}+u_{it} \quad (1.14)$$

$$u_{it} = A_{i0}^{va}+\mu_i t+\varepsilon_{it} \quad (1.15)$$

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 (1.14) is subsumed into the intercept term and/or the linear trend term in (1.13), as its evolution over time is stationary. The interpretation of θ_i in (1.13), as well as that of the deterministic fixed effect and trend terms thus varies depending on whether aggregate investment drives 'intangible investment' or not.[25] 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_{it}^{h} 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 (1.13).[26] Similarly for the population growth term $(\delta+n_{it}+\mu^{\star})$. Under these assumptions, the slope coefficient for measured physical capital θ_i 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 nonstationary panel estimation approach extracts the long-run relationship $\hat{\beta}_i$ (via the cointegrating vector $\hat{\theta}_i$) *without* the requirement of proximity to the steady-state,[27] 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 nonstationary and cointegrated, then their long-run relationship will be extracted via the nonstationary panel econometric method applied, regardless of any short-run dynamics (Pedroni, 2007).

The model in (1.13) is estimated using the aggregate PWT data (version 5.6, 1950-1992).

[25] See Pedroni (2007, p.435/6) for details. The working paper (Pedroni, 2004) expands on this.

[26] 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.

[27] Recall that the convergence equation is the result of an approximation *around the steady-state*.

The nonstationary panel approach depends on a reasonably long time-series dimension, nonstationarity 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 (nonstationary series, cointegrated). With regard to variable stationarity, it is important to note that Pedroni does not postulate that the savings rate proxy is nonstationary *indefinitely* ('global' property). Instead, he suggests that unit root evolution is merely a 'local' behaviour of the series *within-sample*, which is commonly found in macro data (Jones, 1995).

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 FM-OLS — 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): it exploits the fact that under cointegration the bias created by variable endogeneity is of second order, and the first differenced regressors can serve as instruments for this bias to construct bias-adjusted variables (Pedroni, 2007). The approach does away with a great deal of assumptions required in stationary empirics, most notably the proximity to steady state and the strict exogeneity of the regressors.[28]

The resulting cross-country mean for the capital coefficient is .27 — Pedroni (2007) interprets the proximity of this value to macro data on factor shares as vindication of his estimation approach. This value is surprisingly robust to the inclusion of countries previously dropped in the sample selection process: .25 for the sample of 51 countries. This average however falls to around .2 in either sample if data is cross-sectionally demeaned prior to estimation, a method adopted to eliminate *common* TFP. Interestingly, the cross-sectional demeaning of variables was identified as inappropriate in the presence of heterogeneous factor parameters by the same author (Pedroni, 2000). The adoption of the Pesaran (2006) Common Correlated Effects estimator (CCEMG) similarly leads to an average implied capital coefficient of around .2 — for more details on this approach we refer to Chapter 2.

[28]The FMOLS estimates are superconsistent in a cointegrated framework of nonstationary variables. If the nonstationarity assumption is violated, the FMOLS transformation results in 'overdifferencing', which introduces noise while the endogeneity problems remain.

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 argued not to be due to sampling variation and 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 (1.3) assuming

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

The empirical implementation by Pedroni (2007) requires sample countries to satisfy strict data requirements (*full* 43 years of data, nonstationarity, cointegration), which in effect excludes the vast majority of LDCs from the analysis. His preferred regression approach precludes the possibility of a common, nonstationary TFP evolution; once this assumption is relaxed the implied mean capital coefficient drops to around .2. One explanation for this finding may be that following dual economy arguments (Temple, 2005) aggregate economy data represents the wrong basic unit of analysis for growth empirics, which should be based on sector-level analysis of agriculture, manufacturing and services instead. This however cannot explain the direction of the perceived bias in the Pedroni (2007) estimates.

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. The evolution we charted is one away from single cross-section regressions imposing common technology parameters and toward models allowing for parameter heterogeneity and explicit treatment of variable time-series 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.

1.2 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 nonstationarity, 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 & Bond, 1991; Blundell & Bond, 1998), and more recently in its treatment of nonstationarity and cointegration in a panel setup (Pedroni, 1995, 1999, 2000; Kao, 1999; Phillips & Moon, 1999). The latter development has thus far found limited attention in the *mainstream* literature on empirical growth modeling[29] — for instance the chapter on growth econometrics in the recent *Handbook of Economic Growth* by Durlauf et al. (2005) contains only limited discussion of nonstationarity 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 annual data in the production function framework.[30]

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 'nonstationary' processes in *some* countries (C. Nelson & Plosser, 1982; Granger, 1997; Lee et al., 1997; Rapach, 2002; Bai & Ng, 2004; Pedroni, 2007; Canning & Pedroni, 2008). Nonstationarity 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 is a realisation. In case of a nonstationary variable, adding more observations does not help us to get an idea of what the distribution looks like as mean, variance, etc. do not settle at ('converge' to) certain values.

[29] This is suprising given that in their important contribution to the nonstationary panel literature, Phillips and Moon (1999, p.1057) mention the Penn World Tables in their opening paragraph and the same dataset is also highlighted in the opening paragraphs of review chapters on nonstationary panel econometrics by Kao and Chiang (2000), Smith and Fuertes (2007) and Baltagi (2005).

[30] Similarly for the static Mankiw et al. (1992) specification in equation (1.8) (see Pedroni, 2007).

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, p.44) — 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 generalisation. Further, as was pointed out in the previous section, 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). As we shall see, nonstationarity has important implications for estimation and inference.

In the context of the empirical literature on growth and development it is important to note here that the stationarity/nonstationarity property is invariant to time-averaging of variables (Granger, 1988; Granger & Siklos, 1995; Marcellino, 1999). Given that data for educational attainment is only available at five-year intervals (Barro & Lee, 1993, 1996, 2001) many empirical implementations using 'human capital' variables are based on short panels of five-year averages in levels. The above statement confirms that if the underlying variable series are nonstationary, this common practice of using 5-year averages does *not* render the resulting series of averages stationary I(0) and/or may introduce data dependencies which lead to violation of the standard assumptions of the ordinary least squares regression framework.[31]

In conclusion, any macro production function is likely to contain at least some countries with nonstationary 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[32] and for diverse countries such as the PWT or alternative annual

[31] One could even argue that some properties of nonstationarity are still contained in period-averaged observations of a cross-section specification in the style of Mankiw et al. (1992). See also Madsen (2005).

[32] 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 datasets including

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 & Fuertes, 2007): a Fisher-type test of unit roots in a panel (Maddala & Wu, 1999), for instance, investigates the null of *all* country series containing unit roots against the alternative that *at least one* series is stationary.[33] Thus while panel unit root tests have improved power over time-series unit root tests (Im, Pesaran, & Shin, 2003; Baltagi, 2005; Choi, 2007), the matter of their interpretation hinders a straightforward application to the data, with ultimately time-series unit root tests of individual country series providing the 'most relevant information' (Maddala, 1999). A second generation of panel unit root tests has been introduced to account for cross-section dependence in the data (e.g. Bai & 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 nonstationary output on nonstationary 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.[34] 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 nonstationary variable series (Pedroni, 2007; Canning & 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 Data Generating Process or DGP) would be the

observations from LDCs are considerably shorter.

[33] Panel-based tests that provide explicit information on which country series are I(1) and which I(0) (Breuer, McNown, & Wallace, 2001) were found to be highly sensitive to sample selection (Ford, Jackson, & Kline, 2006).

[34] Granger and Newbold (1974) illustrated this in a simulation using two independent random walks.

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 noncointegration and potential for spurious regression does not arise (for these countries).

Formulating the correct hypotheses to empirically test for cointegration in a panel dataset involves severe conceptual difficulties and test results are often inconclusive, especially so in moderate T, N samples (Baltagi & Kao, 2000; Coakley, Fuertes, & Smith, 2001; Pedroni, 2007).[35] In addition, it is unclear how the cointegration tests perform if variable series are nonstationary 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 & Edgerton, 2008) which account for cross-section dependence. These tests however still show comparatively low power for $T = 100$. Thus formally testing variable series for cointegration in a diverse panel of moderate T dimension may leave many questions unanswered.

A conclusion to be drawn from this discussion is that while formal testing for nonstationarity 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 possessing the nonstationary property and adjust our estimation strategy accordingly (Pedroni, 2007). There are two implications: *firstly*, running separate country regressions in levels for the 'nonstationary countries' will only yield valid estimates if all elements of the underlying cointegrating relationship have been included in the empirical model; any nonstationary elements left out will enter the error terms and thus lead to the breakdown of the cointegrating relationship and potentially spurious regression results. *Secondly*, some countries will have stationary variable series, in which case this problem of noncointegration does not arise.

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 nonstationarity, whether analysed at the economy level (Coe & Helpman, 1995; Coe, Helpman, & Hoffmaister, 1997; Kao, Chiang, & Chen, 1999; Engelbrecht, 2002; Bond, Leblebicioglu, &

[35] Note that cointegration is invariant to temporal aggregation (Granger, 1988; Marcellino, 1999).

Schiantarelli, 2007) or the sectoral level (Bernard & Jones, 1996b; Funk & Strauss, 2003). Further, Coakley, Fuertes, and Smith (2006) explicitly state with reference to cross-country production function estimation that technology shocks (treated as unobserved common factors) are plausibly nonstationary. 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 & Helpman, 1991; Aghion & Howitt, 1998). They model knowledge (proxied as in Jones (1995) by 'measured' TFP[36]) 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 characterised as nonstationary and integrated of order one" (Abdih & Joutz, 2006, p.243). Thus the nonstationarity of TFP/nonfactor 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 nonstationarity 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 nonstationary errors, just like in the case of misspecification of factor parameter heterogeneity. 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 idiosyncracies of nonstationary TFP evolution. The evolution of TFP in a nonstationary fashion implies that TFP becomes part of the cointegrating relationship.

We noted above that temporal aggregation of time-series data — e.g. taking five-year averages — does not change the stationarity properties. The use of five-year averages of variable series in empirical growth regressions or production functions is still a common practice in the literature, with only a small number of empirical papers using annual data (Lee et al., 1997; Martin & Mitra, 2002; Pedroni, 2007; Canning & Pedroni, 2008). The typical justification for the use of period-averaged data is that annual data will contain business cycle factors and/or changes in capacity utilization irrelevant for long-run income movements, whereas long-run averages are

[36]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.

said to enable identification of long-run growth effects (Brock & Durlauf, 2001). In contrast, Lee et al. (1997, p.359) note 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." This argument finds further support in the work of Mundlak, Larson, and Butzer (2008) reviewed in Chapter 3. Our general models are therefore intended for empirical application using annual data, forcing the econometrician to deal with time-series properties explicitly.

1.3 Cross-section dependence in macro panel data

> *"Most of the work carried out on panel data has usually assumed some form of cross sectional independence to derive the theoretical properties of various inferential procedures. However, such assumptions are often suspect . . . "*
> Kapetanios et al. (2009, p.2)

Panel data econometrics over the recent years has seen a rising 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 & Sul, 2003; Pesaran, 2006; Pesaran & Tosetti, 2007). 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 & 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) even uses GDP series from the Penn World Table, concluding 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). Long 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 empirical specification in at least three ways (Pesaran & Tosetti, 2007): *firstly*, 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 assigns a unity value to a country pair which share a common border, and a zero otherwise (Pesaran, 2004). This weight matrix is then applied to a spatial error component or spatially lagged dependent variable in the empirical model. In practice, application of most of this work 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. It is important to distinguish this impact of 'geography' on growth from that commonly discussed in the literature: the latter, most prominently the work by Jeffrey Sachs (Sachs & Warner, 1995, 2001; Sachs, 2003), argues that 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 group-specific 'counterfactual estimates'. 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.

Secondly, 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, 1999; Coakley et al., 2006). In either case, violation of the homogeneity assump-

tion leads to dependence in the error terms across countries 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. We replace our general specification in (1.1) with a simple factor model for illustrative purposes

$$y_{it} = \beta_i x_{it} + u_{it} \qquad u_{it} = \alpha_i + \lambda_i f_t + \varepsilon_{it} \tag{1.16}$$

$$x_{it} = \phi_i f_t + \psi_i g_t + v_{it} \tag{1.17}$$

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 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 ($\bar{\lambda}_i \neq 0$ *and* $\bar{\phi}_i \neq 0$) 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 nonstationary 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 (1.17) for f_t and plugging

this into equation (1.16)

$$
\begin{aligned}
y_{it} &= \alpha_i + \beta_i x_{it} + \lambda_i f_t + \varepsilon_{it} && (1.18)\\
&= \alpha_i + \beta_i x_{it} + \lambda_i \phi_i^{-1}(x_{it} - \psi_i g_t - e_{it}) + \varepsilon_{it} \\
&= \alpha_i + \underbrace{(\beta_i + \lambda_i \phi_i^{-1})}_{\varrho_i} x_{it} \underbrace{-\lambda_i \phi_i^{-1} \psi_i g_t - \lambda_i \phi_i^{-1} e_{it} + \varepsilon_{it}}_{\varsigma_{it}} \\
&= \alpha_i + \varrho_i x_{it} + \varsigma_{it} && (1.19)
\end{aligned}
$$

From equation (1.19) 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 (Kapetanios et al., 2009) — β_i is unidentified. Under the specification described, a standard pooled fixed effects or Pesaran and Smith (1995) Mean Group estimator will therefore likely yield an inconsistent estimator (due to residual nonstationarity) of a parameter we are not interested in (due to the identification problem).

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 & Mairesse, 1998).

1.4 Parameter heterogeneity in the face of nonstationarity

> *"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' using an empirical specification characterised by homogeneous factor parameters, coupled with heterogeneous TFP levels and evolution (Parente & Prescott, 1994, 1999; Klenow & Rodriguez-Clare, 1997a; Caselli & Coleman II, 2006; Acemoglu, Antras, & Helpman, 2007; Aiyar & 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 specification has 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 has been argued that variable omission and parameter heterogeneity may be interpreted as examples of deviations 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 & Durlauf, 2001). 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 *nonstationary*, 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 nonstationary errors if 'true' technology parameters are heterogeneous and input variables are nonstationary. With reference to our general production model in equation (1.1) pooled error terms may contain one or more of

$$(\alpha_i - \pi_L)\, L_{it} \qquad (\beta_i - \pi_K)\, K_{it} \qquad (\gamma_i - \pi_M)\, M_{it} \tag{1.20}$$

where π_L, π_K, π_M are the *common* regression coefficients for labour (L), capital stock (K)

and material inputs (M) respectively (all variables in logs), while α_i, β_i, γ_i are the 'true' country-specific technology parameters. Each of the three terms in equation (1.20) is a linear combination of a nonstationary variable/process and thus will itself be nonstationary. 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. 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 nonstationary errors if 'true' TFP is heterogeneous and nonstationary. In our general notation

$$(\boldsymbol{\lambda}_i' \boldsymbol{f}_t - \pi_{TFP}) \tag{1.21}$$

where $\boldsymbol{\lambda}_i \boldsymbol{f}_t$ represents the 'true' country-specific TFP process (nonstationary) and π_{TFP} the estimated TFP evolution specified via the set of common year dummies. Note that if the 'true' TFP process is nonstationary, 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 nonstationary evolution) and thus result in nonstationary country-regression errors. Similarly 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 & Moon, 2000; Smith & Fuertes, 2007). This implies that regardless of whether we have homogeneous cointegration, heterogeneous cointegration or noncointegration, the pooled fixed effects estimator in levels can provide a consistent point estimate of some average long-run correlation between inputs and output. The joint limit theorem of the paper however requires $N/T \to 0$, which limits the applicability of this result for cross-country growth analysis using macro data. This aside, the asymptotic result builds on the assumption of cross-section independence, which as was discussed above is seemingly unjustified in

the macro panel data context. Monte Carlo simulations by Coakley et al. (2006) replicated in Chapter 4 of this thesis however 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 is rather inefficient (Table 4.1, Case C). However, if unobserved factors drive both the input(s) and output variables, the Phillips and Moon (1999) results break down.

In Section 1.2 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 nonstationary process. 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 nonstationary 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 regards to the specification of TFP evolution, we can state that the use of a linear trend may not capture the (potentially) nonstationary evolution and may thus equally lead to a breakdown in the cointegrating relationship, even if factor parameters were modelled correctly.

1.5 Investigating growth and development at the sectoral level

> *"If we ask 'why do they [undeveloped economies] save so little', the truthful answer is not 'because they are so poor' ... The truthful answer is 'because their capitalist sector is so small' (remembering that 'capitalist' here does not mean private capitalist, but would apply equally to state capitalist)."*
> Lewis (1954, p.159)

Although much of the motivation used in the theoretical parts of this thesis is focused on aggregate economy data such as the Penn World Table, the arguments do similarly apply to macro panel datasets at other levels of aggregation, notably regional macro data and sectoral cross-country data. In our empirical implementations in Chapters 2 and 3 we do not investigate aggregate economy data, but instead focus on macro panel data for the agricul-

tural and manufacturing sectors respectively. This choice is motivated by the notion that if cross-country technology and TFP heterogeneity can be shown to matter *at the sectoral level* then it is much more likely to lead to empirical difficulties *in the aggregate economy data*. This econometric argument is further backed up by the economic argument contained in the dual economy model literature — briefly introduced below — which suggests that the use of aggregate economy data for analysis may miss out on one of the most prevalent features of development and growth, namely the structural change undergone in transforming an economy dominated by (subsistence) agriculture to one with a substantial modern, industrial sector.

In the early literature on developing countries a distinction was made between the processes of economic development and of economic growth. Economic development was seen to be a process of structural transformation by which in Lewis' frequently cited phrase an economy which was "previously saving and investing 4 or 5 percent of its national income or less, converts itself into an economy where voluntary savings is running at about 12 to 15 percent of national income" (Lewis, 1954, p.155). An acceleration in the investment rate was only one part of this process of structural transformation; of equal importance was the process by which an economy moved from a dependence on subsistence agriculture to one where an industrial modern sector absorbed an increasing proportion of the labour force (Jorgensen, 1961; Kaldor, 1966; Kindleberger, 1967; Kuznets, 1961; Leibenstein, 1957; Ranis & Fei, 1961; Robinson, 1971). In contrast to these models of "development for backward economies" (Jorgensen, 1961, p.309), where duality between the modern and traditional sectors was a key feature of the model, was the analysis of economic growth in developed economies. Here the processes of factor accumulation and technical progress occur in an economy which is already 'developed', in the sense that it has a modern industrial sector and agriculture has ceased to be a major part of the economy (e.g. Solow, 1956, 1957; Swan, 1956; Cass, 1965).

Much of the early growth modelling work proceeded without close connection to observed data. The models were in Solow's classic exposition of growth theory inspired by stylised 'Kaldor' facts (Kaldor, 1957). As Solow (1970, p.2) notes, "[t]here is no doubt that they are stylized, though it is possible to question whether they are facts". The dual economy models of structural transformation used case studies (e.g. Paauw & Fei, 1973) and facts at least as stylised as those in the Solow-Swan growth context. Empirical studies employed a vast array of explanatory variables of growth, while methodological, statistical, and conceptual difficulties on top of

sample heterogeneity made it difficult to draw reliable conclusions from the existing literature (Levine & Renelt, 1991). The key papers which brought modelling and data together were the contributions of Barro (1991) and Mankiw et al. (1992), which initiated a major revival in the Solow-Swan model and effectively merged the concerns of economic development with those of growth.

The literature begun in the early 1990s has yielded a large array of models in which there has been increasing interaction between the theory and the empirics (see discussion in Aghion & Howitt, 1998; Durlauf & Quah, 1999; Temple, 1999; Easterly, 2002; Durlauf et al., 2005). It remains true that the empirical analysis continues to be dominated by the empirical version of the aggregate Solow-Swan model (Temple, 2005) with much of the empirical debate focusing on the roles of factor accumulation versus technical progress (Chen, 1997; Easterly & Levine, 2001; Baier, Dwyer, & Tamura, 2006; Klenow & Rodriguez-Clare, 1997a, 1997b; Lipsey & Carlaw, 2001; Young, 1995). While there is some new theoretical and empirical work using a dual economy model, this is largely absent from textbooks on economic growth and has not been a central focus of attention for the empirical growth literature (Temple, 2005).

Table 1.1: Share of Agriculture in GDP (in %)

Agricultural VA † (in % of GDP; Decadal Medians)					
	1960s	1970s	1980s	1990s	2000s
Canada & US		4.3	3.0	2.3	1.3
Europe (Euro area)		6.2	4.3	2.9	2.2
Latin America & Caribbean	14.0	12.8	10.2	7.5	6.6
Middle East & North Africa	21.7	15.2	15.2	15.2	12.4
Australia & New Zealand		9.0	6.7	5.3	4.0
East Asia & Pacific	37.8	32.0	27.6	19.0	13.2
Sub-Saharan Africa	26.2	21.5	20.1	19.4	17.5
South Asia	42.3	38.6	31.6	27.5	21.6

Employment in Agriculture ‡ (% of total employment; Means, Medians for 2000s ♯)					
	1960	1970	1980	1990	2000s
United States	6.6	4.3	3.5	2.8	2.6
Europe	31.0	21.1	15.9	12.2	4.8
Latin America & Caribbeans	49.0	42.0	34.2	25.4	16.8
Australia & New Zealand	11.9	8.7	7.3	6.3	6.0
Eastern Asia	76.8	70.9	66.9	64.8	45.4
Africa	79.6	75.8	68.7	62.8	

† World Bank (2008) World Development Indicators. ‡ ILO decadal estimates 1950-1990, 'economically active population in agriculture'. ♯ World Bank (2008) WDI; here: 'employment in agriculture' and Europe = Euro Area. 2000s includes the most recently available data, which differs somewhat by region but typically includes data up to 2006.

A primary reason for the focus has been the availability of data. The Penn World Table dataset — most recently Heston et al. (2006) — and the Barro-Lee data on human capital (Barro & Lee, 1993, 1996, 2001) have supplied macro-data which ensure that the aggregate Solow-Swan model can be readily estimated. Our point in case is that the sectoral makeup of economies matters, and that the aggregate empirical analysis of countries with diverse economic structure misses out on important differences. In Table 1.1 we provide some statistics relating to the 'importance' (in terms of GDP and employment shares) of the agriculture sector in a number of countries/country groupings.

The contrasts in the employment numbers across groups and over time are particularly pronounced. Overall, the data implies considerable heterogeneity in the sectoral makeup and structural change of the economies summarised in these groupings, and our separate empirical analysis for agriculture and manufacturing is intended to accommodate these heterogeneities, rather than sweep them under the carpet of economy-level aggregation.

1.6 Concluding remarks

In this chapter 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 as well as the emerging nonstationary panel literature over the past two decades. A selective review of the literature in Section 1.1 highlighted how factor and non-factor parameter heterogeneity can be accommodated within our general modelling framework.

The core argument of this chapter, presented in Sections 1.2, 1.3 and 1.4, relates to the interplay between parameter heterogeneity and hitherto neglected time-series and cross-section dependence properties pertinent in macro panel data commonly used for cross-country growth and development analysis. In standard regression models, implictly assuming stationary variables, it is deemed unnecessary to allow for countries to have fully heterogeneous empirical production function representations (with regard to TFP *and* factor parameters), when instead flexible TFP specification can 'mop up' empirical variation in the data and allow for convenient country rankings in terms of TFP levels and growth rates. However, once nonstationarity of the observable and unobservable processes (including TFP) is treated as a distinct

possibility, the failure to account for cross-country differences in production technology leads to the breakdown of the cointegrating relationship and thus potentially yields spurious regression results. We further highlighted the difficulties arising from the assumption of cross-section independence, which is standard for the micro-panel estimators (e.g. Fixed Effects, Difference and System GMM) frequently applied to macro panel data. A common factor model specification like our own indicates that technology parameters on observable inputs will be difficult to identify in these standard regression setups.

Our next step in the investigation of an encompassing approach to growth empirics will be to take these arguments to the data and to compare and contrast results for standard (micro-)panel estimators employed in the literature with those from empirical approaches addressing parameter heterogeneity, variable nonstationarity and cross-section dependence.

Appendix to Chapter 1

1.A 'Finding productivity': empirical extraction of TFP

There is a sizeable literature concerned with conceptual difficulties over what TFP actually represents and how it can be measured most appropriately (Young, 1994, 1995; Chen, 1997; Easterly & Levine, 2001; Baier et al., 2006). The analysis of *TFP growth* was brought about by studies using the time-series dimension of data more intensively, in contrast to studies employing period-averaged cross-country regressions where *TFP levels* are emphasised by construction. The following review will show that conceptual and empirical issues arise out of this debate.

The assumption that underlying factor parameters are identical across countries has led to the emergence of a vast empirical literature assuming heterogeneous TFP levels and evolution (Parente & Prescott, 1994, 1999; Klenow & Rodriguez-Clare, 1997a; Caselli & Coleman II, 2006; Acemoglu et al., 2007; Aiyar & Dalgaard, 2008). On the theoretical side, barriers to adoption of new technologies are commonly regarded as the crucial factor in production in their differential effect on TFP across countries (Parente & Prescott, 1994, 1999). On the empirical side, there are two approaches to extracting TFP from observed data: via econometric estimation of aggregate production functions, or using a growth accounting approach. The former is the central focus of this thesis and we have introduced the various assumptions and issues related to this approach in great detail in the main section of this chapter. We now turn to the various types of growth accounting for TFP investigation. In practice, it can be suggested that the latter approach dominates the literature. This is rather curious, given that "in essence there is no real difference between the Cobb-Douglas production function estimation and the growth accounting approach as far as the underlying methodology is concerned" (Chen, 1997, p.21).

In the standard *regression approach*, for an aggregate value-added Cobb-Douglas production function as in our general model in equation (1.2) we have for each country

$$Y_t = \alpha^{va} L_t + \beta^{va} K_t + A_0^{va} + \mu_t^{va} + u_t \tag{1.22}$$

where μ_t^{va} is the value-added TFP evolution which only captures disembodied, exogenous TFP growth. This 'Hicks-neutral' technological change μ_t^{va} increases the efficiency of both labour and capital — the proverbial 'manna from heaven' (Chen, 1997) or 'free lunch' (Lipsey & Carlaw, 2004). Any technical progress embodied in the factors is assumed to have been properly specified and accounted for in the construction of input variables.

In the *growth accounting approach* TFP growth is derived as a residual from observed output and factor input growth, commonly from a firm-level, sector-level or country-level time-series. Under standard assumptions aggregate value-added is decomposed into contributions of factor-inputs and TFP growth, imposing factor shares of capital and labour in output η^L and η^K:[37]

$$\frac{\dot{Y}}{Y} = \eta^L \frac{\dot{L}}{L} + \eta^K \frac{\dot{K}}{K} + \Delta\text{TFP} \qquad \Leftrightarrow \qquad \Delta\text{TFP} = \frac{\dot{Y}}{Y} - \eta^L \frac{\dot{L}}{L} - \eta^K \frac{\dot{K}}{K} \tag{1.23}$$

A variant on this approach is *levels accounting*, also referred to as *development accounting*, which uses cross-country data on inputs and output in levels at one point in time (Caselli, 2005), commonly the year which maximises data coverage.

The standard practice in either case is to impose a capital coefficient of $\eta^K = .3$ (e.g. Klenow & Rodriguez-Clare, 1997a; Hall & Jones, 1999; Caselli, 2005) based on "US time-series data on the capital-share, whose long-run (and roughly constant) average value is 1/3" (Caselli, 2005, p.686) and the assumption of constant returns to scale $\eta^L + \eta^K = 1$ — note that this capital coefficient is applied to *all* countries in the sample, thus rejecting any notion of cross-country parameter heterogeneity (Caselli, 2008b). The imposition of .3 has become second nature in this approach, to the extent that some celebrated contributors to the literature do not even mention which value they impose and from which source (Barro, 1999).

The accounted TFP growth is in theory identical to the above μ_t^{va} — disembodied, Hicks-neutral exogenous technical progress. In practice however, one needs to keep in mind that TFP is a residual, such that it represents a 'catch-all' for output growth that cannot be explained by factor accumulation. Thus if TFP growth is recovered via growth accounting its

[37] There are more complex methods for TFP extraction, most commonly along the lines of Caves, Christensen, and Diewert (1982), described in Harrigan (1999).

coefficient "need not represent only technological change and may not represent technological change at all" (Baier et al., 2006, p.27) since measurement error, violations of assumptions and incorrect variable construction can cause considerable bias. Any measurement error in output, labour or capital enters the residual term and thus TFP growth. This may have considerable impact since factor inputs need to account correctly for embodied technical change, which given the difficulty of distinguishing between embodied and disembodied technical progress seems impossible (Chen, 1997; Lipsey & Carlaw, 2001; Baier et al., 2006). Violations of the assumptions of constant returns to scale (externalities, spillovers), and private and social marginal product equality can add to further accounting error (Young, 1995; Barro, 1999). Which types of capital to include in the aggregate stock, the appropriate treatment of depreciation as well as capacity utilisation pose further hurdles in the accurate construction of factor variables, resulting in measurement error (Chen, 1997; Caselli, 2005). As a result, it is now widely recognised that TFP growth derived from growth accounting "does not really measure technical change" (Caselli, 2008a).

Empirical studies using TFP growth accounting have a long tradition since Abramowitz (1956), Kendrick (1956) — who coined the term Total Factor Productivity — and Solow (1957). During the 1990s an academic debate raged over the importance of factor accumulation vis-à-vis technical progress in the development of the East Asian Newly Industrialised Economies (NIEs). In a series of papers Alwyn Young (1994, 1995) showed that growth accounting yielded low TFP growth once increased labour force participation, rapid structural change, increased capital intensity, and increased average human capital had been taken into account. His conclusion was that factor accumulation played a vital and technical progress a minor role in these nations' spectacular economic growth. The work by Easterly and Levine (2001) seriously challenged this finding.[38] Chen (1997) however argues that success or failure to detect TFP growth in empirical studies depends on the measurement of factor inputs and the definition of TFP — both of which are fraught with difficulty. As was developed above, the TFP evolution term μ_t^{va} only captures disembodied technical change. Chen (1997) shows that more recent empirical work (for instance Young, 1995) has gone to great lengths to adjust for quality changes in the process of factor input aggregation, resulting in lower disembodied TFP growth values. The

[38] They strengthen their support for technical progress as the growth engine by pointing to empirical studies investigating the timing of growth and savings/capital formation, which suggest that the causality runs from the former to the latter but not vice-versa (Carroll & Weil, 1994; Blomström, Lipsey, & Zejan, 1996). Bond et al. (2007) have since criticised the Granger-causality testing procedure used in these papers.

most recent empirical findings by Baier et al. (2006) seem to back this argument.[39] Still, in contrast to country-specific growth accounting using time-series data where TFP is found to play a relatively minor role (Caselli, 2008a) "[t]he consensus view in development accounting is that Efficiency [TFP] plays a very large role [in explaining income differences]." (Caselli, 2005, p.681), typically around 50% across all countries.

Our general production function and convergence equation frameworks challenge the methodology these papers have used to 'find' productivity: *firstly*, as we argued above, it seems sensible to allow for maximum flexibility in the structure of the empirical TFP terms; if (as we prefer) TFP represents a 'measure of our ignorance' (Abramowitz, 1956), then it makes sense to allow for differential TFP across countries and time, with the latter unconstrained with regard to stationarity.

Secondly, it further makes sense to keep an open mind about the commonality of TFP: while early empirical models such as that of Mankiw et al. (1992) or Islam (1995) assumed common TFP growth for all countries, later studies preferred to specify differential TFP evolution across countries. We believe the arguments for commonality (non-rival nature of knowledge) and idiosyncracy (patents, tacit knowledge, learning-by-doing) call for an empirical specification which does not rule out either by construction. Common factor models allow for this type of setup.

Thirdly, we believe that the findings of the 'TFP-industry' ('it's not factor accumulation', Easterly & Levine, 2001) are a necessary artefact of the rigid empirical framework employed. We use the following metaphor about racing cars and lap-times to illustrate this point: if all drivers were to use identical racing cars, any difference in their lap-times could be attributed to the drivers' individual ability.[40] If, however, some racing cars were high-tech machines while others were rickety bangers, we would be concerned to attribute lap-time differences as being *exclusively* determined by driver ability. This setup is in effect the approach of the mainstream growth empirics literature: since production technology is assumed to be the same

[39]The conclusion to be drawn from Young's work is thus that disembodied technical progress has not been important in the growth path of the Four Asian Tigers, while embodied technical change has gone unmeasured (Chen, 1997). In the same vein, Lipsey and Carlaw (2001, p.43, emphasis added) conclude that "whatever TFP does measure ... it emphatically does not measure *all* of technological change".

[40]Luck and differential environmental conditions would arguably not come into the equation if we had each of them do repeated laps over time.

(racing cars), any differences in long-run cross-country development are attributed to unobserved growth determinants/TFP (driver ability).

Following Durlauf et al. (2001) and Pedroni (2007), we argue for an empirical specification that allows for parameter heterogeneity across countries and a shift away from the widespread focus on TFP analysis and toward an integrated treatment of the production technology *in its entirety*: "a full understanding of cross-country growth differences will need to explain why this parameter heterogeneity exists." (Durlauf et al., 2001, p.935).

1.B A preoccupation with value-added based models

> *"... aggregate value added is clearly appropriate for focusing on the* uses *of output. But ... value added is not generally appropriate for studying productivity growth — that is, attempting to understand the* sources *of output change — since it is not in general a valid production measure."*
> Basu and Fernald (1997, p.254, emphasis in original)

Most of the existing literature either estimates production function/convergence equations or carries out growth accounting exercises using *aggregate value-added* (GDP) as dependent variable (usually in per capita terms). In the 'real world', however, firms use intermediate inputs to produce goods valued at gross output value, and not in units of value-added (Oulton & O'Mahony, 1994; Hulten, 2000; Cobbold, 2003). Being an index number, value-added implicitly assumes perfect competition and constant returns to scale.[41] This implies that both labour and capital need to be (additively) separable from other inputs in the production process — i.e. it is required that the elasticity of substitution between materials and other inputs is zero (Basu & Fernald, 1997). Further, it is assumed that technical change and innovation only affects the value-added equation but not intermediate input use (Gollop, 1979; Baily, 1986; Cobbold, 2003; Baptist & Teal, 2008). Cobbold (2003, p.10) points out that

> *"[m]odern productivity improvement techniques are aimed at improving the efficiency with which both intermediate and primary inputs are used. For example, in the manufacturing sector, just-in-time production, statistical process control, computer-aided design and manufacturing, and other developments reduce error rates and cut down on sub-standard rejected production."*

[41] As Basu and Fernald (1995) develop the construction of real value-added assumes that the marginal product of intermediate inputs is observable from factor payments. As in the cost vs. revenue share concerns (Hall, 1990) raised above if competition is imperfect the marginal product *exceeds* factor payments. Constant returns in the gross output equation is assumed during a normalisation necessary in the construction of VA: $\Delta y \equiv (1 - s_M)\Delta v + s_M \Delta m \Rightarrow \Delta v = (\Delta y - s_M \Delta m)(1 - s_M)^{-1}$ (Basu & Fernald, 1995, p.171f).

Violation of the above assumptions potentially leads to spurious results: with imperfect competition, value-added based specifications suffer from omitted variable bias regardless of the level of aggregation (Basu & Fernald, 1995, 1997). The variation between value-added and gross-output based results seems to increase for estimation at the sub-sector level where intermediate inputs tend to make up a larger share of gross output than at the aggregate sector or economy level (Cobbold, 2003). Another, and by no means final, source of bias in a value-added specification comes from outsourcing: increased use of outsourcing will raise the growth rate of value-added even if there is no change in gross output productivity (Baptist & Teal, 2008).[42]

It seems rather intuitive that a value-added based approach may lead to biased TFP growth estimates in the face of efficiency improvements in the use of intermediate inputs. A gross-output based measure for TFP growth is thus deemed a better indicator of the full extent of disembodied technological change and a number of empirical studies at lower levels of aggregation (manufacturing subsectors) have adopted this approach (Oulton & O'Mahony, 1994; Harrigan, 1999; Hulten, 2000). Within the mainstream empirical growth literature, as pointed out above, the use of the PWT dataset has predetermined a focus on aggregate economy analysis based on value-added specifications, rather than a sector-based approach using gross-output and intermediate inputs.

Our general empirical production function framework does not preclude any possible impact of misspecification with regard to unit of output in cross-country growth analysis. The question whether the level of aggregation of the data (economy-level, sector-level) 'favours' the use of an empirical specification (convergence regression, production function) is beyond the scope of this thesis. For the present purposes, we suggest that empirical production analysis should compare the results for gross-output and value-added based specifications, so as to investigate whether there are significant differences.

[42] If a firm outsources an internal process then it will reduce its capital, labour and intermediate/material (K, L, M) inputs according to the production function that applies to this sub-process. The outsourced component is then purchased back as intermediate input. Thus the firm will have lower K and L, but higher M,... with the same level of output. As a result its value-added will have changed.

Chapter 2

Analysing Heterogeneity in Global Production Technology and TFP: The Case of Manufacturing

Abstract:

In this chapter we carry out an empirical investigation of the impact of parameter heterogeneity, cross-section dependence and variable nonstationarity on production function estimation using a long macro panel dataset for manufacturing. We develop a two-stage estimation strategy that can be applied to 'diverse' panels where some country time-series display nonstationarity while others do not: the Augmented Mean Group (AMG) estimator allows for parameter heterogeneity while at the same time accounting for unobserved common effects, similar to the Pesaran (2006) Common Correlated Effects estimators. Our empirical results indicate considerable changes in the estimated parameters when moving from pooled models to country regressions, between equations in levels and first differences, and between estimators which account for unobserved common factors and others that do not. Diagnostic testing provides strong support for technology heterogeneity across countries and we further find that modelling Total Factor Productivity (TFP) as possessing both common and country-idiosyncratic elements is supported by the data: manufacturing innovation has the ability to 'lift all boats', but to a different extent across countries. We link our measure of common TFP evolution to important historical events over the thirty-three-year sample period. In the light of these findings the interpretation of regression intercepts as TFP levels collapses and we suggest an alternative measure which allows for technology heterogeneity. Our results are supported when we adopt a nonstationary panel econometric approach allowing for reverse causality.

> *"...within the neoclassical framework of a common aggregate Cobb-Douglas production function an appeal to unobserved factors of production or different rates of technology absorption alone is unlikely to be sufficient to explain the persistent cross-country differences in per capita income that are observed. Rather ..., empirically, aggregate production technologies differ significantly among countries, and ... within the neoclassical framework accounting for these heterogeneities is necessary in order to explain observed patters of per capita income divergence."*
> Pedroni (2007, p.431/2)

In the previous chapter we introduced two general empirical frameworks to analyse cross-country production data at the macro level. A selective review of the growth empirics literature showed that existing empirical approaches (using aggregate or sectoral data) are encompassed by our suggested models.

In this chapter we turn to a practical implementation of the general empirical frameworks set out in Chapter 1. Informed by our previous discussion of nonstationarity, cross-section dependence and parameter heterogeneity, we use production data for the manufacturing sectors of 48 countries to investigate how results amd diagnostics change across a range of estimators and specifications, and whether this provides any insights into the validity of assumptions implicit in the empirical strategy. We introduce the various steps of analysis in the following.

Section 2.1 re-introduces the general empirical framework for cross-country production functions, highlights its features and discusses how this could be empirically implemented: we present standard and novel estimation strategies focused on unbalanced panel data sets with a substantial time-series element, covering diverse sets of developing, developed and emerging economies. We develop a two-stage regression approach which is argued to suit the analysis of data characterised by variable nonstationarity in some, but not all countries or groups. In addition we propose the recently developed Common Correlated Effects Mean Group estimator (Pesaran, 2006) as a viable cross-check for results from our two-stage approach. Both estimation strategies allow for technology parameter heterogeneity, nonstationary variables and cross-section dependence, based on an unobserved common factor modelling approach. Furthermore, these empirical estimators can deal with the identification problem arising from the commonality of unobservables in driving the evolution of both output and inputs, a characteristic that leads to serious difficulties in standard panel estimators.

Section 2.2 then uses an unbalanced macro panel dataset for manufacturing (UNIDO, 2004) to estimate production functions (48 countries, 1970 to 2002). We begin with an analysis of the time-series properties of the annual data, using tests from the time-series and panel data literatures. Our main point is to show that while we cannot provide evidence for nonstationarity in *all* country variable series, we are confident in our claim that country time-series *cannot all* be seen as stationary. The pooled production function estimations provide evidence of the impact of these time-series properties as well as the empirical misspecification of observable and unobservable processes. Our results illustrate considerable changes when moving between equations in levels and first difference, and from pooled models to country regressions. Based on residual diagnostic tests we discuss how these patterns can be attributed to the interplay of nonstationarity, cross-section dependence and parameter heterogeneity in the estimation equation(s). Interestingly, the choice between gross output and value-added based empirical models does not seem to matter at this level of aggregation.

In Section 2.3 we carry out a number of formal parameter heterogeneity tests to provide further evidence for the central finding of our empirical results. Section 2.4 then turns to a discussion of TFP level estimates in the light of the parameter heterogeneity finding and offers a new procedure to adjust TFP level estimates so as to allow for comparison across countries and between pooled and heterogeneous parameter estimators. A final empirical section investigates the robustness of our results to the use of nonstationary panel econometric methods thus testing for the impact of reverse causality on our results. A brief conclusion summarises the findings and discusses their implications for cross-country growth empirics.

2.1 Empirical implementation of a production function for cross-country analysis

We adopt the production function model and assumptions described in detail in Chapter 1. For sake of generality and to highlight the properties of some of the estimators discussed in this section we generalise the model as follows: for $i = 1, \ldots, N$ and $t = 1, \ldots, T$, let

$$y_{it} = \boldsymbol{\beta}_i' \boldsymbol{x}_{it} + u_{it} \qquad u_{it} = \alpha_i + \boldsymbol{\lambda}_i' \boldsymbol{f}_t + \varepsilon_{it} \tag{2.1}$$

$$x_{mit} = \pi_{mi} + \boldsymbol{\delta}_{mi}' \boldsymbol{g}_{mt} + \rho_{1mi} f_{1mt} + \ldots + \rho_{nmi} f_{nmt} + v_{mit} \tag{2.2}$$

$$\text{where} \qquad m = 1, \ldots, k \qquad \text{and} \qquad \boldsymbol{f}_{\cdot mt} \subset \boldsymbol{f}_t$$

$$\boldsymbol{f}_t = \boldsymbol{\varrho}' \boldsymbol{f}_{t-1} + \boldsymbol{\epsilon}_t \qquad \text{and} \qquad \boldsymbol{g}_t = \boldsymbol{\kappa}' \boldsymbol{g}_{t-1} + \boldsymbol{\epsilon}_t \tag{2.3}$$

For the analysis of manufacturing in this chapter y_{it} represents gross output or value-added and $\boldsymbol{x}_{it}$ is a vector of observable factor-inputs including labour, capital stock and in the gross-output specification materials (all in logarithms).[1] Technology parameters $\boldsymbol{\beta}_i$ can differ across countries but are assumed constant over time.[2] For unobserved TFP we employ the combination of a country-specific TFP level α_i and a set of common factors $\boldsymbol{f}_t$ with country-specific factor loadings $\boldsymbol{\lambda}_i$ — TFP is thus in the spirit of a 'measure of our ignorance' (Abramowitz, 1956) and operationalised via an unobserved common factor representation. In equation (2.2) we further add an empirical representation of the k observable input variables, which are modelled as linear functions of unobserved common factors $\boldsymbol{f}_t$ and $\boldsymbol{g}_t$, with country-specific factor loadings respectively. The model setup thus introduces cross-section dependence in the observables and unobservables. As can be seen, some of the unobserved common factors driving the variation in y_{it} in equation (2.1) also drive the regressors in (2.2). This setup leads to endogeneity whereby the regressors are correlated with the unobservables of the production function equation (u_{it}), making $\boldsymbol{\beta}_i$ difficult to identify. Equation (2.3) specifies the evolution of the unobserved factors.

[1] As a reminder, parameter estimates and interpretation will differ between a value-added based and gross-output based empirical specification, but if we assume constancy of the material-output ratio we can transform results to make them directly comparable: $\boldsymbol{\beta}_i^{va} = \boldsymbol{\beta}_i/(1 - \boldsymbol{\gamma}_i)$ (Söderbom & Teal, 2004).

[2] The latter assumption is clearly restrictive, but given the focus on cross-country technology heterogeneity against the background of data restrictions in the time-series dimension we cannot relax this assumption for the heterogeneous regression models. For the pooled models we ran separate regressions using pre- and post-1985 subsamples for the value-added based models. Parameter estimates for the POLS, CCEP and FD-OLS estimators are virtually identical for the two sub-periods. Separate period estimates for the FE estimator differ somewhat but the 95% confidence intervals still show considerable overlap.

We maintain the following assumptions for the general production function model and the data it is applied to:

A.1 The $\boldsymbol{\beta}_i$ parameters are unknown random coefficients with fixed means and finite variances. Similarly for the unknown factor loadings, i.e. $\boldsymbol{\lambda}_i = \boldsymbol{\lambda} + \boldsymbol{\eta}_i$ where $\boldsymbol{\eta}_i \sim \text{iid}(0, \Omega_\eta)$.[3]

A.2 Error terms $\varepsilon_{it} \sim N(0, \sigma^2)$, where σ^2 is finite. Similarly for v_{mit} and ϵ_t.

A.3 Observable inputs $\boldsymbol{x}_{it}$ and output y_{it}, as well as the unobserved common factors $\boldsymbol{f}_t$ and $\boldsymbol{g}_t$ are <u>not</u> *a priori* assumed to be stationary variables/processes.

A.4 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.5 There is an overlap between the unobserved common factors driving output and the regressors ($\boldsymbol{f}_{\cdot mt} \subset \boldsymbol{f}_t$), creating difficulties for the identification of the technology parameters $\boldsymbol{\beta}_i$.

We have adopted the production function framework, rather than a Mankiw et al. (1992) type convergence regression equation. This choice is made with the empirical parts of this and the following chapter in mind, which employ data from the manufacturing and agriculture sectors respectively. In keeping with data realities, we assume an unbalanced panel dataset where some countries display nonstationary processes in inputs and output while others do not. The two most important features of this setup are the potential nonstationarity of observables and unobservables (y_{it}, $\boldsymbol{x}_{it}$, $\boldsymbol{f}_t$, $\boldsymbol{g}_t$), as well as the potential heterogeneity in the impact of observables and unobservables on output across countries (α_i, $\boldsymbol{\beta}_i$, $\boldsymbol{\lambda}_i$). Taken together these properties have important bearings on estimation and inference in macro panel data which are at the heart of this thesis.

In the remainder of this section we first discuss the investigation of variable time-series and cross-section correlation properties. We then present the empirical estimators considered in this thesis, including the novel Augmented Mean Group estimator we develop. Our discussion will be organised to highlight the treatment of time-series properties of the data and of heterogeneity in the impact of observables and unobservables for each of the estimators.

[3] The assumption of random coefficients is for convenience. Based on the findings by Pesaran and Smith (1995, footnote 2, p.81) the coefficients could alternatively be fixed but differing across groups.

2.1.1 Data time-series and cross-section correlation properties

A detailed discussion of available unit root and cointegration tests in the time-series and panel setup is beyond the scope of this chapter. A panel of considerable time-series dimension T opens up the opportunity to use both country-specific tests from the time-series literature as well as panel-based tests — see Enders (2004), Hamilton (1994) or Maddala and Kim (1998) for a discussion of the former, and Choi (2007), Basile, Costantini, and Destefanis (2005), Gengenbach, Palm, and Urbain (2008) or Westerlund and Breitung (2009) for a discussion of the latter. Standard time-series unit root tests are well-known to have low power, such that as the autoregressive parameter approaches unity the test is less and less able to distinguish between a true unit root (the null hypothesis of the test) and a false null.

> *"One of the primary reasons behind the application of unit root tests ... to a panel of cross-section units was to gain statistical power and to improve on the poor power of their univariate counterparts."* (Breitung & Pesaran, 2005, p.2)

Many panel unit root and cointegration tests discussed in the literature however cannot shake off an inherent difficulty in terms of interpretation, whereby the null of nonstationarity (or stationarity or cointegration or noncointegration) for *all* countries in the panel is contrasted with an alternative that *at least one* country is stationary (or nonstationary or noncointegrated or cointegrated). Panel unit root test results are often highly sensitive to the number of lags (of the differenced dependent variable) included and it is unclear whether the lag structure should be the same across countries or whether information criteria should inform on individual countries' ideal lag structure. An alternative testing approach which allows to pinpoint stationary and nonstationary countries (Breuer et al., 2001) was found to be highly sensitive to sample selection (Ford et al., 2006), limiting its applicability. More recently it was also suggested that correlation across units of the panel may bias results from 'first generation' panel unit root tests (Pesaran, 2007). Since the seminal contribution by Bai and Ng (2004) the 'second generation' literature for these tests which allows for cross-section dependence has been developing rapidly, with most recent contributions integrating the impact of structural breaks (Westerlund, 2006) and changes in the volatility of the innovations (Hanck, 2008).

Rather than making the data fit the requirements of the estimators, as practiced in Pedroni (2007) for the nonstationary panel econometric approach, we postulate the use of estimation methods which are robust to the potential for nonstationarity and cointegration within some, but not all countries in the panel. This approach is less dependent on making assumptions

about the data which are difficult to test in relatively short panels.[4] Further recall that unit root evolution should not be seen as continuing indefinitely ('global' property) but represents a 'local' behaviour of variables *within-sample* (Pedroni, 2007).

The formal analysis of cross-section dependence (CSD) in macro panels has only been introduced over the past few years, and while the spatial econometric literature has long had measures for spatial autocorrelation (e.g. Moran's I, following Moran, 1950) there are only limited means to measure cross-section correlation in panels *without* imposing significant structure on the nature of the relationship. Following Coakley et al. (2006) two common measures applied to investigate the presence of cross-section correlation are based on averaged correlation coefficients and principal component analysis (PCA). For the former we can simply correlate the variable or regression residual series for each country with that of every other country in the sample and then compute the mean correlation coefficient of all $N(N-1)$ pairwise correlations. Since correlation coefficients are bounded $[-1, 1]$ it is commonly more meaningful to obtain the mean *absolute* correlation coefficient. For the PCA approach we can investigate the principal components[5] driving the variable or residual series across the sample countries. Computing the proportion of the variance accounted for by the first or first and second principal components can then act as an indication for the underlying factor structure of the data or whether cross-section correlation in the residuals is particularly prevalent.

At present one of the few formal tests for cross-section correlation in macro panels is the CD test introduced by Pesaran (2004). In contrast to the spatial econometric approach discussed above, this simple test does not require the specification of some type of weight matrix and is applicable to a wide range of macro panel models. It is robust to the presence of nonstationary processes, parameter heterogeneity or structural breaks, and was shown to perform well even in small N, T samples. The test statistic is based on the normalised average of all pair-wise correlation coefficients of country variable series or residual series (for details see Section 3.C in the Appendix of Chapter 3).

[4] Here and in the following we judge the length of a panel with the eyes of a time-series econometrician, for whom annual data for, say, 25 to 35 years represents a relatively short T-dimension.

[5] Principal Component Analysis can be best thought of as revealing the internal structure of the data in a way which best explains its variance. The principal components represent a series of uncorrelated linear combinations of the variable time-series (usually multiple variables, but in our case one variable or residual series for multiple countries) that contain most of the variance.

Due to the problems of data coverage encountered in the manufacturing dataset analysed in this chapter we do not carry out any diagnostic tests for cross-section correlation. The formal tests for CSD were introduced at this point to highlight that "[w]hen 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 & Edgerton, 2008, p.666, emphasis added). Our prior belief as expressed in our general empirical model is that cross-section dependence is to be expected in our data. Data properties for the agriculture data analysed in Chapter 3 allows us to implement the formal testing procedures introduced here.

2.1.2 Empirical estimators

As was highlighted in Chapter 1 this thesis emphasises the importance of parameter and factor loading heterogeneity across countries. The following 2×2 matrix indicates how the various estimators discussed and implemented below account for these matters.

		Factor loadings: homogeneous	heterogeneous
Technology parameters:	homogeneous	POLS, FE, FD-OLS	CCEP
	heterogeneous	MG, RCM, GM-FMOLS	AMG, ARCM, CCEMG

We begin our discussion with the standard panel estimators which impose homogeneity on both technology parameters and factor loadings — pooled OLS (POLS), Fixed Effects (FE), First Difference OLS (FD-OLS) — and then move on to the Mean Group (MG) and Swamy Random Coefficient Model (RCM) estimators which allow for technology parameter heterogeneity. The Pedroni (2000) Group-Mean Fully-Modified OLS (GM-FMOLS) estimator is also discussed briefly. Next up we introduce the Pesaran (2006) Common Correlated Effects estimator in its pooled version (CCEP), which allows for factor loading heterogeneity. Finally our own Augmented Mean Group (AMG) and Augmented Swamy (ARCM) estimators, as well as the Mean Group version of the CCE estimator (CCEMG) are discussed in great detail. These augmentation approaches allow for heterogeneity in technology parameters *and* factor loadings. Our discussion of estimators thus follows a path of gradual relaxation of homogeneity assumptions on technology parameters and factor loadings within the general empirical framework laid out

in equations (2.1) to (2.3).

All regression models discussed in this section assume well-behaved error properties for consistency — we raise where this is likely not to hold from a theoretical perspective. Our empirics then investigate residual diagnostics for each specification in some detail.

Pooled estimators

A typical starting point for empirical analysis of panel data are the pooled OLS (POLS) and Fixed Effects (FE) estimators. The latter can be implemented either via LSDV or variables in deviations from countries' mean values.[6] These models provide estimates of the common technology parameters, common TFP evolution over time, and (in the Fixed Effects case) country-specific TFP levels. The regression models for these estimators are

$$\text{POLS} \qquad y_{it} = a + \boldsymbol{b}'\boldsymbol{x}_{it} + \sum_{t=2}^{T} c_t D_t + e_{it} \tag{2.4}$$

$$\text{FE} \qquad y_{it} = a + \boldsymbol{b}'\boldsymbol{x}_{it} + \sum_{t=2}^{T} c_t D_t + \sum_{i=2}^{N} a_j + e_{it} \tag{2.5}$$

where $\boldsymbol{b}$ is the vector of *common* parameter coefficients on the observable inputs, D_t is a vector of $(T-1)$ year dummies with correspondent parameters c_t, a is the intercept and the a_j are $(N-1)$ additional intercepts in the FE case.

The above statements hinge on the underlying assumption of common technology parameters $\boldsymbol{\beta}_i = \boldsymbol{\beta}$ (and common factor loadings $\boldsymbol{\lambda}_i = \boldsymbol{\lambda}$) in the general empirical model. If technology parameters are assumed to differ across countries the pooled estimates $\boldsymbol{b}$ provide estimates of the means of the $\boldsymbol{\beta}_i$, appealing to the random coefficient property.[7] Similarly for the factor loadings. As was shown in Section 1.3 of Chapter 1 there is however a serious identification problem with regard to the underlying $\boldsymbol{\beta}$ (or mean of the $\boldsymbol{\beta}_i$) if unobserved common factors drive both inputs *and* output.

[6] ($\tilde{X}_{it} = X_{it} - T^{-1}\sum_t X_{it}$, henceforth: mean-deviations.

[7] If country technology parameters were distributed in two discrete 'clubs' the pooled estimates would provide a weighted mean of the underlying $\boldsymbol{\beta}_i$ where the subsample-sizes act as weights — see Chapter 4 for simulations of this setup.

These approaches furthermore neglect any influence of variable time-series properties on the consistency of the estimates (see our discussion in Section 1.2 of Chapter 1): if variable series are nonstationary these estimators require output y and factor-inputs $\boldsymbol{x}$ to be homogeneously cointegrated. Once variable series are nonstationary and technology parameters *differ* across countries the pooled regressions in levels lead to nonstationary error terms. Under standard assumptions the nonstationarity in the error terms leads to the breakdown of the cointegrating relationship and manifests itself in spurious regression results (Engle & Granger, 1987).

Given the breakdown of cointegration in the presence of nonstationary error terms, we would expect *any* pooled panel estimation to yield spurious results if variables were nonstationary and parameters heterogeneous across countries. A recent development in econometric theory (Phillips & Moon, 1999) however implies that this conclusion needs qualification. We do believe that the empirical results observed question the applicability/validity of the latter findings in the type of macro panels investigated here and refer to Appendix 2.C for further discussion. Crucially, once residuals are nonstationary the conventional standard errors on $\hat{\boldsymbol{b}}$ will be invalid, and the associated standard t-statistics will diverge *away from zero* with probability one asymptotically (Kao, 1999).

The year dummies in the above panel models can capture the impact of unobserved common factors provided the factor loadings are the same across countries ($\boldsymbol{\lambda}_i = \boldsymbol{\lambda}$). Regardless of whether variables and common factors are stationary or nonstationary, the failure to account for potential heterogeneity in the factor loadings will lead to inconsistency and incorrect inference for these standard panel econometric models in the present empirical setup (Phillips & Sul, 2003; Pesaran, 2006) — we refer to the discussion below and our simulations in Chapter 4 for more details.

Thus the pooled OLS and Fixed Effects regressions in equations (2.4) and (2.5) *by construction* yield nonstationary error terms if the regressors are nonstationary and production technology differs across countries. Simply differencing the model equation renders *all elements* stationary, such that we can apply OLS (FD-OLS) to compute an estimate of the common technology parameters or an unweighted mean of the country-specific cointegrating coefficients (Smith & Fuertes, 2007), depending on the assumptions about the underlying technology parameters $\boldsymbol{\beta}_i$.

Common or mean TFP evolution can be implemented via a set of $T-1$ year dummies. The underlying common factors (in the levels equation) can be either stationary or nonstationary, since differencing renders them I(0). The guiding principle for this approach is expressed by Smith and Fuertes (2004, p.40, change in notation for consistency):

> *"... if it is known that [the regression error terms are I(1)] in most cases the sensible procedure would be to use first differenced data which will produce $\sqrt{T}$-consistent (individual OLS) estimates of $\boldsymbol{b}_i$ and $\sqrt{NT}$-consistent (pooled OLS) estimates of a common $\boldsymbol{b}$ or mean of $\boldsymbol{b}_i$."*

Two matters are particularly noteworthy: firstly, as is clear from the above quote the FD-OLS estimates do *not* represent short-run coefficients. The first differencing is a means to deal with the nonstationarity of levels residuals, not to limit the dynamics. Secondly, the FD-OLS estimates represent common technology parameters or means of individual country parameters, depending on the underlying assumptions about the true technology parameters $\boldsymbol{\beta}_i$. A similar argument can be made for the factor loadings — note here that our Monte Carlo simulations suggest that the FD-OLS augmented with year dummies is *not* subject to the identification problem encountered by its levels counter-parts. FD-OLS may be biased for other reasons, which we highlight and address in Chapter 4. The pooled regression model in first differences is thus

$$\text{FD-OLS} \qquad \Delta y_{it} = \boldsymbol{b}'\Delta \boldsymbol{X}_{it} + \sum_{t=2}^{T} c_t \Delta D_t + \Delta e_{it} \tag{2.6}$$

where we have technology parameters $\boldsymbol{b}$ and $T-1$ year dummies in *first differences*[8] ΔD_t with corresponding parameters c_t.

The emphasis of this estimation approach is by design on common or average TFP growth across countries — simple country-specific regressions would ascribe all of the unexplained growth to TFP evolution in the country, whereas a pooled regression can make use of the cross-section variation in the data to identify common or mean evolution across countries. The inclusion of year dummies is widespread in the econometric analysis of micro-panels; as was suggested above the inclusion of year dummies in a pooled regression model in first differences can go some way to account for cross-section dependence and to solve the identification problem for $\boldsymbol{\beta}$ raised in Section 1.3 of Chapter 1.

[8] Including dummies in first differences may seem awkward, this is however easily implemented via the `xtivreg` command in `Stata` with the options `fd` (and `regress` to run OLS rather than 2SLS). The advantage of this practical implementation using *first differenced* dummies is that the associated parameters c_t describe a *level* evolution over time t, whereas parameter estimates on standard *level* dummies in a growth regression provide a vector of year-on-year growth rates.

Standard Mean Group estimators

Pesaran and Smith (1995) introduced the *Mean Group* (MG) estimator for the study of stationary panels. This estimator represents a simple average of country-regression estimates

$$\text{MG} \qquad y_{it} = a_i + \boldsymbol{b}_i' \boldsymbol{X}_{it} + c_i t + e_{it} \tag{2.7}$$
$$\hat{\boldsymbol{b}}_{MG} = N^{-1} \sum_i \hat{\boldsymbol{b}}_i$$

While the individual country estimates $\hat{\boldsymbol{b}}_i$ may not be entirely reliable due to the moderate T dimension in macro panels, the Mean Group estimate $\hat{\boldsymbol{b}}_{MG}$ is a robust indicator of the central tendency of the parameter coefficients across countries.[9] By construction the MG estimator does not impose parameter homogeneity on the model (Smith & Fuertes, 2007) and thus allows for formal testing of heterogeneity. The coefficient on the country trend c_i can capture omitted variables and processes which evolve in a linear fashion.

In a nonstationary panel this estimator represents an unweighted average of estimated cointegrating coefficients and thus requires the existence of a cointegrating relationship within each country equation (Phillips & Moon, 2000). If cointegration is homogeneous ($\boldsymbol{b}_i = \boldsymbol{b} \forall\, i$) the MG estimator is consistent but inefficient (relative to the pooled model). In the case of heterogeneous cointegration, the individual $\hat{\boldsymbol{b}}_i$ is estimated consistently in the time-series regression as $T \to \infty$. Subsequent averaging over N provides a $T\sqrt{N}$-consistent estimate of the mean of the cointegrating relations across countries (Smith & Fuertes, 2007).

The validity of these statements hinges on each country regression correctly specifying the cointegrating relationship. With reference to cross-country production function estimation, we need to stress that correct specification of the TFP evolution in each country regression is a crucial requirement for the MG estimator to provide an unbiased estimate of the mean of the $\boldsymbol{b}_i$. TFP is potentially nonstationary (Lee et al., 1997; Bond et al., 2007; Canning & Pedroni, 2008) and thus may form part of the cointegrating vector. If nonstationary TFP is misspecified this will lead to noncointegration in the individual country regression — in this case the MG estimator yields potentially spurious results.

[9] Since a number of options exist to average individual country estimates (unweighted mean, median, outlier-robust mean), the MG estimator is in effect more versatile and can provide robust averages across sample countries.

The most general estimation approach would be to include year dummies in each country regression, allowing for idiosyncratic TFP processes ($\boldsymbol{\lambda}_i' \boldsymbol{f}_t$, in our model). Subsequent comparison of the dummy coefficients across countries could provide insights as to whether the dynamics of TFP are common or idiosyncratic. Of course, this is not possible, given the dimensionality problem in the individual country regression. Transforming all variables into deviations from the cross-section mean at time t (henceforth: 'demeaning') is sometimes raised as equivalent to year dummies in accounting for common dynamic processes, with particular reference to the MG estimator (Lee et al., 1997). Note however that cross-sectional demeaning is equivalent to year dummies *if and only if* all regression coefficients are homogeneous across countries (Pedroni, 1999) — otherwise, using data in deviations from the cross-sectional mean introduces additional (nonstationary) terms into the country-regression residuals (Coakley et al., 2006).[10] Inclusion of a linear trend in each country regression to capture the impact of unobserved factors and thus TFP evolution saves on degrees of freedom, but leads to noncointegration if the 'true' TFP process is nonstationary and part of the cointegrating relationship. The MG estimator can only provide a consistent estimate of the average cointegrating relationship across countries if the (idiosyncratic or common) TFP process(es) are modeled correctly and countries cointegrate heterogeneously. If the TFP process is nonstationary and follows a common factor structure, we cannot detect it by simply using country regressions.

A closely related estimator which provides a variation on weighting country estimates is the *Swamy Random Coefficients Model* (RCM) estimator. This represents a feasible GLS estimator, which is equivalent to using a weighted average of the individual OLS country estimates (Swamy, 1970) — the weights are measures of precision of the individual country estimate. Our critique of the MG estimator naturally extends to the Swamy RCM estimator.

For both Mean Group and Swamy estimators, it is thus crucial that individual country equations cointegrate in the presence of nonstationary variable series. Identifying the cointegrating relationship may prove difficult, as the literature suggests a large number of candidate variables. Note that the cointegrating relationship by definition only includes nonstationary variables — stationary variables omitted from the model will need to be dealt with separately, since in time-series of moderate length their impact is assumed to be non-negligible.

[10] If variables are nonstationary these additional terms (and thus the residuals) will also be nonstationary. If variables are stationary but serially correlated these terms will lead to serial correlation in the residuals (Pesaran & Smith, 1995).

In analogy to the pooled regression models we can suggest the Mean Group and Swamy RCM estimators with variables in first differences. In these models variable nonstationarity no longer plays any role since regressors and unobserved factors are assumed to be stationary *by construction*. However, the presence of unobserved common factors, now $\Delta \boldsymbol{f}_t$, and thus cross-section dependence in the unobservables continue to influence estimation and inference (Phillips & Sul, 2003; Pesaran, 2006). In essence, whether adopting linear trends or cross-sectionally demeaned data, these country regression estimators (in levels or first difference) have no means to account for the presence of flexible (non-linear) common factors with heterogeneous impact across countries and thus suffer from an omitted variable problem.

An alternative estimator based on the Mean Group principle is the nonstationary panel econometric approach by Pedroni (2000). This makes use of the nonstationarity and cointegration properties of the data and averages country regression estimates using the 'Group Mean Fully-Modified OLS' (Group-Mean FMOLS) estimator. The FMOLS estimator originates from the time-series literature (Phillips & Hansen, 1990) and requires that the country variable series in the sample are nonstationary and cointegrated. The 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): it exploits the fact that under cointegration the bias created by variable endogeneity is of second order, and the first differenced regressors can serve as instruments for this bias to construct bias-adjusted variables (Pedroni, 2007). Country-specific intercept and trend terms can account for stationary variables and processes omitted from the model. The empirical strategy is thus to select a sample to suit the requirements of the estimation method since otherwise the desirable properties of the Pedroni (2000) estimator (prime amongst these superconsistency) cannot be assumed to hold. A more detailed discussion of this estimator is beyond the scope of this chapter, although we can highlight that the method assumes cross-section independence for consistency.

We also want to highlight that the notion of heterogeneous technology parameters has important bearings on the *interpretation of regression intercepts* as TFP level estimates: pooled estimators allow for the inclusion of country fixed-effects, which are routinely interpreted as TFP-level estimates. This interpretation is valid since all estimates (on factor-inputs, intercepts, year dummies or trends, etc.) are obtained from a single linear regression, such that the *ceteris paribus* property applies. Abstracting from TFP growth over time for simplicity

of illustration, assume a single-input empirical production function with country fixed effects $y_{it} = a_i + b\,x_{it} + e_{it}$. Holding the common slope parameter b fixed, the relative ordering and magnitude of the intercept terms a_i represents the 'shifter' for the production function associated with the TFP level. Once we allow for heterogeneous technology parameters across countries $y_{it} = a_i + b_i x_{it} + e_{it}$, this *ceteris paribus* property of regression parameters breaks down: *the intercept terms can no longer be interpreted in isolation from the slope coefficients*, since geometrically a steeper slope (larger positive value for b_i) forces the regression intercept downwards. TFP still acts as a shifter, but it can no longer be compared across countries unless slopes happen to be the same ($b_i = b_j$). We discuss this matter in more detail using empirical examples in Section 2.4 below.

Finally, it bears reminding the reader that *none of the above estimators* (pooled, MG or GM-FMOLS) can identify $\boldsymbol{\beta}$ (or the mean of the $\boldsymbol{\beta}_i$) in the case where the same unobservable factors drive both output *and* inputs: as we have shown in Section 1.3 of Chapter 1, the technology parameter is not separately identified from the impact of the common factors, as represented by the factor loadings $\boldsymbol{\lambda}_i$. This statement holds regardless of whether 'true' technology parameters are heterogeneous or common, factor loadings are heterogeneous or common,[11] and whether y and $\boldsymbol{x}$ are cointegrated or not — *unless* the cross-country average loadings are zero ($\bar{\boldsymbol{\lambda}} = 0$). We now turn to a number of estimators which address this identification problem.

Pooled estimators allowing for cross-section dependence

Building on the principal component analysis approach adopted in Coakley, Fuertes, and Smith (2002), the pooled version of the Common Correlated Effects estimator (CCEP) developed by Pesaran (2006) accounts for unobserved common factors and cross-sectional dependence in the panel by including N sets of cross-section averages of *all* observable variables ($\bar{y}_t$, $\bar{\boldsymbol{x}}_t$) in a pooled Fixed Effects regression equation (i.e. interacted with country dummies). The unobserved common factors are thus allowed to change over time and to have differential impact

[11] It can be shown quite easily that neither parameter nor factor loading heterogeneity are necessary for this result: provided $\boldsymbol{\lambda} \neq 0$ we have

$$\begin{aligned} y_{it} &= \alpha_i + \boldsymbol{\beta}\boldsymbol{x}_{it} + \boldsymbol{\lambda}\boldsymbol{f}_t + \varepsilon_{it} \qquad \text{and for simplicity} \qquad \boldsymbol{x}_{it} = \boldsymbol{\rho}\boldsymbol{f}_t + v_{it} \\ \Leftrightarrow y_{it} &= \alpha_i + \boldsymbol{\beta}\boldsymbol{x}_{it} + \boldsymbol{\lambda}\boldsymbol{\rho}^{-1}(\boldsymbol{x}_{it} - v_{it}) + \varepsilon_{it} \\ \Leftrightarrow y_{it} &= \alpha_i + (\boldsymbol{\beta} + \boldsymbol{\lambda}\boldsymbol{\rho}^{-1})\,\boldsymbol{x}_{it} + (\varepsilon_{it} + \boldsymbol{\lambda}\boldsymbol{\rho}^{-1} v_{it}) \end{aligned}$$

such that we cannot identify $\boldsymbol{\beta}$ separately from $(\boldsymbol{\beta} + \boldsymbol{\lambda}\boldsymbol{\rho}^{-1})$.

on individual countries. Technology parameters on observables are assumed to be common ($\boldsymbol{b}$) in this regression model. Country fixed effects are required due to the asymptotic theory of the estimator, but are operationally viewed as nuisance parameters. While the original Pesaran (2006) study introducing the CCE estimators assumed *stationarity* in the unobserved common factors (in our general model: $\boldsymbol{f}_t$ and $\boldsymbol{g}_t$, and thus also $\boldsymbol{x}_{it}$ and y_{it}) an extension by Kapetanios et al. (2009) has established that the augmentation approach yields a consistent estimate of $\boldsymbol{\beta}$ or the mean of the $\boldsymbol{\beta}_i$ when observables and unobservables are *nonstationary*.

The pooled CCE estimator is thus

$$\begin{aligned} \text{CCEP} \qquad y_{it} &= a + \boldsymbol{b}'\boldsymbol{x}_{it} + \sum_{j=2}^{N} a_j D_j \\ &\quad + \sum_{j=1}^{N} d_{1i}\{\bar{y}_t D_j\} + \sum_{j=1}^{N} \boldsymbol{d}'_{2i}\{\bar{\boldsymbol{x}}_t D_j\} + e_{it} \end{aligned} \tag{2.8}$$

where the second line contains the cross-section averages at time t for each of the variables interacted with a set of N country dummies — due to this interaction the coefficients d_{1i} and $\boldsymbol{d}'_{2i}$ are *country-specific*, as indicated by the subscript i. Thus the cross-section averages can take different parameters for each country i and the estimation equation has a total of $k + (k+2)N$ regressors.[12]

The intuition why this rather simple augmentation can do away with the impact of unobserved common factors (with heterogeneous factor loadings) is as follows: assuming homogeneous technology parameters we take the empirical production function model in equation (2.1), compute the cross-section average in each time period t and solve for the unobserved common factors

$$y_{it} = \boldsymbol{\beta}'\boldsymbol{x}_{it} + \alpha_i + \boldsymbol{\lambda}'_i \boldsymbol{f}_t + \varepsilon_{it} \tag{2.9}$$

$$\bar{y}_t = \bar{\boldsymbol{\beta}}'\bar{\boldsymbol{x}}_t + \bar{\alpha} + \bar{\boldsymbol{\lambda}}'\bar{\boldsymbol{f}}_t \qquad \text{since} \qquad \bar{\varepsilon}_t = 0 \tag{2.10}$$

$$\Leftrightarrow \quad \bar{\boldsymbol{f}}_t = \bar{\boldsymbol{\lambda}}^{-1}(\bar{y}_t - \bar{\alpha} - \bar{\boldsymbol{\beta}}'\bar{\boldsymbol{x}}_t) \tag{2.11}$$

Equation (2.11) indicates that provided the average impact of each factor across all countries is non-zero ($\bar{\boldsymbol{\lambda}} \neq 0$), the use of cross-section averages of the dependent ($\bar{y}_t$) and independent

[12] k for the covariates, $(k+1)N$ for the cross-section averages of the dependent variable and the covariates interacted with country dummies, and N for country fixed effects.

variables ($\bar{\boldsymbol{x}}_t$) can act as a representation of the unobserved common factors $\boldsymbol{f}_t$, since as the cross-section dimension (N) becomes large $\bar{\boldsymbol{f}}_t \rightarrow \boldsymbol{f}_t$ in probability (Pesaran, 2006). To allow for heterogeneity in the factor loadings this representation must be implemented in the fashion outlined above. Crucially, the augmentation of the regression equation with cross-section averages allows for separate identification of $\boldsymbol{\beta}$ in the case where the same factors drive output and inputs (Kapetanios et al., 2009, p.6). Recall that unless unobserved common factors are somehow accounted for, $\boldsymbol{\beta}$ is *not* identified separately from the country-specific factor loadings $\boldsymbol{\lambda}_i$ — refer to Section 1.3 in Chapter 1 for more details.

Pesaran (2006) also shows that the asymptotic consistency of this estimator is based on any weighted cross-section aggregates ($\bar{Z}_t = \sum_i w_i Z_{it}$), provided the weights w_i satisfy the conditions

$$w_i = O\left(\frac{1}{N}\right) \qquad \sum_{i=1}^{N} w_i \lambda_i \neq 0 \qquad \sum_{i=1}^{N} |w_i| < K \tag{2.12}$$

where K is a finite positive constant (Coakley et al., 2006).

Note that the focus of the Common Correlated Effects estimator is the estimation of consistent $\hat{\boldsymbol{b}}_{CCEP}$ (by accounting for the impact of unobserved common factors) and not the nature of the unobserved common factors or their factor loadings: we cannot obtain an explicit estimate for the unobserved factors $\boldsymbol{f}_t$ or the factor loadings $\boldsymbol{\lambda}_i$, since the *average* impact of the factors ($\bar{\boldsymbol{\lambda}}$) is unknown, s.t. the $\boldsymbol{\lambda}_i$ are treated as nuisance parameters. Further strategies have been suggested to shed light on these issues (Kapetanios et al., 2009).

Mean Group estimators allowing for cross-section dependence

We now introduce a novel estimator, the Augmented Mean Group (AMG) estimator, which accounts for cross-section dependence by inclusion of a 'common dynamic effect' in the country regression. This variable is extracted from the year dummy coefficients of a pooled regression in first differences and represents the levels-equivalent *mean evolution of unobserved common factors across all countries.*[13] Provided that the unobserved common factors form part of the country-specific cointegrating relation, the augmented country regression model encompasses

[13] In practice we implement the first stage of FD-OLS with year dummies in first differences, thus resulting in a level series for the unobserved average common factors, rather than the annual growth rates as would be the case for levels year dummies in a model where inputs and output are in first differences.

the cointegrating relationship, which is allowed to differ across countries.[14]

$$\text{AMG — Stage (i)} \qquad \Delta y_{it} = \boldsymbol{b'}\Delta \boldsymbol{x}_{it} + \sum_{t=2}^{T} c_t \Delta D_t + e_{it} \tag{2.13}$$

$$\Rightarrow \hat{\boldsymbol{c}}_t \equiv \hat{\mu}_t^{\bullet}$$

$$\text{AMG — Stage (ii)} \qquad y_{it} = a_i + \boldsymbol{b}_i'\boldsymbol{x}_{it} + c_i t + d_i \hat{\mu}_t^{\bullet} + e_{it} \tag{2.14}$$

$$\hat{\boldsymbol{b}}_{AMG} = N^{-1} \sum_i \hat{\boldsymbol{b}}_i$$

The first stage represents a standard FD-OLS regression with $T-1$ year dummies in first differences, from which we collect the year dummy coefficients which are relabelled as $\hat{\mu}_t^{\bullet}$. This process is extracted from the pooled regression *in first differences* since nonstationary variables and unobservables are believed to bias the estimates in the pooled levels regressions. There is also some evidence that the identification problem for $\boldsymbol{\beta}$ we discussed earlier is addressed successfully in the FD-OLS. In the second stage this variable is included in each of the N standard country regressions which also include linear trend terms to capture omitted idiosyncratic processes which evolve in a linear fashion over time. Alternatively (not shown) we can subtract $\hat{\mu}_t^{\bullet}$ from the dependent variable, which implies the common process is imposed on each country with unit coefficient.[15] In either case the AMG estimates are then derived as averages of the individual country estimates, following the Pesaran and Smith (1995) MG approach. Based on the results of our Monte Carlo simulations in Chapter 4 we posit that the inclusion of $\hat{\mu}_t^{\bullet}$ allows for the *separate* identification of $\boldsymbol{\beta}_i$ or $\mathbb{E}[\boldsymbol{\beta}_i]$ and the unobserved common factors driving output and inputs ($\boldsymbol{\lambda}_i$), like in the CCE case discussed above. Note that permutations of the Augmented Mean Group estimator which could be implemented in a similar fashion but with more desirable small sample properties may well be developed in further research. For instance, it might be possible to transform the AMG into a 'one-step' estimation procedure whereby an equivalent to $\hat{\mu}_t^{\bullet}$, its parameter and the heterogeneous slope parameter coefficients were estimated *jointly* but in an iterative process, as already implemented in case of the Bai, Kao, and Ng (2009) 'continuous-updated' (CUP) estimators. Specification of a pooled version of the AMG constructed along similar lines may also be of interest.

As an illustration, assume that the cointegrating relationship in country j is made up of gross output, labour, capital, materials and a hitherto undetected phenomenon, say manna from

[14] As we will see below this matches the assumption of the Pesaran (2006) CCEMG estimator (Pedroni, 2007).

[15] Note further that unity is also the cross-country average parameter expected for the coefficients on $\hat{\mu}_t^{\bullet}$: $\bar{d} = N^{-1}\sum_i d_i = 1$.

heaven, all of which are nonstationary I(1) variables. Assume further that our country regression in levels is misspecified, as we failed to include the unobserved manna process. As a result the regression errors are nonstationary and our country regression is potentially spurious. However, what if we include the common dynamic process obtained from the pooled regression in first differences? This represents an average of all the processes we failed to include in the pooled equation in first differences... including manna. Based on this interpretation, a two-stage estimation procedure allows us to provide an estimate of the unobserved factors in the second stage by using information collected in the first stage. We can thus identify the cointegrating relationship in the *augmented* country estimation in levels. Note that we use one omitted process (manna) for illustrative purposes: there is no reason why $\hat{\mu}_t^{\bullet}$ could not represent an estimate for several omitted processes or factors that are part of the cointegrating relationship. An alternative justification for the inclusion of the common dynamic process in the country regression is the assumption that some TFP evolution is truly common to all countries (e.g. non-rival knowledge).

Similar to applying $\hat{\mu}_t^{\bullet}$ in the country equations in levels, we can use $\Delta\hat{\mu}_t^{\bullet}$ in the country equations in first differences. Again we have the option to include this term as an explicit regressor or impose it on the country with unit coefficient. The second stage regressions in the latter case are then

$$\Delta\text{AMG Stage (ii)} \qquad \Delta y_{it} - \Delta\hat{\mu}_t^{\bullet} = \boldsymbol{b}_i'\Delta\boldsymbol{x}_{it} + c_i + \Delta e_{it} \tag{2.15}$$

$$\hat{\boldsymbol{b}}_{\Delta AMG} = N^{-1}\sum_i \hat{\boldsymbol{b}}_i$$

where c_i is a country-specific drift term to capture other omitted variables. By construction all variables and processes in this equation are stationary.

Just like applying this augmentation strategy for the standard MG estimator, we can augment the Swamy RCM estimator to yield the Augmented Random Coefficient Model (ARCM) estimator in levels and first differences.

Conceptually very similar to this augmentation approach is the the Mean Group version of the Pesaran (2006) Common Correlated Effects (CCEMG) estimator. In contrast to its pooled cousin in equation (2.8) this accounts for unobserved common factors in the panel by in-

cluding cross-section averages of the dependent and independent variables *in the individual country regression*, thus allowing for technology parameter heterogeneity across countries. The unobserved common factors ($\boldsymbol{f}_t$, and thus the observables y_{it}, $\boldsymbol{x}_{it}$) can be nonstationary processes (Kapetanios et al., 2009), and can have differential impact on individual countries ($\boldsymbol{\lambda}_i$). Formally, we obtain N country regression equations

$$\text{CCEMG} \qquad y_{it} = a_i + \boldsymbol{b}_i' \boldsymbol{x}_{it} + d_{1i}\bar{y}_t + \boldsymbol{d}_{2i}'\bar{\boldsymbol{x}}_t + e_{it} \tag{2.16}$$

$$\hat{\boldsymbol{b}}_{CCEMG} = N^{-1} \sum_i \hat{\boldsymbol{b}}_i$$

where $\bar{y}_t$ and $\bar{\boldsymbol{x}}_t$ are the cross-section averages at time t for the dependent and independent variables respectively, a_i is an intercept. Optionally a country-specific trend term can be added. The coefficients on $\bar{y}_t$ and $\bar{\boldsymbol{x}}_t$ are not investigated since the underlying factor loadings $\boldsymbol{\lambda}_i$ are treated as nuisance parameters. The same applies to the country-regression intercepts a_i (α_i as nuisance parameters). The identification problem for $\boldsymbol{\beta}$ can be shown not to arise for this estimation approach (Kapetanios et al., 2009). We can construct the N CCE country regressions in first differences in analogy. Like the augmented MG and RCM estimators we introduced, the levels CCEMG estimator requires the unobserved common process to be part of the cointegrating relationship in each country (Pedroni, 2007).

Comparing augmented Mean Group estimators

We have shown the intuition for the CCE estimators when discussing the CCEP above. In the following we want to indicate how our own augmented estimators relate to the Pesaran (2006) estimators. We argue that our augmentation approach uses an *explicit* rather than implicit estimate for $\boldsymbol{f}_t$ from the pooled first stage regression in first differences. From (2.1) we obtain

$$\Delta y_{it} = \beta_i \Delta \boldsymbol{x}_{it} + \boldsymbol{\lambda}_i' \Delta \boldsymbol{f}_t + \Delta u_{it} \tag{2.17}$$

Pooled estimation of this model is carried out using a set of as $T-1$ year dummies in first differences ($\sum_{t=2}^{T} \Delta D_t$) to yield the common or mean evolution of unobservables in levels[16] across countries over time $\hat{\mu}_t^{\bullet}$ — see equation (2.6). Our pooled estimate is some function $h(\cdot)$ of the unobserved common factors $\boldsymbol{f}_t$: $\hat{\mu}_t^{\bullet} = h(\bar{\boldsymbol{\lambda}} \boldsymbol{f}_t)$. It is clear from this discussion that as

[16] Recall that the differencing of the regression equation is carried out to avoid spurious correlation brought about by nonstationary errors in a misspecified levels model with nonstationary observables and unobservables. If we difference the year dummies as well, their coefficients describe a levels process, rather than year-on-year growth.

in the Pesaran (2006) CCE approach we require that the average impact of the unobserved common factors across countries is non-zero ($\bar{\boldsymbol{\lambda}} \neq 0$). Plugging this estimate back into the model in equation (2.1) yields

$$y_{it} = \alpha_i + \boldsymbol{\beta}_i \boldsymbol{x}_{it} + \boldsymbol{\lambda}_i \, h(\bar{\boldsymbol{\lambda}} \boldsymbol{f}_t) + u_{it} \tag{2.18}$$

From equation (2.18) we can see that provided there are no issues related to Δu_{it} in the first stage regression (see Chapter 4 for counter-examples) the estimate $\hat{\mu}_t^{\bullet}$ obtained can be included in the second stage to account for unobserved common factors $\boldsymbol{f}_t$ and allow for their heterogeneous impact on y_{it}. Compared with the CCEMG approach we *do* get an economically meaningful construct from this setup: the common dynamic process $\hat{\mu}_t^{\bullet} = h(\bar{\boldsymbol{\lambda}} \boldsymbol{f}_t)$ represents common TFP evolution over time, whereby common is defined either in the literal sense of the term, or as the sample mean of country-specific TFP evolution. The country-specific regression coefficient on the common dynamic process, d_i in equation (2.14), represents the implicit factor loading on common TFP.

2.1.3 Summary

This section introduced a number of approaches for empirical production function estimation with macro panel data. We discussed the assumptions and implications of pooled estimators and standard estimators allowing for parameter heterogeneity. We also highlighted a novel estimation approach which allows for both parameter heterogeneity and unobserved common factors with heterogeneous factor loadings — the Augmented Mean Group estimator. It is suggested that the AMG approach is closely related to the Pesaran (2006) CCEMG estimator, however a formal investigation of the AMG's properties and its general validity would go beyond the scope of this chapter. These issues are addressed in Chapter 4 of this thesis.

2.2 Data and main empirical results

2.2.1 Data

For our empirical analysis we concentrate on aggregate sectoral data for manufacturing from developed and developing countries for 1970 to 2002 (UNIDO, 2004). Our sample represents an unbalanced panel of 48 countries with an average of 24 time-series observations (1,162 ob-

servations in levels, 1,094 in first difference regressions — gross output based sample; for the value-added based sample we have slightly more observations, 1,194 and 1,128 for levels and first difference regressions respectively).[17] For a detailed discussion and descriptive statistics see Appendix 2.A. Note that *all* of the results presented are strikingly robust to the use of a reduced sample constructed with application of a set of rigid 'cleaning' rules.[18]

Our choice of *sectoral* data is driven by the notion/belief that it is inappropriate to view the *aggregate economy* as the basic unit of analysis when investigating growth and development against the background of structural change: sectors as diverse as manufacturing and agriculture are likely to be characterised by different production technologies, such that separate estimation is desirable. A detailed discussion of the influence of sectoral diversity across and structural change within countries in empirical growth analysis is beyond the scope of this thesis and is left for future work. With regard to our analysis of technology heterogeneity, the manufacturing sector can be argued to be *more homogeneous* in its production technology across diverse sets of countries than a representative production technology for the aggregate economy.

2.2.2 Time-series properties of the data

We carry out a range of stationarity and nonstationarity tests for individual country time-series as well as the panel as a whole, results for which are presented in Appendix 2.B. The tests conducted include 'naïve' AR(1) regressions in a pooled model, country-specific unit root tests, and panel unit root tests of the first and second generation.

Ultimately, in case of the present data dimensions and characteristics, and given all the problems and caveats of individual country unit root tests as well as panel unit root tests, we can suggest *most conservatively* that nonstationarity cannot be ruled out in this dataset. Investigation of the time-series properties of the data was not intended to select a subset of countries which we can be reasonably certain display nonstationary variable series like in Pedroni (2007); instead, our aim was to indicate that the sample (possibly due to the limited time-series di-

[17] Empirical analysis is carried out using `Stata` versions 9 and 10, `GAUSS` version 9 and `WinRATS` version 6.2.

[18] In fact the analysis was first entirely based on this 'cleaned' sample, but since results were robust to the use of the 'uncleaned' sample we preferred the larger sample size of the latter.

mension) is likely to be made up of a mixture of some countries with stationary and others with nonstationary variable series.

2.2.3 Pooled regressions

We estimate pooled models with variables in levels or first differences, including $T-1$ year dummies or country-specific period-averages *à* la Pesaran (2006) to account for unobserved common factors. In all pooled regressions the slope coefficients on the factor-inputs and the year dummies are restricted to be *the same* across all countries. Our results presented in Tables 2.1 and 2.2, in each case for a gross output and value-added specification in the upper and lower panels respectively, are for the following estimators: for the data in levels we apply **[1]** the pooled OLS estimator (POLS); **[2]** the pooled fixed effects estimator (FE); **[3]** the Pesaran (2006) common correlated effects estimator in its pooled version (CCEP); and **[4]** the pooled OLS estimator for data in first differences (FD-OLS) — additional information is contained in the notes to the table. Estimations in columns [1], [2], and [4] of both tables include sets of $(T-1)$ year dummies.

We first discuss the results for the empirical specification without any restrictions on the returns to scale — Table 2.1. Estimates for the factor input parameters in all regressions are statistically significant at the 5% level or 1% level. The POLS results in column [1] suggest that leaving out country-specific intercepts yields severely biased results: the VA-equivalent capital coefficient from the gross-output specification in the upper panel at .9 is considerably inflated, with the capital coefficient from the VA specification in the lower panel somewhat more moderate at .8. Inclusion of country intercepts in [2] reduces these coefficient estimates somewhat. The same parameter in the CCEP results in [3] is yet lower still, around .6. In both the FE and CCEP estimators the fixed effects are highly significant (F-tests are reported in the Table footnote). For all three estimators in levels the regression diagnostics suggest serial correlation in the error terms, while constant returns to scale are rejected at the 1% level of significance in all cases except POLS in the value-added specification. Note that for the FE estimator the data rejects CRS in favour of increasing returns — an unusual finding. The OLS regressions in first differences in [4] yields somewhat different technology estimates: the capital coefficient is now around .3 in both specifications (VA-equivalent), thus in line with observed macro data on factor shares in income (Mankiw et al., 1992). CRS cannot be

rejected, the AR(1) tests show serial correlation for this model, which is to be expected given that errors are now in first differences. There is however evidence of some higher order autocorrelation, especially in the VA-specification. Note that we obtain identical results for models in [1], [2] and [4] if we use data in deviation from the cross-sectional mean ('demeaned data', results not presented) instead of using a set of year dummies. Replacing year dummies with cross-sectionally demeaned data is only valid if parameters are homogeneous across countries (Pedroni, 1999, 2000) — in the pooled regressions we force this homogeneity onto the data, such that identical results are to be expected.

Under intercept and technology parameter heterogeneity, given nonstationarity in (some of) the country variable series the pooled FE estimates in [3] asymptotically converge to the 'long-run average' relation at speed $\sqrt{N}$ (Phillips & Moon, 1999) provided $T/N \to 0$ (joint asymptotics) and cross-section independence. In the present sample, however, nonstationary error terms and unobserved common factors seem to influence the results considerably: a capital coefficient of around .7 (in VA-equivalent terms) is more than twice the magnitude of the macro data on factor shares in income. As an aside, note that t-values are invalid for the estimations *in levels* if error terms are nonstationary (Coakley et al., 2001; Kao, 1999). The CCEP estimator accounts for cross-section dependence and yields, as the residual analysis in Section 2.2.7 below suggests, stationary error terms. The difference estimator in [4] converges to the common cointegrating vector $\boldsymbol{\beta}$ or the mean of the individual country cointegrating relations, $E(\boldsymbol{\beta}_i)$, at speed $\sqrt{TN}$ (Smith & Fuertes, 2007). Further investigation regarding alternative specification with parameter heterogeneity (CCEMG estimator) as well as residual diagnostics will be necessary to judge the bias of the present CCEP results. Given our priors regarding the magnitude of the capital coefficient we regard the estimators using differenced data in [4] as *preferred pooled estimators (unrestricted).*

As the model in first differences does not reject constant returns to scale (CRS), we impose CRS on all models and investigate the outcome — given that we are analysing a 'global' production function, the assumption of constant returns should be far from controversial and may help the data to differentiate between capital and material inputs more readily than is evidenced in the previous results. Table 2.2 presents the results for the pooled regression models with CRS imposed.[19] The imposition of CRS does not alter the pooled regression results in levels

[19] This is implemented by expressing all variables in per worker terms and dropping the labour variable.

Table 2.1: Pooled regressions (unrestricted returns to scale)

Static pooled regressions (unrestricted)
regression equations augmented with year dummies or period-means (Pesaran, 2006)

(i) Gross output regressions

estimator / *dependent variable*	[1] **POLS** lO	[2] **FE**♮ lO	[3] **CCEP**♮ lO	*estimator* / *dependent variable*	[4] **FD-OLS** ΔlO
regressors				*regressors*	
labour	0.0106 [2.09]*	0.1209 [12.83]**	0.1757 [14.61]**	**Δlabour**	0.1662 [4.40]**
capital	0.0452 [6.15]**	0.1923 [21.08]**	0.1720 [15.13]**	**Δcapital**	0.0675 [2.17]*
materials	0.9500 [101.69]**	0.7356 [81.17]**	0.7228 [65.11]**	**Δmaterials**	0.7879 [30.86]**
intercept	0.4157 [9.45]**	0.4883 [3.86]**	-0.5225 [2.68]**		
returns to scale				*returns to scale*	
F-Test for CRS (p)	8.0 (.01)	65.3 (.00)	53.1 (.00)	F-Test for CRS (p)	0.7 (.42)
VA-equivalent†				*VA-equivalent*†	
labour (VA)	0.213 [2.47]*	0.457 [14.27]**	0.637 [16.50]**	labour (VA)	0.784 [5.62]**
capital (VA)	0.904 [12.70]**	0.727 [25.01]**	0.620 [19.84]**	capital (VA)	0.318 [2.09]**
AR tests‡				*AR tests*‡	
AR(1) (p)	18.2 (.00)	12.8 (.00)	5.1 (.00)	AR(1) (p)	-2.8 (.00)
AR(2) (p)	18.1 (.00)	7.9 (.00)	0.4 (.70)	AR(2) (p)	-2.1 (.04)
AR(3) (p)	17.9 (.00)	10.5 (.00)	-0.5 (.60)	AR(3) (p)	0.9 (.35)
RMSE	.107	.043	.030	RMSE	.033
obs (countries)	1,162 (48)	1,162 (48)	1,162 (48)	obs (countries)	1,094 (94)

(ii) Value-added regressions

estimator / *dependent variable*	[1] **POLS** lY	[2] **FE**♮ lY	[3] **CCEP**♮ lY	*estimator* / *dependent variable*	[4] **FD-OLS** ΔlY
regressors				*regressors*	
labour	0.2100 [12.08]**	0.4402 [17.43]**	0.6009 [19.48]**	**Δlabour**	0.6849 [7.43]**
capital	0.7896 [67.34]**	0.7174 [32.17]**	0.6144 [22.76]**	**Δcapital**	0.3463 [3.40]**
intercept	1.1510 [7.99]**	-0.0470 [0.13]	-0.5160 [0.98]		
returns to scale				*returns to scale*	
F-Test for CRS (p)	0.0 (.96)	82.2 (.00)	57.3 (.00)	F-Test for CRS (p)	0.1 (.72)
AR tests‡				*AR tests*‡	
AR(1) (p)	18.5 (.00)	12.9 (.00)	8.9 (.00)	AR(1) (p)	-3.3 (.00)
AR(2) (p)	18.0 (.00)	9.4 (.00)	4.4 (.00)	AR(2) (p)	-2.7 (.01)
AR(3) (p)	17.6 (.00)	8.0 (.00)	3.7 (.00)	AR(3) (p)	1.2 (.25)
RMSE	.462	.130	.102	RMSE	.103
obs (countries)	1,194 (48)	1,194 (48)	1,194 (48)	obs (countries)	1,128 (94)

Notes: Values in parentheses are absolute t-statistics, based on White heteroskedasticity-consistent standard errors. For the CCEP estimator we include sets of cross-section period averages of gross output (lO) [or value-added (lY)], labour (lL), capital stock (lK) and materials (lM), all in logs (estimates not reported) — see main text for description of this estimator. All other models include $T-1$ year dummies in levels or FD (estimates not reported).
We indicate statistical significance at the 5% and 1% level by * and ** respectively.
The null hypotheses for the CRS tests are $\hat{\alpha}+\hat{\beta}+\hat{\gamma}=1$ and $\hat{\alpha}^{va}+\hat{\beta}^{va}=1$ for gross output and value-added specifications respectively.
RMSE reports the root mean squared errors.
♮ The F-tests in the FE and CCEP regressions reject the null that fixed effects do not differ across countries. $F(47,1079)=118.57\,(p=.00)$ (GO), $F(47,1112)=274.48\,(p=.00)$ (VA) for FE; $F(47,919)=8.96\,(p=.00)$ (GO), $F(47,1000)=4.38\,(p=.00)$ (VA) for CCEP.
† These can be derived as $\hat{\alpha}/(1-\hat{\gamma})$ for labour and in analogy for capital (Söderbom & Teal, 2004).
‡ Arellano and Bond (1991) AR test on the regression residuals. H_0: no serial correlation in the residuals.

Table 2.2: Pooled regressions (CRS)

Static pooled regressions (CRS imposed)
regression equations augmented with year dummies or period-means (Pesaran, 2006)

(i) Gross output regressions

	[1]	[2]	[3]		[4]
estimator	**POLS**	**FE**♮	**CCEP**♮	*estimator*	**FD-OLS**
dependent variable	lo	lo	lo	*dependent variable*	Δlo
regressors				*regressors*	
capital pw	0.0478 [6.56]**	0.1765 [19.24]**	0.1640 [15.56]**	**Δcapital pw**	0.0491 [1.74]
materials pw	0.9493 [102.18]**	0.7380 [79.15]**	0.7401 [67.29]**	**Δmaterials pw**	0.7875 [30.75]**
intercept	0.4695 [11.19]**	1.2411 [14.05]**	-0.3355 [1.78]		
VA-equivalent†				*VA-equivalent*†	
capital pw (VA)	0.942 [13.44]**	0.673 [23.15]**	0.631 [21.03]**	capital pw (VA)	0.231 [1.72]
AR tests‡				*AR tests*‡	
AR(1) (p)	18.7 (.00)	10.9 (.00)	6.8 (.00)	AR(1) (p)	-2.8 (.00)
AR(2) (p)	18.7 (.00)	8.4 (.00)	2.1 (.03)	AR(2) (p)	-2.1 (.04)
AR(3) (p)	18.6 (.00)	7.1 (.00)	1.0 (.34)	AR(3) (p)	1.0 (.33)
RMSE	.107	.044	.035	RMSE	.033
obs (countries)	1,162 (48)	1,162 (48)	1,162 (48)	obs (countries)	1,094 (48)

(ii) Value-added regressions

	[1]	[2]	[3]		[4]
estimator	**POLS**	**FE**♮	**CCEP**♮	*estimator*	**FD-OLS**
dependent variable	ly	ly	ly	*dependent variable*	Δly
regressors				*regressors*	
capital pw	0.7895 [72.97]**	0.6752 [29.89]**	0.5823 [23.38]**	**Δcapital pw**	0.3195 [3.61]**
intercept	1.1474 [8.47]**	2.3786 [9.98]**	0.2489 [0.63]		
AR tests‡				*AR tests*‡	
AR(1) (p)	18.5 (.00)	13.0 (.00)	9.6 (.00)	AR(1) (p)	-3.3 (.00)
AR(2) (p)	18.0 (.00)	9.6 (.00)	4.7 (.00)	AR(2) (p)	-2.7 (.01)
AR(3) (p)	17.5 (.00)	8.2 (.00)	3.4 (.00)	AR(3) (p)	1.2 (.24)
RMSE	.462	.135	.113	RMSE	.103
obs (countries)	1,194 (48)	1,194 (48)	1,194 (48)	obs (countries)	1,128 (48)

Notes: Values in parentheses are absolute t-statistics, based on White heteroskedasticity-consistent standard errors. For the CCEP estimator we include sets of cross-section period-averages of gross output per worker (lo) [or value-added per worker (ly)], capital stock per worker (lk) and materials per worker (lm), all in logs (estimates not reported) — see text for description of this estimator. All other models include $T-1$ year dummies (estimates not reported). We indicate statistical significance at the 5% and 1% level by * and ** respectively.
RMSE reports the root mean squared errors.
♮ The F-tests in the FE and CCEP regressions reject the null that fixed effects do not differ across countries.
$F(47, 1080) = 111.81$ ($p = .00$) (GO), $F(47, 1113) = 254.20$ ($p = .00$) (VA) for FE; $F(47, 968) = 7.00$ ($p = .00$) (GO), $F(47, 1049) = 5.35$ ($p = .00$) (VA) for CCEP.
† These can be derived as $\hat{\alpha}(\text{VA}) = \hat{\alpha}/(1-\hat{\gamma})$ for labour and in analogy for capital (Söderbom & Teal, 2004).
‡ Arellano and Bond (1991) AR test on the regression residuals. H_0: no serial correlation in the residuals.

to a great extent: POLS in [1], FE in [2] and CCEP in [3] yield capital coefficient estimates of around .9, .7 and .6 respectively; in the gross output specification our preferred ΔOLS estimator in [4] now has a slightly less precisely identified capital coefficient, which is however within the 95% confidence intervals of the estimate in the unrestricted equation.[20] The VA-equivalent capital coefficient as a result has dropped to around .23. In the value-added specification, the ΔOLS estimates show very similar results for the unrestricted and restricted returns to scale models.

Our pooled regression analysis suggests that time-series properties of the data play an important role in estimation: the pooled OLS and FE levels regressions, where some or all country variable series may be I(1), yield very high parameter coefficients on capital, which translate into VA-equivalent capital coefficients of between .6 to .9. Our regression results in Appendix 2.D seem to suggest that dynamic mis-specification is not driving these results. We would suggest that the bias is the result of nonstationary errors, which are introduced into the pooled equation by the imposition of parameter homogeneity on heterogeneous country equations — formal analysis of regression residuals and parameter heterogeneity is carried out in Sections 2.2.7 and 2.3 respectively. In contrast, the OLS regressions where variables are in first differences and thus stationary (FD-OLS) yielded more sensible capital parameters. This pattern of results fits the case of level series being integrated of order one in at least some of the countries in our sample. The results for the CCEP are somewhat surprising, given our findings when we relax parameter homogeneity.

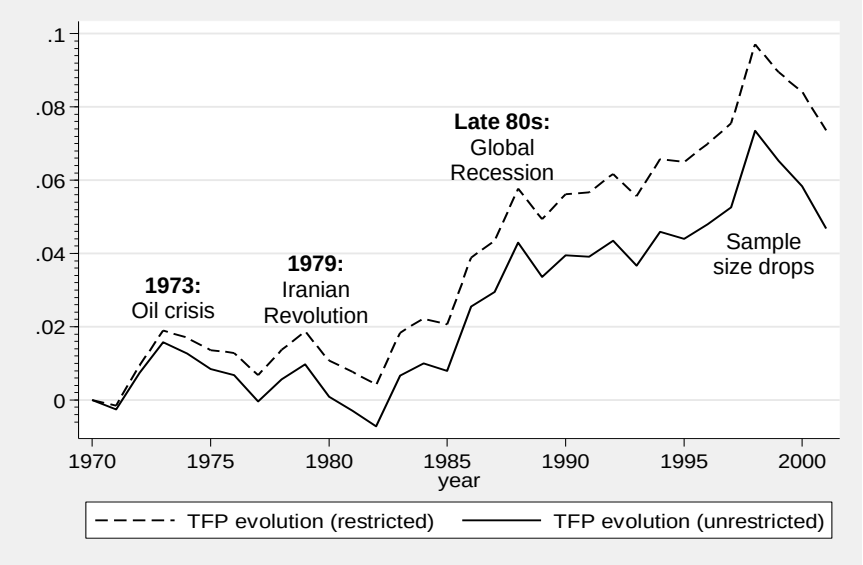
2.2.4 Common TFP

Following our argumentation above, the FD-OLS regression represents the only pooled regression model which estimates a cross-country average relationship *safe from difficulties introduced by nonstationarity*. We therefore make use of the year dummy coefficients derived from our preferred pooled regressions with CRS imposed (FD-OLS, column [4] in Table 2.2) to obtain an estimate of the 'common dynamic process' $\hat{\mu}_t^{\bullet}$, which represents an estimate of the common TFP evolution (mean of the unobserved common factors evaluated at the average impact across countries). Figure 2.1 illustrates the evolution path of this common dynamic process.[21]

[20] We also carry out a Wald test for parameter equality and cannot reject the null that they are the same: $\chi^2(1) = 0.00$, $p = .98$.

[21] In the graph we omit the data-point for 2002 for which there are only two observations. This implied 'globally common evolution' is for a gross-output specification. In order to provide a graph translated into a

Figure 2.1: Evolution of the 'Common dynamic process' $\hat{\mu}_t^{\bullet}$



This graph is for the unrestricted and CRS specifications of the gross-output based model. The graph shows severe slumps following the two oil shocks in the 1970s, while the 1980s and 1990s indicate considerable upward movement.[22] The version of this graph for results from the VA-specification shows the same patterns over time (not reported). Recall that we favour the 'measure of ignorance' interpretation of TFP, such that a decline in global manufacturing TFP as evidenced in the 1970s should not be interpreted as a decline in knowledge, but a worsening of the global manufacturing *environment*. A simpler explanation may be that our variable deflation[23] does not adequately capture all the price changes in general, and material input price changes vis-à-vis output prices in particular, occurring in the post-oil shock periods.

2.2.5 Country regressions

In the following we relax the assumption implicit in the pooled regressions that all countries possess the same production technology, and allow for country-specific slope coefficients on factor inputs. At the same time, we maintain that common shocks and/or cross-sectional dependence have to be accounted for in some fashion. We contrast models that allow for unobserved common factors in different ways, with models which ignore these matters. A total

VA specification, we first need to scale the year dummy coefficients by $1/(1-\hat{\gamma})$ to account for material inputs. The resulting TFP evolution path (not presented) is very similar to that derived from the VA specification.

[22] Note that these graphs are 'data-specific': for years where data coverage is good, this can be interpreted as 'global', whereas for years from 2000 (only 10 countries have data for 2001, only 2 for 2002 which is omitted from the graph) this interpretation collapses.

[23] Country deflators applied to monetary values in nominal local currency units [LCU] turn the data into real LCU values, which are subsequently translated into real US$ at 1990 value by using the LCU-US$ exchange rate for that specific year.

of four specifications are investigated:

$$(a) \quad y_{it} = a_i + \boldsymbol{b}_i' \boldsymbol{x}_{it} + c_i t + e_{it} \tag{2.19}$$

$$(b) \quad y_{it} - \hat{\mu}_t^{\bullet} = a_i + \boldsymbol{b}_i' \boldsymbol{x}_{it} + c_i t + e_{it} \tag{2.20}$$

$$(c) \quad y_{it} = a_i + \boldsymbol{b}_i' \boldsymbol{x}_{it} + c_i t + d_i \hat{\mu}_t^{\bullet} + e_{it} \tag{2.21}$$

$$(d) \quad y_{it} = a_i + \boldsymbol{b}_i' \boldsymbol{x}_{it} + \boldsymbol{d}_{1i}' \bar{y}_t + \boldsymbol{d}_{2i}' \bar{\boldsymbol{x}}_t + e_{it} \tag{2.22}$$

In specification **(*a*)** we do not account for any cross-section dependence and/or common TFP — this represents standard country regressions augmented with linear country-trends. Our novel augmented estimator has two variants: in **(*b*)** the 'common dynamic process' $\hat{\mu}_t^{\bullet}$ is imposed with a unit coefficient, whereas in **(*c*)** it is included as additional regressor. The augmentation with cross-section averages is represented by option **(*d*)**. Similarly for specifications with variables in first difference and the VA versions.

An alternative, econometric interpretation of the different specifications would be as follows: for options **(b) – (d)** the inclusion of an estimate for unobserved common factors in the model — either via $\hat{\mu}_t^{\bullet}$ or cross-section averages — can account for nonstationary processes omitted from the individual country-regression and enables country equations with nonstationary inputs and output to cointegrate. In these cases we require the unobserved common factors to be part of the cointegrating relation (Pedroni, 2007). Option **(a)** in contrast does not account for this process and/or cross-section dependence and we would therefore expect for (some of) the country regressions to yield spurious estimation results unless the cointegrating vector is adequately represented by this specification.

These four specifications are implemented as follows: **[1]** the standard Pesaran and Smith (1995) Mean Group (MG) and **[2]** the Swamy (1970) Random Coefficient Model (RCM) estimators are representatives of option (a); the novel Augmented Mean Group (AMG) estimator can either **[3]** have the common dynamic process ($\hat{\mu}_t^{\bullet}$) imposed with unit coefficient like in option (b), or **[4]** have it included as additional regressor like in option (c). Similarly for the Augmented RCM (ARCM) **[7]**, **[8]**. The Mean Group version of the Common Correlated Effects estimator (CCEMG) represents option (d). We test two alternative specifications, namely **[5]** as defined by Pesaran (2006), and **[6]** augmented with a country-specific linear trend

term.[24] In the first instance we deliberately do not use 'demeaned' data, since cross-sectional demeaning is only valid if all model parameters are homogeneous across countries, but creates bias in case of parameter heterogeneity (Pedroni, 2000; Smith & Fuertes, 2007). We briefly discuss the results for 'demeaned' data at the end of this section, referring to Table 2.23 in Appendix 2.E.

Unweighted averages of country parameter estimates are presented for regressions in levels and first differences in Tables 2.3 and 2.4 respectively, with gross-output models in the upper panel and value-added models in the lower panel of each table. In all cases we have imposed constant returns to scale on the country regression equation, in line with the findings from our preferred pooled model in first differences — this decision and its implications are discussed in more detail below.

The concept of 'mean-group' estimates — be they from MG, CCEMG or the weighted variant of the Swamy RCM — suggests that while *individual* country regression estimates may be unreliable, by averaging across the estimates we obtain a more reliable measure of the average relationship across groups/countries (Pesaran & Smith, 1995). The t-statistics for the country-regression averages reported in all tables for the country regression results are measures of dispersion for the sample of country-specific estimates, adopted from Pesaran, Smith, and Yamagata (2008).[25] In these and the following tables we further provide the Pedroni (1999) 'panel t-statistic' $(1/\sqrt{N})\sum_i t_i$, constructed from the country-specific t-statistics (t_i) of the parameter estimates, which indicates the precision of the individual country estimates for capital, materials and country-specific linear trend terms. From our results we can see that the individual country estimates for the *levels* specification are on average more precisely estimated than those in the specifications in *first differences*. In the results for value-added based models this pattern is repeated.

We begin our discussion with the averaged estimates in the levels specification in Table 2.3. Our first observation regarding these results is that across *all* specifications the means of the

[24] In conversation with Hashem Pesaran during September 2008 he suggested that country-specific trends could be viewed as common factors, and that they should therefore already be accounted for by the cross-section averages included in model [**5**].

[25] In case of the MG-type estimators these are computed from the standard errors $se(\hat{\beta}_{MG}) = [\sum_i(\hat{\beta}_i - \hat{\beta}_{MG})^2]^{-1/2}$. For the Swamy estimators, see the footnote of the table for details. We will deal with formal tests for parameter heterogeneity in section 2.3.

capital coefficients are considerably lower than in the pooled levels models: between .2 and .5 (VA-equivalent) rather than between .6 and .9 in the pooled levels models in Table 2.2. Comparing results for the levels specification with those for the specification in first differences in Table 2.4 reveals that estimates from the heterogeneous models in levels and first differences follow very similar patterns. Recall that for the pooled models estimates for technology parameters in levels and first differences showed radically different results. Parameter estimates are thus similar in all specifications, gross-output or value-added, levels or first-differences.

In all cases it seems that the technology parameters are estimated reasonably precisely, suggested by the reported t-statistic and the Pedroni (1999) panel t-test. A considerable number of country trends/drifts are significant at the 10% level, although much more so for the levels than for the first difference specifications — the fact that their mean is *not* statistically significantly different from zero in all of the 'augmented' regressions (AMG, ARCM) is easily explained: these country-specific trends have positive and negative magnitudes for different countries. Results presented are robust to alternative specifications for the country-level deterministics (additional squared trend in the levels models, additional trend in the models in first difference).[26] Coefficients on the common dynamic process $\hat{\mu}_t^{\bullet}$ in models [4] and [8] for all specifications are uniformly high and close to their theoretical value of unity.[27]

The two-stage approach adopted by the augmented estimators (AMG, CCEMG) naturally faces concerns about the inclusion of estimated variables ($\hat{\mu}_t^{\bullet}$, $\bar{y}_t$, $\bar{x}_t$) in the second-stage equations. Pagan (1984, p.242) suggests that "[m]ost two-step estimators are consistent and many of these are also efficient", although inference is most commonly affected. As discussed above, inference for the MG-type estimators in this thesis is based on the Pesaran et al. (2008) t-statistics which focus on the central tendency of the country estimates, rather than the country regression standard errors and t-statistics. This aside, the Monte Carlo simulations carried out in Chapter 4 suggest that the AMG and CCEMG estimators have regression standard errors (average of reported standard errors across M replications) very much in line with the empirical standard

[26] Given that results were almost identical albeit somewhat less precise we did not pursue these specifications further (residual and heterogeneity tests).

[27] The year dummy coefficients which make up $\hat{\mu}_t^{\bullet}$ represent an average of the unobservables *and* their country-specific factor loadings across countries. As an illustration, if we regress y on $Z = (\hat{\beta}'X)$ (with $\hat{\beta}$ estimated earlier) this would yield a parameter estimate of unity on this new 'regressor', since Z represents the combination of a variable *and* its coefficient. If $\hat{\mu}_t^{\bullet}$ is included as a covariate in our first stage regression (instead of the $T-1$ year dummies) it enters with a unit coefficient due to this representation and similarly in the Mean Group averages of the country regressions. The only difference to our $Z = \hat{\beta}'X$ example is that both (latent) elements of $\hat{\mu}_t^{\bullet}$ (factors, loadings) are treated as parameters (Bai & Ng, 2008).

errors (standard deviation of the M $\hat{\beta}$ parameter estimates) in most scenarios considered. The theoretical explanation for this finding is developed in Bai and Ng (2008), where they point out that second stage standard errors need not be adjusted for first stage estimation uncertainty if $\sqrt{T}/N \to 0$, as is the case here.

Table 2.3: Country regression averages (levels, CRS imposed)

Country regressions in levels
static specification with CRS imposed; estimates presented are unweighted means of the country coefficients

(i) Gross output regressions

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
estimator	**MG**	**RCM**	**AMG**	**AMG**	**CCEMG**	**CCEMG**	**ARCM**	**ARCM**
specification	(a)	(a)	(b)	(c)	(d)	(d)	(b)	(c)
dependent variable♯	lo	lo	lo-$\hat{\mu}_t^{\bullet}$	lo	lo	lo	lo-$\hat{\mu}_t^{\bullet}$	lo
regressors								
capital pw	0.0525 [1.85]	0.0731 [2.38]*	0.0695 [2.75]**	0.0815 [3.17]**	0.1216 [5.25]**	0.0839 [3.22]**	0.0825 [3.00]**	0.0923 [3.31]**
materials pw	0.7090 [26.31]**	0.7305 [25.18]**	0.7274 [27.32]**	0.7303 [28.31]**	0.7193 [22.92]**	0.7091 [24.96]**	0.7487 [26.24]**	0.7480 [26.96]**
common trend				1.0083 [4.22]**				0.9464 [3.63]**
country trend	0.0044 [3.97]**	0.0034 [2.82]**	0.0008 [0.76]	0.0003 [0.26]		0.0024 [2.17]*	-0.0001 [0.05]	-0.0001 [0.07]
intercept	2.8429 [8.61]**	2.4251 [6.76]**	2.4713 [8.24]**	2.3099 [7.38]**	0.8193 [1.71]	1.4514 [3.02]**	2.1392 [6.56]**	2.0359 [6.01]**
Panel t-statistics, trends†								
capital pw	8.40	10.59	9.98	11.64	19.02	9.37	12.37	13.48
material pw	103.13	109.32	116.50	114.56	118.90	89.52	121.59	121.65
country trend	20.61	19.57	19.11	12.92		14.06	17.10	12.75
# of sign. trends (at 10%)	30	33	26	24		26	26	23
VA-equivalent coefficients‡								
capital/worker (VA) mean	0.180 [1.95]	0.271 [2.53]*	0.255 [2.96]**	0.302 [3.35]**	0.433 [6.14]**	0.288 [3.49]**	0.328 [3.24]**	0.366 [3.49]**
RMSE	.029	.030	.029	.025	.025	.022	.029	.026
obs (countries)	1,162 (48)	1,162 (48)	1,162 (48)	1,162 (48)	1,162 (48)	1,162 (48)	1,162 (48)	1,162 (48)

(ii) Value-added regressions

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
estimator	**MG**	**RCM**	**AMG**	**AMG**	**CCEMG**	**CCEMG**	**ARCM**	**ARCM**
specification	(a)	(a)	(b)	(c)	(d)	(d)	(b)	(c)
dependent variable♯	ly	ly	ly-$\hat{\mu}_t^{va\,\bullet}$	ly	ly	ly	ly-$\hat{\mu}_t^{va\,\bullet}$	ly
regressors								
capital pw	0.1789 [2.25]*	0.2691 [3.13]**	0.2896 [3.95]**	0.2982 [3.70]**	0.4663 [6.76]**	0.3125 [3.72]**	0.3557 [4.49]**	0.3529 [4.08]**
common trend				0.8787 [4.39]**				0.8751 [3.88]**
country trend	0.0174 [5.95]**	0.0148 [4.77]**	0.0001 [0.04]	0.0023 [0.56]		0.0108 [3.09]**	-0.0018 [0.66]	0.0003 [0.06]
intercept	7.6528 [9.05]**	6.7191 [7.35]**	6.3823 [8.42]**	6.2431 [7.40]**	0.8961 [0.89]	4.7860 [3.66]**	5.7380 [6.98]**	5.7620 [6.34]**
Panel t-statistics, trends†								
capital pw	12.42	15.45	16.55	17.43	31.86	16.15	20.34	19.68
country trend	27.65	26.65	20.40	12.12		16.08	19.14	10.99
# of sign. trends (at 10%)	39	36	28	19		24	27	19
RMSE	.100	.101	.097	.091	.100	.088	.098	.093
obs (countries)	1,194 (48)	1,194 (48)	1,194 (48)	1,194 (48)	1,194 (48)	1,194 (48)	1,194 (48)	1,194 (48)

Notes: All variables are in logs. lo — log output per worker, ly — log value-added per worker. $\hat{\mu}_t^{\bullet}$ and $\hat{\mu}_t^{va\,\bullet}$ are derived from the year dummy coefficients of a pooled regression (CRS imposed) in first differences (FD-OLS) as described in the text.
We omit reporting the parameters on the cross-section averages for the CCEMG estimators (columns [5] and [6]) to save space.
Values in parentheses are absolute t-statistics. These were obtained by regressing the N country estimates on an intercept term, except for the Swamy t-stats, which are provided by `xtrc` in `Stata` and represent $\sum_i(\hat{\Sigma}+\hat{\mathbf{V}}_i)$ where $\hat{\Sigma}$ is a measure of dispersion of the country OLS estimates and $\hat{V}$ is the variance of the N OLS estimates scaled by $\sum x_i^2$.
'Specification' refers to equations (2.19) to (2.22). We indicate statistical significance at the 5% and 1% level by * and ** respectively.
RMSE reports the root mean squared errors.
♯ We subtract the common dynamic process $\hat{\mu}_t^{\bullet}$ from log output (log value-added) per worker for country i in models [3] and [7].
† We report the Pedroni (2000) panel-t statistics $(1/\sqrt{N})\sum_i t_i$ where t_i is the country-specific t-statistic of the parameter estimate.
‡ This is obtained using a non-linear combination of the capital and materials coefficients accounting for the precision of these estimates.

Closer inspection of the VA-equivalent capital coefficients suggests the following patterns: *firstly*, estimation approaches that do not account for unobserved common factors or cross-section dependence have parameter estimates around .2. *Secondly*, for the 'augmented' estimators which include a 'common dynamic process' in the estimation equation (either explicitly or implicitly with unit coefficient) the capital coefficients are around .3. *Thirdly*, the results for the CCEMG with and without additional country trend differ considerably, with the former close to all other augmented regression results and the latter slightly larger, around .45.

Table 2.4: Country regression averages (FD, CRS imposed)

Country regressions in first differences
static specification with CRS imposed; estimates presented are unweighted means of the country coefficients

(i) Gross output regressions

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
estimator	**ΔMG**	**ΔRCM**	**ΔAMG**	**ΔAMG**	**ΔCCEMG**	**ΔCCEMG**	**ΔARCM**	**ΔARCM**
specification	(a)	(a)	(b)	(c)	(d)	(d)	(b)	(c)
dependent variable♯	Δlo	Δlo	Δlo-$\Delta\hat{\mu}_t^{\bullet}$	Δlo	Δlo	Δlo	Δlo-$\Delta\hat{\mu}_t^{\bullet}$	Δlo
regressors								
Δcapital pw	0.0476	0.0544	0.0659	0.0763	0.0836	0.0547	0.0675	0.0667
	[2.01]*	[1.92]	[2.85]**	[3.31]**	[3.98]**	[2.30]*	[2.45]*	[2.43]*
Δmaterials pw	0.7634	0.7736	0.7744	0.7672	0.7579	0.7575	0.7834	0.7779
	[42.75]**	[37.14]**	[41.69]**	[40.19]**	[36.21]**	[39.46]**	[36.63]**	[35.67]**
common drift	-	-	-	1.0363	-	-	-	0.8633
	-	-	-	[4.93]**	-	-	-	[3.64]**
country drift	0.0020	0.0028	-0.0014	-0.0016	-	0.0013	-0.0003	0.0001
	[1.67]	[1.85]	[1.17]	[1.42]	-	[0.93]	[0.23]	[0.04]
Panel t-statistics, trends†								
capital pw	4.6	5.8	6.3	5.8	7.8	4.5	7.5	7.3
materials pw	90.8	114.8	102.6	99.8	86.6	86.2	118.0	119.0
country trends	7.3	8.0	6.3	6.3	-	7.9	6.4	5.9
# of sign. drifts (at 10%)	8	13	6	7	-	12	6	6
VA-equivalent coefficients‡								
capital/worker (VA) mean	0.201	0.240	0.292	0.328	0.345	0.226	0.312	0.300
	[2.03]*	[1.92]	[2.90]**	[3.30]**	[3.97]**	[2.30]*	[2.47]*	[2.43]*
RMSE	.029	.030	.028	.026	.025	.025	.029	.028
obs (countries)	1,094 (48)	1,094 (48)	1,094 (48)	1,094 (48)	1,094 (48)	1,094 (48)	1,094 (48)	1,094 (48)

(ii) Value-added regressions

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
estimator	**ΔMG**	**ΔRCM**	**ΔAMG**	**ΔAMG**	**ΔCCEMG**	**ΔCCEMG**	**ΔARCM**	**ΔARCM**
specification	(a)	(a)	(b)	(c)	(d)	(d)	(b)	(c)
dependent variable♯	Δly	Δly	Δly-$\Delta\hat{\mu}_t^{va\bullet}$	Δly	Δly	Δly	Δly-$\Delta\hat{\mu}_t^{va\bullet}$	Δly
regressors								
Δcapital pw	0.1642	0.2085	0.2734	0.2834	0.3837	0.2577	0.3053	0.2969
	[1.91]	[2.13]*	[3.48]**	[3.77]**	[5.62]**	[3.48]**	[3.37]**	[3.38]**
common drift	-	-	-	1.0497	-	-	-	0.8938
	-	-	-	[5.71]**	-	-	-	[4.21]**
country drift	0.0161	0.0171	-0.0011	-0.0020	-	0.0123	-0.0001	0.0017
	[5.54]**	[4.42]**	[0.44]	[0.50]	-	[3.76]**	[0.03]	[0.34]
Panel t-statistics, trends†								
capital pw	4.8	6.9	8.3	7.7	12.0	7.2	10.7	10.4
country trends	9.53	13.95	5.77	7.46	-	7.68	6.18	7.26
# of sign. drifts (at 10%)	14	24	6	10	-	11	7	10
RMSE	.097	.099	.094	.090	.090	.088	.096	.097
obs (countries)	1,128 (48)	1,128 (48)	1,128 (48)	1,128 (48)	1,128 (48)	1,128 (48)	1,128 (48)	1,128 (48)

Notes: Variables are in first differences of logs. Δlo — output per worker growth rate, Δly — value-added per worker growth rate. See Table 2.3 for further information.

In Figures 2.3 and 2.4 in Appendix 2.E we plot the density estimates for the sample of country-specific technology parameters estimated in the levels regressions presented in Table 2.3, using standard kernel methods with automatic bandwidth selection. Both sets of plots indicate that the distribution of these parameter estimates is fairly symmetric around their respective means, such that no significant outliers drive our results. These plots cannot be taken as evidence that technology parameters are heterogeneous across countries; we address this issue in Section 2.3. In each plot we note the presence of country estimates for the capital coefficient below zero and above unity. Previous empirical analysis averaging over individual country regressions has frequently found that while country estimates were widely dispersed and at times economically implausible, averages represented very plausible estimates (Boyd & Smith, 2002; Baltagi, Bresson, Griffin, & Pirotte, 2003). Pedroni (2007, p.440) calls for caution when interpreting the estimates for *any individual country* since the "long-run signals contained in [limited] years of data may be relatively weak" (see also Smith & Fuertes, 2007, p.32).

We also briefly review the results for country regressions where the data was first transformed into deviations from the cross-sectional mean to account for any common dynamic processes. As was noted previously, this transformation is only valid if the parameters are homogeneous across countries (Pedroni, 2000). The averaged results for this exercise are presented in Table 2.23 in Appendix 2.E. Here, the gross output specifications (levels, FD) yield capital coefficients of between .44 and .48 (VA-equivalent). In stark contrast, the value-added specifications yield capital coefficients between .13 and .20. We take these results as an indication of technology parameter heterogeneity in the production function, since as was discussed above the transformation of variables into deviations from the cross-section mean introduces nonstationarity into the errors if the underlying production technology differs across countries.

2.2.6 The importance of constant returns to scale

Further investigation reveals that the imposition of constant returns to scale, justified by the CRS test for the pooled regression with variables in first differences, plays an important role in our story. We repeat all country regressions in levels, but with all variables in 'levels' form, rather than in per worker terms. Results are presented in Table 2.5 below. Results for the specifications in first differences are reported in Table 2.24 in Appendix 2.E.

Table 2.5: Country regressions (unrestricted returns to scale)

Country regressions in levels
static specification, no restriction on returns to scale; estimates presented are unweighted means of the country coefficients

(i) Gross output regressions

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
estimator	**MG**	**RCM**	**AMG**	**AMG**	**CCEMG**	**CCEMG**	**ARCM**	**ARCM**
specification	(a)	(a)	(b)	(c)	(d)	(d)	(b)	(c)
dependent variable‡	lO	lO	lO-$\hat{\mu}_t^{\bullet}$	lO	lO	lO	lO-$\hat{\mu}_t^{\bullet}$	lO
regressors								
labour	0.2143	0.1908	0.1815	0.1806	0.2004	0.2226	0.1677	0.1690
	[5.54]**	[4.62]**	[5.44]**	[4.80]**	[6.86]**	[6.83]**	[4.67]**	[4.22]**
capital	0.1206	0.0860	0.1794	0.1969	0.0764	-0.0017	0.1361	0.1250
	[0.82]	[0.57]	[1.43]	[1.33]	[1.11]	[0.02]	[1.06]	[0.82]
materials	0.7315	0.7464	0.7474	0.7380	0.7172	0.7210	0.7627	0.7570
	[26.05]**	[24.90]**	[27.95]**	[25.31]**	[24.53]**	[27.93]**	[26.75]**	[24.51]**
common process				0.9259				0.9279
				[3.44]**				[3.21]**
country trend	0.0040	0.0024	-0.0004	0.0005		0.0045	-0.0018	-0.0007
	[1.36]	[0.78]	[0.15]	[0.16]		[1.44]	[0.60]	[0.20]
intercept	0.9570	1.7807	-0.4315	-0.7703	1.3273	2.7179	0.5327	0.8579
	[0.27]	[0.48]	[0.14]	[0.21]	[0.81]	[1.08]	[0.17]	[0.23]
Returns to scale								
CRS test (p)	0.24 (.63)	0.03 (.87)	0.82 (.36)	0.61 (.43)	0.01 (.92)	0.35 (.55)	0.29 (.59)	0.11 (.74)
Panel t-statistics, trends								
labour	19.03	21.05	17.1	16.57	17.09	16.83	19.92	19.33
capital	3.50	4.53	6.76	5.69	13.74	1.75	8.33	6.87
materials	103.31	110.20	111.30	104.13	89.71	80.90	118.91	115.95
country trends	20.60	19.49	18.65	15.16		14.28	17.36	14.60
# of sign. trends (at 10%)	27	27	27	23		26	27	25
VA-equivalent coefficients‡								
labour (VA) mean	0.798	0.752	0.719	0.689	0.709	0.798	0.707	0.696
	[7.25]**	[5.90]**	[7.51]**	[6.82]**	[7.59]**	[9.35]**	[6.23]**	[5.82]**
capital (VA) mean	0.449	0.339	0.710	0.751	0.270	-0.006	0.574	0.514
	[0.84]	[0.58]	[1.50]	[1.43]	[1.18]	[0.02]	[1.09]	[0.86]
RMSE	.026	.027	.025	.022	.021	.020	.026	.024
obs (countries)	1,162 (48)	1,162 (48)	1,162 (48)	1,162 (48)	1,162 (48)	1,162 (48)	1,162 (48)	1,162 (48)

(ii) Value-added regressions

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
estimator	**MG**	**RCM**	**AMG**	**AMG**	**CCEMG**	**CCEMG**	**ARCM**	**ARCM**
specification	(a)	(a)	(b)	(c)	(d)	(d)	(b)	(c)
dependent variable‡	lY	lY	lY-$\hat{\mu}_t^{va\bullet}$	lY	lY	lY	lY-$\hat{\mu}_t^{va\bullet}$	lY
regressors								
labour	0.7976	0.7612	0.6789	0.6638	0.6769	0.6793	0.6663	0.6684
	[8.80]**	[7.82]**	[8.79]**	[8.70]**	[7.19]**	[8.83]**	[7.93]**	[8.01]**
capital	0.1770	0.3074	0.3216	0.3453	0.5179	0.3144	0.4201	0.4295
	[0.68]	[1.15]	[1.47]	[1.52]	[5.13]**	[1.18]	[1.85]	[1.81]
common trend				1.0254				0.9019
				[5.27]**				[4.14]**
country trend	0.0105	0.0094	-0.0066	-0.0077		0.0090	-0.0069	-0.0053
	[1.53]	[1.30]	[0.91]	[0.96]		[0.81]	[0.91]	[0.62]
intercept	8.5557	5.6880	6.5479	6.0760	-0.0126	4.8980	4.2023	3.9300
	[1.48]	[0.95]	[1.34]	[1.19]	[0.01]	[0.88]	[0.83]	[0.74]
Returns to scale								
CRS test (p)	0.02 (.90)	0.10 (.75)	0.00 (.99)	0.00 (.96)	3.98 (.05)	0.00 (.98)	0.2 (.66)	0.22 (.64)
Panel t-statistics, trends								
labour	30.2	34.0	27.2	26.7	26.5	27.0	31.3	31.3
capital	7.2	9.3	10.2	9.7	26.5	9.9	12.2	11.8
country trends	20.2	19.5	16.8	14.9		18.8	16.3	13.6
# of sign. trends (at 10%)	24	28	25	20		30	25	22
RMSE	.091	.093	.088	.082	.086	.076	.090	.084
obs (countries)	1,194 (48)	1,194 (48)	1,194 (48)	1,194 (48)	1,194 (48)	1,194 (48)	1,194 (48)	1,194 (48)

Notes: All variables are in logs. lO — log output, lY — log value-added. $\hat{\mu}_t^{\bullet}$ and $\hat{\mu}_t^{\bullet,va}$ are derived from the year dummy coefficients of a pooled regression (no restriction on returns to scale) in first differences (FD-OLS) as described in the text.
We omit reporting the parameters on the cross-section averages for the CCEMG estimators (columns [5] and [6]) to save space.
Values in parentheses are absolute t-statistics. These were obtained by regressing the N country estimates on an intercept term, except for the Swamy t-stats, which are provided by `xtrc` in `Stata` and represent $\sum_i(\hat{\Sigma}+\hat{\mathbf{V}}_i)$ where $\hat{\Sigma}$ is a measure of dispersion of the country OLS estimates and $\hat{V}$ is the variance of the N OLS estimates scaled by $\sum x_i^2$.
'Specification' refers to equations (2.19) to (2.22). We indicate statistical significance at the 5% and 1% level by * and ** respectively.
RMSE reports the root mean squared errors.
♯ We subtract the common dynamic process $\hat{\mu}_t^{\bullet}$ from log output (log value-added) for country i in models [3] and [7].
† We report averaged t-statistics for country-specific t-statistics t_i of the factor estimates.
‡ This is obtained using a non-linear combination of the capital and materials coefficients accounting for the precision of these estimates.

The failure to impose constant returns to scale leads to a loss of precision in the capital estimates in virtually all specifications, for the regressions in levels and first differences. The labour coefficients are however around .7 in all cases and only one out of 16 levels models rejects a parameter test for constant returns to scale — the latter result is not surprising given the sizeable standard errors on the capital coefficients. In the first difference models in the Appendix CRS is never rejected. Despite these reaffirming diagnostics, the sharp difference between the impact of imposing CRS in the country regressions (considerable loss of precision) and the pooled regressions (limited change) is somewhat puzzling and merits further investigation in the future.

2.2.7 Residual diagnostics

The nature of our data prevents us from implementing a number of specification tests from the nonstationary panel econometric literature: first of all, existing procedures used to investigate cross-section dependence (PCA, mean absolute correlation, Pesaran (2004) CD statistic for residual diagnostics) rely on fairly balanced panels without missing observations. Secondly, the present panel cointegration tests for moderate T panels — i.e. *not* Westerlund and Edgerton (2008) — do not yet allow for cross-section dependence, while first generation test routines are similarly developed with balanced panels of full time-series in mind. It is unclear how these can be implemented for a macro panel which includes data from developing countries and thus missing observations and considerable imbalance. We therefore need to resort to a mix of diagnostic tests for pooled regressions and averaged test statistics following the Fisher (1932) combination principle.

We begin with normality and homoskedasticity tests for regression residuals in Table 2.6. Using the D'Agostino, Balanger, and D'Agostino Jr. (1990) and Cameron and Trivedi (1990) tests virtually all of our pooled regression results reject the joint null hypothesis of normal, and normal and homoskedastic residuals respectively. In contrast, the same statistics transformed into Fisher statistics (p_λ) suggest that all of the country regression residuals can be assumed as jointly normal or jointly normal and homoskedastic.

Table 2.6: Testing Residuals for Normality and Homoskedasticity

Regression residuals — normality, homoskedasticity♯
H_0 for respective test statistics: no heteroskedasticity, regular skewness, regular kurtosis;♭
we report the χ^2 or the Fisher p_λ statistic and corresponding p-values

	Cameron & Trivedi tests									**D'Agostino *et al* normality test**			
	Pooled regression residuals: Gross output specifications (CRS)†												
	Joint test		Heterosked.		Skewness		Kurtosis			Joint test		Skewness	Kurtosis
	χ^2	p	χ^2	p	χ^2	p	χ^2	p		χ^2	p	p	p
POLS	176.1	.01	139.5	.01	36.4	.36	0.2	.64		33.6	.00	.04	.00
FE	1355.6	.01	1162.0	.49	180.2	.00	13.4	.00		69.2	.00	.00	.00
CCEP	-	-	-	-	-	-	-	-		106.2	.00	.12	.00
FD-OLS	541.3	.00	491.9	.00	43.1	.11	6.3	.01		214.2	.00	.00	.00
	Pooled regression residuals: Value-added specifications (CRS)†												
	Joint test		Heterosked.		Skewness		Kurtosis			Joint test		Skewness	Kurtosis
	χ^2	p	χ^2	p	χ^2	p	χ^2	p		χ^2	p	p	p
POLS	194.3	.00	95.4	.01	95.3	.00	3.6	.06		51.2	.00	.00	.16
FE	1324.2	.16	1194.0	.49	98.1	.08	32.2	.00		53.0	.00	.00	.00
CCEP	-	-	-	-	-	-	-	-		84.7	.00	.12	.00
FD-OLS	281.7	.00	237.2	.00	36.8	.26	7.8	.01		279.4	.00	.00	.00

	Cameron & Trivedi tests									**D'Agostino *et al* normality test**					
	Country regression residuals: Gross output specifications (CRS)†														
	Joint test		Heterosked.		Skewness		Kurtosis			Joint test		Skewness		Kurtosis	
	p_λ	p	p_λ	p	p_λ	p	p_λ	p		p_λ	p	p_λ	p	p_λ	p
MG	80.0	.88	70.4	.98	69.0	.98	39.7	1.00		51.0	1.00	55.4	1.00	43.2	1.00
AMG(i)	81.7	.85	70.9	.97	67.5	.99	41.8	1.00		46.0	1.00	47.7	1.00	41.8	1.00
AMG(ii)	64.7	.99	56.7	1.00	59.0	1.00	42.9	1.00		48.7	1.00	50.6	1.00	43.9	1.00
CCE(i)	46.5	1.00	44.5	1.00	42.3	1.00	50.7	1.00		45.2	1.00	41.0	1.00	46.5	1.00
CCE(ii)	41.7	1.00	39.0	1.00	47.7	1.00	43.9	1.00		34.4	1.00	32.4	1.00	39.8	1.00
	Country regression residuals: Value-added specifications (CRS)†														
	Joint test		Heterosked.		Skewness		Kurtosis			Joint test		Skewness		Kurtosis	
	p_λ	p	p_λ	p	p_λ	p	p_λ	p		p_λ	p	p_λ	p	p_λ	p
MG	93.6	.55	83.8	.81	72.3	.97	43.5	1.00		56.2	1.00	57.6	1.00	47.2	1.00
AMG(i)	97.3	.44	80.1	.88	81.9	.85	46.6	1.00		56.0	1.00	57.0	1.00	47.2	1.00
AMG(ii)	81.4	.86	71.9	.97	63.8	1.00	50.0	1.00		52.4	1.00	54.2	1.00	46.7	1.00
CCE(i)	96.8	.46	71.8	.97	86.1	.76	43.0	1.00		58.7	1.00	63.4	1.00	45.3	1.00
CCE(ii)	65.7	.99	59.4	1.00	56.1	1.00	43.9	1.00		51.6	1.00	51.0	1.00	47.6	1.00

Notes: ♯ We report heteroskedasticity and normality test statistics for residual series from
(a) pooled estimators: POLS — pooled OLS, FE — pooled OLS with N country dummies, CCEP — Pesaran (2006) common correlated effects estimator (pooled version), FD-OLS — pooled OLS with variables in first differences; POLS, FE and FD-OLS are augmented with $T-1$ year dummies; the regression results from which these residuals are taken are presented in Table 2.2; and
(b) heterogeneous parameter estimators: MG — Pesaran and Smith (1995) Mean Group estimator, AMG(i) — Augmented Mean Group estimator ($\hat{\mu}_t^\bullet$ imposed with unit coefficient), AMG(ii) — dto. ($\hat{\mu}_t^\bullet$ included as additional regressor), CCE(i) — Pesaran (2006) common correlated effects estimator (mean group version), CCE(ii) — dto. with additional country trend; MG, AMG(i), AMG(ii), and CCE(ii) contain country-specific trend terms; the regression results from which these residuals are taken are presented in Table 2.3.
The Cameron and Trivedi (1990) decomposition test analyses residual heteroskedasticity, skewness and kurtosis. The D'Agostino et al. (1990) test investigates skewness and kurtosis. Both tests provide joint hypothesis tests. If a test rejects skewness or kurtosis we can no longer assume normality for the residual distribution.
♭ 'Regular' in this case is taken to mean in line with the normal distribution.
† For the pooled regressions we report χ^2 statistics and corresponding p-values. For the country regression we compute the Fisher statistic $p_\lambda = -2\sum_i log(p_i)$, where p_i is the p-value of the *country-specific* test statistic. The Fisher statistic p_λ is distributed $\chi^2(2N)$.

Tables 2.7 and 2.8 present panel unit root test results for pooled regression and country regression errors respectively. This analysis broadly suggests that residuals from CCEP and pooled OLS in first differences are stationary, whereas those from POLS and FE are likely to be nonstationary. Country regression residuals are more likely to be stationary across all specifications: nonstationarity is rejected for up to 4 lags in the first generation Maddala and Wu (1999) test, and still for up to 3 lags in the Pesaran (2007) CIPS test (CIPS*), where we have specified that the residual series are not 'demeaned' prior to testing.

Table 2.7: Testing Residuals for Stationarity (i)

Pooled regression residuals — stationarity tests♯
H_0 for each test: nonstationary residual series;♭
Maddala and Wu (1999) (MW) and Pesaran (2007) (CIPS) test results;
CIPS* does *not* apply cross-sectional demeaning prior to the testing procedure

Residuals from gross output specifications (CRS imposed)†										
	lags: 0		lags: 1		lags: 2		lags: 3		lags: 4	
MW	p_λ	p	p_λ	p	p_λ	p	p_λ	p	p_λ	p
POLS	193.16	.00	169.00	.00	91.82	.60	105.65	.24	110.42	.15
FE	188.54	.00	216.55	.00	127.73	.02	129.72	.01	98.44	.41
CCEP	397.64	.00	370.93	.00	203.45	.00	148.20	.00	155.512	.00
FD-OLS	1164.92	.00	632.24	.00	393.81	.00	297.95	.00	140.95	.00
CIPS	Z[t-bar]	p	Z[t-bar]	p	Z[t-bar]	p	Z[t-bar]	p	Z[t-bar]	p
POLS	-1.88	.03	-0.36	.36	1.27	.90	4.22	1.00	6.33	1.00
FE	-1.49	.07	-0.61	.27	1.65	.95	3.98	1.00	6.80	1.00
CCEP	-10.08	.00	-8.23	.00	-4.12	.00	2.40	1.00	7.08	1.00
FD-OLS	-20.88	.00	-10.26	.00	-1.47	.07	3.97	1.00	8.89	1.00
CIPS*	Z[t-bar]	p	Z[t-bar]	p	Z[t-bar]	p	Z[t-bar]	p	Z[t-bar]	p
POLS	-2.32	.01	-1.41	.08	1.51	.93	1.19	.88	3.09	1.00
FE	-1.98	.02	-2.00	.02	1.39	.92	2.86	1.00	4.21	1.00
CCEP	-10.08	.00	-8.23	.00	-4.12	.00	2.40	.99	7.08	1.00
FD-OLS	-22.28	.00	-14.19	.00	-8.51	.00	-5.00	.00	1.72	.96

Residuals from value-added specifications (CRS imposed)†										
	lags: 0		lags: 1		lags: 2		lags: 3		lags: 4	
MW	p_λ	p	p_λ	p	p_λ	p	p_λ	p	p_λ	p
POLS	151.96	.00	176.19	.00	108.60	.18	74.32	.95	72.70	.96
FE	146.64	.00	140.02	.00	94.70	.52	72.66	.96	92.21	.59
CCEP	323.84	.00	306.74	.00	196.42	.00	136.74	.00	127.54	.02
FD-OLS	2033.95	.00	1142.64	.00	580.42	.00	455.09	.00	314.48	.00
CIPS	Z[t-bar]	p	Z[t-bar]	p	Z[t-bar]	p	Z[t-bar]	p	Z[t-bar]	p
POLS	-0.96	.17	0.52	.70	3.59	1.00	8.14	1.00	9.82	1.00
FE	-1.76	.04	-1.13	.13	0.78	.78	5.17	1.00	6.53	1.00
CCEP	-8.32	.00	-7.53	.00	-2.74	.00	1.41	.92	8.26	1.00
FD-OLS	-27.11	.00	-20.16	.00	-6.73	.00	-1.17	.12	4.23	1.00
CIPS*	Z[t-bar]	p	Z[t-bar]	p	Z[t-bar]	p	Z[t-bar]	p	Z[t-bar]	p
POLS	-1.34	.09	-1.53	.06	1.67	.95	3.81	1.00	4.81	1.00
FE	-1.18	.12	-0.44	.33	2.28	.99	4.40	1.00	4.96	1.00
CCEP	-8.32	.00	-7.53	.00	-2.74	.00	1.41	.92	8.26	1.00
FD-OLS	-28.18	.00	-20.99	.00	-12.63	.00	-8.92	.00	-1.35	.09

Notes: † We apply the Pesaran (2007) (CIPS) and Maddala and Wu (1999) (MW) panel unit root tests; the former accounts for possible cross-section dependence in the variable tested (here: regression residuals). We also present the results for the CIPS where cross-sectional demeaning of the data prior to the testing procedure is prevented (marked as CIPS*).
See Table 2.6 for a description of the regressions from which these residuals are taken.

A cautious conclusion from these diagnostic tests would be that we are more confident about the country regression residuals possessing desirable properties (stationarity, normality, homoskedasticity) than we are for their pooled counterparts.

Table 2.8: Testing Residuals for Stationarity (ii)

Country regression residuals — stationarity tests[♯]
H_0 for each test: stationary residual series;[♭]
Maddala and Wu (1999) (MW) and Pesaran (2007) (CIPS) test results;
CIPS* does *not* apply cross-sectional demeaning prior to the testing procedure

Gross output specifications (CRS)†										
	lags: 0		lags: 1		lags: 2		lags: 3		lags: 4	
MW	p_λ	p	p_λ	p	p_λ	p	p_λ	p	p_λ	p
MG	429.01	.00	386.48	.00	227.16	.00	166.57	.00	138.59	.00
RCM	384.54	.00	359.93	.00	211.93	.00	154.20	.00	127.73	.02
AMG(i)	444.71	.00	388.14	.00	251.08	.00	168.88	.00	132.25	.01
AMG(ii)	551.52	.00	468.54	.00	299.32	.00	195.87	.00	145.25	.00
CCE(i)	581.95	.00	585.88	.00	329.75	.00	285.80	.00	201.71	.00
CCE(ii)	699.77	.00	624.26	.00	330.92	.00	242.79	.00	177.41	.00
ARCM(i)	397.22	.00	355.32	.00	226.16	.00	145.11	.00	127.07	.02
ARCM(ii)	485.06	.00	444.00	.00	275.65	.00	176.20	.00	132.18	.01
CIPS	Z[t-bar]	p	Z[t-bar]	p	Z[t-bar]	p	Z[t-bar]	p	Z[t-bar]	p
MG	-9.63	.00	-9.28	.00	-2.65	.00	0.96	.83	8.05	1.00
RCM	-9.07	.00	-8.99	.00	-1.85	.03	1.10	.86	7.25	1.00
AMG(i)	-10.15	.00	-9.94	.00	-5.18	.00	1.62	.95	7.69	1.00
AMG(ii)	-12.23	.00	-11.32	.00	-5.62	.00	0.14	.56	5.63	1.00
CCE(i)	-12.88	.00	-11.18	.00	-5.61	.00	0.37	.65	8.03	1.00
CCE(ii)	-15.28	.00	-11.66	.00	-4.62	.00	0.49	.69	7.45	1.00
ARCM(i)	-9.35	.00	-8.20	.00	-4.37	.00	2.61	1.00	7.74	1.00
ARCM(ii)	-11.39	.00	-10.98	.00	-6.10	.00	0.66	.74	5.72	1.00
CIPS*	Z[t-bar]	p	Z[t-bar]	p	Z[t-bar]	p	Z[t-bar]	p	Z[t-bar]	p
MG	-10.17	.00	-8.95	.00	-4.76	.00	-2.47	.01	0.42	.66
RCM	-9.18	.00	-8.31	.00	-4.25	.00	-2.07	.02	0.76	.78
AMG(i)	-10.69	.00	-9.09	.00	-5.48	.00	-2.79	.00	-0.14	.45
AMG(ii)	-13.35	.00	-11.26	.00	-7.05	.00	-3.66	.00	-0.35	.36
CCE(i)	-14.15	.00	-13.28	.00	-7.93	.00	-6.08	.00	-2.25	.01
CCE(ii)	-16.60	.00	-14.48	.00	-8.23	.00	-5.37	.00	-1.67	.05
ARCM(i)	-9.73	.00	-8.22	.00	-4.62	.00	-1.80	.04	0.25	.60
ARCM(ii)	-11.98	.00	-10.61	.00	-6.22	.00	-2.82	.00	0.21	.58

Value-added specifications (CRS)†										
	lags: 0		lags: 1		lags: 2		lags: 3		lags: 4	
MW	p_λ	p	p_λ	p	p_λ	p	p_λ	p	p_λ	p
MG	360.10	.00	365.52	.00	236.82	.00	274.77	.00	136.90	.00
RCM	336.11	.00	349.73	.00	227.42	.00	285.17	.00	132.23	.01
AMG(i)	351.49	.00	347.73	.00	234.17	.00	160.73	.00	134.66	.01
AMG(ii)	434.12	.00	403.79	.00	254.07	.00	172.11	.00	137.53	.00
CCE(i)	399.32	.00	393.45	.00	270.81	.00	176.23	.00	149.47	.00
CCE(ii)	482.29	.00	440.86	.00	271.00	.00	184.36	.00	165.30	.00
ARCM(i)	330.71	.00	327.87	.00	217.22	.00	148.04	.00	176.24	.00
ARCM(ii)	384.88	.00	362.08	.00	224.68	.00	159.38	.00	162.32	.00
CIPS	Z[t-bar]	p	Z[t-bar]	p	Z[t-bar]	p	Z[t-bar]	p	Z[t-bar]	p
MG	-8.22	.00	-7.64	.00	-3.90	.00	0.56	.71	7.78	1.00
RCM	-7.83	.00	-6.98	.00	-3.72	.00	0.50	.69	7.56	1.00
AMG(i)	-8.97	.00	-7.19	.00	-2.90	.00	2.59	1.00	7.90	1.00
AMG(ii)	-10.98	.00	-9.86	.00	-4.77	.00	1.25	.89	7.60	1.00
CCE(i)	-9.20	.00	-9.08	.00	-3.90	.00	1.96	.97	8.62	1.00
CCE(ii)	-11.78	.00	-9.57	.00	-4.05	.00	1.54	.94	6.26	1.00
ARCM(i)	-8.55	.00	-7.32	.00	-2.93	.00	1.50	.93	6.98	1.00
ARCM(ii)	-10.35	.00	-9.13	.00	-4.72	.00	2.43	.99	7.35	1.00
CIPS*	Z[t-bar]	p	Z[t-bar]	p	Z[t-bar]	p	Z[t-bar]	p	Z[t-bar]	p
MG	-8.72	.00	-8.50	.00	-5.12	.00	-3.31	.00	0.01	.50
RCM	-8.18	.00	-8.07	.00	-4.82	.00	-2.91	.00	0.40	.65
AMG(i)	-8.72	.00	-8.13	.00	-4.87	.00	-2.14	.02	-0.12	.45
AMG(ii)	-10.60	.00	-9.76	.00	-5.83	.00	-2.45	.01	0.20	.58
CCE(i)	-9.59	.00	-8.97	.00	-5.96	.00	-2.35	.01	-0.03	.49
CCE(ii)	-11.78	.00	-10.45	.00	-6.26	.00	-3.35	.00	-1.21	.11
ARCM(i)	-8.23	.00	-7.55	.00	-4.35	.00	-1.48	.07	-0.14	.44
ARCM(ii)	-9.59	.00	-8.61	.00	-4.78	.00	-1.68	.05	0.16	.57

Notes: See Tables 2.6 and 2.7 for details.

2.2.8 Discussion of main empirical results

Our analysis investigated the changing parameter estimates across a number of empirical specifications and estimators. Our pooled estimators in levels are suggested to be severely biased, given the diagnostic tests and the fact that their capital coefficients range from .6 to .8 (VA-equivalent), far in excess of the macro evidence of around 1/3. There may be a number of possible sources for this bias: firstly, if parameter heterogeneity prevails and variable series are indeed nonstationary the imposition of common technology across countries leads to nonstationary errors and thus noncointegration. Secondly, the results may be the outcome of an identification problem, which arises if the same unobserved common factors drive both output and inputs (see Section 1.3 in Chapter 1). In this case the pooled estimates (with the exception of the CCEP) cannot identify the technology parameters $\boldsymbol{\beta}$ separately from the impact of unobserved factors $\boldsymbol{\lambda}_i$. The fact that CCEP yields very similar results to POLS and FE suggests that interplay of parameter heterogeneity and variable stationarity plays an important role.

The first difference estimator (FD-OLS) in contrast has rather sound diagnostics and yields sensible parameter coefficients. The empirical properties of this estimator are somewhat unclear, especially with regard to the identification issue, and our simulation study in Chapter 4 can throw some light on this issue.

The heterogeneous parameter estimators (averaged country regression estimators) uniformly yield lower implied capital coefficients more in line with the aggregate economy factor income share data. Across levels and first difference, gross-output and value-added specifications there seems to be a consistent pattern emerging whereby the standard heterogeneous estimators (MG, RCM) obtain qualitatively different results from the augmented heterogeneous estimators (AMG, ARCM, CCEMG). Diagnostic tests cannot differentiate between these two groups of estimators, however we can argue that the MG and RCM estimators are likely to be biased for two reasons: firstly, they use an overly simplified representation for TFP evolution (linear trend) which requires stationarity. Secondly, as was discussed above, these estimators are argued to suffer from an identification problem whereby the technology parameters $\boldsymbol{\beta}_i$ cannot be identified separately from the impact of unobserved common factors on output $\boldsymbol{\lambda}_i$ in the case where the same unobserved common factors drive output and inputs.

Turning to the Augmented estimators, we suggest that the combination of a 'common dynamic process' and a linear country trend is confirmed by the data: a considerable number of country trends are statistically significant, while the cross-country average coefficient for $\hat{\mu}_t^{\bullet}$ is close to unity in the model where it is included as additional regressor. The CCEMG estimator provides results broadly in line with those for the AMG and ARCM, with the latter two on the whole more consistent across specifications. The comparison between the country regression results presented above and the results for 'demeaned' variables in the Appendix indicate that the unobserved common factors exert differential impact across countries, thus meriting the adoption of approaches which allow for heterogeneous factor loadings (and thus TFP).

The estimate for globally common or mean TFP evolution $\hat{\mu}_t^{\bullet}$ shows serious dips following the two oil crisis shocks in 1973 and 1979, with almost continuous growth thereafter from the mid-1980s to the late 1990s — this development path is arguably in line with historical events and anecdotal evidence regarding the global evolution of manufacturing over the thirty-three-year sample period.

As a further robustness test we also carry out pooled and country regressions adopting a dynamic empirical specification — results do not change considerably in comparison to the static analysis presented here (see Section 2.D in the Appendix).

The empirical results suggest that heterogeneity in technology parameters and unobservables matters greatly for our estimations. Our next step in the analysis of manufacturing production is therefore to formally test the parameter heterogeneity finding, which will be the subject of Section 2.3. We also want to investigate the properties of the unobservable processes, TFP, and Section 2.4 will provide some important insights in this regard. Finally, the issue of variable endogeneity and the potential for 'reverse causality' in the empirical production function is addressed in Section 2.5.

2.3 Testing parameter heterogeneity

The individual country coefficients emerging from the above regressions imply considerable parameter heterogeneity across countries. However, this apparent heterogeneity may be due to sampling variation and the relatively limited number of time-series observations in each country individually (Pedroni, 2007). We therefore carry out a number of parameter heterogeneity tests for the results from the various CCEMG and augmented MG/RCM estimations.

As a *first test*, we compute the residuals in case of parameter homogeneity for each country

$$\begin{aligned} H_{het} &\equiv o_{it}^{\bullet} - \bar{b}\, k_{it} - \bar{c}\, m_{it} - \bar{\mu} t - \bar{A}_0 \\ H_{het} &\equiv o_{it} - \bar{b}\, k_{it} - \bar{c}\, m_{it} - \bar{\mu} t - \bar{d}\hat{\mu}_t^{\bullet} - \bar{A}_0 \end{aligned}$$

where $\bar{b}$, $\bar{c}$ and $\bar{d}$ (for Augmented models) are the *mean* estimates for capital per worker (k), materials per worker (m) and the common dynamic process taken from the results in Table 2.3 (Table 2.4 for the regressions with variables in first differences), with $\bar{\mu}$ the average country trend term and $\bar{A}_0$ the average intercept term (the latter is not important for this analysis).[28] The common dynamic process is either subtracted from the output variable ($o_{it}^{\bullet}$) or included as indicated above. Similarly for the other models, the VA specifications and the specifications in first differences. In a second step, we regress the residuals created on the input variables, a country trend or drift term and country- and year-dummies in a pooled regression

$$H_{het} = \pi_b\, k_{it} + \pi_c\, m_{it} + \pi_d\, t\, (+ \sum_i \pi_{e,i}) \tag{2.23}$$

The rationale behind this test is as follows: if factor input parameters were truly heterogeneous across countries, we would expect the pooled regression to produce statistically significant coefficients ($\pi_j \neq 0$). Results are presented in Tables 2.9 and 2.10.

As can be seen the levels regressions imply that capital parameter homogeneity is rejected, while the materials coefficients are more likely to be homogeneous (in the gross output specification). In the VA-specification capital parameter homogeneity is rejected in all models. In contrast the tests for the first difference specifications on the whole do not provide much evidence for heterogeneity, with all covariates insignificant with the exception of the case of

[28] This testing procedure was suggested by Steve Bond.

Table 2.9: Parameter Heterogeneity — Pooled Tests (levels)

Parameter Heterogeneity — Pooled Tests (levels) pooled fixed effects regressions; dependent variable: residuals H_{het} as defined in main text								
(i) Gross output regressions								
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
estimator	**MG**	**RCM**	**AMG**	**AMG**	**CCEMG**	**CCEMG**	**ARCM**	**ARCM**
specification	(a)	(a)	(b)	(c)	(d)	(d)	(b)	(c)
dependent variable♯	H_{het}	H_{het}	H_{het}	H_{het}	H_{het}	H_{het}	H_{het}	H_{het}
regressors								
capital pw	0.1200 [10.23]**	0.0993 [8.47]**	0.1007 [8.77]**	0.0910 [7.76]**	0.0538 [4.66]**	0.0912 [7.94]**	0.0877 [7.64]**	0.0780 [6.79]**
materials pw	0.0224 [1.64]	0.0009 [0.06]	0.0119 [0.88]	0.0011 [0.08]	0.0157 [1.18]	0.0265 [1.99]*	-0.0094 [0.70]	-0.0092 [0.68]
country trends	-0.0027 [7.69]**	-0.0016 [4.70]**	-0.0020 [5.78]**	0.0014 [4.06]**	0.0004 [1.22]**	-0.0019 [5.36]**	-0.0012 [3.46]**	-0.0010 [2.89]**
intercept terms	all sign at 1%	all sign at 1%	all sign at 1%	all sign at 1%	all sign at 1%	all sign at 1%	all sign at 1%	all sign at 1%
obs	1,162	1,162	1,162	1,162	1,162	1,162	1,162	1,162

(ii) Value-added regressions								
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
estimator	**MG**	**RCM**	**AMG**	**AMG**	**CCEMG**	**CCEMG**	**ARCM**	**ARCM**
specification	(a)	(a)	(b)	(c)	(d)	(d)	(b)	(c)
dependent variable♯	H^{va}_{het}	H^{va}_{het}	H^{va}_{het}	H^{va}_{het}	H^{va}_{het}	H^{va}_{het}	H^{va}_{het}	H^{va}_{het}
regressors								
capital pw	0.4704 [15.94]**	0.4242 [14.38]**	0.3733 [12.95]**	0.3511 [11.90]**	0.197 [6.55]**	0.3517 [12.00]**	0.3072 [10.65]**	0.3083 [10.68]**
country trends	-0.0112 [11.72]**	-0.0026 [2.73]**	-0.0088 [9.07]**	0.0039 [4.07]**	-0.0018 [1.87]**	-0.0087 [9.13]**	-0.0068 [7.04]**	-0.0071 [7.33]**
intercept terms	all sign at 1%	all sign at 1%	all sign at 1%	all sign at 1%	all sign at 1%	all sign at 1%	all sign at 1%	all sign at 1%
obs	1,194	1,194	1,194	1,194	1,194	1,194	1,194	1,194

Notes: All variables are in logs. Values in brackets are absolute t-statistics. 'Specification' refers to equations (2.19) to (2.22). The models underlying the construction of H_{het} are presented in Table 2.3 above. We indicate statistical significance at the 5% and 1% level by * and ** respectively.

Table 2.10: Parameter Heterogeneity — Pooled Tests (FD)

Parameter Heterogeneity — Pooled Tests (FD) pooled fixed effects regressions; dependent variable: residuals H_{het}								
(i) Gross output regressions								
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
estimator	**ΔMG**	**ΔRCM**	**ΔAMG**	**ΔAMG**	**ΔCCEMG**	**ΔCCEMG**	**ΔARCM**	**ΔARCM**
specification	(a)	(a)	(b)	(c)	(d)	(d)	(b)	(c)
dependent variable♯	H_{het}	H_{het}	H_{het}	H_{het}	H_{het}	H_{het}	H_{het}	H_{het}
regressors								
capital pw	-0.0102 [0.36]	-0.0169 [0.60]	-0.0264 [0.95]	-0.0367 [1.32]	-0.0663 [2.35]*	-0.0416 [1.50]	-0.0280 [1.01]	-0.0274 [0.99]
materials pw	0.0226 [0.91]	0.0124 [0.50]	0.0163 [0.67]	0.0236 [0.98]	0.0198 [0.81]	0.0145 [0.60]	0.0073 [0.30]	0.0122 [0.50]
drift terms	only 1 sign.	only 1 sign.	only 1 sign.	only 1 sign.	only 1 sign.	only 1 sign.	only 1 sign.	only 1 sign.
obs	1,094	1,094	1,094	1,094	1,094	1,094	1,094	1,094

(ii) Value-added regressions								
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
estimator	**ΔMG**	**ΔRCM**	**ΔAMG**	**ΔAMG**	**ΔCCEMG**	**ΔCCEMG**	**ΔARCM**	**ΔARCM**
specification	(a)	(a)	(b)	(c)	(d)	(d)	(b)	(c)
dependent variable♯	H^{va}_{het}	H^{va}_{het}	H^{va}_{het}	H^{va}_{het}	H^{va}_{het}	H^{va}_{het}	H^{va}_{het}	H^{va}_{het}
regressors								
capital pw	0.0989 [1.11]	0.0546 [0.61]	0.0157 [0.18]	0.0069 [0.08]	-0.0791 [0.88]	-0.0202 [0.22]	-0.0162 [0.18]	-0.0106 [0.12]
drift terms	only 2 sign.	only 2 sign.	only 2 sign.	only 2 sign.	only 2 sign.	only 2 sign.	only 2 sign.	only 2 sign.
obs	1,128	1,128	1,128	1,128	1,128	1,128	1,128	1,128

Notes: See Table 2.9 for details. The results for the country regressions in first difference tested here for parameter heterogeneity are presented in Table 2.4.

CCEMG in first differences. Note that the kernel densities for the technology parameters underlying the above heterogeneity tests do not differ considerably between levels and FD specifications (FD densities not reported). This stark difference is therefore likely to be driven by the impact of nonstationarity on the test.

Secondly, we report the Swamy (1970) $\hat{S}$ statistic from the gross output and VA regressions in levels and first differences in Table 2.11.[29] For a detailed discussion of this test see Pesaran and Yamagata (2008). 'Specification' refers to the country regression models in equations (2.19) to (2.21), as in our regression results section. Note that the test for the equation in levels is testing heterogeneity of *all* parameters, including the intercepts; since the assumption of heterogeneous TFP levels is rather uncontroversial, this test does not adquately address our interest in the homogeneity of technology parameters. We therefore also provide a test for the levels specification where the intercept terms have been dispensed with via transformation of the data into mean-deviations. Estimates for this specification are of course identical to those of the untransformed levels equation (see references in footnote 29).

Table 2.11: Parameter Heterogeneity — Swamy (1970) Tests

Swamy (1970) tests H_0: homogeneity across countries; p-values are reported in parentheses; 'MD' refers to variables in deviations from country-specific means ('within' transformation)						
Levels and mean-deviation (MD) specifications†						
	RCM		**ARCM**			
Specification	(a) Levels	(a)' MD	(b) Levels	(b)' MD	(c) Levels	(c)' MD
Gross Output	30,598.1 (.00)	2,198.2 (.00)	41,427.1 (.00)	2,119.7 (.00)	45,871.5 (.00)	2,598.4 (.00)
Value-Added	51,123.0 (.00)	1,531.6 (.00)	62,499.9 (.00)	1,440.4 (.00)	69,157.0 (.00)	1726.7 (.00)

Specifications in first differences (FD)						
	RCM		**ARCM**			
Specification	(a) FD		(b) FD		(c) FD	
Gross Output	359.5 (.00)		332.3 (.00)		453.4 (.00)	
Value-Added	191.1 (.00)		153.5 (.00)		258.7 (.00)	

Notes: Swamy $\hat{S}$ is distributed χ^2 with $k(N-1)$ degrees of freedom. † Data in mean-deviations.

The Swamy $\hat{S}$ test rejects parameter heterogeneity for all specifications tested. In general, this test was developed for panels where N is large relative to T. Using Monte Carlo experiments, Pesaran and Yamagata (2008) show that in case of a panel of $T = 30, N = 50$ the test has power but tends to over-reject — a tendency which becomes worse with the number of parameters included in the model.[30] Further, as Pedroni (2007) points out, the Swamy-based tests

[29] The levels and FD tests are taken from the regressions in columns [2], [7] and [8] of Tables 2.3 and 2.4.

[30] The adjusted Swamy statistic $\tilde{\Delta}_{adj}$ developed by the same authors, although appropriately sized, suffers

are not designed for nonstationary panel data.

Thirdly, we produce Wald statistics, as suggested by Canning and Pedroni (2008)

$$W_\theta = \sum_i \frac{(\hat{\theta}_i - \bar{\theta})^2}{\mathbb{V}ar(\theta_i)} \qquad W_\theta \sim \chi^2(N)$$

where $\hat{\theta}_i$ is the parameter coefficient from the country regression, $\mathbb{V}ar(\theta_i)$ its variance, and $\bar{\theta}$ is the unweighted average parameter estimate across all countries. If parameters are similar across countries, the statistic will be low, whereas if parameters are heterogeneous W_θ will be larger. The validity of this test depends on T being moderate to large. The null for this test is that *all* countries have the same parameter value. Table 2.12 presents the summed Wald statistics for the entire sample, as well as an indication of the share of country-specific tests rejecting the null of equality between country estimate and full sample mean estimate (for both the levels and FD specifications).

The Wald tests reject homogeneity for the factor parameters derived from the levels models in case of both the gross-output and value-added specifications. The statistics are particularly large for the trend terms in the levels specifications, thus rejecting homogeneity emphatically, which is not always the case for the drift terms in the first-difference specifications. Turning to the share of countries rejecting parameter homogeneity, it can be seen that roughly half of all countries reject homogeneity for all covariates in the levels specifications. This share falls to less than one third in the models in first differences.

Fourthly, following Pedroni (2007), we produce an F-statistic for the standard and augmented MG and RCM regression models (Pesaran & Yamagata, 2008, p.52)

$$\begin{aligned} F &= \left(\frac{RSS_{hom} - RSS_{het}}{RSS_{het}}\right)\left(\frac{df_D}{df_N}\right) \\ F &\sim F(df_N, df_D) \qquad df_N = k \times (N-1) \qquad df_D = N(\bar{T} - k - 1) \end{aligned}$$

where k is the number of parameters in each country-regression and RSS_{hom} and RSS_{het} are the sums of the squared residuals of the homogeneous and heterogeneous models respec-

from low power in a sample of $T = 30, N = 50$, in particular if errors are non-normal.

Table 2.12: Parameter Heterogeneity — Wald Tests (levels and FD)

Wald Tests (levels)
Full sample Wald statistics (W_θ) and share of countries rejecting H_0 of parameter homogeneity

(i) Gross output regressions

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
estimator	**MG**	**RCM**	**AMG**	**AMG**	**CCEMG**	**CCEMG**	**ARCM**	**ARCM**
specification	(a)	(a)	(b)	(c)	(d)	(d)	(b)	(c)
full sample W_θ								
capital pw (k)	667.3**	406.4**	517.1**	553.9**	397.7**	318.6**	323.9**	298.8**
materials pw (m)	599.4**	368.3**	518.4**	626.5**	588.6**	375.0**	325.8**	329.7**
country trends (t)	932.2**	603.5**	802.0**	264.7**		298.0**	530.3**	245.4**
Country-specific $W_{\theta,i}$								
share rejecting H_0: k	56%	46%	44%	54%	58%	40%	44%	50%
share rejecting H_0: m	56%	54%	44%	54%	58%	40%	44%	56%
share rejecting H_0: t	54%	56%	54%	50%		48%	54%	48%
obs (countries)	1,162 (48)	1,162 (48)	1,162 (48)	1,162 (48)	1,162 (48)	1,162 (48)	1,162 (48)	1,162 (48)

(ii) Value-added regressions

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
estimator	**MG**	**RCM**	**AMG**	**AMG**	**CCEMG**	**CCEMG**	**ARCM**	**ARCM**
specification	(a)	(a)	(b)	(c)	(d)	(d)	(b)	(c)
full sample W_θ								
capital pw (k)	578.1**	388.1**	474.0**	542.3**	515.5**	548.0**	320.1**	354.9**
country trends	1,044.3**	720.9**	781.8**	267.7**		374.4**	636.7**	194.5**
Country-specific $W_{\theta,i}$								
share rejecting H_0: k	52%	46%	52%	56%	54%	46%	50%	54%
share rejecting H_0: t	65%	56%	58%	44%		42%	52%	40%
obs (countries)	1,194 (48)	1,194 (48)	1,194 (48)	1,194 (48)	1,194 (48)	1,194 (48)	1,194 (48)	1,194 (48)

Wald Tests (FD)
Full sample Wald statistics (W_θ) and share of countries rejecting H_0 of parameter homogeneity

(i) Gross output regressions

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
estimator	**MG**	**RCM**	**AMG**	**AMG**	**CCEMG**	**CCEMG**	**ARCM**	**ARCM**
specification	(a)	(a)	(b)	(c)	(d)	(d)	(b)	(c)
full sample W_θ								
capital pw (k)	130.2**	119.0**	102.2**	116.0**	101.7**	134.9**	92.8**	113.6**
materials pw (m)	263.2**	160.3**	248.2**	290.0**	176.2**	206.2**	156.2**	174.3**
country drifts	87.2**	81.7**	69.0*	112.9**		109.7**	67.0*	68.0*
Country-specific $W_{\theta,i}$								
share rejecting H_0: k	23%	23%	27%	29%	25%	25%	23%	21%
share rejecting H_0: m	23%	29%	27%	29%	25%	25%	27%	29%
share rejecting H_0: drift	15%	13%	13%	13%		23%	17%	13%
obs (countries)	1,094 (48)	1,094 (48)	1,094 (48)	1,094 (48)	1,094 (48)	1,094 (48)	1,094 (48)	1,094 (48)

(ii) Value-added regressions

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
estimator	**MG**	**RCM**	**AMG**	**AMG**	**CCEMG**	**CCEMG**	**ARCM**	**ARCM**
specification	(a)	(a)	(b)	(c)	(d)	(d)	(b)	(c)
full sample W_θ								
capital pw (k)	166.1**	147.1**	117.1**	139.8**	125.2**	141.5**	115.9**	132.4**
country drifts	80.5**	73.5*	51.7	109.5**		74.7**	54.0	86.8**
Country-specific $W_{\theta,i}$								
share rejecting H_0: k	25%	33%	33%	31%	35%	29%	31%	33%
share rejecting H_0: drift	19%	17%	15%	21%		15%	15%	21%
obs (countries)	1,128 (48)	1,128 (48)	1,128 (48)	1,128 (48)	1,128 (48)	1,128 (48)	1,128 (48)	1,128 (48)

Notes: See main text for construction of the Wald statistic W_θ. 'Specification' refers to the empirical models presented in equations (2.19) to (2.22). The models underlying the construction of the Wald statistics are presented in Table 2.3 above. In the full sample tests $W_\theta \sim \chi^2(48)$, with 5% and 1% critical values 65.17 and 73.70 respectively ($W_\theta = \sum_i W_{\theta,i}$); for country-specific tests ($W_{\theta,i}$) we apply the 10% critical value of 2.7. The null hypothesis in all cases is parameter homogeneity. For W_θ we indicate statistical significance at the 5% and 1% level by * and ** respectively.

tively — in the former case the mean coefficient estimates are imposed. This tests the full parameter heterogeneity versus the full homogeneity case. We do not compute F-tests for the CCEMG models, as the period-average parameters are not meant to be identical in the country-regression case.

Table 2.13: Parameter Heterogeneity — F-Tests

Panel F-tests Full sample statistics; H_0: model parameters are homogeneous						
(i) Gross output regressions						
	[1]	[2]	[3]	[4]	[5]	[6]
estimator	**MG**	**RCM**	**AMG**	**AMG**	**ARCM**	**ARCM**
specification	(a)	(a)	(b)	(c)	(b)	(c)
levels						
F	283.1 (.00)	199.7 (.00)	209.0 (.00)	207.1 (.00)	169.2 (.00)	151.4 (.00)
distr	$F(188, 922)$	$F(188, 922)$	$F(188, 922)$	$F(235, 874)$	$F(188, 922)$	$F(235, 874)$
first differences						
F	2.4 (.00)	1.9 (.00)	2.5 (.00)	1.7 (.00)	2.0 (.00)	2.2 (.00)
distr	$F(141, 902)$	$F(141, 902)$	$F(141, 902)$	$F(188, 854)$	$F(141, 902)$	$F(188, 854)$
(ii) Value-added regressions						
	[1]	[2]	[3]	[4]	[5]	[6]
estimator	**MG**	**RCM**	**AMG**	**AMG**	**ARCM**	**ARCM**
specification	(a)	(a)	(b)	(c)	(b)	(c)
levels						
F	413.7 (.00)	334.1 (.00)	339.9 (.00)	279.4 (.00)	287.5 (.00)	232.3 (.00)
distr	$F(141, 1002)$	$F(141, 1002)$	$F(141, 1002)$	$F(188, 954)$	$F(141, 1002)$	$F(188, 954)$
first differences						
F	2.0 (.00)	1.5 (.00)	2.6 (.00)	2.4 (.00)	1.6 (.00)	1.8 (.00)
distr	$F(94, 950)$	$F(94, 950)$	$F(94, 950)$	$F(141, 902)$	$F(94, 950)$	$F(141, 902)$

Notes: See main text for construction of the Panel F statistic. 'Specification' refers to the empirical models presented in equations (2.19) to (2.22). The models underlying the construction of the F statistics are presented in Table 2.3 and 2.4 above. The null hypothesis in all cases is parameter homogeneity.

The F tests are valid for fixed N, when the regressors are strictly exogenous and the error variances are homoskedastic (Pesaran et al., 2008).[31]

All of the test results presented in Table 2.13 reject parameter homogeneity for the factor input variables at the 1% level of significance. It is intuitive why the test statistics may emphatically reject the null: if the homogeneity restriction is incorrect, the country regressions do not cointegrate under the null, such that the regression errors will be nonstationary. As a result the F-statistic will quickly diverge and reject the null (Pedroni, 2007).

Like in the Swamy $\hat{S}$ Test we are faced with the problem that the tests evaluate the *full* regres-

[31] In the levels specifications, $k = 4$ includes technology parameters, intercept and trend terms ($k = 3$ in the VA case); in the first difference ones $k = 3$ includes technology parameters and drift terms ($k = 2$ in the value-added specification). N is the number of countries, 48, and $\bar{T}$ represents the average time series length, s.t. $N\bar{T} = 1,162$ (VA:1,194) in the levels and $N\bar{T} = 1,094$ (1,128) in the FD case.

sion model for the null of parameter homogeneity, which is not sensible in the levels regression case since heterogeneous intercepts are commonly accepted in the literature. In order to bypass this problem we also computed F-statistic for the levels MG and Augmented MG cases where the intercepts have been dispensed with by taking all variables in deviations from the country-mean — all of these reject parameter homogeneity at the 1% level.[32]

Taken together the various diagnostic tests we carried out in this section do give a strong indication that parameter homogeneity is rejected. The differences in the results for levels and first difference specifications however indicate that nonstationarity may drive some of the results reported. Nevertheless, even if heterogeneity were not very significant in qualitative terms, our contrasting of pooled and country regression results has shown that it nevertheless matters greatly for correct empirical analysis in the case of nonstationary variable series.

Further parameter heterogeneity tests were considered for this analysis: Pesaran and Yamagata (2008) compare their own version of Swamy's test of parameter homogeneity (denoted $\tilde{\Delta}$) with the 'traditional' Swamy test and F-Test we computed above, a Hausman-type comparison of fixed effects and mean group estimates and the Phillips and Sul (2003) G-test. Their Monte Carlo experiments suggest that all of these tests have low power in panels with the dimensions we observe ($N = 48, T \approx 24$) and we therefore did not pursue any of these in the present analysis.

2.4 Estimating TFP

Most production function estimations found in the literature are just the first step in an empirical analysis that concentrates on the determinants of Total Factor Productivity (TFP). In our analysis we have thus far focused on the factor parameter estimates and their magnitudes and robustness across specifications. We do however also want to provide some estimates for TFP and its evolution. Two alternative approaches are open to us: firstly, we can rely on the parameter estimates from our country regressions; secondly, we can carry out productivity levels accounting by imposing the country-specific factor parameter estimates on the data and

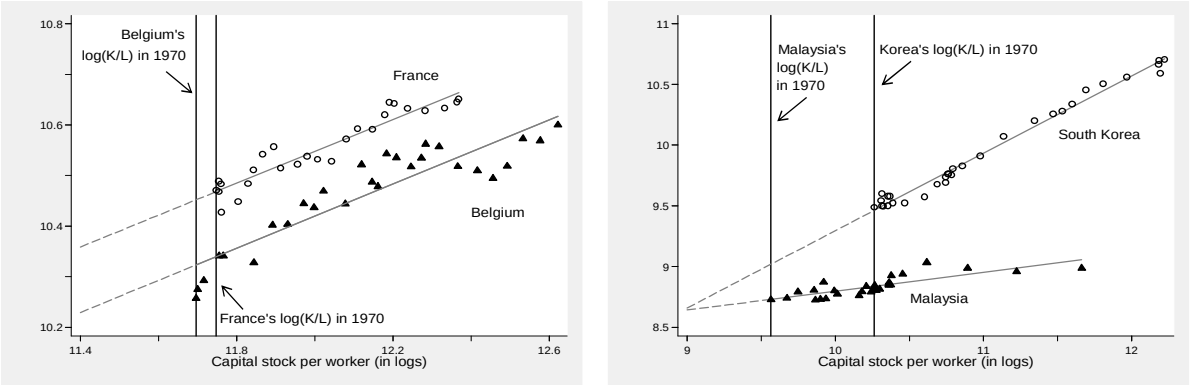
[32]Gross-output: MG $F = 47.6$, distribution $F(141, 970)$, $p = .00$; AMG, specification (b), $F = 13.1$, distribution $F(141, 970)$, $p = .00$; AMG, specification (c), $F = 13.6$, distribution $F(188, 922)$, $p = .00$. Value-added: MG $F = 19.2$, distribution $F(94, 1050)$, $p = .00$; AMG, specification (b), $F = 16.1$, distribution $F(94, 1050)$, $p = .00$; AMG, specification (c), $F = 12.8$, distribution $F(141, 1002)$, $p = .00$.

collecting the residuals as 'measure of our ignorance'. We discuss both approaches in the following.

2.4.1 Deriving TFP via regression estimates

From the country regressions we can obtain estimates for the intercept, technology parameters, idiosyncratic and common trend coefficients or the parameters on the cross-section averages for AMG and CCEMG specification respectively. One could be tempted to view the coefficients on the intercepts as TFP level estimates, just like in the pooled fixed effects case. However, once we allow for heterogeneity in the slope coefficients, the interpretation of the intercept as estimate for base-year TFP level is no longer valid. Consider the example in Figure 2.2.

Figure 2.2: Regression intercepts and TFP level estimates



We provide scatter plots for 'adjusted' log value-added per worker (y-axis) against log capital per worker (x-axis) as well as a fitted regression line for these observations in each of the following four countries: in the left panel France (circles) and Belgium (triangles), in the right panel South Korea (circles) and Malaysia (triangles). The 'adjustment' is based on the country-specific estimates from the AMG regression (value-added version) in specification (c) and constructed in the following steps:[33] consider the fitted values of this estimation

$$\hat{y}_{it} = \hat{b}_i \log(\mathrm{K/L})_{it} + \hat{a}_i + \hat{c}_i t + \hat{d}_i \hat{\mu}_t^{\bullet} \tag{2.24}$$

where $\hat{b}_i$, $\hat{a}_i$, $\hat{d}_i$ and $\hat{c}_i$ are country-specific estimates for log capital per worker, the intercept, the common dynamic process and a linear trend term. In order to illustrate our case, we want to obtain a simple linear relationship between value-added and capital where the contribution

[33] The results applied are those presented in the lower panel of Table 2.3, column [4].

of TFP growth has already been accounted for. We compute

$$y_{it}^{adj} = y_{it} - \hat{c}_i t - \hat{d}_i \hat{\mu}_t^{\bullet} \tag{2.25}$$

We then plot this variable against log capital per worker for each country separately. This procedure in essence provides a two-dimensional visual equivalent of the estimates for the capital coefficient (slope) and the TFP level (intercept) in the augmented country regression. The left panel of Figure 2.2 shows two countries (France, Belgium) with virtually identical capital coefficient estimates (slopes). The in-sample fitted regression line is plotted as a solid line, the out-of-sample extrapolation toward the y-axis is plotted in dashes. The country-estimates for the intercepts can be interpreted as TFP level, since the countries have very similar capital coefficients ($b_i \approx b_j$). In this case, the graph represents the linear model $y_{it}^{adj} = \hat{a}_i + \hat{b}\log(K/L)_{it}$, where $\hat{a}_i$ possesses the *ceteris paribus* property. In contrast, the right panel shows two countries which exhibit very different capital coefficient estimates (slopes). In this case $\hat{a}_i$ cannot be interpreted as possessing the *ceteris paribus* quality since $\hat{b}_i \neq \hat{b}$: *ceteris non paribus*! In the graph we can see that Malaysia has a considerably higher intercept term than South Korea, even though the latter's observations lie above those of the former at any given point in time. This illustrates that once technology parameters in the production function differ across countries the regression intercept can no longer be interpreted as a TFP-level estimate.

We now want to suggest an alternative measure for TFP-level which is robust to parameter heterogeneity. Referring back to the scatter plots in Figure 2.2, we marked the base-year level of log capital per worker by vertical lines for each of the four countries. We suggest to use the locus where the solid (in-sample) regression line hits the vertical base-year capital stock level as an indicator of TFP-level in the base year. In the value-added specification these *adjusted* base-year and final-year TFP-levels are thus

$$\hat{b}_i \log(\mathrm{K/L})_{0,i} + \hat{a}_i \qquad \text{and} \qquad \hat{b}_i \log(\mathrm{K/L})_{0,i} + \hat{a}_i + \hat{c}_i \tau + \hat{d}_i \hat{\mu}_{\tau}^{\bullet} \tag{2.26}$$

respectively, where $\log(K/L)_{0,i}$ is the country-specific base-year value for capital per worker (in logs), τ is the total period for which country i is in the sample and $\hat{\mu}_{\tau}^{\bullet}$ is the accumulated *common* TFP growth for this period τ with the country-specific parameter d_i. It is easy to see that the intercept-problem discussed above only has bearings on TFP-*level* estimates. A

similar formula applies for the CCEMG estimator[34]

$$\hat{b}_i \log(\mathrm{K/L})_{0,i} + \hat{a}_i + \hat{d}_{1i}\,\bar{y}_0 + \hat{d}_{2i}\,\bar{k}_0 \tag{2.27}$$

Focusing exclusively on the TFP-levels in the base-year, Table 2.25 in Appendix 2.E presents the rank (by magnitude) for adjusted TFP levels derived from the AMG estimator in specification (*c*), i.e. with $\hat{\mu}_t^{\bullet}$ included as additional regressor, and the standard CCEMG estimator (in specification (*d*)). We further show the country ranking based on a pooled fixed effects estimation, such as that presented in Table 2.3, column [2]. As can be seen in the right half of the table, the rankings differ considerably between the AMG and CCEMG results on the one hand and standard fixed effects results on the other (median absolute rank difference (MARD): 10, respectively), but not between the AMG and CCEMG results (MARD: 1).

We present adjusted TFP base-year and final-year levels for these AMG and CCEMG models in Figure 2.5 in Appendix 2.E. Note that in our sample base- and final-year differ across countries (see Table 2.16 in Appendix 2.A). The countries in both charts are arranged in order of magnitude of their final-year adjusted TFP levels *in the AMG(ii) model*, for which results are shown in the left bar-chart. With the exception of a small number of countries (e.g. MLT — Malta, SGP — Singapore) the general ordering of countries by final-year TFP levels is very similar in the two specifications: countries such as Finland, Canada, the United States or Ireland can be found toward the top of the ranking, with Bangladesh, India, Sri Lanka and Poland closer to the bottom.

2.4.2 Using accounting strategies for TFP determination

The primary interest of the Pesaran (2006) CCE estimation approach is to obtain robust estimates for the 'betas', the parameters on the observables, save from the influence of unobserved common factors, and *not* to determine the nature of these factors (Kapetanios et al., 2009). If the analysis is interested in obtaining estimates of the latter along with country factor loadings, Pesaran suggests to first compute the following residuals

$$\hat{r}_{it} = y_{it} - \hat{\boldsymbol{\theta}}_i \boldsymbol{x}_{it} \tag{2.28}$$

[34] In analogy for the end-year TFP level where $\bar{y}_0$ and $\bar{k}_0$ are replace with $\bar{y}_T$ and $\bar{k}_T$.

where applied to our production function context y is the country-level output variable, $\boldsymbol{x}$ is the vector of factor-inputs for each country and $\hat{\boldsymbol{\theta}}_i$ is the matrix of country-specific technology parameter estimates *obtained from the CCEMG regression.* In a second step these residuals are then investigated using Principle Component Analysis (PCA).[35] However, serious difficulties arise from this two-step strategy with regard to our own application: our panel is unbalanced with missing observations and therefore the same applies to the residuals $\hat{r}$ created. Standard versions of PCA, i.e. readily coded algorithms in `STATA` and other statistical programs, however rely on investigation of variables (here: $\hat{r}$) which have *the same dimensions.* As a result the application of PCA to the residual series fails, as there is no one year for which we have data for all countries.

2.5 Reverse causality

In the analysis of empirical production functions the issue of variable endogeneity is typically of great concern, requiring means and ways to instrument for factor inputs. A popular approach to address this problem is to convert macro-panel data into a short panel of (typically) 5-year averages and to employ empirical estimators originally developed for large N, small T micro-panels (Arellano & Bond, 1991; Blundell & Bond, 1998). In Chapter 1 we discussed the Caselli et al. (1996) empirics for Penn World Table data which adopt this approach. For analysis of other macroeconomic relationships it is also not uncommon to see these estimators used for *annual* data where T becomes moderate to large, despite the difficulties arising from overfitting (Bowsher, 2002; Roodman, 2009). We already highlighted the problems arising for these GMM-type estimators when model parameters differ across countries (Pesaran & Smith, 1995), however a more detailed discussion is beyond the scope of this thesis. In the following we instead adopt the empirical approach by Pedroni (2007) and estimate country regressions by Fully-Modified OLS, whereupon parameter estimates are averaged across countries ('Group-Mean FM-OLS' — Pedroni, 2000). Provided the variables are nonstationary *and* cointegrated the FMOLS parameter estimates are super-consistent.

In the upper panel of Table 2.14 we present averaged parameter estimates for the full sample value-added-based FMOLS regressions, where column [**1**] represents the standard Group-Mean

[35] The discussion presented is the outcome of a conversation by the author with Hashem Pesaran during the 'Unobserved Common Factors' conference in Maastricht in September 2008.

FMOLS, and columns **[2]** and **[3]** augment the country FMOLS equation with the 'common dynamic process' $\hat{\mu}_t^{\bullet}$. Columns **[4]** and **[5]** apply FMOLS to a country regression with the CCE-augmentation before averaging the parameter estimates.

Table 2.14: Country regressions using FMOLS

Averaged Fully Modified OLS estimates
We apply the Phillips and Hansen (1990) FM-OLS estimator to each country;
estimates reported are unweighted means of country estimates

Value-added specification: Full sample					
estimator: **FMOLS-**	**MG**	**AMG**	**AMG**	**CCEMG**	**CCEMG**
dependent variable	ly	ly-$\hat{\mu}_t^{\bullet,va}$	ly	ly	ly
	[1]	[2]	[3]	[4]	[5]
regressors					
capital pw	0.1663	0.2659	0.2937	0.5544	0.3042
	[1.99]	[3.32]**	[3.18]**	[8.05]**	[3.33]**
common process			0.8977		
			[3.49]**		
country trends	0.0171	0.0004	0.0014		0.0108
	[5.50]**	[0.12]	[0.29]		[2.95]**
intercept	-4.6095	6.5985	7.1405	5.6737	-1.6007
	[1.63]	[2.69]**	[1.52]	[5.57]**	[0.57]
Panel-t statistics, trends†					
capital pw	18.29	14.73	15.36	40.59	15.88
trends	24.94	18.93	12.71		20.70
# of sign. trends (at 10%)	37	25	23		28
RMSE	.099	.096	.090	.103	.088
obs (countries)	1,152 (48)	1,152 (48)	1,152 (48)	1,122 (48)	1,122 (48)

Value-added specification: I(1) sample					
estimator: **FM-**	**MG**	**AMG**	**AMG**	**CCEMG**	**CCEMG**
dependent variable	ly	ly-$\hat{\mu}_t^{\bullet,va}$	ly	ly	ly
	[1]	[2]	[3]	[4]	[5]
regressors					
capital pw	0.0816	0.2675	0.2784	0.5528	0.2485
	[1.27]	[4.11]**	[3.08]**	[7.38]**	[3.13]**
common process			0.8034		
			[4.61]**		
country trends	0.0179	-0.0012	0.0019		0.0108
	[5.64]**	[0.40]	[0.41]		[2.56]*
intercept	-4.9092	6.1012	5.2980	5.3936	-2.3070
	[1.59]	[2.63]*	[1.50]	[4.60]**	[0.76]
Panel-t statistics, trends†					
capital pw	11.45	10.37	9.97	34.96	10.16
trends	23.28	14.63	10.56		17.10
# of sign. trends (at 10%)	23	15	15		16
RMSE	.071	.068	.065	.080	.062
obs (countries)	622 (26)	622 (26)	622 (26)	593 (26)	593 (26)

Notes: The results in [1] are for the Pedroni (2000) Group-Mean FMOLS estimator; the results in the remaining columns allow for cross-section dependence using either $\hat{\mu}_t^{\bullet,va}$ or cross-section averages in the FMOLS country regressions. In all cases the estimates presented are the unweighted means of the FMOLS country estimates. The lower panel uses observations from only those countries for which variables were determined to be nonstationary (via ADF and KPSS testing). Due to the instrumentation in the FMOLS procedure we lose one observation for each time-series. RMSE reports the root mean squared errors.

In the lower panel of the same table our sample only includes 26 countries which 'pass' two nonstationarity test (KPSS, ADF) for value-added per worker and capital stock per worker

respectively. Due to the unbalanced nature of our panel we are prevented from testing for cointegration between these two variables.

The standard Pedroni (2000) approach yields insignificant capital estimates, whereas augmentation with the common dynamic process yields statistically significant estimates very close to those arising from our previous AMG regressions. The inclusion of $\hat{\mu}_t^{\bullet,va}$ thus yields the same parameter estimates in the Group-Mean Fully-Modified OLS and Mean Group OLS approaches. In analogy to the results in Table 2.3 the CCE-augmented average estimate is somewhat larger at .55, although inclusion of a linear country-trend yields an average estimate for capital of around .3.

These results are robust to restricting the sample to countries for which value-added and capital stock per worker 'pass' the nonstationarity tests: for the Pedroni (2000) estimator the capital stock is insignficant, whereas the inclusion of the common dynamic process replicates our earlier findings. We take these results as a vindication that the production process represents a three-way cointegration between output, factor-inputs and some unobserved common factors.

2.6 Concluding Remarks

This chapter set out to provide an empirical implementation of the issues discussed in the first chapter, namely the interplay of parameter heterogeneity, cross-section dependence and variable time-series properties in the estimation of production functions using macro panel data. We discussed a number of standard approaches to empirical implementation and introduced our novel Augmented Mean Group estimator (AMG), which is conceptually close to the Mean Group version of the Pesaran (2006) Common Correlated Effects estimator (CCEMG). Both of these aproaches introduce the notion of a globally common, unobserved factor (or factors) interpreted either as common TFP or an average of country-specific TFP evolution. While in the CCEMG this 'common dynamic process' is only implicit, the AMG approach uses an explicit estimate for this process in the augmentation of country-regressions.

We have shown that the adoption of a heterogeneous technology model in combination with a common factor representation for TFP and for the evolution of observable input variables en-

ables us to conceptualise and model a number of characteristics which are likely to be prevalent in manufacturing data from a diverse sample of developing and developed countries:

- *firstly*, we allowed for cross-country heterogeneity in the impact of observables on output. Empirical results are confirmed by the formal testing procedures in Section 2.3 to suggest that 'technology parameters' in manufacturing production indeed differ across countries. This result is in support of the earlier findings by Durlauf (2001) and Pedroni (2007) using aggregate economy data: if the production technology differs in cross-country manufacturing, aggregate economy technology is unlikely to be homogeneous. The result of production technology heterogeneity across countries has immediate implications for standard empirical approaches to TFP analysis: it leads to the breakdown of the interpretation of regression intercept terms as TFP-levels estimates, as we have shown in Section 2.4. We therefore introduce a new procedure to compute 'adjusted' TFP-level estimates which is robust to parameter heterogeneity and can thus be compared across countries and between pooled and heterogeneous parameter models. Further analysis highlighted the significant differences between 'adjusted' TFP-level estimates derived from our preferred heterogeneous parameter estimators on the one hand and the standard pooled fixed effects estimator on the other. Finally, the finding of cross-country technology heterogeneity further questions the validity of standard growth and development accounting practices which impose *common* coefficients on capital stock to extract country-specific measures of TFP.

- *Secondly*, we allowed for unobserved common factors to drive output, but with differential impact across countries. These common factors are visualised by our 'common dynamic process', which follows patterns over the 1970-2002 sample period that match historical events and anecdotal evidence of the global developments in manufacturing. The interpretation of this 'common dynamic process' $\hat{\mu}_t^{\bullet}$ would be that for the manufacturing sector *similar factors drive production in all countries*, albeit to a different extent. This is equivalent to suggesting that the 'global tide' of innovation can 'lift all boats', and that technology transfer from developed to developing countries is possible but dependent on the country's production technology and absorptive capacity, among other things.

- *Thirdly*, our empirical setup allows for a type of endogeneity which is arguably very intuitive, namely that some of the unobservables driving output are also driving the evolution of inputs. This leads to an identification problem, whereby standard panel estimators

cannot identify the parameters on the observable inputs as distinct from the impact of unobservables. Our discussion has highlighted the ability of the Pesaran (2006) CCE estimators and our own AMG approach to deal with this problem successfully. Furthermore, our additional analysis in Section 2.5 confirms that the empirical results are robust to the use of a panel time-series econometric approach. The Pedroni (2000) Group-Mean FMOLS approach suggests that failure to account for unobserved common factors when analysing cross-country manufacturing production leads to the breakdown of the empirical estimates, whereas the inclusion of the 'common dynamic process' yields results very close to those from the AMG and CCEMG. From an econometric point of view, it would seem that the AMG and CCEMG estimators allow us to specify the cointegrating vector to be made up of observable inputs, output and (some of) the unobserved factors. Unobserved *stationary* processes are captured by the country intercepts and linear trends.

As was indicated throughout this chapter, we will conduct more formal testing of the AMG estimator in Chapter 4, where we compare and contrast its performance with that of the CCE estimators using Monte Carlo simulations

Appendix to Chapter 2

2.A Data construction and descriptives

Data for output, value-added, material inputs and investment in manufacturing, all in current local currency units (LCU), are taken from the UNIDO Industrial Statistics 2004 (UNIDO, 2004), where material inputs were derived as the difference between output and value-added. The labour data series is taken from the same source, which covers 1963-2002. The capital stocks are calculated from investment data in constant US$ (see below) following the 'perpetual inventory' method developed in Klenow and Rodriguez-Clare (1997a).

In order to make data in monetary values internationally comparable, it is necessary to transform all values into a common unit of analysis. We follow the strategy suggested by Martin and Mitra (2002) and derive all values in 1990 US$,[36] using current LCU and exchange rate data from UNIDO, and GDP deflators from the UN Common Statistics database (UN, 2005), for which data are available from 1970-2003.[37] Since our model is for a small open economy, we prefer using a single market exchange rates (LCU-US$ exchange rate for 1990) to purchasing-power-parity (PPP) adjusted exchange rates, since the latter are more appropriate when non-traded services need to be accounted.

The resulting panel is unbalanced and has gaps within individual country time-series. We have a total of $n = 1,194$ observations ($n = 1,162$ for the gross-output specification) from $N = 48$ countries, which have a time-series dimension between $T = 11$ and $T = 33$, with average $T = 24$. This is the sample we use in the empirical analysis throughout this chapter. Table 2.15 provides the descriptive statistics for the raw variables and variables in logs used in our regressions, further country-specific information is contained in Table 2.16.

[36] Martin and Mitra (2002) apply a single exchange-rate (that for 1990) to the whole data series, whereas for instance Larson, Butzer, Mundlak, and Crego (2000) apply the annual exchange rate. The latter approach is deemed less appropriate, since the variable series would also capture international price and exchange rate movements.

[37] Current LCU values are deflated with country-specific GDP deflators, the resulting real LCU values are then made internationally comparable by using the LCU-US$ market exchange rate for 1990.

As a robustness check we also produced a 'cleaned' dataset where we applied mechanical 'cleaning rules' in order to address the most serious issues of measurement error,[38] which created a sample of $n = 872$ observations for $N = 38$ countries. The empirical results for this sample are surprisingly close to those from the larger sample (available on request).

Table 2.15: Descriptive statistics

Variables in level terms						
Variable	**obs**	**mean**	**median**	**std. dev.**	**min.**	**max.**
levels						
output	1,162	1.40 E+11	2.41 E+10	4.06 E+11	7.21 E+07	3.15 E+12
value-added	1,194	5.47 E+10	9.04 E+09	1.78 E+11	1.76 E+07	1.50 E+12
labour	1,162	1,442,446	5.02 E+05	2,948,201	5,552	1.97 E+07
capital	1,162	1.31 E+11	2.61 E+10	3.16 E+11	5.78 E+07	2.27 E+12
materials	1,162	8.44 E+10	1.55 E+10	2.27 E+11	5.45 E+07	1.65 E+12
logs						
output	1,162	23.74	23.91	2.12	18.09	28.78
value-added	1,194	22.70	22.93	2.15	16.68	28.04
labour	1,162	12.88	13.13	1.79	8.62	16.79
capital	1,162	23.67	23.98	2.22	17.87	28.45
materials	1,162	23.31	23.46	2.11	17.81	28.13
annual growth rate						
output	1,094	4.4%	3.9%	9.8%	-43.7%	100.1%
value-added	1,128	3.9%	3.5%	12.3%	-78.3%	92.7%
labour	1,094	1.7%	0.8%	8.1%	-38.8%	78.1%
capital	1,094	4.2%	3.1%	4.4%	-2.4%	47.8%
materials	1,094	4.6%	3.7%	10.6%	-39.3%	103.1%

Variables in per worker terms						
variable	**obs**	**mean**	**median**	**std. dev.**	**min.**	**max.**
levels						
output	1,162	71,470	51,673	52,945	4,711	286,509
value-added	1,194	25,305	17,867	19,385	1,660	91,011
capital	1,162	75,522	45,865	71,512	2,007	346,064
materials	1,162	46,261	33,958	34,881	2,596	195,497
logs						
output	1,162	10.86	10.85	0.86	8.46	12.57
value-added	1,194	9.78	9.79	0.92	7.41	11.42
capital	1,162	10.79	10.73	1.00	7.60	12.75
materials	1,162	10.42	10.43	0.85	7.86	12.18
annual growth rate						
output	1,094	2.6%	2.6%	8.3%	-44.5%	85.0%
value-added	1,128	2.2%	2.5%	10.8%	-90.3%	74.4%
capital	1,094	2.5%	2.6%	7.9%	-68.0%	45.4%
materials	1,094	2.9%	2.7%	9.5%	-53.8%	87.9%

Notes: We report the descriptive statistics for output, labour, capital and material inputs for the gross-output regression sample ($n = 1,162$) and value-added for the VA regression sample ($n = 1,194$) — descriptive statistics for capital and labour in the latter specification do not differ noticeably from those reported above and we therefore do not report them separately. Monetary values are in real US$ (1990). Labour is in number of workers.

[38] We used the capital-to-materials ratio (K/M) to define a rule, bounded as $0.02 < K/M < 2$, and then dropped countries for which we had less than ten observations.

Table 2.16: Sample of countries and number of observations

Country	**Code**	**Gross-output**		**Value-added**				**I(1)***
		levels	*FD*	*levels*	*FD*	$t=1$	$t=T$	
Australia	AUS	20	17	20	17	1970	1993	✓
Austria	AUT	30	28	30	28	1970	2000	✓
Belgium	BEL	15	14	28	27	1970	1997	✓
Bangladesh	BGD	14	12	14	12	1970	1992	✓
Bolivia	BOL	11	10	11	10	1987	1997	
Barbados	BRB	26	25	26	25	1970	1995	
Canada	CAN	21	20	21	20	1970	1990	✓
Chile	CHL	25	24	25	24	1974	1998	✓
Colombia	COL	30	29	30	29	1970	1999	✓
Cyprus	CYP	33	32	33	32	1970	2002	✓
Ecuador	ECU	30	29	30	29	1970	1999	
Egypt	EGY	26	25	26	25	1970	1995	
Spain	ESP	26	25	26	25	1970	1995	
Finland	FIN	28	26	28	26	1970	2000	✓
Fiji	FJI	25	24	25	24	1970	1994	✓
France	FRA	26	25	26	25	1970	1995	
United Kingdom	GBR	23	22	23	22	1970	1992	✓
Guatemala	GTM	16	15	16	15	1973	1988	✓
Hungary	HUN	26	25	26	25	1970	1995	
Indonesia	IDN	26	25	26	25	1970	1995	✓
India	IND	32	31	32	31	1970	2001	
Ireland	IRL	22	21	22	21	1970	1991	
Iran	IRN	24	22	24	22	1970	2001	
Israel	ISR	13	12	13	12	1989	2001	
Italy	ITA	30	28	31	30	1970	2000	
Korea	KOR	32	31	32	31	1970	2001	
Sri Lanka	LKA	20	17	20	17	1970	2000	✓
Luxembourg	LUX	23	22	23	22	1970	1992	✓
Morocco	MAR	17	16	17	16	1985	2001	✓
Mexico	MEX	16	14	16	14	1984	2000	✓
Malta	MLT	32	31	32	31	1970	2001	✓
Malaysia	MYS	28	25	28	25	1970	2001	
Netherlands	NLD	24	23	24	23	1970	1993	✓
Norway	NOR	32	31	32	31	1970	2001	✓
New Zealand	NZL	21	20	21	20	1970	1990	
Panama	PAN	30	28	30	28	1970	2000	
Philippines	PHL	26	25	26	25	1970	1995	
Poland	POL	14	11	31	30	1970	2000	✓
Portugal	PRT	31	30	31	30	1970	2000	✓
Senegal	SEN	16	14	17	14	1970	1990	
Singapore	SGP	33	32	33	32	1970	2002	✓
Sweden	SWE	18	17	18	17	1970	1987	✓
Swaziland	SWZ	24	22	24	22	1970	1995	
Tunisia	TUN	21	19	21	19	1970	1997	
Turkey	TUR	27	25	27	25	1970	1997	
United States	USA	26	25	26	25	1970	1995	✓
Venezuela	VEN	26	24	26	24	1970	1998	✓
Zimbabwe	ZWE	27	26	27	26	1970	1996	
Obs		1,162	1,094	1,194	1,128		644	

Notes: * This refers to the sample used in Section 2.5 — this includes 26 countries which cannot reject nonstationarity (for ly, lk) in country ADF regressions and reject stationarity in country KPSS regressions.

2.B Time-series properties of the data

'Naïve' AR(1) regressions

As a first indicator of the time-series properties of the data we carry out AR(1) regressions

$$ln\, Z_{it} = \rho\, ln\, Z_{i,t-1} + v_{it} \qquad v_{it} = \varepsilon_{it} \left\{ + \eta_t \sum_{t=2}^{T} D_t \right\}$$

for Z_{it} as output, value-added, capital, labour and materials. This approach imposes a homogeneous autoregressive parameter ρ on all country series, which is clearly restrictive, but provides crude insights into the data properties. Year dummies are included in all four regression models (pooled OLS, fixed effects, Difference GMM and Systems GMM) to account for possible common development in the variable series (η_t, not constrained to linearity). The results for *levels* and *per capita* transformations of the variables, presented in Table 2.17, provide evidence that the data series are highly persistent and for a considerable number of countries are likely to constitute unit root processes. Note the collapse of the Difference GMM estimator in the regressions with year dummies: the poor performance of this estimator with highly persistent data is well-known and led to the development of the System GMM estimator (Blundell & Bond, 1998). Note further that GMM regressions with year dummies are subject to overfitting bias, as indicated by the Sargan test p-values close to unity (Bowsher, 2002).

Time-series stationarity tests

Since the time dimension of the panel is sizeable ($T \in [11, 33]$, average $T = 24$), we carry out Augmented Dickey and Fuller (1979) (ADF), Phillips and Perron (1988) (PP) and Kwiatkowski, Phillips, Schmidt, and Shin (1992) (KPSS) tests for the variable series within each individual country.We use this combination of tests since the ADF and PP tests have the null of nonstationary variable series, whereas the KPSS test has the null of stationary variable series. The PP test uses nonparametric methods to account for potential serial correlation in the errors and thus does not reduce the number of time-series observations like the ADF test, which uses lagged differences for the same aim. The results for variables in *levels* and in *per worker terms* are shown in Table 2.18: we report the share of countries (in %) for which the null hypothesis (stationarity or nonstationarity as indicated) is rejected.

Table 2.17: Naïve AR regressions

Naïve autoregressions for all variables
AR(1) coefficients and diagnostics ($N = 48$)

Variables in levels								
	Estimation equation without year dummies				*Estimation equation with year dummies*			
	POLS	**FE**	**DiffGMM**	**SysGMM**	**POLS**	**FE**	**DiffGMM**	**SysGMM**
log output	1.0018	0.9465	0.9296	1.0026	1.0005	0.9307	0.6884	0.9923
	[5275.89]**	[95.71]**	[35.18]**	[2296.26]**	[1728.81]**	[46.60]**	[8.49]**	[237.75]**
AR(1) p	0.61	0.37	0.00	0.00	0.87	0.32	0.00	0.00
AR(2) p	0.27	0.04	0.30	0.32	0.25	0.20	0.39	0.77
Sargan p			0.15	0.13			1.00	1.00
log VA	1.0017	0.9339	0.9061	1.0025	1.0000	0.9121	0.7054	0.9917
	[4838.26]**	[72.60]**	[25.69]**	[1673.16]**	[1453.11]**	[41.10]**	[7.91]**	[174.76]**
AR(1) p	0.57	0.07	0.00	0.00	0.32	0.04	0.00	0.00
AR(2) p	0.24	0.02	0.16	0.16	0.22	0.03	0.18	0.21
Sargan p			0.07	0.10			1.00	1.00
log labour	1.0011	0.9232	0.7506	1.0031	1.0000	0.9345	0.7887	0.9847
	[3322.24]**	[55.87]**	[8.74]**	[901.75]**	[1146.26]**	[53.90]**	[15.61]**	[234.19]**
AR(1) p	0.01	0.28	0.01	0.01	0.02	0.60	0.00	0.00
AR(2) p	0.08	0.44	0.03	0.02	0.22	0.21	0.02	0.01
Sargan p			0.24	0.09			1.00	1.00
log capital	1.0017	0.9907	0.9897	1.0028	1.0002	0.993	1.0238	0.9824
	[4890.52]**	[121.07]**	[134.56]**	[3695.91]**	[2434.64]**	[84.98]**	[23.12]**	[193.69]**
AR(1) p	0.00	0.00	0.07	0.07	0.00	0.00	0.07	0.04
AR(2) p	0.00	0.01	0.61	0.60	0.00	0.01	0.60	0.97
Sargan p			0.42	0.30			1.00	1.00
log materials	1.0019	0.9415	0.9131	1.0027	1.0006	0.9289	0.7285	0.9886
	[4944.34]**	[91.47]**	[33.65]**	[2175.35]**	[1534.16]**	[47.53]**	[13.45]**	[189.79]**
AR(1) p	0.79	0.24	0.00	0.00	0.96	0.24	0.00	0.00
AR(2) p	0.04	0.54	0.79	0.76	0.02	0.64	0.41	0.31
Sargan p			0.06	0.04			1.00	1.00

Variables in per worker terms								
	Estimation equation without year dummies				*Estimation equation with year dummies*			
	POLS	**FE**	**DiffGMM**	**SysGMM**	**POLS**	**FE**	**DiffGMM**	**SysGMM**
log output	1.0024	0.9431	0.9996	1.0038	1.0004	0.8553	0.7418	1.0033
per worker	[4761.23]**	[63.81]**	[34.38]**	[875.63]**	[1158.22]**	[29.38]**	[9.75]**	[142.10]**
AR(1) p	0.16	0.1	0.00	0.00	0.15	0.12	0.00	0.00
AR(2) p	0.15	0.01	0.64	0.64	0.33	0.20	0.97	0.99
Sargan p			0.07	0.09			1.00	1.00
log VA	1.0022	0.9127	0.9886	1.0046	0.9994	0.8232	0.6234	0.9985
per worker	[4072.13]**	[36.85]**	[31.48]**	[432.10]**	[661.64]**	[19.55]**	[9.02]**	[123.20]**
AR(1) p	0.03	0.02	0.00	0.00	0.03	0.02	0.00	0.00
AR(2) p	0.07	0.07	0.61	0.61	0.09	0.28	0.55	0.64
Sargan p			0.15	0.18			1.00	1.00
log capital	1.0023	0.9929	1.0613	1.0031	1.0001	0.9241	0.7796	1.0003
per worker	[4064.26]**	[69.76]**	[28.24]**	[875.76]**	[1230.09]**	[39.36]**	[6.60]**	[136.71]**
AR(1) p	0.19	0.42	0.00	0.00	0.31	0.29	0.00	0.00
AR(2) p	0.65	0.15	0.02	0.02	0.53	0.27	0.04	0.02
Sargan p			0.07	0.08			1.00	1.00
log materials	1.0027	0.9275	0.9836	1.005	1.0003	0.8444	0.7109	0.9976
per worker	[3865.35]**	[61.59]**	[30.75]**	[700.96]**	[1048.78]**	[32.08]**	[9.29]**	[118.99]**
AR(1) p	0.20	0.11	0.00	0.00	0.18	0.16	0	0
AR(2) p	0.19	0.03	0.53	0.52	0.41	0.37	0.96	0.96
Sargan p			0.10	0.13			1.00	1.00

Notes: All variables are in logs. POLS — pooled OLS estimator with single intercept; FE — pooled OLS estimator with N intercepts; DiffGMM, SysGMM — Arellano and Bond (1991) and Blundell and Bond (1998) GMM estimators. All estimators are implemented using `xtabond2` in `Stata` so as to obtain the residual tests. For POLS and FE the Arellano and Bond (1991) AR-tests are carried out for the *levels* residuals, whereas for DiffGMM and SysGMM they are carried out for the residuals in first differences (following (Bond, 2002)). The two GMM estimators are implemented using [t-3] and earlier lagged values as instruments, but stacking the entire instrument matrix to avoid overidentification bias (which still cannot be avoided as the Sargan p-value of unity indicates (Bowsher, 2002)). Average length of T is around 23 (min: 12, max 32).

For the majority of countries the ADF and PP tests cannot reject nonstationarity, whereas the majority of country KPSS tests reject the null of level stationarity — these results are stronger for variables in per worker terms than for those in levels. The overall pattern of test results is reversed when we run ADF and KPSS tests on variables in *first-difference*, indicating that variable series are indeed 'difference-stationary'.

Table 2.18: Time-series unit root tests — rejection frequency

Unit root tests
share of country tests which reject H_0 (stationarity or nonstationarity as indicated);
no adjustment for cross-sectional dependence

Testing for levels-stationarity								
Test	H_0	H_1	*variable*	output	value-added	labour	capital	materials
			variables in levels					
ADF without trend	nonstationary	levels-stationary		17%	21%	21%	10%	19%
PP without trend	nonstationary	levels-stationary		19%	10%	17%	17%	17%
KPSS without trend	levels-stationary	nonstationary		85%	81%	73%	98%	79%
			variables in levels (in per worker terms)					
ADF without trend	nonstationary	levels-stationary		4%	8%		6%	2%
PP without trend	nonstationary	levels-stationary		8%	17%		8%	10%
KPSS without trend	levels-stationary	nonstationary		77%	69%		71%	77%

Testing for trend-stationarity								
Test	H_0	H_1	*variable*	output	value-added	labour	capital	materials
			variables in levels					
ADF with trend	nonstationary	trend-stationary		21%	29%	19%	31%	25%
PP with trend	nonstationary	trend-stationary		15%	10%	10%	4%	10%
KPSS with trend	trend-stationary	nonstationary		31%	25%	42%	46%	31%
			variables in levels (in per worker terms)					
ADF with trend	nonstationary	trend-stationary		23%	25%		31%	25%
PP with trend	nonstationary	trend-stationary		27%	13%		6%	15%
KPSS with trend	trend-stationary	nonstationary		31%	23%		31%	27%

Testing for difference-stationarity								
Test	H_0	H_1	*variable*	output	value-added	labour	capital	materials
			variables in first differences					
ADF with drift	nonstationary	stationary		94%	85%	90%	83%	88%
PP with drift	nonstationary	stationary		90%	94%	71%	29%	90%
KPSS with drift	stationary	nonstationary		4%	0%	4%	10%	2%
			variables in first differences (in per worker terms)					
ADF with drift	nonstationary	stationary		94%	90%		83%	83%
PP with drift	nonstationary	stationary		96%	94%		75%	94%
KPSS with drift	stationary	nonstationary		2%	2%		2%	4%

Notes: All variables are in logs. We report the share of countries (out of $N = 48$) for which the respective unit root test is rejected at the 5% level of significance. All unit root tests for variables in levels contain an intercept term in the estimating equation. ADF refers to the augmented Dickey-Fuller test, which has the null of nonstationarity. PP refers to the Phillips and Perron (1988) unit root test, which has the null of nonstationarity. KPSS refers to the Kwiatkowski et al. (1992) unit root test, which has the null of (trend-)stationarity. Lag-augmentation or bandwidth selection in these tests to account for serial correlation in the variables is allowed to vary by country. For the ADF test we determined 'ideal' lag-augmentation using the Akaike Information Criterion (AIC). The PP test uses $int(4(T/100)^{2/9})$ lags throughout, for the KPSS tests an automated bandwidth selection following Hobijn et al. (1998) is used. For the latter we use the `kpss` command in `Stata` written by Kit Baum.

Our dataset is an unbalanced panel with missing observations — properties that may affect the unit root tests. A simulation exercise by Ryan and Giles (1998) suggests that (with respect to the ADF tests) filling the gaps with the last known observation produces more powerful unit root tests in comparison with tests where gaps were ignored. They also point out that regular Dickey-Fuller critical values remain valid for either approach. We carried out ADF tests for the altered dataset (levels, first differences) and obtained very similar patterns of rejection as when testing original data with gaps (not reported). Thus the results from these country unit-root tests are a further indication of the potential for integrated processes in our data.

Table 2.19: First generation panel unit root tests

Im, Pesaran & Shin (1997) panel unit root tests — IPS(i)♯

H_0: unit root process; augmentation with country-specific lag length (average reported);
no adjustment for cross-sectional dependence, tests for which H_0 is rejected are indicated

output		**value-added**		**labour**			**capital**			**materials**	
lags	[t-bar]	lags	[t-bar]	lags	[t-bar]		lags	[t-bar]		lags	[t-bar]
1.42	-1.57	1.96	-1.54	1.48	-1.78	*reject*	1.50	-1.92	*reject*	1.65	-1.67
output/worker		**VA/worker**					**capital/worker**			**materials/worker**	
lags	[t-bar]	lags	[t-bar]				lags	[t-bar]		lags	[t-bar]
1.44	-0.92	1.65	-1.03				1.71	-0.97		1.83	-1.05

Maddala and Wu (1999) panel unit root tests — MW♯

H_0: unit root process; augmentation with lags as indicated;
no adjustment for cross-sectional dependence

output			**value-added**			**labour**			**capital**			**materials**		
lags	p_λ	(p)	lags	p_λ	(p)	lags	p_λ	(p)	lags	p_λ	(p)	lags	p_λ	(p)
0	129.37	(.01)	0	125.69	(.02)	0	142.42	(.00)	0	274.01	(.00)	0	126.47	(.02)
1	126.57	(.02)	1	109.99	(.16)	1	141.35	(.00)	1	67.05	(.99)	1	133.62	(.01)
1.42	69.12	(.98)	1.96	85.44	(.77)	1.48	114.54	(.10)	1.50	55.16	(1.00)	1.65	66.65	(.99)
2	114.75	(.09)	2	124.13	(.03)	2	105.34	(.24)	2	80.86	(.87)	2	134.85	(.01)
3	74.36	(.95)	3	56.97	(1.00)	3	88.76	(.69)	3	87.84	(.71)	3	108.31	(.18)
output/worker			**VA/worker**						**capital/worker**			**materials/worker**		
lags	p_λ	(p)	lags	p_λ	(p)				lags	p_λ	(p)	lags	p_λ	(p)
0	107.76	(.19)	0	102.23	(.31)				0	54.16	(1.00)	0	102.35	(.31)
1	70.70	(.98)	1	84.92	(.78)				1	60.09	(1.00)	1	77.74	(.91)
1.44	30.08	(1.00)	1.65	63.13	(1.00)				1.71	32.32	(1.00)	1.83	32.45	(1.00)
2	75.85	(.94)	2	65.26	(.99)				2	34.04	(1.00)	2	77.85	(.91)
3	61.41	(1.00)	3	44.65	(1.00)				3	66.85	(.99)	3	87.42	(.72)

Notes: ♯ All variables are in logs. The IPS(i) and MW statistics are constructed as t-bar $= N^{-1}\sum_i t_i$ for the former and $p_\lambda = -2\sum_i log(p_i)$ for the latter, where t_i are the country ADF statistics (t-values) and p_i the corresponding p-values. For the Im et al. (1997) test the critical values (-1.73 for 5%, -1.69 for 10% significance level — distribution is approximately t) are reported in Table 2, Panel A of their paper (we used $T = 25$, $N = 50$, the model without trend). For the Maddala and Wu (1999) test the critical values are distributed $\chi^2(2N)$. The IPS(i) test uses the 'ideal' lag-length as determined via the AIC (see notes to Table 2.18.

Panel stationarity tests

Next we applied panel unit root tests to the data. It is important to stress that rejection of the unit root null hypothesis does not imply that the panel is stationary, but rather that the variable series does not follow a unit root process *in all countries*. We first present the results for a test by Im et al. (1997) and the Maddala and Wu (1999) panel unit root test, both of which do not account for cross-sectional dependence in the variables. Results in Table 2.19 show that for the variable series in 'levels' these tests cannot agree on the level of integration prevalent in the data. For the per worker variable series, however, neither test can reject the null hypothesis that *all* countries have I(1) series.

Over the past decade panel unit root tests which explicitly allow for cross-sectional dependence in the variable series have been developed. These include a simple augmentation to the Im et al. (1997) panel unit root test (Im et al., 2003) (IPS(ii)), and the Pesaran (2007) test based on cross-section augmented ADF regressions (CIPS). For the former we are required to

Table 2.20: Second generation panel unit root tests

Im, Pesaran & Shin (2003) panel unit root tests — IPS(ii)♯

H_0: unit root process; augmentation with lags as indicated

output			value-added			labour			capital			materials		
lags	W[t-bar]	(p)	lags	W[t-bar]	(p)	lags	W[t-bar]	(p)	lags	W[t-bar]	(p)	lags	W[t-bar]	(p)
0	3.88	(1.00)	0	1.73	(.96)	0	1.00	(.84)	0	4.07	(1.00)	0	4.11	(1.00)
1	1.82	(.97)	1	1.51	(.94)	1	-0.68	(.25)	1	2.84	(1.00)	1	2.86	(1.00)
2	2.11	(.98)	2	2.50	(.99)	2	0.43	(.66)	2	2.28	(.99)	2	2.26	(.99)
3	3.13	(1.00)	3	3.23	(1.00)	3	0.96	(.83)	3	1.67	(.95)	3	2.74	(1.00)
output/worker			**VA/worker**						**capital/worker**			**materials/worker**		
lags	W[t-bar]	(p)	lags	W[t-bar]	(p)				lags	W[t-bar]	(p)	lags	W[t-bar]	(p)
0	0.30	(.62)	0	-0.68	(.25)				0	1.23	(.89)	0	0.16	(.56)
1	-0.10	(.46)	1	0.31	(.62)				1	-0.39	(.35)	1	-0.14	(.44)
2	-0.45	(.33)	2	1.30	(.90)				2	1.21	(.89)	2	-0.82	(.21)
3	-0.25	(.40)	3	1.74	(.96)				3	0.65	(.74)	3	-0.93	(.18)

Pesaran (2007) panel unit root tests — CIPS♯

H_0: unit root process; augmentation with lags as indicated

output			value-added			labour			capital			materials		
lags	Z[t-bar]	(p)	lags	Z[t-bar]	(p)	lags	Z[t-bar]	(p)	lags	Z[t-bar]	(p)	lags	Z[t-bar]	(p)
0	-1.22	(.11)	0	-1.85	(.03)	0	2.39	(.99)	0	5.11	(1.00)	0	0.29	(.62)
1	0.01	(.51)	1	0.06	(.52)	1	1.26	(.90)	1	3.79	(1.00)	1	0.89	(.81)
1.42	1.13	(.87)	1.96	3.54	(1.00)	1.48	3.74	(1.00)	1.50	4.55	(1.00)	1.65	3.68	(1.00)
2	2.65	(1.00)	2	2.30	(.99)	2	4.21	(1.00)	2	3.96	(1.00)	2	1.05	(.85)
3	7.04	(1.00)	3	3.59	(1.00)	3	4.76	(1.00)	3	7.64	(1.00)	3	4.21	(1.00)
output/worker			**VA/worker**						**capital/worker**			**materials/worker**		
lags	Z[t-bar]	(p)	lags	Z[t-bar]	(p)				lags	Z[t-bar]	(p)	lags	Z[t-bar]	(p)
0	-1.08	(.14)	0	-2.55	(.01)				0	1.92	(.97)	0	0.57	(.72)
1	2.91	(1.00)	1	-0.73	(.23)				1	1.33	(.91)	1	3.74	(1.00)
1.44	5.98	(1.00)	1.65	3.77	(1.00)				1.71	5.92	(1.00)	1.83	9.62	(1.00)
2	5.02	(1.00)	2	2.37	(.99)				2	4.60	(1.00)	2	5.96	(1.00)
3	8.73	(1.00)	3	5.48	(1.00)				3	7.34	(1.00)	3	8.08	(1.00)

Notes: ♯ All variables are in logs. For the IPS(ii) tests we were forced to reduce the sample to $n = 754$ observations ($N = 29, T = 26$) due to the requirement of a balanced panel for the `Stata` implementation of this test. The data is also interpolated in this case. For the Pesaran (2006) CIPS tests we can use the original unbalanced panel dataset with missing observations. In the third row for each variable in the lower panel we present the CIPS test statistic for 'ideal' lag augmentation of the underlying ADF regression (based on information criteria); the value for lags reported in this case is the average lag length across countries.

use a *balanced* panel; we therefore use a balanced subset of the sample where missing values have been interpolated, which considerably reduces our sample size ($T, N = 26, 29$; $n = 754$) compared to the *unbalanced* panel ($T_{max} = 33$, $N = 48$; $n = 1,162$).

We reiterate the difficulties relating to panel unit root tests (Smith & Fuertes, 2004, 2007), in particular in a (relatively) short, unbalanced panel with gaps like in our own case. In most cases these testing procedures cannot reject the null of nonstationarity. If we apply the CIPS test to data in first differences, we reject nonstationarity throughout if we augment with up to two lags (or none); for more lags the p-value for most variable tests jumps to unity (not reported).

2.C The Phillips & Moon (1999) result

Given the breakdown of cointegration in the presence of nonstationary error terms — a standard result in time-series econometrics (Engle & Granger, 1987) — we would expect *any* pooled panel estimation to yield spurious results if variables were nonstationary and parameters heterogeneous across countries. A recent development in econometric theory however implies that this conclusion needs 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 input variables and output *provided N and T are large enough* (see also Kao, 1999; Phillips & Moon, 2000; Smith & Fuertes, 2007).[39]

This implies that regardless of whether we have homogeneous cointegration, heterogeneous cointegration or noncointegration, the pooled fixed effects estimator in levels can provide a consistent point estimate of some average long-run correlation between inputs and output — provided both the time-series and cross-sectional dimension of the panel are large enough. In case of heterogeneous cointegration in all countries, the pooled FE estimator however does not yield the sample mean of the cointegrating vectors across countries, but a *different* cross-country average. Similarly if country equation errors are nonstationary for other reasons, e.g. due to a misspecified common factor process (TFP). The estimates derived using the pooled FE estimator represent weighted averages of the country-specific cointegrating vectors. The weights are determined by the share of country i's variation in the regressors in the total variation across all countries. The joint limit theorem of the paper further requires $N/T \to 0$. When the latter condition is violated, the spurious regression bias can dominate and the pooled regression results remain distorted.

Provided $N/T \to 0$, it is difficult to contrast the economic interpretation of the Phillips and Moon (1999) estimate for the 'long-run average' with that of an unweighted mean of the cointegrating vector parameters across countries (e.g. the MG estimator which yields the 'average long-run'). In the case where *all* countries have nonstationary variable series and cointegrate

[39] The intuition for this result is as follows: one can think of a single time-series regression of two nonstationary variables which do not cointegrate as taking a single draw from the distribution of the 'true' parameter (be it zero or non-zero). We only have a single shot at it, so that we would not expect to pick exactly the mean of the distribution. Increasing the time-series dimension T will not help in this case, as we still just take a single draw from the 'true' parameter distribution. Pooling N time-series from different countries then represents taking not one but N draws from this parameter distribution. The larger N, the more precise our estimate becomes of the mean of the 'true' parameter distribution. [This intuition was suggested by Francis Teal.]

both estimators yield consistent estimates of *some* average long-run relationships between the variables, but differ in the way they arrive at these averages, due to the options opened by two-way asymptotics. How would the Phillips and Moon (1999) effect hold up in a panel of countries with a *mixture* of stationary and nonstationary variable series? Further, in this case the estimates would represent an amalgamate of unweighted means of country parameters for the 'stationary countries', and weighted means of country parameters for the nonstationary ones — thus making economic interpretation of the result even more difficult.

In our own empirics, given $N/T \nrightarrow 0$ and suggested cross-section dependence, we argue that the Phillips and Moon (1999) result does not affect the bias in the pooled regression. Our Monte Carlo simulations in Chapter 4 will provide some additional insights into the validity of the Phillips and Moon (1999) result under different empirical setups.

2.D Robustness check: dynamic specification

The serial correlation tests for the models presented above indicate some form of residual serial correlation (to a greater or lesser extent) in virtually every pooled model and this may also be the case for our country regressions. We note that the Arellano and Bond (1991) AR tests cannot distinguish between stationary, but serially correlated, and nonstationary errors. In order to investigate whether serial correlation may seriously impact the regression results obtained, we re-specify the general model in a dynamic form so as to allow for serial correlation explicitly. The underlying assumption is that in equation (2.1) we have $\varepsilon_{it} = \rho_i \varepsilon_{i,t-1} + v_{it}$. Following Bond (2002) we can rewrite the general model as

$$y_{it} = \rho_i y_{i,t-1} + \boldsymbol{\beta}_i' \boldsymbol{x}_{it} - \rho_i \boldsymbol{\beta}_i' \boldsymbol{x}_{i,t-1} + (1-\rho_i)\alpha_i + \boldsymbol{\lambda}_i' \boldsymbol{f}_t - \rho_i \boldsymbol{\lambda}_i' \boldsymbol{f}_{t-1} + v_{it} \tag{2.29}$$

This type of autoregressive distributed lag (ARDL) model is often expressed in form of an (equivalent) Error Correction Model (ECM)

$$\begin{aligned} \Delta y_{it} &= -(1-\rho_i) y_{i,t-1} + (1-\rho_i)\boldsymbol{\beta}_i' \boldsymbol{x}_{i,t-1} + \boldsymbol{\beta}_i' \Delta \boldsymbol{x}_{it} \\ &\quad + (1-\rho_i)\alpha_i + (1-\rho_i)\boldsymbol{\lambda}_i' \boldsymbol{f}_{t-1} + \boldsymbol{\lambda}_i' \Delta \boldsymbol{f}_t + v_{it} \end{aligned} \tag{2.30}$$

The coefficients on the terms in first differences in (2.30) are interpreted as the 'short-run' dynamics, whereas from the coefficients on the lagged levels we can derive the 'long-run' or 'equilibrium' relationship in the data. Note that this is *not* a growth regression — equation (2.30) is equivalent to equation (2.29). We prefer the ECM to the ARDL version as computation of long-run coefficients only involves two parameter estimates, rather than three in the ARDL model. Our representation in (2.30) indicates 'common factor restrictions' whereas in the below regressions we allow for more dynamic flexibility. Heterogeneity of the parameter coefficients in the regressions is naturally only possible in the country-regressions.

Pooled regressions

We begin with the results for the pooled regression model restricted to CRS in Table 2.21. We do not estimate the dynamic version of our First Difference OLS regression as this would entail instrumentation for the lagged dependent variable, which was found to be cumbersome and subject to a weak instrument problem.

The dynamic versions of pooled OLS (POLS), fixed effects (FE) and Pesaran (2006) CCEP estimators in essence obtain very similar long-run factor parameter estimates as those implied by their static equivalents in Table 2.2, with capital coefficients ranging between .6 and .8. The estimates for POLS yield insignificant capital coefficients in both gross-output and value-added specifications, and the the severe bias evident in the former manifests itself in a long-run material coefficient exceeding unity. In contrast, FE and CCEP yield statistically significant parameters in the ECM equation.

We also include the results for the Arellano and Bond (1991) test for serial correlation in the error terms. It is important to emphasize again that the ECM is not a growth regression, such that the error terms are not in first differences, but as equation (2.30) shows in levels, such that the AR(1) test is relevant and we would not expect any evidence for this type of serial correlation.[40] As the p-values indicate (H_0: no serial correlation) some of the models reported show evidence of AR(1) and AR(2) errors — a characteristic which *should* have been addressed by the dynamic specification.[41] This specific serial correlation test was developed for

[40] Tests are implemented in `Stata` using the `abar` command, which applies the test to the untransformed (levels) residuals of the regression.

[41] AR(3) on the other hand can only be ruled out if we introduced additional *lagged* differences for capital (and in the gross-output versions for materials).

Table 2.21: Pooled regressions (Error Correction Model)

Pooled dynamic regressions (CRS) dynamic specification (Error Correction Model) with CRS imposed; regression equations augmented with year dummies or period-means (Pesaran, 2006)

	Gross output models			**Value-added models**		
	[1]	[2]	[3]	[4]	[5]	[6]
estimator	**POLS**	**FE**	**CCEP**	**POLS**	**FE**	**CCEP**
dependent variable‡	Δlo	Δlo	Δlo	Δly	Δly	Δly
regressors						
lagged log output pw	-0.0522 [4.89]**	-0.3003 [13.67]**	-0.4326 [14.75]**			
lagged log VA pw				-0.0143 [2.03]*	-0.2780 [12.38]**	-0.3747 [14.13]**
lagged log capital pw	-0.0022 [0.84]	0.0406 [5.09]**	0.0585 [5.00]**	0.0109 [1.81]	0.1661 [7.11]**	0.22767 [8.24]**
lagged log materials pw	0.0557 [4.80]**	0.2284 [12.95]**	0.3498 [14.19]**			
capital growth rate	0.0456 [1.65]	0.0487 [3.49]**	0.0485 [2.93]**	0.3299 [3.70]**	0.3225 [7.93]**	0.3416 [7.58]**
materials growth rate	0.7981 [32.63]**	0.7945 [66.02]**	0.8095 [58.85]**			
intercept	0.0101 [0.77]	0.4309 [6.13]**	-0.0997 [0.50]	0.0178 [0.49]	0.8848 [4.68]**	0.06028 [0.14]
long-run coefficients†						
capital	-0.043 [0.84]	0.135 [5.84]**	0.135 [5.50]**	0.766 [4.14]**	0.597 [9.49]**	0.608 [9.91]**
materials	1.067 [17.35]**	0.761 [32.18]**	0.809 [29.79]**			
long-run VA equivalent†						
capital (VA)	0.635 [2.26]*	0.565 [7.21]**	0.707 [8.35]**			
AR tests‡						
AR(1) (p)	-2.6 (.01)	-1.1 (.29)	-2.4 (.02)	-3.3 (.00)	-1.9 (.06)	-1.2 (.25)
AR(2) (p)	-1.6 (.11)	0.0 (.98)	-2.35 (.02)	-2.5 (.01)	-0.5 (.60)	-2.4 (.02)
AR(3) (p)	1.5 (.13)	2.2 (.03)	0.1 (.93)	1.4 (.16)	2.6 (.01)	0.4 (.69)
RMSE	.033	.030	.023	.102	.095	.082
Observations	1,094 (48)	1,094 (48)	1,094 (48)	1,128 (48)	1,128 (48)	1,128 (48)

Notes: Values in parentheses are absolute t-statistics. * and ** indicate statistical significance at the 5% and 1% level respectively. † Long-run coefficients are computed from the pooled estimates using the Delta method (available via the `nlcom` command in `Stata`).
‡ Arellano and Bond (1991) serial correlation test.

microeconomic datasets with large N and short T, and we take these results as an indication that these tests may not be performing as expected in a large T context. This aside, the two FE models indicate AR(3). Residual diagnostics (not reported) suggest that all three pooled models 'fail' the D'Agostino et al. (1990) residual normality tests in the gross-output and value-added data, POLS furthermore fails the Cameron and Trivedi (1990) normality and homoskedasticity test (not applicable for CCEP, FE turns out to have homoskedastic errors). In conclusion the pooled estimates and residual diagnostics do not change in any considerable way from their static model versions and given the greater complexity of the former we prefer our static results.

Country regressions

Table 2.22 presents averaged coefficients from the country regressions in dynamic specification — we use the same set of estimators as in the static case in Table 2.3.[42] All models are estimated in the Error Correction Model form, and for the AMG estimators we adopt the 'common dynamic process' derived from year dummy coefficients in the respective pooled static equations in first differences. CRS is imposed in all regression models. We do not estimate the country regressions in first differences for the same reason as in the pooled dynamic models. This aside we also do not estimate the gross-output based version, since computation of comparable long-run coefficients for VA-equivalent capital would involve too many nonlinear combinations of estimated parameters.

The averaged coefficients on lagged value-added in levels is statistically significant and has the correct sign in all models. In contrast the averaged coefficients on lagged capital stock in levels are insignificant in the CCEMG models and the Swamy RCM, with averaged short-run coefficients on capital stock only significant for one of the AMG and the CCEMG models. All averaged model parameters however possess the correct signs. Residual diagnostics are next to identical to those from the static models.

There are two ways to arrive at the average long-run coefficients for capital: we can compute them based on the averaged ECM estimates (average first, then compute long-run) or simple average country-specific long-run coefficient estimates (compute long-run for each country i and then average these across countries). Results for the former are presented in the first row for 'long-run coefficients', while the mean for the latter approach — according to Pesaran and Smith (1995, p.82) this represents "the appropriate measure for the long-run effect" of inputs on output — is given in the subsequent row. The mean long-run capital coefficients for the standard CCEMG in [5], and the AMG and ARCM in [4] and [8] respectively are now all slightly higher at around .4, thus still broadly in line with their static equivalents. Note that the Pedroni (2000) panel-t statistics are highest for long-run capital in these three models. On the whole, we see these dynamic country regressions as providing additional support for our findings in the static specifications.

[42] Note that for the AMG where $\hat{\mu}_t^{\bullet}$ is included as additional regressor we treated this as an additional covariate, rather than a deterministic term. Dynamic CCEMG includes cross-section averages of the dependent variable, as well as of the lagged levels and first differenced capital stock per worker.

Table 2.22: Country regression averages (dynamic model, CRS imposed)

Country regressions (ECM specification) dynamic specification with CRS imposed; estimates presented are unweighted means of the country coefficients								
Value-added regressions								
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
estimator	**MG**	**RCM**	**AMG**	**AMG**	**CCEMG**	**CCEMG**	**ARCM**	**ARCM**
specification	(a)	(a)	(b)	(c)	(d)	(d)	(b)	(c)
dependent variable♯	Δly	Δly	Δly-$\Delta\hat{\mu}_t^{\bullet}$	Δly	Δly	Δly	Δly-$\Delta\hat{\mu}_t^{\bullet}$	Δly
regressors								
value-added pw $(t-1)$	-0.6020 [11.29]**	-0.4985 [8.34]**	-0.6396 [17.28]**	-0.6825 [12.43]**	-0.5511 [9.68]**	-0.7276 [8.86]**	-0.5330 [11.58]**	-0.5766 [9.34]**
capital pw $(t-1)$	0.2233 [2.27]*	0.2018 [1.94]	0.1511 [2.37]*	0.3559 [3.35]**	0.2005 [1.86]	0.0589 [0.20]	0.1720 [2.41]*	0.2679 [2.39]*
common process $(t-1)$				0.6153 [2.29]*				0.7486 [2.58]**
Δcapital pw (t)	0.1173 [1.05]	0.1522 [1.25]	0.1546 [1.44]	0.2623 [2.24]*	0.3475 [3.19]**	0.3579 [2.69]**	0.2070 [1.74]	0.2446 [1.92]
Δcommon process (t)				-0.0978 [0.45]				0.1229 [0.52]
country trend	0.0050 [2.18]*	0.0041 [1.62]	-0.0004 [0.16]	-0.0064 [1.96]		0.0052 [1.39]	-0.0008 [0.33]	-0.0068 [1.82]
intercept	3.2907 [3.78]**	2.6032 [2.76]**	4.4500 [6.34]**	2.6316 [2.63]*	5.5285 [1.30]	11.3752 [1.19]	3.2598 [4.12]**	2.6042 [2.43]*
Panel-t statistic, trends†								
country trend	11.26	11.76	11.85	10.17		11.64	11.8	9.09
# of sign. trends (at 10%)	22	21	24	14		18	25	15
long-run coefficients‡								
capital pw (from above)	0.3709 [9.45]**	0.4048 [2.13]*	0.2363 [7.09]**	0.5214 [15.32]**	0.3637 [1.16]	0.0810 [1.66]	0.3227 [2.45]*	0.4646 [2.58]**
capital pw (mean)	0.3168 [2.89]**	0.3403 [5.24]**	0.1933 [1.90]	0.4052 [3.10]**	0.3825 [2.50]*	0.5800 [1.68]	0.3153 [4.50]**	0.4007 [5.41]**
Panel t-statistic	9.48	8.48	4.76	11.23	16.16	3.98	8.72	11.55
RMSE	.076	.078	.077	.066	.067	.062	.164	.071
obs (countries)	1,128 (48)	1,128 (48)	1,128 (48)	1,128 (48)	1,128 (48)	1,128 (48)	1,128 (48)	1,128 (48)

Notes: All variables are in logs. ly — log value-added per worker. $\hat{\mu}_t^{\bullet}$ and $\hat{\mu}_t^{va\,\bullet}$ are derived from the year dummy coefficients of a static pooled regression (CRS imposed) in first differences (FD-OLS) as described in the text.
We omit reporting the parameters on the cross-section averages for the CCEMG estimators (columns [5] and [6]) to save space.
Values in parentheses are absolute t-statistics. These were obtained by regressing the N country estimates on an intercept term, except for the Swamy t-stats, which are provided by `xtrc` in `Stata` and represent $\sum_i(\hat{\Sigma}+\hat{\mathbf{V}}_i)$ where $\hat{\Sigma}$ is a measure of dispersion of the country OLS estimates and $\hat{V}$ is the variance of the N OLS estimates scaled by $\sum x_i^2$.
'Specification' refers to equations (2.19) to (2.22). We indicate statistical significance at the 5% and 1% level by * and ** respectively.
♯ We subtract the common dynamic process $\hat{\mu}_t^{\bullet}$ from lagged value-added per worker (in logs) for country i, and common dynamic process $\hat{\mu}_t^{\bullet}$ from in models [3] and [7].
† We report averaged t-statistics for country-specific t-statistics t_i of the factor estimates.
‡ This is obtained using a non-linear combination of the capital and materials coefficients accounting for the precision of these estimates. The first estimate ('from above') simply calculates the long-run capital coefficient based on the sample averages on capital pw $(t-1)$ and value-added pw $(t-1)$ reported in the upper section of the table. The second estimates ('mean') use the long-run coefficients computed for each country i and averages them.

2.E Additional tables, graphs and figures

Table 2.23: Country regressions — data in deviation from x-section mean

Country regressions (demeaned data) static specification with CRS imposed; estimates presented are unweighted means of the country coefficients					
Gross output regressions					
	[1]	[2]		[3]	[4]
estimator	**MG**	**RCM**	*estimator*	**MG**	**RCM**
dependent variable	$l\check{o}$	$l\check{o}$	*dependent variable*	$\Delta l\check{o}$	$\Delta l\check{o}$
regressors			*regressors*		
capital pw	0.1125 [5.62]**	0.1128 [5.09]**	**Δcapital pw**	0.0989 [5.51]**	0.0997 [4.67]**
materials pw	0.7429 [26.79]**	0.7616 [25.83]**	**Δmaterials pw**	0.7845 [38.97]**	0.7902 [34.35]**
trend term	0.0000 [0.04]	0.0000 [0.00]	**drift term**	-0.0008 [1.08]	0.0000 [0.03]
intercept	-0.0101 [0.36]	-0.0011 [0.04]			
Panel-t-statistics, trends†			*Panel-t-statistics, trends*†		
capital pw	14.0	17.2	Δcapital pw	10.5	13.4
materials pw	98.8	111.0	Δmaterials pw	96.4	108.5
country trends	15.6	14.5	drift terms	4.1	4.0
# of sign. trends (at 10%)	25	26	# of sign. trends (at 10%)	1	2
VA-equivalent‡			*VA-equivalent*‡		
capital-VA	0.437 [6.62]**	0.473 [5.90]**	capital-VA	0.459 [5.74]**	0.475 [4.89]**
RMSE	.030	.031		.029	.030
obs (countries)	1,162 (48)	1,162 (48)		1,094 (48)	1,094 (48)

Value-added regressions					
	[1]	[2]		[3]	[4]
estimator	**MG**	**RCM**	*estimator*	**MG**	**RCM**
dependent variable	$l\check{y}$	$l\check{y}$	*dependent variable*	$\Delta l\check{y}$	$\Delta l\check{y}$
regressors			*regressors*		
capital pw	0.1325 [2.30]*	0.2005 [3.21]**	**Δcapital pw**	0.1401 [2.33]*	0.1483 [2.17]*
trend term	0.0204 [9.29]**	0.0192 [8.33]**	**drift term**	0.0193 [8.47]**	0.0214 [6.57]**
intercept	9.5765 [79.04]**	9.6035 [77.55]**			
Panel-t-statistics, trends†			*Panel-t-statistics, trends*†		
capital pw	10.7	13.17	Δcapital pw	5.96	6.75
country trends	50.7	48.55	drift terms	11.45	20.65
# of sign. trends (at 10%)	43	43	# of sign. trends (at 10%)	18	44
RMSE	.105	.106		.097	.099
obs (countries)	1,194 (48)	1,194 (48)	obs (countries)	1,128 (48)	1,128 (48)

Notes: All variables are in logs and deviation from the cross-section mean. $l\check{o}$ — log output per worker, $l\check{y}$ — log value-added per worker (in deviation from cross-section mean).
Values in parentheses are absolute t-statistics. These were obtained by regressing the N country estimates on an intercept term, except for the Swamy t-stats, which are provided by `xtrc` in `Stata` and represent $\sum_i(\hat{\Sigma}+\hat{\mathbf{V}}_i)$ where $\hat{\Sigma}$ is a measure of dispersion of the country OLS estimates and $\hat{V}$ is the variance of the N OLS estimates scaled by $\sum x_i^2$.
We indicate statistical significance at the 5% and 1% level by * and ** respectively.
† We report averaged t-statistics for country-specific t-statistics t_i of the factor estimates.
‡ This is obtained using a non-linear combination of the capital and materials coefficients accounting for the precision of these estimates.

Table 2.24: Country regressions (FD, unrestricted returns to scale)

Country regressions in first differences
static specification, no restriction on returns to scale; estimates presented are unweighted means of the country coefficients

(i) Gross output regressions

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
estimator	**MG**	**RCM**	**AMG**	**AMG**	**CCEMG**	**CCEMG**	**ARCM**	**ARCM**
specification	(a)	(a)	(b)	(c)	(d)	(d)	(b)	(c)
dependent variable♯	ΔlO	ΔlO	ΔlO-$\Delta\hat{\mu}_t^{\bullet}$	ΔlO	ΔlO	ΔlO	ΔlO-$\Delta\hat{\mu}_t^{\bullet}$	ΔlO
regressors								
Δlabour	0.1931 [6.35]**	0.1766 [5.00]**	0.1623 [5.73]**	0.1675 [5.42]**	0.1659 [5.79]**	0.1919 [5.47]**	0.1543 [4.65]**	0.1624 [4.60]**
Δcapital	0.0543 [0.40]	0.0075 [0.05]	0.1230 [0.98]	0.1293 [0.86]	0.0761 [1.06]	-0.1267 [0.82]	0.0617 [0.46]	0.0368 [0.23]
Δmaterials	0.7617 [39.79]**	0.7715 [34.92]**	0.7741 [39.28]**	0.7643 [38.40]**	0.7570 [39.52]**	0.7449 [34.35]**	0.7830 [34.72]**	0.7763 [34.14]**
Δcommon process				0.9982 [4.41]**				0.8614 [3.38]**
country drift	0.0042 [1.72]	0.0032 [1.08]	0.0001 [0.05]	-0.0003 [0.10]		0.0007 [0.23]	-0.0004 [0.15]	-0.0006 [0.20]
Returns to scale								
CRS test (p)	0.01 (.94)	0.11 (.74)	0.26 (.61)	0.19 (.67)	0.01 (.99)	1.71 (.19)	0.01 (.99)	0.03 (.87)
Panel-t-statistics, trends†								
labour	12.7	16.4	10.9	12.2	10.9	11.5	15.0	15.7
capital	0.0	0.8	1.9	1.5	5.3	1.1	2.8	2.2
materials	85.2	112.6	96.2	92.0	82.7	74.9	115.2	116.4
country trends	8.2	8.4	6.8	7.6		7.7	6.9	7.0
# of sign. trends (at 10%)	10	11	8	10		11	11	7
VA-equivalent coefficients‡								
labour (VA) mean	0.810 [8.13]**	0.773 [6.13]**	0.719 [7.33]**	0.711 [6.97]**	0.683 [7.07]**	0.753 [7.50]**	0.711 [5.69]**	0.726 [5.69]**
capital (VA) mean	0.228 [0.41]	0.033 [0.05]	0.544 [0.99]	0.549 [0.88]	0.313 [1.10]	-0.497 [0.85]	0.285 [0.46]	0.165 [0.23]
RMSE	.029	.030	.028	.026	.024	.024	.029	.027
obs (countries)	1,094 (48)	1,094 (48)	1,094 (48)	1,094 (48)	1,094 (48)	1,094 (48)	1,094 (48)	1,094 (48)

(ii) Value-added regressions

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
estimator	**MG**	**RCM**	**AMG**	**AMG**	**CCEMG**	**CCEMG**	**ARCM**	**ARCM**
specification	(a)	(a)	(b)	(c)	(d)	(d)	(b)	(c)
dependent variable♯	ΔlY	ΔlY	ΔlY-$\Delta\hat{\mu}_t^{va\bullet}$	ΔlY	ΔlY	ΔlY	ΔlY-$\Delta\hat{\mu}_t^{va\bullet}$	ΔlY
regressors								
Δlabour	0.8559 [9.62]**	0.8115 [8.02]**	0.7444 [9.59]**	0.7453 [9.95]**	0.7038 [9.91]**	0.7720 [10.37]**	0.7142 [7.91]**	0.7304 [8.30]**
Δcapital	-0.1816 [0.62]	-0.0025 [0.01]	0.0044 [0.02]	-0.0070 [0.03]	0.2231 [0.92]	-0.1634 [0.50]	0.1478 [0.51]	0.1137 [0.37]
Δcommon process				1.0460 [5.60]**				0.8763 [4.06]**
country drift	0.0187 [2.58]*	0.0195 [2.24]**	0.0016 [0.22]	-0.0019 [0.25]		0.0173 [1.93]	0.0018 [0.21]	0.0040 [0.44]
Returns to scale								
CRS test (p)	2.03 (.15)	0.53 (.47)	1.49 (.22)	1.30 (.25)	0.11 (.74)	1.88 (.17)	0.33 (.57)	0.36 (.55)
Panel-t-statistics, trends†								
labour	24.41	29.73	21.6	22.08	21.3	22.35	27.07	28.57
capital	-1.03	0.50	0.81	0.43	8.94	-0.19	2.48	2.15
country trends	9.17	10.23	6.61	7.47		8.18	7.03	7.84
# of sign. trends (at 10%)	12	15	8	10		12	10	12
RMSE	.095	.098	.092	.088	.085	.083	.095	.092
obs (countries)	1,128 (48)	1,128 (48)	1,128 (48)	1,128 (48)	1,128 (48)	1,128 (48)	1,128 (48)	1,128 (48)

Notes: All variables are in logs. ΔlO — output growth rate, ΔlY — value-added growth rate. $\hat{\mu}_t^{\bullet}$ and $\hat{\mu}_t^{\bullet,va}$ are derived from the year dummy coefficients of a pooled regression (no restriction on returns to scale) in first differences (FD-OLS) as described in the main text. We omit reporting the parameters on the cross-section averages for the CCEMG estimators (columns [5] and [6]) to save space.
Values in parentheses are absolute t-statistics. These were obtained by regressing the N country estimates on an intercept term, except for the Swamy t-stats, which are provided by `xtrc` in `Stata` and represent $\sum_i(\hat{\Sigma}+\hat{\mathbf{V}}_i)$ where $\hat{\Sigma}$ is a measure of dispersion of the country OLS estimates and $\hat{V}$ is the variance of the N OLS estimates scaled by $\sum x_i^2$.
'Specification' refers to equations (2.19) to (2.22). We indicate statistical significance at the 5% and 1% level by * and ** respectively. RMSE reports the root mean squared errors.
♯ We subtract the common dynamic process $\hat{\mu}_t^{\bullet}$ from log output (log value-added) for country i in models [3] and [7].
† We report averaged t-statistics for country-specific t-statistics t_i of the factor estimates.
‡ This is obtained using a non-linear combination of the capital and materials coefficients accounting for the precision of these estimates.

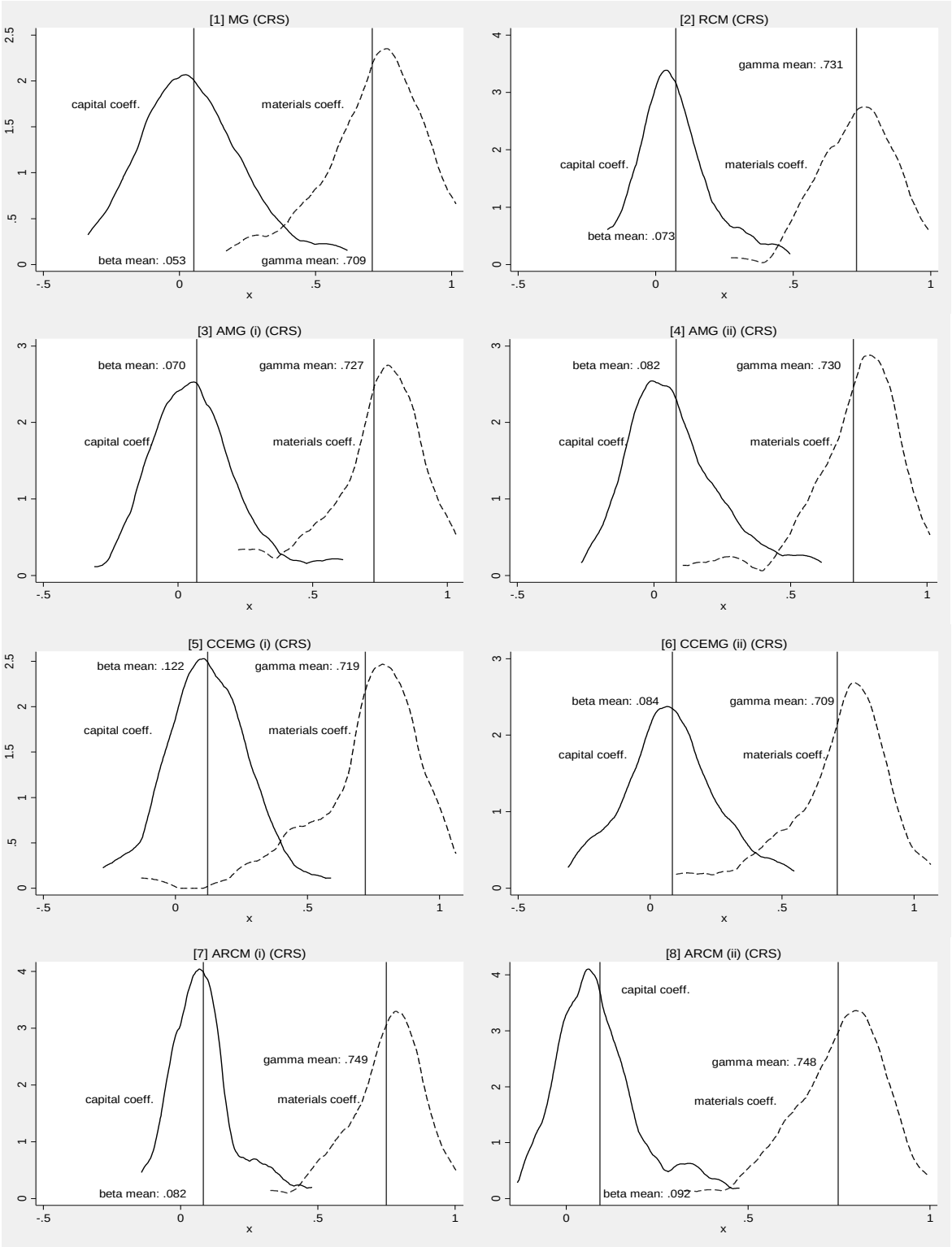
Table 2.25: Country rankings by TFP-level

TFP-level estimates
Comparison of country ranking (by magnitude) and difference across estimators

	Rank			Absolute Rank Difference		
	[1]	[2]	[3]	abs([2]-[1])	abs([3]-[1])	abs([3]-[2])
Country	**FE**	**AMG†**	**CCE‡**	**AMG-FE**	**CCE-FE**	**CCE-AMG**
AUS	12	14	13	2	1	1
AUT	22	10	10	12	12	0
BEL	24	12	12	12	12	0
BGD	34	45	45	11	11	0
BOL	4	29	23	25	19	6
BRB	33	32	32	1	1	0
CAN	6	5	4	1	2	1
CHL	3	21	19	18	16	2
COL	11	31	31	20	20	0
CYP	20	28	29	8	9	1
ECU	41	34	36	7	5	2
EGY	48	44	44	4	4	0
ESP	5	15	16	10	11	1
FIN	14	7	6	7	8	1
FJI	43	35	34	8	9	1
FRA	18	8	7	10	11	1
GBR	19	13	14	6	5	1
GTM	17	30	30	13	13	0
HUN	45	41	42	4	3	1
IDN	44	47	48	3	4	1
IND	46	48	47	2	1	1
IRL	2	17	20	15	18	3
IRN	23	23	24	0	1	1
ISR	21	2	9	19	12	7
ITA	15	4	3	11	12	1
KOR	13	25	27	12	14	2
LKA	31	46	46	15	15	0
LUX	39	6	5	33	34	1
MAR	32	42	39	10	7	3
MEX	9	19	17	10	8	2
MLT	8	38	38	30	30	0
MYS	35	39	40	4	5	1
NLD	27	9	8	18	19	1
NOR	25	11	11	14	14	0
NZL	29	16	15	13	14	1
PAN	37	22	22	15	15	0
PHL	30	37	35	7	5	2
POL	47	40	41	7	6	1
PRT	36	24	26	12	10	2
SEN	38	27	28	11	10	1
SGP	10	26	25	16	15	1
SWE	16	3	2	13	14	1
SWZ	40	43	43	3	3	0
TUN	42	36	37	6	5	1
TUR	7	18	18	11	11	0
USA	1	1	1	0	0	0
VEN	28	20	21	8	7	1
ZWE	26	33	33	7	7	0
Median				**10.3**	**10.1**	**1.0**

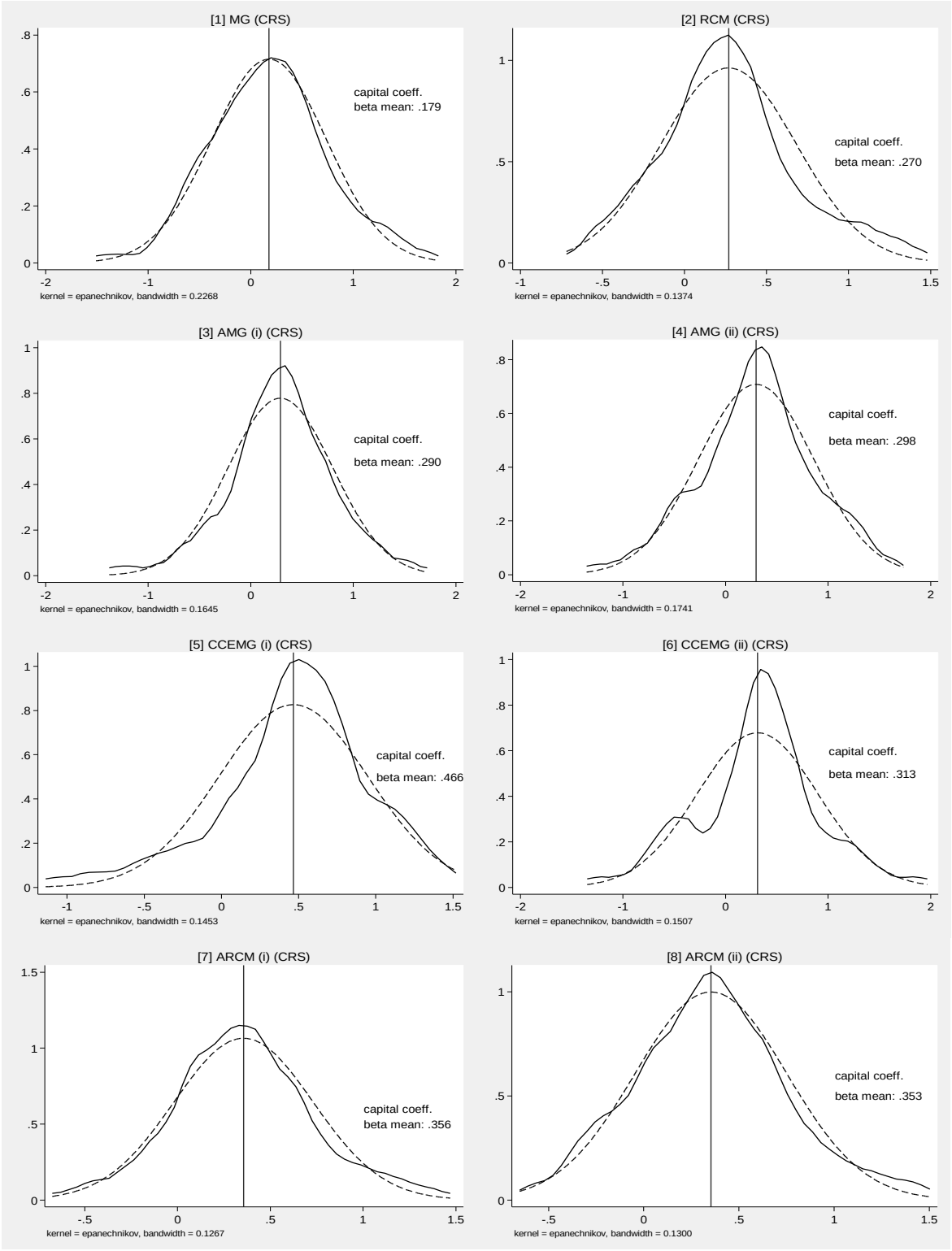
Notes: The table provides the respective TFP level ranking (by magnitude) for each country derived from the three regression models, as well as the absolute rank differences between them.
† AMG refers to the Augmented Mean Group estimator in specification (c), Table 2.3 column [4] (value-added based version), with the TFP-level adjustment detailed in Section 2.4.1.
‡ CCE refers to the Pesaran (2006) CCEMG estimator in specification (d), ibid. column [5], with the TFP-level adjustment detailed in Section 2.4.1.

Figure 2.3: Kernel densities for technology parameter estimates
(gross-output)



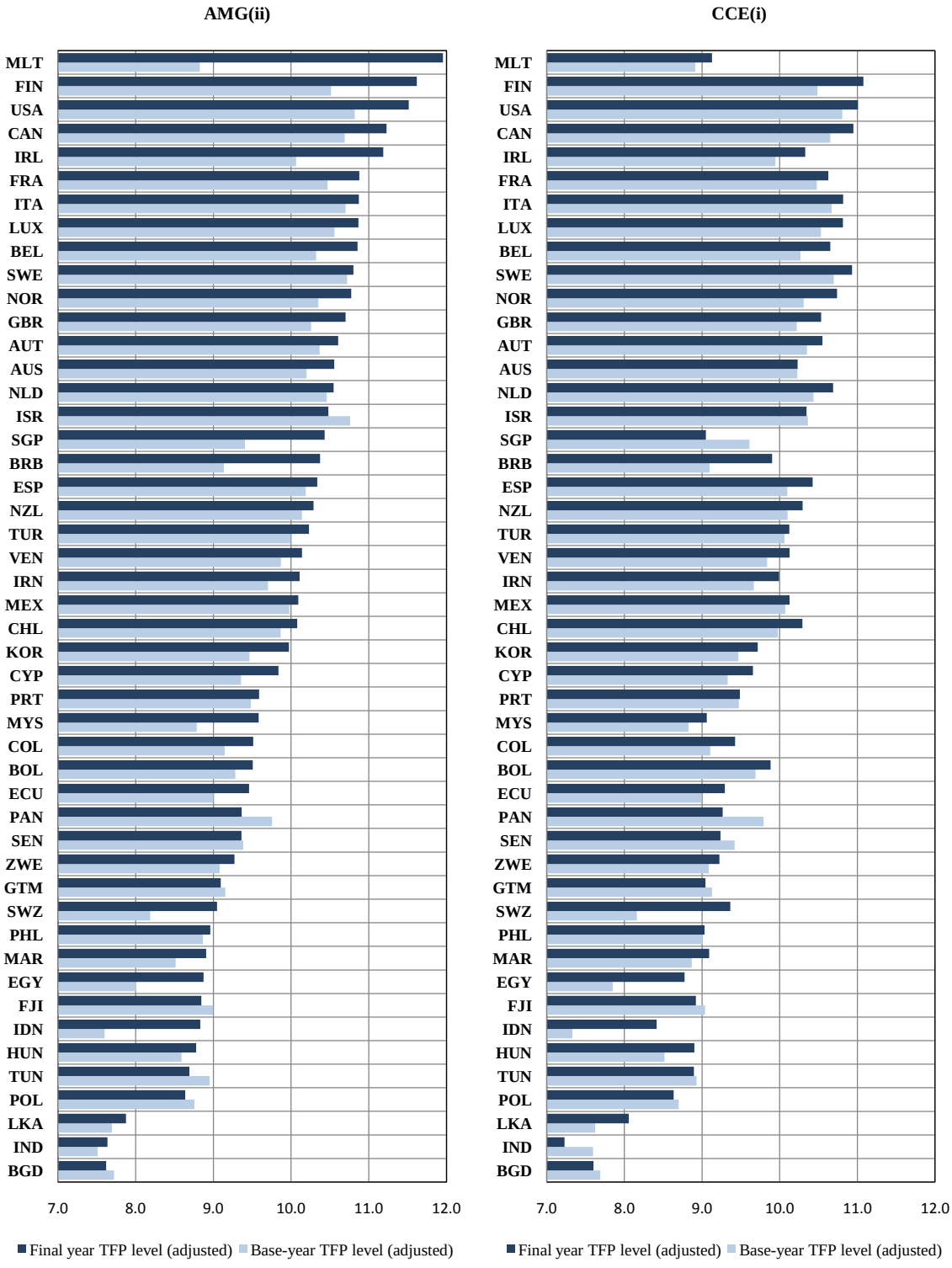
Notes: Each plot shows the distribution (kernel density) of the capital ($\hat{\beta}_i$) and material ($\hat{\gamma}_i$) coefficients from the country regressions (gross-output based analysis, CRS imposed, numbering as columns in Table 2.3 in the main section).

Figure 2.4: Kernel densities for technology parameter estimates
(value-added)



Notes: Each plot shows the distribution (kernel density) of the capital ($\hat{\beta}_i^{va}$) coefficients from the country regressions (value-added based analysis, CRS imposed, numbering as columns in Table 2.3 in the main section). The dashed line represents the normal distribution, the vertical line marks the mean of the parameter distribution.

Figure 2.5: TFP levels — adjusted from AMG(ii) and CCE(i) estimates*



Notes: VA specification with CRS imposed, $*$ AMG(ii) has $\hat{\mu}_t^{\bullet}$ *included* as additional regressor; CCE(i) refers to the standard CCEMG estimator; the estimates from which these 'adjusted' TFP levels were constructed can be found in Table 2.3, columns [4] and [5] respectively. Countries are ranked by final period TFP-level based on the AMG(ii) estimates. Levels adjustment as described in the main text.

Chapter 3

Modelling Spatial Heterogeneity in Agricultural Productivity Analysis

Abstract:

In this chapter we investigate a 'global' production function for agriculture, using FAO data for 128 countries from 1961-2002. Our review of the empirical literature in this field highlights that existing cross-country studies largely neglect variable time-series properties, parameter heterogeneity and the potential for heterogeneous Total Factor Productivity (TFP) processes across countries. We motivate the case for technology heterogeneity in agricultural production and present statistical tests indicating nonstationarity and cross-section dependence in the data. Our empirical approach deals with these difficulties by adopting the Pesaran (2006) Common Correlated Effects estimators, which we extend by using alternative weight-matrices to model the nature of the cross-section dependence. We furthermore investigate returns to scale of production and production dynamics. Our 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 or distance, 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'.

> *"[A]ssumptions of a common production function, and perfect and competitive factor markets ... get in the way of understanding international differences in productivity — particularly differences between advanced and underdeveloped economies."*
> R. Nelson (1968, p.1229)

> *"Techniques developed in advanced countries were not generally directly transferable to less developed countries with different climates and resource endowments."*
> Ruttan (2002, p.162)

Ever since Hayami and Ruttan (1970) introduced the use of panel data to estimate cross-country production functions for the agriculture sector, academic studies have emphasised the conceptual desirability of 'technology heterogeneity' across the diverse range of agro-climatic environments across countries. The literature further highlighted the potential for barriers to technology transfer between countries which are specific to the agricultural sector, in particular the problems of transfers between the developed countries of the temperate 'North' and the developing countries of the arid or equatorial 'South': innovations in the former did not seem to yield the desired productivity boost in the latter context. In practice, however, empirical investigation was typically based on models which impose technology homogeneity across the diverse sample of countries under analysis, or only allowed for heterogeneity by splitting the sample into crude geographical groups of countries. Further to this choice of empirical specification, many studies opted to impose constant returns to scale on their regressions.

In this chapter we extend the insights gained from the emerging literature on multi-factor models in nonstationary panels (Bai & Ng, 2004; Coakley et al., 2006; Pesaran, 2006; Kapetanios et al., 2009) to cross-country empirical productivity analysis in the agricultural sector — to the best of our knowledge this is the first empirical study to do so. We adopt a common factor model approach and estimate production functions for a panel of 128 developing and developed countries using annual data from 1961 to 2002 (FAO, 2007). Our focus is on the changes in the parameter estimates and diagnostic tests when we move between pooled and heterogeneous estimators, between methods which ignore cross-section dependence and those which accommodate it, and between approaches that put different emphasis on the time-series properties of long T panels. This aside, the nature of the data for agriculture allows us to investigate the cross-section dependence properties in a formal manner.

These innovations aside we particularly focus on three issues: *firstly*, we do not restrict the returns to scale of production, given that with a fixed input (land) part of the production function we do not have any priors about whether returns can assumed to be constant, decreasing or increasing.

Secondly, having introduced the Pesaran (2006) common correlated effects estimators in Chapter 2, we set out to identify what these 'common effects' might actually represent in the case of agriculture. In the standard CCE approach the nature of the cross-sectional dependence across countries is unspecified; in our extension to this approach, which we believe represents a further original contribution to the empirical literature, we apply a number of weighting schemes in the construction of the cross-section averages, based on neighbourhood, geographical distance and agro-climatic distance. The original idea to model heterogeneity in the unobserved common factors in a crude manner was to run AMG first stage estimations (see below) for different geographical regions (Africa, Europe, East Asia, etc.) and then to use the respective 'common dynamic factors' in the second stage country regressions — however an implementation extending the Pesaran (2006) CCE estimators allows for much more flexibility in the hypothesised structure of cross-section dependence across countries.[1] The application of weight matrices in essence imposes structure on the factor loadings across countries, which in the standard CCE estimators are left unconstrained. The first two of our extensions investigated essentially mimic standard spatial econometric approaches using geographical distance measures, the third represents a more complex cross-section correlation. The economic interpretation of our extension would be that the set of unobserved factors influencing productivity in each country is the same for its neighbours, countries in close(r) proximity or countries with similar agro-climatic environment respectively. Thus, countries that are *not* neighbours, are geographically *more distant* or have a very *dissimilar* agro-climatic conditions are argued to to be driven by different sets of factors.

Thirdly, we check the robustness of our empirical results by comparing them with those derived from a dynamic specification of the production function. Again, the majority of studies in the literature concentrate on static models.

[1] Conversation with Hashem Pesaran at the Conference on "Factor Structures for Panel and Multivariate Time Series Data", Maastricht, September 2008.

Our empirical analysis thus investigates the interplay and salience of time-series properties of the data, parameter heterogeneity, returns to scale assumptions and the presence as well as potential structure of cross-section dependence in the estimation of cross-country production functions in agriculture. The remainder of the chapter is organised as follows: Section 3.1 briefly reviews the existing literature on cross-country production function estimation for agriculture. Section 3.2 introduces a number of graphs to highlight agro-climatic heterogeneity across our sample of 128 countries and provides the motivation for the empirical approach. Section 3.3 reintroduces the common factor model adopted, develops our extension to the Pesaran (2006) CCE estimators and introduces the data. Section 3.4 presents and discusses the empirical results,[2] before we conclude our findings in Section 3.5.

3.1 Literature review

Following the seminal contribution by Hayami and Ruttan (1970) and a later update by the same authors (Hayami & Ruttan, 1985), the literature on agricultural productivity analysis across countries using panel data or 'repeated cross-sections' can be broadly distinguished by two aspects. The first of these does not relate to methodological approach, but to the dataset used, whereas the second major aspect relates to the empirical restrictions placed on production technology: whether countries are allowed to have differential technology parameters, TFP levels and TFP evolution, and whether constant returns to scale are imposed. While our short overview of the literature claims by no means to be exhaustive, we believe that the indicative literature presented in Table 3.1 and briefly discussed below does represent the breadth of the empirical field at the present time.

Most cross-country studies on agriculture use the data collated by the Food and Agriculture Organisation (FAO) which provides output and input variables for a large number of countries from 1961 onwards, but relies on tractors and agricultural machinery as a proxy for agricultural capital. Examples include Craig, Pardey, and Roseboom (1997), Cermeño, Maddala, and Trueblood (2003) and Bravo-Ortega and Lederman (2004). The alternative to this is a dataset developed by researchers at the World Bank which provides agricultural fixed capital stock data for a maximum of 57 developing and developed countries (although in practice not all

[2] The analysis of data time-series and cross-section dependence properties is presented in Appendices 3.B and 3.C respectively.

countries are included in empirical analysis due to data availability for other variables) from 1967-1992 (Larson et al., 2000).[3] An updated version of the dataset provides agricultural fixed capital stock from 1972-2000 for 30 countries.[4] The use of the two World Bank datasets is to our knowledge limited to Mundlak et al. (1997, 2008), as well as Martin and Mitra (2002) and Gutierrez and Gutierrez (2003).

Differences in empirical specification make it difficult to compare parameter estimates on factor-inputs across studies, especially if different 'non-standard' inputs (e.g. schooling, infrastructure, literacy, life expectancy) are included in the regressions across studies. Nevertheless it seems that with the exception of Martin and Mitra (2002) *all* empirical studies using the World Bank dataset(s) obtain very high capital coefficients, typically between .35 and .6 — the Martin and Mitra (2002) paper arrives at a much lower coefficient of .12.[5] In contrast, all studies using the FAO data with tractors proxying for fixed capital stock obtain capital coefficients in the range .05 to .2.

The literature on intercountry production analysis has in practice constrained parameter heterogeneity to the modelling of TFP: cross-country heterogeneity in factor-input parameters (what we refer to as 'technology parameters') has either been ignored (Hayami & Ruttan, 1970; Craig et al., 1997; Mundlak et al., 1997, 2008; Martin & Mitra, 2002) or approached by splitting the sample into 'homogeneous groups', e.g. by level of development or continent (Hayami & Ruttan, 1985; Cermeño et al., 2003; Gutierrez & Gutierrez, 2003). Although many of these studies stress the importance of allowing for technology differences across countries or country groups, none of them investigates this in an approach which allows for *all* parameters in the empirical specification to differ. Our common factor model specification for agricultural production thus represents the first attempt in the literature at modeling both heterogeneity in factor-input parameters and TFP evolution across countries.

[3] This dataset is also noteworthy for its construction of 'orchard capital', which has thus far only been considered in the work by Mundlak, Larson, and Butzer (1997). Other variables, such as sectoral value added, arable land and agricultural labour force are taken from World Bank, FAO and the ILO sources (see Martin & Mitra, 2002, for details).

[4] At present the updated data is not publicly available, although the World Bank team headed by Don Larson seem happy to provide the data if approached.

[5] Given that the methods applied are very similar this discrepancy may be caused by the alternative deflation strategy applied in Martin and Mitra (2002). Similar to the practice in the FAO data, the authors advocate the use of a single LCU-US$ exchange rate in favour of the practice of using annual exchange rates as implemented in the Larson et al. (2000) data.

Closely linked to this issue is the returns to scale assumption reflected by the empirical specification adopted. Imposition of constant returns to scale in agriculture is at times argued to be inherently problematic given the presence of a fixed factor input (land) in the production function. However, this does not prevent constant returns to all (fixed and variable) inputs, which implies decreasing returns to variable inputs (Graham & Temple, 2006). We would therefore postulate a more accommodating approach whereby none of the priors (constant, decreasing, increasing returns to scale) is allowed to influence/dominate the results. For cross-country production analyses in agriculture findings of increasing, decreasing or constant returns to scale (all of which are present in the literature, as can be seen in Table 3.1) are typically either justified by referring to micro-data studies of production ('similar coefficients to the study by...') or arguing for the impact of structural change (labour shifting out of agriculture) within countries during the sample period.

Hayami and Ruttan (1985), for instance, report increasing returns to scale for a subsample of *developed* countries (DC),[6] while their *developing* country (LDC) sample cannot reject constant returns. They argue that increasing returns are linked to the indivisibility of fixed capital, which has played an increasingly important role in the substitution of labour in developed countries ('labour-saving technology'). The result for their LDC sample is said to be the outcome of increasing population pressure on land over the sample period: efforts to increase productivity were therefore directed toward saving land by applying more inputs that acted as land-substitutes, such as fertilizer, chemicals or improved seeds ('land-saving technology'). Since these inputs are highly divisible, the authors argue, it is not surprising to encounter constant returns in this subsample. Although some LDCs also witnessed labour-saving technological change, it is argued that this must have been dominated by the scale-neutral impact of land-saving technical change.

Mundlak et al. (1997) in contrast argue that the "contribution of inputs to growth should be judged by their contribution to output *under a constant technology*, attributing the rest of the growth to technical change [TFP]" (paper summary, emphasis added). With implicit reference to results by Hayami and Ruttan they claim that their own specification "succeeds in capturing the impact of cross-country differences in technology and thus eliminates the spurious result

[6] Only if production data is deflated by the number of farms, i.e. everything is expressed in *per farm terms*. They remark that their national aggregate data displays constant returns for both subsamples — unfortunately results are not presented.

of increasing returns to scale" (p.13). It is important to point out that once the empirical implementation allows for heterogeneous technology parameters across countries, the *ceteris paribus* property of regression parameters alluded to by Mundlak et al. (1997) breaks down. We refer to the discussion on empirical TFP levels in Chapter 2 of this thesis.

The work by Gutierrez and Gutierrez (2003) to the best of our knowledge represents the only cross-country study of agricultural production which accounts for time-series properties of the data, using nonstationary panel econometric methods (Kao et al., 1999). They find that the World Bank dataset (1970-1992, 47 countries) indeed contains nonstationary variables, with capital, land and labour seemingly cointegrated. Their preferred panel estimator yields a *very* high capital coefficient of .58 for the full sample.[7] Their analysis rejects parameter homogeneity when they split the data into 'tropical' and 'temperate' countries. Echoing the Hayami and Ruttan (1985) study they find that constant returns is rejected in the latter subsample, where elasticities sum to 1.08.

Virtually all of the studies reviewed allow for some unobservables to be accounted for in the analysis through the inclusion of year dummies or by converting the data into deviations from the cross-sectional mean. As is recognised in the present literature on unobserved heterogeneity in panel time series (Coakley et al., 2006), these practices implicitly specify the impact of common factors to be *the same across all countries*. If the true impact is heterogeneous, the remaining error dependence can result in misleading inference and even inconsistency in the standard fixed effects panel estimators favoured in this literature (Phillips & Sul, 2003). Furthermore, if common factors drive both the regressors and the error terms this will lead to inconsistent panel estimators due to the correlation between the regressors and the error components (Pesaran, 2006).

In conclusion, the present empirical literature on cross-country production functions for agricultural largely neglects parameter heterogeneity, variable time-series properties and the potential for unobserved factors with heterogeneous impact across countries. In most cases the empirical specification is static and the dynamics of the production process are not considered.

[7] In excess of 50% higher than the next capital coefficient in magnitude, .36 in Mundlak et al. (1997).

Table 3.1: Selective Review of the Literature

	Hayami & Ruttan (1970)	**Hayami & Ruttan (1985)**	**Craig et al (1997)**	**Mundlak et al (1997)‡**	**Martin & Mitra (2002)**	**Cermeño et al (2003)**	**Gutierrez & Gutierrez (2003)**	**Bravo-Ortega & Lederman (2004)**	**Mundlak et al (2008)**
Data									
Countries	36	43	98	37	49	84	47	86	30
Observations	108	129	588	777	1248	2604	1081	2993	870
Year(s)/Period	1955, 1960, 1965	1960, 1970, 1980	1961-1990	1970-1990	1967-1992	1961-1991	1970-1992	1961-1997	1972-2000
Data source	own, annual	own, annual	FAO, 5-year avg.	WB, annual	WB, annual	FAO, annual	WB, annual	FAO, annual	WB, annual
Specification									
Model	pooled	pooled	pooled	pooled	pooled	pooled by region	pooled	pooled	pooled
Dynamics	static	static	static	static	static	dynamic	static	static	static
Estimation method	POLS	POLS	POLS	2FE	POLS	POLS	panel DOLS	POLS	2FE
Covariates:									
'standard'	L, N, Live, F, K	L, N, Live, F, K †	L, N, Live, F, K	L, N, K, F	L, N, K	L, N, K, Live, F	L, N, K	L, N, P, K, F, Live,	L, N, Live, K, F
'non-standard'	School, Tech Edu	School, Tech Edu	Lit, Life, Infra (all quality-adjusted)	School, Dev, Yield, Prices				irrigated land	
Year dummies	✓	✓	✓	implicit	N trends	✓	✓	✓	implicit
Country dummies			✓	implicit	✓	✓	✓	✓	implicit
Returns to scale	CRS imposed	unrestricted	unrestricted	unrestricted	CRS imposed	unrestricted	unrestricted	CRS imposed	CRS imposed
Findings									
labour	.40	.50	.16	.08	.64		.29	.11	.12
land	.07	.03	.40	.42	.24	.02 – .47	.14	.04	.33
pastures								.53	
irrigated land								.03	
capital/machinery	.11	.07	.05	.36	.12	.02 – .08	.58	.04	.34
livestock	.29	.31	.35			.03 – .40		.25	.08
traction animals			-.06						
fertilizer	.14	.15	.04	.08		.00 – .10		.05	.13
TFP			no difference across i or over t	accounted via 2FE transformation	sign. diff. across i, 1.4-3.5% growth pa	sign. diff. across i and over t	accounted via DOLS method	sign. trend term (1%pa)	accounted via 2FE transformation
Returns to scale	(CRS imposed)	not tested	DRS	CRS not rejected	not tested	differs across regions	CRS	(CRS imposed)	(CRS imposed)
Reference	Table 2, (17)	Table 6-2, (Q14)	Table 1, (1)	Table 4	Table 1, footnote	Tables 1-4, (7)	Table 2, DOLS	Table 6(A), (2)	Table 2, Within

Notes: The table compares Cobb-Douglas production function estimates from studies as indicated. Since many studies report results for more than one specification we concentrate on the most general model for 'standard' production inputs with panel data. The specific reference for each study's entry is stated at the bottom of the table. The studies by Hayami and Ruttan (1985) and Gutierrez and Gutierrez (2003) also provide separate estimations by country groups (developed, developing), finding production technology differences to be considerable and negligible respectively.
Datasets: The Hayami and Ruttan (1970, 1985) studies use data from a wider range of sources which includes the FAO, OECD and other international institutions. FAO — the various FAOSTAT databases provided by the Food and Agriculture Organisation of the UN, WB — the Larson et al. (2000) dataset and its updated version.
Output: agricultural output (expressed in some type of constant 'international' monetary terms) in all studies.
Inputs: L — agricultural labour, N — (crop) land, Live — livestock, F — fertilizer, K — capital stock/machinery/tractors, P — pastures, School — school enrolment, Tech Edu — technical education, Lit — literacy, Life — life expectancy, Dev — development level, Yield — peak yield, Infra — various infrastructure variables, Prices — various price indicators.
Estimators: POLS — pooled OLS (we indicate separately whether the regression equation contains country fixed effects), 2FE — Two-way Fixed Effects, DOLS — Dynamic OLS, following Kao et al. (1999). This estimator imposes a homogeneous cointegration vector but allows for country-specific short-run dynamics.
†The reported coefficients are derived from data deflated by the number of farms. Regressions for national aggregates are not presented. ‡A slightly different version of this paper is published as Mundlak et al. (1999).

3.2 Technology Heterogeneity and the Analysis of Agricultural Production

3.2.1 Heterogeneity in the Agro-Climatic Environment

In this section we attempt to create an impression of the differential agro-climatic environment across different countries. Our focus is here on the differences in the share of arable land[8] in climatic zones across regions of the globe. The data is taken from FAO (2007) and Gallup, Mellinger, and Sachs (1999). The sample is made up of 128 developing and developed countries; the measures presented are essentially time-invariant. Detailed information about data construction and sample coverage is presented in Appendix 3.A.

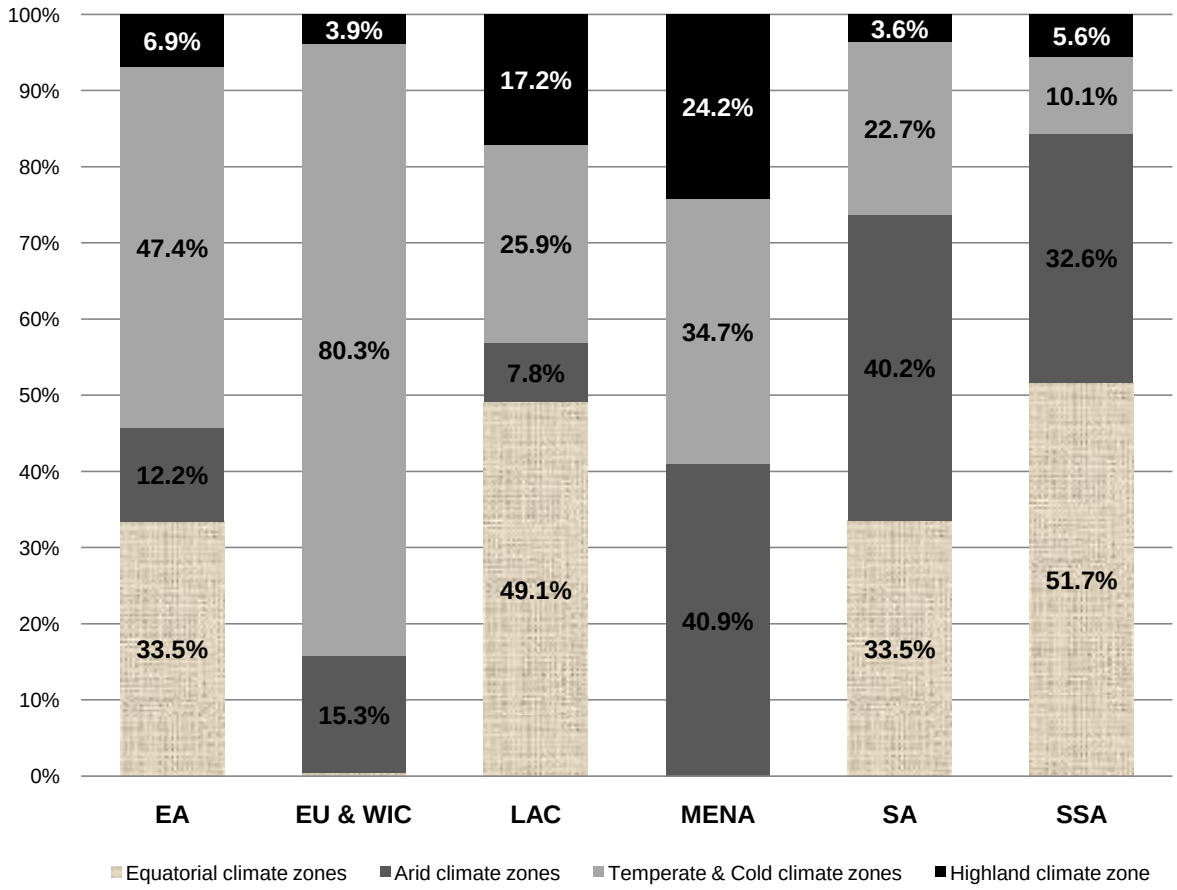
In Figure 3.1 we present the average share of agricultural land in the four major climatic zones (equatorial, arid, temperate & cold, highland) for a number of country groups.[9] The means are computed as weighted means within regions, where the weights are the size of arable land for each country. The bar chart provides a strong sense of the heterogeneity in agro-climatic environment across regions: Sub-Saharan Africa (SSA), for instance, at 52% has the highest average share of land in the equatorial zone, higher even than the Latin America & Caribbean group (LAC), which includes Brazil. In contrast, this agro-climatic zone is (virtually) absent from the developed (EU & WIC) and MENA countries. Developed countries are dominated by agricultural land in the temperate & cold zone (>80%), which in turn makes up only around 10% of land in the average Sub-Saharan African country — the lowest share across all regions. It would thus be difficult to refer to any of these groupings as 'similar' to each other, as far as agro-climatic environment is concerned.

In order to get an alternative, more refined visualisation of these geographical differences across individual countries, rather than crude groupings, we compute a 'Jaffe-measure' (Jaffe, 1986) for agro-climatic environment. This measure is computed for every possible country-pair and can be interpreted as a multi-variate correlation coefficient which varies between zero and unity, where a low (high) value indicates little (high) similarity in the distribution of cultivated land across climatic zones in the two countries (Pardey et al., 2007). From the Matthews (1983) data

[8] Instead of using the full definition of 'arable land and permanent crop area', throughout this chapter we refer to land or 'arable land' for simplicity.

[9] See Appendix 3.7 for the makeup of each grouping.

Figure 3.1: Geographical regions and climate zones — weighted means



Notes:
This figure shows the weighted group mean for the share of arable land in each climate zone, where the weights are the area of arable and permanent crop land. The groups are East Asia (EA), Europe & other Western Industrialised Countries (EU & WIC), Latin America & Caribbean (LAC), Middle East & North Africa (MENA), South Asia (SA) and Sub-Saharan Africa (SSA).

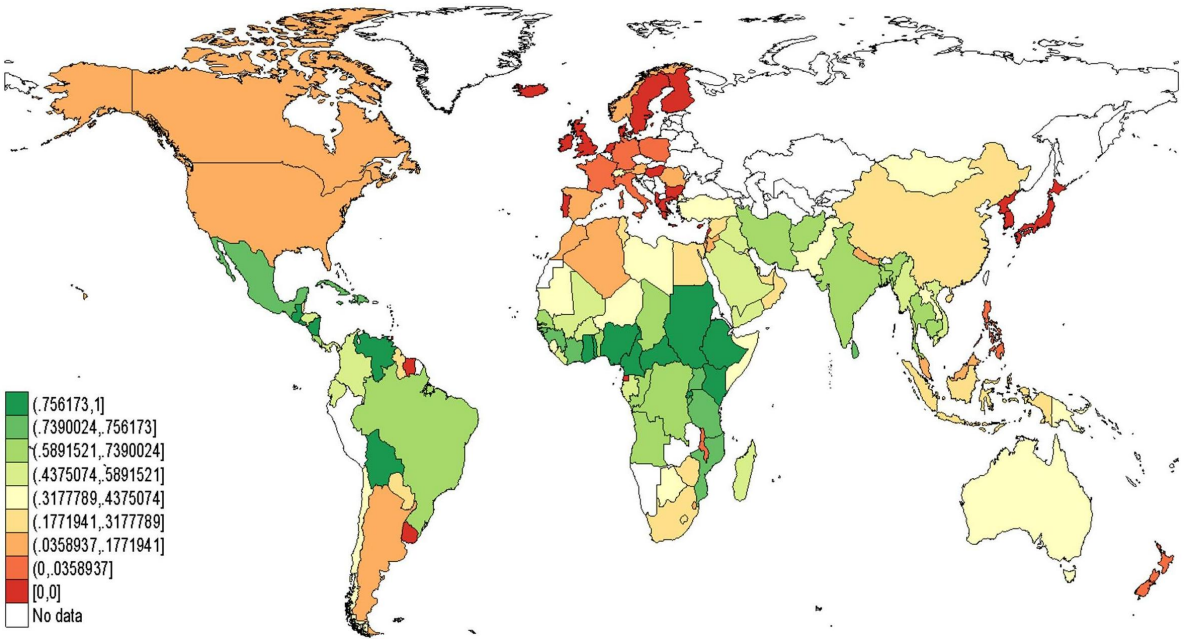
we obtain the share of cultivated land within each of twelve climatic zones ($h_i = (h_{i1}, \ldots, h_{i12})$), such that for each country i the values in the twelve zones sum up to unity ($\sum_m h_{im} = 1$). The Jaffe measure for 'agro-climatic distance' between countries i and j is then

$$\omega_{ij} = \frac{\sum_m h_{im} h_{jm}}{\left(\sum_m h_{im}^2\right)^{1/2} \left(\sum_m h_{jm}^2\right)^{1/2}}$$

In Figure 3.2 we plot the Jaffe-measures for Kenya: countries with dark green shading, such as Ethiopia, Sudan or Ghana, as well as some Central American countries have an agro-climatic makeup (distribution of agricultural land across climate zones) very similar to that of Kenya, whereas countries with lighter, yellow and orange shading are progressively different. At the other end of the scale, countries in red, such as the UK, Ireland or Japan have an agro-climatic environment which is entirely different from the Kenyan one.

We may now find the tales of failed agricultural technology transfer from developed to developing countries (Ruttan, 2002; Gutierrez, 2002) rather unsurprising: many of the former have very different agro-climatic characteristics from the latter thus making it rather intuitive that technology between (and even within) these two groups differs and the potential for successful technology transfer *without adaptation* may be limited.

Figure 3.2: Agro-climatic 'distance' — the view from Kenya



Notes:
Kenya's cultivated land: 40% is located in zone Aw (Equatorial savannah, dry winters), 19% in zone BS (steppe), 17% in zone BW (desert) and 25% in zone H (highland climate). Source: Matthews (1983), in Gallup et al. (1999).

3.2.2 The Case for Technology Heterogeneity and Cross-section Dependence in the Data

The previous section highlighted the stark differences in the agro-climatic environment for agricultural production across different countries. In our econometric implementation we translate this into the need to allow for *technology heterogeneity* across countries.[10] The assumption of a homogeneous production function may mask or distort important insights into development, as the comment by R. Nelson (1968) at the beginning of this chapter suggests. Hayami and Ruttan (1970) state that Nelson's comment (referring to the analysis of manufacturing data)

[10] An alternative approach to technology heterogeneity may be to specify a more general homogeneous production function, for instance containing separate variables for the share of land in various climatic zones. This would however merely capture the impact of climatic differences contained in the land variable, whereas the 'agro-climatic distance effect' approach of our common factor approach implies that climatic factors drive the unobservables (TFP) in agricultural production (the standard Pesaran (2006) approach is agnostic over the driving force).

equally applies to the analysis of agricultural production. Their approach further highlights the implications of and anecdotal evidence for technology being endogenous to the prevailing 'economic system', namely the "differential diffusion of agricultural technology... [and] of the scientific and technical capacity to invent and develop new mechanical, biological, and chemical technology specifically adapted to the factor endowments and prices in a particular country or region" (Hayami & Ruttan, 1970, p.898). In addition to technology heterogeneity, this highlights potential limits to technological spillover from innovations between countries with very different agro-climatic makeup and resource endowment: we indicated the vast differences for the former in our visualisation of agro-climatic 'distance' in the case of Kenya. Much of agricultural technology has to be viewed as location-specific, with attempts at direct technology transfer from one agro-climatic region to another largely doomed to failure (Gutierrez, 2002; Ruttan, 2002). By the early 1960s it had become increasingly clear that "[t]echniques developed in advanced countries were not generally directly transferable to less developed countries with different climates and resource endowments" (Ruttan, 2002, p.162). However, this insight does not seem to have affected the specification of empirical models in any way.

We incorporate these ideas into our empirical model in three ways: *firstly*, we allow for parameter heterogeneity in the observable variables in our empirical model. At present no intercountry production analysis in agriculture has considered this specification.

Secondly, we adopt a common factor approach to model TFP, which we interpret in the Abramowitz (1956) fashion as 'measure of ignorance': all time-variant processes, inputs, 'factors' and characteristics which are not captured by the observable covariates in our model are arguably captured by unobserved common factors, which are allowed to possess heterogeneous 'factor loadings' across countries. This common factor representation enables us to account for cross-section dependence between countries — this dependence can be hypothesised to arise from common shocks and/or spillover effects, the latter determined for instance by trade, policy or technology (Costantini & Destefanis, 2009). The importance of accounting for cross-section dependence in estimation and inference has become a major theme in the recent nonstationary panel econometric literature (Bai & Ng, 2004; Pesaran, 2004; Breitung & Pesaran, 2005; Coakley et al., 2006; Baltagi, Bresson, & Pirotte, 2007) and this chapter represents our second empirical example of these themes.

Thirdly, we use the above arguments of potential barriers to technology spillovers to motivate an extension to the Pesaran (2006) CCE estimators. In the standard setup these estimators account for the impact of unobserved factors by augmenting the regression equation with a vector of cross-section averages at time t for each of the variables, whereby the averages are constructed from equally weighted observations across all countries. The economic interpretation of this approach would be that the set of unobserved factors which influences productivity is common to all countries — this is not to say that their impact is the same (in fact their heterogeneity is the major contribution of the Pesaran (2006) approach), it is just required that across all countries the impact of each factor is on average non-zero. In our extension we chose to impose some more structure on this framework by investigating three alternative scenarios:

(i) **The 'Neighbourhood Effect'**: many empirical studies have argued that the economic performance of contiguous neighbours to country i has a significant effect on the latter's total factor productivity, and attempted to measure the impact of this spillover empirically by specifying spatially-lagged dependent variables in a production function model (e.g. Ertur & Koch, 2007). In our empirical models we will allow for a conceptually similar relationship in the data using the alternative CCE estimation approach.

(ii) **The 'Gravity Model Effect'**: gravity models are common in empirical trade analysis where they suggest geographical distance as a powerful determinant of the magnitude of economic exchange between countries (e.g. Frankel & Romer, 1999; Redding & Venables, 2004). We adopt this approach to hypothesise that distance between countries (a crude proxy for factors such as climatic, soil, cultural, ethnic and socio-economic differences) can explain the effects of unobserved heterogeneity across countries. Existing examples in the spatial econometric literature are reviewed in Magrini (2004), Abreu, de Groot, and Florax (2005) and Bode and Rey (2006).

(iii) **The 'Agro-Climatic Distance Effect'**: much of the existing literature on intercountry production functions in agriculture particularly highlights the differential agro-climatic characteristics across countries and even explicitly links the failure of technology transfer between diverse countries to heterogeneity in resource endowment and climate. Our third alternative to the standard CCE approach hypothesises that countries with similar agro-climatic makeup are influenced by similar unobserved factors.

In the following section we show somewhat more formally how these ideas are introduced into our econometric model.

3.3 Empirical Model and Implementation

3.3.1 Common factor model representation

As in the manufacturing analysis we adopt an unobserved common factor model as our empirical framework. For $i = 1, \ldots, N$, $t = 1, \ldots, T$, let

$$y_{it} = \boldsymbol{\beta}_i' \boldsymbol{x}_{it} + u_{it} \qquad u_{it} = \alpha_i + \boldsymbol{\lambda}_i' \boldsymbol{f}_t + \varepsilon_{it} \tag{3.1}$$

$$x_{mit} = \pi_{mi} + \boldsymbol{\delta}_{mi}' \boldsymbol{g}_{mt} + \rho_{1mi} f_{1mt} + \ldots + \rho_{nmi} f_{nmt} + v_{mit} \tag{3.2}$$

$$\text{where} \qquad m = 1, \ldots, k \qquad \text{and} \qquad \boldsymbol{f}_{\cdot mt} \subset \boldsymbol{f}_t$$

$$\boldsymbol{f}_t = \boldsymbol{\varrho}' \boldsymbol{f}_{t-1} + \boldsymbol{e}_t \qquad \text{and} \qquad \boldsymbol{g}_t = \boldsymbol{\kappa}' \boldsymbol{g}_{t-1} + \boldsymbol{\epsilon}_t \tag{3.3}$$

We assume a production function with observed inputs $\boldsymbol{x}_{it}$ (labour, agricultural capital stock, livestock, fertilizer, land under cultivation) and observed net output y_{it} (all in logarithms). Technology parameters on the inputs can vary across countries ($\boldsymbol{\beta}_i$). 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 factor loadings that can differ across countries ($\boldsymbol{\lambda}_i'$). We also introduce an empirical representation of the observed inputs in equation (3.2) 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 $\boldsymbol{f}_{nmt}$, whereby the latter as indicated represent a subset of the factors driving output in equation (3.1). The intuition is that some unobserved factors driving agricultural production are likely to similarly drive (at least in parts) the evolution of the inputs. As was indicated in the earlier chapters, this overlap of common factors creates severe difficulties for the identification of the technology parameters $\boldsymbol{\beta}_i$ (see Remark 4 in Kapetanios et al., 2009) and our empirical estimators set out to address this issue: previous simulation exercises (Coakley et al., 2006; Kapetanios et al., 2009) as well as our own investigations presented in Chapter 4 indicate that the newly-developed CCE-type estimators are able to accommodate this type of endogeneity in the estimation equation to arrive at consistent parameter estimates for common $\boldsymbol{\beta}$ coefficients or the means of heterogeneous $\boldsymbol{\beta}_i$. Equation (3.3) indicates that the factors are persistent over time, which allows for the setup to accommodate nonstationarity in

the factors ($\boldsymbol{\varrho} = 1$, $\boldsymbol{\kappa} = 1$) and thus the observables. It further allows for various combinations of cointegration: between output y and inputs $\boldsymbol{x}$, and between output $\boldsymbol{y}$, inputs $\boldsymbol{x}$ and (some of) the unobserved factors $\boldsymbol{f}_t$. The latter alternative is of particular interest, since it will allow us to specify a cointegrating relationship *without even knowing its individual elements.*

The above empirical framework allows for the maximum level of flexibility, with regard to parameter heterogeneity, cross-section dependence induced by the common factors, and non-stationarity in the observables and unobserved factors. We refer to Chapter 2 for a detailed discussion of the assumptions of this empirical model.

3.3.2 Empirical estimators

For a detailed discussion of various pooled and country-regression approaches refer to Section 2.1.2 in Chapter 2. In the following we restrict ourselves to introducing alternatives or extensions to the estimators discussed earlier.

In this chapter we use the two-way fixed effects (2FE) estimator as defined in Coakley et al. (2006), which is equivalent to the fixed effects estimator augmented with $T-1$ year dummies. The former is defined as

$$\text{2FE} \qquad \breve{y}_{it} = \boldsymbol{b}'\breve{\boldsymbol{x}}_{it} + \breve{e}_{it} \tag{3.4}$$

where $\breve{Z}_{it} = Z_{it} - \bar{Z}_i - \bar{Z}_t + \bar{Z}$ with $\bar{Z}_i = T^{-1}\sum_{t=1}^{T} Z_{it}$ (country average), $\bar{Z}_t = N^{-1}\sum_{i=1}^{N} Z_{it}$ (cross-section average) and $\bar{Z} = (NT)^{-1}\sum_{t=1}^{T}\sum_{i=1}^{N} Z_{it}$ (full sample average). This estimator imposes homogeneity on the input-parameters and was developed under the assumption of cross-section independence. However, the transformation into cross-sectional deviations allows for *homogeneous* cross-section dependence whereby unobserved common factors influence all countries in the same way (Coakley et al., 2006).

In this chapter we further introduce extensions to the Pesaran (2006) CCE estimators which we develop below using the CCEP as illustrative example — the extensions to the CCEMG are constructed in analogy. As a reminder, the CCEP represents a pooled fixed effects equation which is augmented with cross-section averages of the dependent and independent variables,

in a fashion such that their impact is allowed to vary across countries. Formally,

$$\text{CCEP} \qquad y_{it} = a + \boldsymbol{b'X}_{it} + \sum_{j=2}^{N} a_j D_j + \sum_{j=1}^{N} d_{1i}\{\bar{y}_t D_j\} + \sum_{j=1}^{N} \boldsymbol{d}'_{2i}\{\bar{\boldsymbol{x}}_t D_j\} + e_{it} \tag{3.5}$$

where the first three terms represent a standard fixed effects estimator (the D_j are country dummies) and the following two terms represent the augmentation with cross-section averages at time t ($\bar{Z}_t = N^{-1}\sum_{i=1}^{N} Z_{it}$), in each case interacted with a set of N country dummies D_j. The augmentation creates $k+1$ matrices of dimensions $NT \times N$ where k is the number of observed variables in the model.

In the standard CCE estimators all countries are given the same 'importance' in the construction of $\bar{Z}_t$ (arithmetic mean). In an extension to this standard practice we experiment with a number of weight-matrices to develop alternative CCE estimators which impose some more structure on the unobserved factor loadings $\boldsymbol{\lambda}_i$. We present three variants:

(i) The Neighbourhood CCE approach — for country i the cross-section averages (means) for y and $\boldsymbol{x}$ are constructed from the values for i's *contiguous* neighbours.

(ii) The Gravity CCE approach — for country i the observations for all other countries $j = 1, \ldots, N-1$ are weighted by the inverse of the population-weighted *distance* between i and j before computing the cross-section aggregate.

(iii) The Agro-Climatic CCE approach — for country i the observations for countries $j = 1, \ldots, N-1$ are weighted by a Jaffe-type measure for agro-climatic distance between i and j before computing the cross-section aggregate.[11]

Our conceptual justification for these variants of the CCE estimators is the situation where the average of factor loadings across countries is non-zero, but systematic patterns drive the data. In the distance case we implicitly test the hypothesis that country i is driven by unobserved common factors which are the same in countries *in close proximity*, but is much less affected by other factors which drive countries further away from it. The neighbourhood case represents an extreme extension of this argument whereby *only countries which share a common border* are driven by the same factors. Finally, in the agro-climatic case we test the hypothesis

[11] For the Jaffe (1986) measure see Section 3.2.1 for details. For the distance and agro-climatic distance we normalise the weights (across i) so that they sum to unity.

that *countries* with similar agro-climatic environment are affected by a shared set of common factors, but that these countries are not (or only to a very limited degree) affected by a separate set of common factors, which in turn influences countries in very different agro-climatic environments.

Our empirical analysis also investigates the Augmented Mean Group (AMG) estimator we introduced in the previous chapter. In contrast to our analysis for manufacturing in Chapter 2 the AMG yields average estimates much closer to the standard MG than the CCEMG, and the diagnostic tests hint at its misspecification. While Monte Carlo simulations presented in Chapter 4 reveal that for many standard setups the AMG performs equally well as the CCEMG, we have also identified circumstances in which the former is biased. This is the case when its first stage, pooled OLS with variables in First Differences, is biased due to feedbacks from output to inputs.

3.3.3 Data

The principal data source for our empirical analysis is the Food and Agriculture Organisation's *FAOSTAT* panel database (FAO, 2007), from which we obtain annual observations for agricultural net output, economically active labour force in agriculture, number of tractors used in agriculture, arable land and permanent crop land and fertilizer use in 128 countries from 1961 to 2002 (average $T = 40.3$).

Additional time-invariant data on geographical distance between countries and contiguity (neighborhood) is taken from CEPII (2006), and data on the share of agricultural land by climatic zone from Matthews (1983) available in Gallup et al. (1999). Data construction is discussed in Appendix 3.A which also contains the descriptive statistics.

3.4 Empirical Results

3.4.1 Time-Series Properties and Cross-Section Dependence

We carry out a set of stationarity and nonstationarity tests for individual country time-series as well as the panel as a whole, results for which are presented in Appendix 3.B. Ultimately, in

case of the present data dimensions and characteristics, and given all the problems and caveats of individual country and panel unit root tests, we can suggest *most conservatively* that non-stationarity cannot be ruled out in this dataset. Investigation of the time-series properties of the data was not intended to select a subset of countries which we can be reasonably certain display nonstationary variable series like in Pedroni (2007); instead, our aim was to indicate that the sample (possibly due to the limited time-series dimension, as argued in Pedroni, 2007) is likely to be made up of a mixture of some countries with stationary and others with non-stationary variable series.

The results for the cross-section dependence (CSD) analysis are presented in Appendix 3.C. Our analysis provides strong evidence for the presence of cross-section dependence within the full sample dataset, based on average variable cross-country correlation coefficients, principal component analysis and the Pesaran (2004) CD test. This result holds for both the individual variables as well as residuals from regressions which do not address cross-section correlation: the pooled OLS (POLS), two-way fixed effects (2FE) and Mean Group (MG) estimators.[12] The CD test for residuals from standard CCEMG regressions cannot reject cross-section independence, indicating that this approach successfully deals with the presence of unobserved common factors. It is noteworthy that the average cross-section correlation for residuals from the pooled OLS regression with variables in first differences (FD-OLS) drops considerably and like in the CCEMG case cross-section independence cannot be rejected. This result is surprising — Coakley et al. (2006) did not investigate the FD-OLS estimator in their study and we further highlight its performance in our discussion in Chapter 4. We further carry out the same testing procedures for various subsamples of the data, based on geographic and climatic categories. The standard CCEMG estimator seems to achieve the elimination of cross-section dependence even within and across narrower subsamples, while most other estimation approaches (now including FD-OLS) are subject to considerable residual correlation within and across country groupings. CD tests and mean absolute residual correlation are reported in the results section below for each of the empirical specifications considered.

[12] POLS is augmented with $T-1$ year dummies, MG country regressions with a linear trend term.

3.4.2 Pooled estimation results

We present the estimation results for the pooled specifications in Table 3.2. The dependent variable and the independent variables are expressed in per worker terms, such that the addition of the labour variable indicates deviation from constant returns to scale. In the lower panels of the table we report the implied returns to scale and labour coefficients as well as various diagnostic test results. Recall for the following discussion that the 2FE estimator represents the empirical implementation of choice in the present literature.

We first discuss parameter coefficients: most empirical models presented indicate large decreasing returns to scale in agriculture. In common with many studies using the FAO data, the coefficients on capital (tractors) are relatively low across all models, ranging from .04 (FD-OLS) to .13 (POLS). The land coefficients are high and relatively stable across specifications (between .21 and .33), whereas the fertilizer coefficients range from .01 to .17. Livestock again has a rather large coefficient across all specifications (.22 to .37).

Regarding the implied labour coefficients, we find very low magnitudes across all specifications, with the standard and distance-weighted CCEP even providing nonsensical negative parameter values. Certainly the most striking pattern in these results is the general magnitude of the implied decreasing returns to scale: based on this analysis we can conclude that in a pooled specification the data in all but two of the weighted CCE estimators rejects constant returns emphatically, with input elasticities in the commonly favoured 2FE estimator adding up to around .80. This finding may reflect a global production function with strongly decreasing returns to scale, possibly due to the presence of a fixed factor (median land per worker growth rate for the sample: 0.0%); alternatively, it may reflect empirical misspecification.

Turning to the diagnostics, we can see that with the exception of the models in first differences, *all* models seem to display serial correlation in the residuals. Unit root tests indicate that the CCEP-type and first difference estimators seem to yield stationary residuals, in contrast to the standard panel estimators in levels (POLS, 2FE) for which nonstationary residuals cannot be rejected. Recall that t-statistics are invalid in the presence of nonstationary errors (Kao, 1999). Mean absolute residual correlations for POLS and 2FE are relatively high, at around .4, whereas this measure drops to around .16 in all other regression models. Nevertheless the

Pesaran (2004) CD-test for cross-section dependence yields very mixed results: only the residuals for the agro-climate CCEP and the FD-OLS estimator suggest cross-section independence. Further specification tests emphatically reject residual normality and homoskedasticity in all models. These diagnostics indicate that the commonly preferred 2FE estimator has serially correlated errors, which are nonstationary, nonnormal, heteroskedastic and correlated across countries. Note that input parameter estimates for this estimator are reasonably close to those in Craig et al. (1997), the closest match for this dataset and specification.

In line with much of the literature, we run regressions with CRS imposed to provide a reference point — as just reported this imposition is *rejected* by the data in most models — and report the results in Table 3.3. Interestingly the parameter estimates for tractors, livestock (moderate increase) and fertilizer are relatively stable compared with the unrestricted models. In contrast the land coefficients shoot up (in the 2FE case from .29 to .40) in all specification save for POLS. Implied labour coefficients remain relatively stable for POLS and 2FE but rise substantially for the difference and CCEP regressions. Serial correlation tests and the unit root tests display rejection patterns next to identical to those in the unrestricted case. With regard to the CD test, 2FE and CCEP errors now indicate cross-section independence. Again the normality and homoskedasticity tests reject across the all estimators. Existing results for the FAO data and a comparable empirical specification in Bravo-Ortega and Lederman (2004) show similar patterns (albeit with an even higher land-coefficient).

We can see from these results that the the erroneous imposition of CRS (the unrestricted models reject CRS) has led to larger magnitudes for either the land or the implied labour coefficient — in the commonly preferred 2FE model the implied labour coefficient has merely risen from .03 to .06 — albeit with diagnostics that reject the most common regression assumptions for this estimator (residuals which are stationary, serially uncorrelated, normal and homoskedastic). Note that our empirical setup implies that factor-input parameters are unidentified in the POLS and 2FE *if* the same unobserved common factors drive output *and* inputs — refer to Section 1.3 of Chapter 1.

Table 3.2: Pooled regressions (unrestricted returns to scale)

Pooled regressions dependent variable: [1] & [3]-[6] log output per worker, [2] dto. in 2FE transformation♮, [7] Δlog output per worker, [8] dto. in 2FE transformation♮								
weight matrix‡	[1] **POLS**	[2] **2FE**♮	[3] **CCEP** none	[4] **CCEP** neighbour	[5] **CCEP** distance	[6] **CCEP** agro-climate	[7] **FD-OLS**	[8] **FD-2FE**♮
labour†	-0.0586	-0.1908	-0.3194	-0.0659	-0.2684	-0.2157	-0.3797	-0.1946
Δ in [7]&[8]	[17.50]**	[14.79]**	[2.35]**	[0.39]	[2.52]*	[1.70]	[3.60]**	[3.54]**
tractors pw	0.1315	0.0577	0.0740	0.0898	0.0765	0.0440	0.0383	0.0718
Δ in [7]&[8]	[24.12]**	[13.76]**	[4.86]**	[4.13]**	[4.02]**	[2.68]**	[3.21]**	[5.71]**
livestock pw	0.2189	0.3579	0.3601	0.2984	0.3669	0.3715	0.2861	0.3012
Δ in [7]&[8]	[28.61]**	[30.85]**	[7.00]**	[3.95]**	[10.03]**	[7.26]**	[6.87]**	[7.80]**
fertilizer pw	0.1690	0.0726	0.0255	0.0497	0.0277	0.0258	0.0086	0.0103
Δ in [7]&[8]	[28.08]**	[23.63]**	[4.67]**	[4.97]**	[2.59]*	[4.68]**	[2.48]*	[3.00]**
land pw	0.2529	0.2936	0.2415	0.2854	0.3321	0.2452	0.2101	0.2831
Δ in [7]&[8]	[30.42]**	[21.07]**	[2.34]*	[2.86]**	[3.17]**	[3.23]**	[2.71]**	[4.33]**
We included $(T-1)$ year dummies in equations [1] and (in first differences) [7].								
Constant	1.9084 [28.93]**		2.8646 [1.78]	0.9223 [0.39]	4.2520 [2.92]**	2.6092 [1.49]		
Observations	5,162	5,162	5,162	5,162	5,162	5,162	5,013	5,013
groups	128	128	128	128	128	128	128	128
average T	40.3	40.3	40.3	40.3	40.3	40.3	39.2	39.2
R-squared	0.91	0.66	0.96	0.94	0.98	0.96	0.16	0.12
RMSE	.433	.148	.069	.083	.048	.069	.084	.085
Returns to scale implications of the parameter estimates ♭								
Implied β_L	0.1691	0.0274	-0.0205	0.2108	-0.0716	0.0978	0.0772	0.1390
Returns	DRS	DRS	DRS	CRS	DRS	CRS	DRS	DRS
Arellano-Bond Serial Correlation Test — H_0: no serial correlation in the residuals								
AR(1) (p)	36.72 (.00)	56.21 (.00)	7.62 (.00)	16.20 (.00)	12.82 (.00)	11.40 (.00)	-9.52 (.00)	-9.67 (.00)
AR(2) (p)	36.19 (.00)	49.01 (.00)	0.78 (.44)	10.96 (.00)	5.05 (.00)	3.80 (.00)	-0.61 (.54)	-0.69 (.49)
AR(3) (p)	35.63 (.00)	43.30 (.00)	-3.04 (.00)	6.66 (.00)	-0.67 (.50)	-0.50 (.61)	0.49 (.62)	0.40 (.69)
Nonstationarity: Pesaran (2007) CIPS test applied to residuals♯ — H_0: residuals are I(1)								
lags	Ztbar (p)	Ztbar (p)	Ztbar (p)	Ztbar (p)	Ztbar (p)	Ztbar (p)	Ztbar (p)	Ztbar (p)
none	-7.94 (.00)	-9.54 (.00)	-40.38 (.00)	-31.21 (.00)	-32.76 (.00)	-33.63 (.00)	-48.92 (.00)	-48.78 (.00)
1 lag	-2.67 (.00)	-3.47 (.00)	-32.11 (.00)	-21.03 (.00)	-23.73 (.00)	-24.07 (.00)	-38.06 (.00)	-37.83 (.00)
2 lags	0.01 (.50)	0.39 (.65)	-24.66 (.00)	-15.41 (.00)	-18.45 (.00)	-16.75 (.00)	-24.40 (.00)	-24.21 (.00)
3 lags	1.79 (.96)	2.05 (.98)	-20.39 (.00)	-11.28 (.00)	-14.49 (.00)	-13.18 (.00)	-14.87 (.00)	-15.32 (.00)
4 lags	3.91(1.00)	4.20 (1.00)	-15.83 (.00)	-5.96 (.00)	-8.72 (.00)	-9.26 (.00)	-8.79 (.00)	-8.94 (.00)
Cross-Section Dependence: Mean (Absolute) Correlation & Pesaran (2004) CD test — H_0: no CSD								
Mean ρ_{ij}	-0.01	0.02	0.01	0.01	0.07	0.00	0.00	0.02
Mean $\lvert\rho_{ij}\rvert$	0.42	0.41	0.18	0.15	0.17	0.16	0.15	0.15
CD (p)	-2.49 (.01)	9.64 (.00)	4.97 (.00)	4.06 (.00)	38.52 (.00)	-0.25 (.80)	0.04 (.97)	12.84 (.00)
Further residual diagnostic tests⋆ — H_0: no heteroskedasticity, regular skewness/kurtosis								
(A) Joint test χ^2 (p)	119.9 (.00)	238.8 (.00)	585.4 (.00)	534.9 (.00)	439.6 (.00)	471.72 (.00)	976.1 (.00)	975.5 (.00)
(B) Joint test χ^2 (p)	1128.1 (.00)	459.4 (.00)	-	-	-	-	1064.8 (.00)	334.6 (.00)

Notes: The values in square brackets are absolute t-statistics of the estimates, based on heteroskedasticity-robust standard errors. * and ** indicate statistical significance at the 5% and 1% level respectively. RMSE reports the root mean squared error.

♮ In the 2FE model we transform all variables as follows: $(z_{it} - \bar{z}_i - \bar{z}_t + \bar{z})$, where

$$\bar{z}_i = T^{-1}\sum_{t=1}^{T} z_{it} \qquad \text{and} \qquad \bar{z}_t = N^{-1}\sum_{i=1}^{N} z_{it} \qquad \text{and} \qquad \bar{z} = (NT)^{-1}\sum_{t=1}^{T}\sum_{i=1}^{N} z_{it}$$

† All variables are in log per worker terms with the exception of the log labour variable: this specification allows for convenient determination of the returns to scale, whereby a positive (negative) significant coefficient on log labour is evidence for increasing (decreasing) returns to scale and an insignificant coefficient implies constant returns to scale.

‡ We apply different $N \times N$ weight matrices to construct the cross-section averages: in [4] averages are constructed for each country i from the values of itself and its neighbours, in [5] we use the inverse of the population-weighted distance between two countries, in [6] the Jaffe correlation coefficient for the 12 agro-climatic zones between two countries. See main text for details.

♭ The implied returns to scale are labeled decreasing (DRS) if the coefficient on labour is negative, statistically significant, and constant (CRS) if this coefficient is insignificant, see previous note. The implied labour coefficient is computed by adding up all the coefficients on the RHS variables (except for labour), subtracting them from unity (the result is the implied labour coefficient if constant returns were to hold) and then adding the coefficient on labour.

♯ The Pesaran (2007) CIPS test results are for the ADF equation with an intercept and the indicated augmentation with lagged first differences. The test results for the ADF equation with intercept and trend follow the same pattern of rejection and are therefor omitted. The CIPS accounts for cross-section dependence in the residuals.

⋆ The D'Agostino et al. (1990) test *(A)* investigates skewness and kurtosis. The Cameron and Trivedi (1990) decomposition test in *(B)* analyses residual heteroskedasticity, skewness and kurtosis. Both tests provide joint hypothesis tests. If a test rejects skewness or kurtosis we can no longer assume normality for the residual distribution. Note that we cannot compute the Cameron and Trivedi (1990) for all CCEP estimators due to the large number of parameters in these models.

Table 3.3: Pooled regressions (CRS imposed)

Pooled regressions dependent variable: [1] & [3]-[6] log output per worker, [2] dto. in 2FE transformation♯, [7] Δlog output per worker, [8] dto. in 2FE transformation♯								
weight matrix‡	[1] **POLS**	[2] **2FE**♯	[3] **CCEP** none	[4] **CCEP** neighbour	[5] **CCEP** distance	[6] **CCEP** agro-climate	[7] **FD-OLS**	[8] **FD-2FE**♯
tractors pw	0.1437	0.0650	0.0989	0.0982	0.0879	0.0787	0.0542	0.0788
Δ in [7]&[8]	[25.96]**	[15.27]**	[5.63]**	[5.35]**	[6.37]**	[4.93]**	[4.00]**	[5.93]**
livestock pw	0.2472	0.4147	0.3869	0.3103	0.4120	0.3994	0.3191	0.3248
Δ in [7]&[8]	[29.59]**	[37.10]**	[7.53]**	[4.49]**	[9.30]**	[8.94]**	[7.84]**	[8.61]**
fertilizer pw	0.1616	0.0647	0.0289	0.0485	0.0333	0.0326	0.0084	0.0100
Δ in [7]&[8]	[26.72]**	[20.94]**	[5.19]**	[5.10]**	[3.39]**	[5.72]**	[2.46]*	[2.97]**
land pw	0.2503	0.3955	0.3326	0.3311	0.4550	0.3198	0.3129	0.3239
Δ in [7]&[8]	[28.68]**	[32.01]**	[3.87]**	[4.74]**	[6.49]**	[4.17]**	[4.68]**	[5.12]**
We included $(T-1)$ year dummies in equations [1] and (in first differences) [7].								
Constant	1.1077		-0.1508	0.2140	0.3488	0.0397		
	[23.50]**		[2.62]**	[1.73]	[6.50]**	[0.65]		
Observations	5,162	5,162	5,162	5,162	5,162	5,162	5,013	5,013
groups	128	128	128	128	128	128	128	128
average T	40.3	40.3	40.3	40.3	40.3	40.3	39.2	39.2
R-squared	0.90	0.64	0.95	0.93	0.98	0.95	0.15	0.11
RMSE	.443	.151	.075	.089	.053	.072	.085	.085
Implications for labour coefficient♭								
Implied β_L	0.1972	0.0601	0.1527	0.2119	0.0118	0.1695	0.3054	0.2625
Arellano-Bond Serial Correlation Test — H_0: no serial correlation in the residuals								
AR(1) (p)	36.67 (.00)	56.81 (.00)	11.17 (.00)	17.96 (.00)	15.46 (.00)	14.10 (.00)	-9.57 (.00)	-9.60 (.00)
AR(2) (p)	36.18 (.00)	49.72 (.00)	4.70 (.00)	12.97 (.00)	7.61 (.00)	6.61 (.00)	-0.57 (.57)	-0.62 (.53)
AR(3) (p)	35.67 (.00)	44.01 (.00)	-0.09 (.93)	8.62 (.00)	1.25 (.21)	1.50 (.13)	0.45 (.65)	0.39 (.69)
Nonstationarity: Pesaran (2007) CIPS test applied to residuals♯ — H_0: residuals are I(1)								
lags	Ztbar (p)	Ztbar (p)	Ztbar (p)	Ztbar (p)	Ztbar (p)	Ztbar (p)	Ztbar (p)	Ztbar (p)
none	-7.77 (.00)	-5.83 (.00)	-34.17 (.00)	-27.74 (.00)	-30.88 (.00)	-30.98 (.00)	-49.10 (.00)	-48.99 (.00)
1 lag	-2.47 (.01)	-0.01 (.50)	-24.10 (.00)	-18.15 (.00)	-21.66 (.00)	-21.41 (.00)	-37.52 (.00)	-37.75 (.00)
2 lags	0.01 (.50)	2.74 (1.00)	-17.44 (.00)	-12.01 (.00)	-16.91 (.00)	-15.48 (.00)	-24.85 (.00)	-23.64 (.00)
3 lags	-0.41 (.34)	3.17 (1.00)	-13.51 (.00)	-9.57 (.00)	-13.04 (.00)	-11.36 (.00)	-15.90 (.00)	-14.73 (.00)
4 lags	0.06 (.52)	3.99 (1.00)	-9.47 (.00)	-4.76 (.00)	-6.54 (.00)	-8.72 (.00)	-7.66 (.00)	-7.84 (.00)
Cross-Section Dependence: Mean (Absolute) Correlation & Pesaran (2004) CD test — H_0: no CSD								
Mean ρ_{ij}	-0.01	0.00	0.00	0.01	0.07	0.00	0.00	0.02
Mean $\lvert\rho_{ij}\rvert$	0.43	0.42	0.19	0.16	0.17	0.17	0.14	0.15
CD (p)	-3.79 (.01)	-0.71 (.48)	1.46 (.14)	5.51 (.00)	38.48 (.00)	-1.28 (.20)	-0.19 (.85)	9.21 (.00)
Further residual diagnostic tests⋆ — H_0: no heteroskedasticity, regular skewness/kurtosis								
(A) Joint test χ^2 (p)	179.4 (.00)	229.3 (.00)	557.5 (.00)	452.1 (.00)	368.4 (.00)	461.7 (.00)	1002.5 (.00)	990.4 (.00)
(B) Joint test χ^2 (p)	1078.9 (.00)	424.5 (.00)	-	-	-	-	967.8 (.00)	268.2 (.00)

Notes: See Table 3.2 for further details.

Regarding the Pesaran (2006) CCEP and our extensions, our results show that the former emphatically rejects CRS and yields summed input elasticities of around .68, with *negative* implied labour coefficient; of the other CCE estimators the distance version behaves in a similar fashion, whereas the other two cannot reject CRS. Serial correlation remains a concern in all four models. Surprisingly the agro-climate version is the only model that cannot reject cross-section independence. Once CRS is (despite previous results) imposed, all models yield somewhat similar coefficients, with the land coefficient again rising considerably — most obviously so in the distance model. Diagnostics remain similar, although residuals in the standard CCEP now cannot reject cross-sectional independence.

In conclusion, our pooled models largely reject constant returns to scale, yield very low values for the implied coefficient on labour and over a range of specification tests indicate a combination of non-normality and heteroskedasticity, cross-section dependence, nonstationarity and/or serial correlation in the residuals. Allowing for heterogeneity in the unobserved common factors (CCEP) although alleviating a potential identification problem does not seem to provide an overall panacea; in fact for the most part the CCEP-type estimators are suggested to continue suffering from cross-section dependence. Our next analytical step is therefore to investigate how the parameter estimates and diagnostic tests change if we allow for technology heterogeneity across countries.

3.4.3 Averaged country regression estimates

We present the results from Mean-Group type estimators in the unrestricted regression model (CRS not imposed) in Table 3.4. For all estimators we present the mean and robust mean across N country parameters — the latter uses weights to reduce the impact of outliers. In practice the median estimates (not reported) are very close to the robust estimates. In our discussion below we focus on the robust means. The t-statistics reported for each average estimate test whether the average parameter is statistically different from zero, following Pesaran et al. (2008).

The MG, as well as our two Augmented Mean Group estimators display strongly decreasing returns, although in the AMG version (i) only at the 10% level of significance. The standard, neighbour and agro-climate CCEMG in contrast have insignificant coefficients on labour, indicating constant returns *in the average country regression*. Around 30 to 40 countries reject CRS at 5% level of significance in each of these models. The distance CCEMG indicates very large and highly significant decreasing returns to scale.

Regarding average parameter estimates on the factor inputs, the MG and AMG estimators yield next to identical results: capital around .07, livestock around .25, fertilizer around .03 and land around .20. In comparison to the parameter differences across models in the *pooled* specifications, the MG and AMG parameter estimates are relatively similar to those in the CCEMG models, with the crucial exception of the returns to scale coefficient and thus the

implied labour elasticity: in the former group of models the labour coefficient is around .10, whereas in the latter group it is closer to .30 (leaving the nonsensical results produced by the distance CCEMG to one side). Within the CCEMG group the standard and the agro-climate estimators yield very similar results.

Turning to the diagnostics, *all* models reject nonstationary errors in the Pesaran (2007) CIPS test. The panel t-statistic following Pedroni (1999) suggests that all coefficients are significant at the 5% level, in contrast to the t-statistic following Pesaran et al. (2008) which we used for our discussion above. Mean absolute error correlation is uniformly low at .15 for all estimators, but cross-section independence is rejected in the MG and AMG(ii), as well as in the alternative CCEMG estimators (marginally in case of the agro-climate CCEMG). For the serial correlation tests we need to take recourse to statistics constructed from country-regression diagnostics. In each case these are Fisher (1932)-type statistics (labelled p_λ), derived from p-values for the test statistics in each country.[13] As can be seen the Durbin 'alternative' and Breusch-Godfrey procedures reject the absence of serially correlated errors, whereas for the Ljung-Watson and Durbin-Watson statistics (see table footnotes) there is some evidence for serially uncorrelated errors in the CCEMG-type estimators. Furthermore, the error normality and homoskedasticity tests (similarly expressed as Fisher-statistics) suggest these properties are rejected in the MG and AMG estimators, and offer conflicting evidence across the CCEMG-type models.

Taking the results for all diagnostic tests into account and recognising the potential problems of the use of time-series tests adjusted for the panel context via computation of the Fisher (1932) statistic, we suggest that evidence in favour of the standard CCEMG and agro-climate CCEMG seems most convincing: compared to the other models these provide some evidence of serially uncorrelated and normal residuals. Since neither of these estimators can reject CRS for the average country production function, we rerun all heterogeneous parameter models with CRS imposed.

[13] Constructed as $p_\lambda = -2\sum_i log(p_i)$, distribution $\chi^2(2N)$ under the null.

Table 3.4: Mean Group type estimators (no restrictions on returns to scale)

Average country regression estimates (means and robust means)
Returns to scale are left unrestricted (CRS, DRS, IRS); N=128, average T=40.3, $n = 5,126$;
dependent variable: [1],[3]-[7] log output per worker, [2] log output per worker less $\hat{\mu}_t^{\bullet}$

	[1] **MG**		[2] **AMG(i)**		[3] **AMG(ii)**		[4] **CCEMG** none		[5] **CCEMG** neighbour		[6] **CCEMG** distance		[7] **CCEMG** agro-climate	
weight matrix†	Mean	Robust	Mean	Robust	Mean	Robust	Mean	Robust	Mean	Robust	Mean	Robust	Mean	Robust
labour	-0.5393 [1.87]	-0.3574 [2.23]*	-0.5438 [1.83]	-0.3039 [1.90]	-0.5438 [1.96]	-0.3664 [2.34]*	-0.2162 [1.38]	-0.1257 [1.04]	-0.309 [1.66]	-0.1814 [1.55]	-0.5309 [4.42]**	-0.5014 [5.85]**	-0.2101 [1.15]	-0.1192 [1.12]
tractors pw	-0.0611 [0.51]	0.0748 [3.31]**	-0.045 [0.40]	0.0797 [3.52]**	-0.0351 [0.36]	0.0721 [3.34]**	0.0489 [1.19]	0.0641 [3.26]**	-0.0301 [0.31]	0.0561 [2.08]*	0.0524 [2.51]*	0.0419 [2.83]**	-0.0088 [0.13]	0.0612 [2.74]**
livestock pw	0.2757 [5.87]**	0.2456 [8.07]**	0.2853 [6.08]**	0.2719 [8.28]**	0.2761 [6.38]**	0.2508 [7.82]**	0.3803 [7.04]**	0.3237 [9.44]**	0.3044 [6.34]**	0.2779 [7.22]**	0.3219 [7.66]**	0.3336 [9.48]**	0.3056 [5.78]**	0.2760 [7.91]**
fertilizer pw	0.0370 [5.10]**	0.0299 [4.86]**	0.0379 [5.10]**	0.0315 [4.91]**	0.0367 [4.99]**	0.0309 [4.64]**	0.0366 [5.39]**	0.0310 [5.02]**	0.0394 [4.97]**	0.0274 [4.59]**	0.0409 [6.15]**	0.0413 [6.74]**	0.0467 [5.85]**	0.0299 [5.26]**
land pw	0.2549 [2.12]*	0.2102 [2.79]**	0.2363 [1.88]	0.1952 [2.60]*	0.2617 [2.06]*	0.1695 [2.27]*	0.0935 [0.82]	0.1999 [2.68]**	-0.0327 [0.27]	0.1623 [2.17]*	0.1235 [1.39]	0.1414 [2.07]*	0.2316 [2.05]*	0.1966 [3.13]**
Common factors			imposed		1.0569 [7.38]**	0.7741 [7.72]**								
Country trend	0.013 [2.69]**	0.0123 [4.82]**	-0.0013 [0.25]	-0.0031 [1.20]	-0.0015 [0.30]	-0.0015 [0.30]								
$\hat{t}_i$ sign. at 10%	78		68		42									
Constant	7.0818 [1.79]	5.366 [2.30]*	7.2221 [1.75]	4.7029 [2.05]*	7.2654 [1.87]	5.6869 [2.51]*	0.9917 [0.38]	0.8964 [0.41]	3.442 [1.49]	2.1018 [1.27]	7.5518 [4.33]**	7.1717 [6.11]**	2.0287 [0.82]	1.6316 [1.09]

Returns to scale implications of the parameter estimates‡

implied β_L	-0.0458	0.0821	-0.0583	0.1178	-0.0832	0.1103	0.2245	0.2556	0.4100	0.2949	-0.0696	-0.0596	0.2148	0.3171
Returns	CRS	DRS	CRS	CRS	CRS	DRS	CRS	CRS	CRS	CRS	DRS	DRS	CRS	CRS
reject CRS	47		47		46		28		36		53		40	

Pedroni (1999) panel t-statistics

$N^{-1/2}\sum_i t_{\hat{\beta}_{L,i}}$	[7.20]**	[6.33]**	[6.83]**	[2.10]*	[2.84]**	[14.04]**	[2.54]**	
$N^{-1/2}\sum_i t_{\hat{\beta}_{tr,i}}$	[6.57]**	[7.11]**	[7.27]**	[8.36]**	[5.41]**	[4.81]**	[7.24]**	
$N^{-1/2}\sum_i t_{\hat{\beta}_{live,i}}$	[21.42]**	[22.22]**	[20.30]**	[21.53]**	[20.09]**	[23.52]**	[18.88]**	
$N^{-1/2}\sum_i t_{\hat{\beta}_{f,i}}$	[9.39]**	[9.22]**	[8.93]**	[7.57]**	[9.49]**	[11.16]**	[9.36]**	
$N^{-1/2}\sum_i t_{\hat{\beta}_{n,i}}$	[8.06]**	[7.36]**	[6.87]**	[3.41]**	[5.11]**	[6.71]**	[7.04]**	
$N^{-1/2}\sum_i t_{\hat{\beta}_{TFP,i}}$			[10.37]**					
$N^{-1/2}\sum_i t_{\hat{t}_i}$	[33.60]**	[28.84]**	[15.70]**					

Serial correlation in the residuals♯ — H_0: no serial correlation (except for DW d)

	p_λ	(p)	p_λ	(p)	p_λ	(p)	p_λ	(p)	p_λ	(p)	p_λ	(p)	p_λ	(p)
Ljung-Box	912.3	(.00)	902.1	(.00)	1005.8	(.00)	47.8	(1.00)	407.0	(.00)	201.3	(1.00)	170.5	(1.00)
Durbin AR(1)	778.9	(.00)	929.0	(.00)	758.0	(.00)	356.6	(.00)	545.3	(.00)	387.6	(.00)	431.5	(.00)
Durbin AR(2)	782.2	(.00)	771.1	(.00)	782.9	(.00)	508.6	(.00)	683.9	(.00)	521.8	(.00)	531.4	(.00)
BGod AR(1)	755.9	(.00)	738.9	(.00)	701.0	(.00)	437.3	(.00)	578.1	(.00)	449.6	(.00)	486.4	(.00)
BGod AR(2)	750.6	(.00)	746.6	(.00)	720.0	(.00)	623.2	(.00)	723.8	(.00)	613.8	(.00)	611.9	(.00)
	< 1.23	> 1.79	< 1.23	> 1.79	< 1.23	> 1.79	< 1.23	> 1.79	< 1.23	> 1.79	< 1.23	> 1.79	< 1.23	> 1.79
DW d statistic #	29	43	26	42	25	47	4	82	13	76	10	70	8	74
(in %)	23%	34%	20%	33%	20%	37%	3%	64%	10%	59%	8%	55%	6%	58%

Nonstationarity: Pesaran (2007) CIPS test applied to residuals — H_0: residuals are I(1)

lags	Ztbar	(p)	Ztbar	(p)	Ztbar	(p)	Ztbar	(p)	Ztbar	(p)	Ztbar	(p)	Ztbar	(p)
none	-35.9	(.00)	-36.6	(.00)	-36.2	(.00)	-43.9	(.00)	-42.0	(.00)	-42.1	(.00)	-44.2	(.00)
1 lag	-27.2	(.00)	-29.2	(.00)	-28.4	(.00)	-33.6	(.00)	-32.3	(.00)	-31.4	(.00)	-35.0	(.00)
2 lags	-20.4	(.00)	-20.1	(.00)	-19.7	(.00)	-23.9	(.00)	-22.5	(.00)	-24.7	(.00)	-25.8	(.00)
3 lags	-17.2	(.00)	-16.2	(.00)	-16.4	(.00)	-18.8	(.00)	-18.9	(.00)	-19.6	(.00)	-21.2	(.00)
4 lags	-10.2	(.00)	-9.7	(.00)	-9.4	(.00)	-14.1	(.00)	-12.3	(.00)	-13.5	(.00)	-16.8	(.00)

Cross-section Dependence: Mean (absolute) correlations and Pesaran (2004) test — H_0: no CSD in the residuals

	ρ_{ij}	$\lvert\rho_{ij}\rvert$	ρ_{ij}	$\lvert\rho_{ij}\rvert$	ρ_{ij}	$\lvert\rho_{ij}\rvert$	ρ_{ij}	$\lvert\rho_{ij}\rvert$	ρ_{ij}	$\lvert\rho_{ij}\rvert$	ρ_{ij}	$\lvert\rho_{ij}\rvert$	ρ_{ij}	$\lvert\rho_{ij}\rvert$
Mean Correl.	0.02	0.15	0.00	0.15	0.00	0.15	0.00	0.15	0.00	0.14	0.07	0.15	0.00	0.14
CD statistic (p)	9.16	(.00)	-1.12	(.26)	-2.47	(.02)	0.21	(.83)	2.24	(.03)	37.16	(.00)	-1.72	(.09)

Further residual diagnostic tests⋆ — H_0: no heteroskedasticity, regular skewness/kurtosis

d'Agostini et al	p_λ	(p)	p_λ	(p)	p_λ	(p)	p_λ	(p)	p_λ	(p)	p_λ	(p)	p_λ	(p)
Fisher stat (joint)	379.2	(.00)	343.4	(.00)	355.4	(.00)	318.8	(.00)	299.8	(.01)	258.3	(.45)	308.3	(.03)
Cam & Tri	p_λ	(p)	p_λ	(p)	p_λ	(p)	p_λ	(p)	p_λ	(p)	p_λ	(p)	p_λ	(p)
Fisher stat (joint)	374.1	(.00)	368.7	(.00)	293.8	(.05)	245.2	(.68)	236.4	(.80)	233.5	(.84)	256.1	(.49)
homosk.	317.6	(.00)	314.9	(.01)	253.3	(.54)	218.8	(.96)	228.2	(.89)	218.8	(.96)	218.8	(.96)
skewness	374.0	(.00)	367.7	(.00)	342.2	(.00)	300.7	(.03)	259.6	(.43)	275.3	(.19)	324.6	(.00)
kurtosis	264.2	(.35)	254.4	(.52)	256.7	(.48)	242.5	(.72)	256.7	(.48)	248.4	(.62)	252.2	(.55)
RMSE	.066		.065		.066		.057		.056		.038		.054	

Notes: All variables are in logs. * and ** indicate statistical significance at the 5% and 1% level respectively. We present the sample mean and the robust sample mean estimates of all model parameters. Terms in brackets are absolute t-statistics (H_0: $N^{-1}\sum_i \hat{\beta}_i = 0$) following Pesaran et al. (2008). An alternative panel t-test by Pedroni (1999) is also provided. The 'common factors' variable refers to $\hat{\mu}_t^{\bullet}$. RMSE reports the root mean squared errors.
† Weight matrix: we use the same approach to construct cross-section averages as in the pooled regressions.
‡ We compute implied labour coefficients and report returns to scale (DRS, CRS are decreasing and constant returns respectively).
♯ The Ljung-Box, Durbin 'alternative' and Breusch-Godfrey procedures test first- and higher-order serial correlation in time-series regression residuals. We convert the p-statistics of the N results into a Fisher-statistic p_λ which is distributed $\chi^2(2N)$. For the Durbin-Watson d statistic we report the number of time-series tests with a d statistic larger than 1.79 (smaller than 1.23) which is deemed to (be unable to) reject the null of first order serial correlation (under the strong assumption of exogenous regressors).
For all other test statistics please refer to definitions in the footnotes of Table 3.2.

Table 3.5: Mean Group type estimators (CRS imposed)

Average country regression estimates (means and robust means)
CRS imposed for all country regressions; N=128, average T=40.3;
dependent variable: [1],[3]-[7] log output per worker, [2] log output per worker less $\hat{\mu}_t^{\bullet}$

weight matrix†	[1] **MG**		[2] **AMG(i)**		[3] **AMG(ii)**		[4] **CCEMG** none		[5] **CCEMG** neighbour		[6] **CCEMG** distance		[7] **CCEMG** agro-climate	
	Mean	Robust	Mean	Robust	Mean	Robust	Mean	Robust	Mean	Robust	Mean	Robust	Mean	Robust
tractors pw	0.0700	0.0814	0.0887	0.0903	0.0780	0.0842	0.0879	0.1093	0.0673	0.0972	0.1317	0.1078	0.1247	0.0977
	[2.08]*	[3.57]**	[2.57]*	[3.71]**	[2.27]*	[3.65]**	[2.50]*	[5.23]**	[1.72]	[4.27]**	[4.67]**	[5.59]**	[3.88]**	[4.48]**
livestock pw	0.3051	0.2917	0.3106	0.3075	0.3099	0.2905	0.3424	0.3214	0.3231	0.3124	0.3726	0.3682	0.3257	0.3298
	[7.80]**	[8.82]**	[7.89]**	[8.90]**	[8.40]**	[8.66]**	[8.33]**	[9.47]**	[7.53]**	[8.44]**	[11.29]**	[11.52]**	[8.02]**	[9.53]**
fertilizer pw	0.0373	0.0328	0.0414	0.0378	0.0376	0.0348	0.0435	0.0358	0.0437	0.0376	0.0393	0.0365	0.0501	0.0352
	[5.29]**	[5.17]**	[5.52]**	[5.58]**	[5.32]**	[5.31]**	[5.76]**	[5.63]**	[5.34]**	[5.26]**	[5.25]**	[5.15]**	[6.83]**	[5.87]**
land pw	0.2884	0.2636	0.2653	0.2234	0.2811	0.1919	0.2104	0.2008	0.2364	0.2623	0.3478	0.2905	0.2465	0.2139
	[4.86]**	[4.55]**	[4.32]**	[3.78]**	[4.30]**	[3.34]**	[3.57]**	[3.57]**	[3.41]**	[4.71]**	[6.25]**	[5.62]**	[4.27]**	[4.02]**
Common factors			imposed		1.0590	0.7808								
					[6.56]**	[6.77]**								
Country trend	0.0100	0.0107	-0.0029	-0.0019	-0.0035	-0.0035								
	[6.18]**	[7.86]**	[1.78]	[1.38]	[1.43]	[1.43]								
$\hat{t}_i$ sign. at 10%	82		65		37									
Constant	-0.3487	-0.2095	-0.2145	-0.1052	-0.2871	-0.1604	-0.2011	0.2874	-0.0917	-0.0603	0.0431	0.1377	-0.0580	-0.0582
	[1.49]	[1.11]	[0.89]	[0.55]	[1.19]	[0.90]	[0.44]	[1.02]	[0.38]	[0.39]	[0.13]	[0.95]	[0.26]	[0.33]
Returns to scale implications of the parameter estimates‡														
implied β_L	0.2992	0.3305	0.2940	0.3410	0.2934	0.3986	0.3158	0.3327	0.3295	0.2905	0.1086	0.1970	0.2530	0.3234
Pedroni (1999) panel *t*-statistics														
$N^{-1/2}\sum_i t_{\hat{\beta}_{tr,i}}$	[11.17]**		[13.16]**		[11.35]**		[13.46]**		[9.15]**		[12.91]**		[12.01]**	
$N^{-1/2}\sum_i t_{\hat{\beta}_{live,i}}$	[24.88]**		[26.03]**		[24.16]**		[24.68]**		[25.75]**		[28.7]**		[22.35]**	
$N^{-1/2}\sum_i t_{\hat{\beta}_{f,i}}$	[10.34]**		[10.94]**		[9.27]**		[9.83]**		[11.33]**		[9.85]**		[11.00]**	
$N^{-1/2}\sum_i t_{\hat{\beta}_{n,i}}$	[13.45]**		[11.83]**		[11.28]**		[9.66]**		[10.93]**		[14.22]**		[9.61]**	
$N^{-1/2}\sum_i t_{\hat{\beta}_{TFP,i}}$					[11.85]**									
$N^{-1/2}\sum_i t_{\hat{t}_i}$	[38.89]**		[33.02]**		[14.83]**									
Serial correlation in the residuals — H_0: no serial correlation														
	p_λ	(p)	p_λ	(p)	p_λ	(p)	p_λ	(p)	p_λ	(p)	p_λ	(p)	p_λ	(p)
Ljung-Box	1579.7	(.00)	1542.7	(.00)	1426.8	(.00)	380.2	(.00)	830.6	(.00)	443.5	(.00)	312.2	(.01)
Durbin AR(1)	1017.6	(.00)	1162.2	(.00)	843.5	(.00)	455.6	(.00)	715.8	(.00)	471.0	(.00)	484.1	(.00)
Durbin AR(2)	964.1	(.00)	952.3	(.00)	812.4	(.00)	502.2	(.00)	757.3	(.00)	557.3	(.00)	560.6	(.00)
BGod AR(1)	914.5	(.00)	908.5	(.00)	780.2	(.00)	513.5	(.00)	663.3	(.00)	523.3	(.00)	534.0	(.00)
BGod AR(2)	853.6	(.00)	855.3	(.00)	751.7	(.00)	573.5	(.00)	699.2	(.00)	628.1	(.00)	631.6	(.00)
	< 1.23	> 1.79	< 1.23	> 1.79	< 1.23	> 1.79	< 1.23	> 1.79	< 1.23	> 1.79	< 1.23	> 1.79	< 1.23	> 1.79
DW d statistic #	42	33	42	32	35	41	18	63	21	58	19	62	17	61
(in %)	33%	26%	33%	25%	27%	32%	14%	49%	16%	45%	15%	48%	13%	48%
Nonstationarity: Pesaran (2007) CIPS test applied to residuals — H_0: residuals are I(1)														
lag-augmentation	Ztbar	(p)	Ztbar	(p)	Ztbar	(p)	Ztbar	(p)	Ztbar	(p)	Ztbar	(p)	Ztbar	(p)
none	-33.57	(.00)	-34.67	(.00)	-34.32	(.00)	-44.26	(.00)	-42.06	(.00)	-41.62	(.00)	-43.34	(.00)
1 lag	-24.56	(.00)	-27.02	(.00)	-25.55	(.00)	-35.24	(.00)	-31.87	(.00)	-31.33	(.00)	-34.59	(.00)
2 lags	-18.29	(.00)	-19.13	(.00)	-17.87	(.00)	-25.55	(.00)	-23.61	(.00)	-24.79	(.00)	-25.46	(.00)
3 lags	-15.74	(.00)	-16.33	(.00)	-15.89	(.00)	-21.35	(.00)	-21.24	(.00)	-19.31	(.00)	-21.65	(.00)
4 lags	-9.45	(.00)	-11.25	(.00)	-9.12	(.00)	-16.54	(.00)	-13.57	(.00)	-12.28	(.00)	-17.13	(.00)
Cross-section Dependence: Mean correlations and Pesaran (2004) test — H_0: no CSD in the residuals														
	ρ_{ij}	$\lvert\rho_{ij}\rvert$	ρ_{ij}	$\lvert\rho_{ij}\rvert$	ρ_{ij}	$\lvert\rho_{ij}\rvert$	ρ_{ij}	$\lvert\rho_{ij}\rvert$	ρ_{ij}	$\lvert\rho_{ij}\rvert$	ρ_{ij}	$\lvert\rho_{ij}\rvert$	ρ_{ij}	$\lvert\rho_{ij}\rvert$
Mean Correl.	0.02	0.15	0.00	0.15	0.00	0.15	0.00	0.15	0.00	0.14	0.06	0.15	0.00	0.14
CD statistic (p)	8.98	(.00)	-0.08	(.93)	-1.56	(.12)	-0.23	(.82)	1.69	(.05)	31.34	(.00)	-1.51	(.13)
Further residual diagnostic tests⋆ — H_0: no heteroskedasticity, regular skewness/kurtosis														
d'Agostini et al	p_λ	(p)	p_λ	(p)	p_λ	(p)	p_λ	(p)	p_λ	(p)	p_λ	(p)	p_λ	(p)
Fisher stat (joint)	383.1	(.00)	344.2	(.00)	343.2	(.00)	337.0	(.00)	334.6	(.00)	263.8	(.36)	336.2	(.00)
Cam & Tri	p_λ	(p)	p_λ	(p)	p_λ	(p)	p_λ	(p)	p_λ	(p)	p_λ	(p)	p_λ	(p)
Fisher stat (joint)	458.5	(.00)	447.7	(.00)	355.5	(.00)	259.5	(.43)	260.7	(.41)	247.1	(.63)	248.2	(.62)
homosk.	392.7	(.00)	385.7	(.00)	310.6	(.01)	218.8	(.96)	221.2	(.94)	218.8	(.96)	218.8	(.96)
skewness	396.9	(.00)	393.5	(.00)	356.4	(.00)	337.8	(.00)	324.2	(.00)	311.0	(.01)	305.6	(.02)
kurtosis	262.3	(.38)	245.4	(.67)	251.2	(.57)	247.9	(.63)	294.5	(.05)	254.3	(.52)	278.3	(.16)
RMSE	.070		.069		.069		.059		.060		.041		.057	

Notes: See Table 3.4 for details.

The CRS country regression averages in Table 3.5 produce very similar results across *all* estimators: it can be seen that the imposition of CRS in heterogeneous regression models does not create the same dramatic changes to the estimates as in the pooled estimator case. It can also be seen that the individual country-estimates for each model now contain fewer outliers, as is evidenced by the precision of the mean estimates and their proximity in parameter value to the robust mean estimates. We focus on the models for which on average

we had not rejected CRS: AMG(i), standard, neighbour and agro-climate CCEMG. Factor elasticities in the AMG(i) are close to the unrestricted model ones, with the obvious exception of implied labour, which with CRS imposed is around .33 rather than .12. The standard and agro-climate CCEMG now match each other's (robust) mean parameters even closer than in the unrestricted case, while the neighbour CCEMG in comparison has an inflated land coefficient. Panel t-statistics are significant at the 1% level for all parameters in all models and CIPS tests continue to reject nonstationarity in the residuals. AMG(i) as well as standard and agro-climate CCEMG residuals appear cross-sectionally independent. In comparison to the unrestricted versions, the models now however seem *more* likely to be subject to serially correlated errors, as evidenced by the fall in the share of countries with a Durbin-Watson statistic above 1.79, as well as the Ljung-Box and other test statistics. The CCEMG now all 'pass' the Cameron and Trivedi (1990) normality/homoskedasticity test, while all models continue to fail the D'Agostino et al. (1990) normality test.

In summary, while parameter estimates for the four models which could not reject CRS did not change considerably (in magnitude) compared to the unrestricted models, their precision has improved while some diagnostic test results have deteriorated. Given that in any finite sample the cointegrating relationship may be affected by short-run dynamics, and furthermore on the back of the serial correlation test results we carry out the country-regressions in a dynamic specification, focusing on long-run coefficients in the MG (as a benchmark) and the four CCEMG estimators. Table 3.6 presents these results.

It can be seen that the robust mean long-run parameter estimates[14] are again virtually identical between the standard and agro-climate CCEMG in columns [2] and [5]. Furthermore, there is only one small deviation from the respective parameter estimates in the static models, namely the increase in the mean fertilizer elasticities. According to the Pedroni (1999) panel t-statistics all parameters are significant at the 1% level in these two models.

[14]These long-run parameters are derived from an error correction model regression, in the MG case

$$\Delta y_{it} = -(1-r_i)\, y_{i,t-1} + (1-r_i)\, a_i + (1-r_i)\, c_i t + (1-r_i)\boldsymbol{b}_i\, \boldsymbol{x}_{i,t-1} + \boldsymbol{b}_i\, \boldsymbol{\Delta x}_{it}$$

where in practice the common factor restrictions indicated are not imposed. $\boldsymbol{x}$ and $\boldsymbol{\Delta x}$ represent all observed regressors in levels and first differences respectively; a_i and c_i capture country-specific TFP level and constant growth rate. Long-run coefficients are then computed as the negative difference of the coefficient on the lagged $\boldsymbol{x}$-variables in levels and the lagged dependent variable.
For the CCEMG regressions we include cross-section averages of the dependent variable ($\Delta\bar{y}_t$), the lagged level of the dependent variable ($\bar{y}_{t-1}$), the set of $\boldsymbol{x}$-variables in lagged levels ($\bar{\boldsymbol{x}}_{t-1}$) and in first difference ($\boldsymbol{\Delta}\bar{\boldsymbol{x}}_t$). No trend is included here and the long-run coefficients are computed as in the MG case.

Table 3.6: Dynamic Specification — MG-type estimators (CRS imposed)

Average country long-run coefficient estimates (means and robust means) CRS imposed for all country regressions; N=128, average T=39.2; dependent variable: growth of output per worker										
	[1]		[2]		[3]		[4]		[5]	
	MG		**CCEMG**		**CCEMG**		**CCEMG**		**CCEMG**	
weight matrix†			none		neighbour		distance		agro-climate	
	Mean	Robust	Mean	Robust	Mean	Robust	Mean	Robust	Mean	Robust
long-run coefficients										
tractors pw	0.346	0.043	0.099	0.088	0.125	0.019	-2.780	-0.116	0.121	0.085
	[1.19]	[1.50]	[1.37]	[3.18]**	[0.24]	[0.14]	[0.78]	[1.39]	[1.67]	[2.85]**
livestock pw	0.405	0.288	0.304	0.307	1.203	-0.381	-5.298	-0.597	0.316	0.327
	[3.33]**	[6.97]**	[3.71]**	[6.82]**	[0.51]	[1.75]	[1.66]	[3.43]**	[3.46]**	[6.67]**
fertilizer pw	0.123	0.046	0.070	0.054	0.576	-0.016	-0.546	-0.015	0.081	0.077
	[1.67]	[4.50]**	[5.41]**	[4.84]**	[0.86]	[0.39]	[0.84]	[0.48]	[6.59]**	[6.42]**
land pw	-4.940	0.188	0.244	0.215	0.623	-0.055	-3.026	0.007	0.102	0.191
	[0.95]	[2.50]*	[2.42]*	[2.73]**	[0.49]	[0.23]	[0.64]	[0.03]	[0.72]	[2.33]*
trend	0.007	0.006								
	[4.00]**	[4.96]**								
$\hat{t}_i$ sign. at 10%	58									
Constant	-0.476	-0.098	-1.242	0.056	-0.053	0.061	0.151	0.002	0.030	0.020
	[1.76]	[0.71]	[2.00]*	[0.16]	[0.15]	[0.34]	[0.65]	[0.01]	[0.09]	[0.10]
Returns to scale implications of the parameter estimates‡										
implied β_L	5.066	0.436	0.284	0.335	-1.526	1.433	12.649	1.721	0.381	0.321
Serial correlation in the residuals — H_0: no serial correlation										
	p_λ	(p)	p_λ	(p)	p_λ	(p)	p_λ	(p)	p_λ	(p)
Ljung-Box	72.9	(1.00)	0.6	(1.00)	21.0	(1.00)	2.1	(1.00)	0.9	(1.00)
Durbin AR(1)	299.3	(.03)	332.4	(.00)	487.5	(.00)	488.3	(.00)	493.2	(.00)
Durbin AR(2)	370.5	(.00)	525.8	(.00)	678.2	(.00)	741.9	(.00)	775.2	(.00)
BGod AR(1)	366.7	(.00)	573.0	(.00)	781.9	(.00)	715.4	(.00)	775.2	(.00)
BGod AR(2)	465.6	(.00)	902.7	(.00)	1072.3	(.00)	1158.5	(.00)	1130.6	(.00)
	< 1.23	> 1.79	< 1.23	> 1.79	< 1.23	> 1.79	< 1.23	> 1.79	< 1.23	> 1.79
DW d statistic #	0	111	0	115	0	124	0	124	0	124
(in %)	0%	87%	0%	90%	0%	97%	0%	97%	0%	97%
Pedroni (1999) panel t-statistics										
$N^{-1/2}\sum_i t_{\hat{\beta}_{tr,i}}$	[6.18]**		[7.72]**		[0.58]		[2.01]*		[5.94]**	
$N^{-1/2}\sum_i t_{\hat{\beta}_{live,i}}$	[15.42]**		[12.06]**		[1.03]		[1.46]		[15.91]**	
$N^{-1/2}\sum_i t_{\hat{\beta}_{f,i}}$	[8.20]**		[7.40]**		[1.43]		[0.97]		[8.98]**	
$N^{-1/2}\sum_i t_{\hat{\beta}_{n,i}}$	[6.48]**		[5.74]**		[1.81]		[0.70]		[6.25]**	
$N^{-1/2}\sum_i t_{\hat{t}_i}$	[19.50]**									
Nonstationarity: Pesaran (2007) CIPS test applied to residuals — H_0: residuals are I(1)										
lag-augmentation	Ztbar	(p)	Ztbar	(p)	Ztbar	(p)	Ztbar	(p)	Ztbar	(p)
none	-48.9	(.00)	-50.0	(.00)	-51.1	(.00)	-50.8	(.00)	-50.8	(.00)
1 lag	-32.1	(.00)	-38.9	(.00)	-38.8	(.00)	-38.8	(.00)	-40.8	(.00)
2 lags	-24.6	(.00)	-30.5	(.00)	-29.2	(.00)	-28.0	(.00)	-31.7	(.00)
3 lags	-20.3	(.00)	-25.2	(.00)	-22.3	(.00)	-22.1	(.00)	-25.8	(.00)
4 lags	-11.5	(.00)	-15.1	(.00)	-14.0	(.00)	-13.6	(.00)	-15.5	(.00)
CSD: Mean correlations and Pesaran (2004) test — H_0: no CSD in the residuals										
	ρ_{ij}	$\lvert\rho_{ij}\rvert$	ρ_{ij}	$\lvert\rho_{ij}\rvert$	ρ_{ij}	$\lvert\rho_{ij}\rvert$	ρ_{ij}	$\lvert\rho_{ij}\rvert$	ρ_{ij}	$\lvert\rho_{ij}\rvert$
Mean Correl.	0.02	0.14	0.00	0.16	0.00	0.14	0.07	0.15	0.00	0.15
CD statistic (p)	7.07	(.00)	-0.22	(.83)	2.57	(.01)	38.31	(.00)	-0.42	(.67)
Further residual diagnostic tests⋆ — H_0: no heteroskedasticity, regular skewness/kurtosis										
d'Agostini et al	p_λ	(p)	p_λ	(p)	p_λ	(p)	p_λ	(p)	p_λ	(p)
Fisher stat (joint)	402.2	(.00)	336.7	(.00)	347.4	(.00)	303.3	(.02)	280.1	(.14)
Cameron & Trivedi	p_λ	(p)	p_λ	(p)	p_λ	(p)	p_λ	(p)	p_λ	(p)
Fisher stat (joint)	246.4	(.66)	199.7	(.99)	194.3	(.99)	210.6	(.98)	248.7	(.62)
RMSE	.055		.044		.041		.029		.039	

Notes: The values in square brackets are absolute t-statistics of the estimates, based on heteroskedasticity-robust standard errors (Pesaran et al., 2008). An alternative 'panel t-test' by Pedroni (1999) is also provided — both are with reference to the long-run coefficients computed for each country regression. * and ** indicate statistical significance at the 5% and 1% level respectively. Long-run coefficients are computed from the lagged levels estimates using the Delta method.
‡ We compute implied (average) labour coefficients by subtracting the averaged long-run coefficients on the factor inputs in this model from unity (given that CRS is imposed).
For the diagnostic tests refer to Table 3.4 for more details. Residuals tested are those from the ECM regressions for each country. RMSE reports the root mean squared errors.

The results for the MG estimator in contrast show considerable differences and crucially cannot arrive at a statistically significant capital estimate. Both the neighbour and distance CCEMG collapse in the dynamic specification and yield largely insignificant long-run coefficients — in this case the panel t-statistics support this insignificance emphatically. We interpret this breakdown in the shift from static to dynamic specification as a strong indication for the invalidity of the weighting strategy adopted. This suggests that neither neighbourhood nor geographical distance drive the impact of the unobserved common factors in this data.

Comparing diagnostic test results across the dynamic models and with their static versions in Table 3.5 we can see that while parameter estimates have not changed a great deal for our two preferred estimators (standard and agro-climate CCEMG), the serial correlation tests now provide more reassuring evidence of serially uncorrelated errors. Furthermore, the agro-climate CCEMG now 'passes' both tests for homoskedasticity and normality. Residuals from both estimators continue to show no evidence of nonstationarity and/or cross-section dependence — the latter in contrast to all other models presented. We provide an interpretation of the similarity between standard and agri-climate CCEMG in the next section.

We also estimated the *pooled* version of the dynamic specification with CRS imposed. Results presented in Table 3.19 in Appendix 3.D are still broadly in line with those from the *static* pooled CRS models. Focusing on the 2FE as the commonly preferred estimator in the literature, the dynamic equation yields very similar long-run coefficients for capital ($\approx$.07), livestock ($\approx$.40) and fertilizer ($\approx$.08), with the land elasticity dropping from .40 in the static to .31 in the dynamic equation. As a result the implied labour coefficient rises from .06 to .13. Crucially, however, the dynamic 2FE rejects cross-section independence, normality/homoskedasticity and the null of no serial correlation in the residuals, thus not improving on the diagnostic results of the static model.

3.5 Interpretation and Conclusions

In this chapter we investigated the impact of variable nonstationarity, technology heterogeneity, cross-section dependence and imposition of returns to scale on the results for cross-country production functions in agriculture for a large sample of countries. Our review of the literature

indicated that the empirical implementations of the intercountry production functions for agriculture are dominated by the two-way fixed effects (2FE) model, with constant returns to scale often imposed without formal testing. Variable and residual analysis indicates that all of the aforementioned properties may be present in the data for 128 countries in our empirical analysis. Our estimation results display considerable differences in the estimated parameters when moving from pooled models to averaged country regressions, and between equations ignoring and accounting for cross-section dependence, and we use diagnostic testing to determine the most favourable specification and estimator(s).

We draw the following conclusions from our empirical analysis:

(i) Parameter heterogeneity plays an important role in cross-country productivity investigation for agriculture. The use of pooled models is largely rejected by the diagnostic tests we apply. These tests further indicate that the standard pooled estimator applied in the literature (2FE) is seriously misspecified.

(ii) The imposition of constant returns to agricultural production (CRS) on pooled regression equations — e.g. in the commonly preferred two-way fixed effects model — is rejected by the data and further leads to qualitatively different empirical results. In contrast, the imposition of CRS on individual country regressions does not change average parameter coefficients considerably compared with the unrestricted results: thus the observed decreasing returns in the pooled models is an artefact of empirical misspecification.

(iii) The presence of unobserved common factors with heterogeneous factor loadings *does* matter, as is highlighted by diagnostic tests comparing the MG and CCEMG-type estimators: the standard CCEMG estimator yields very favourable test statistics and shows robustness across various specifications (static, dynamic; unrestricted, CRS imposed). As was suggested a number of times in this thesis the estimators which do not account for cross-section correlation are unable to identify the technology parameters $\boldsymbol{\beta}$ or mean of $\boldsymbol{\beta}_i$ separately from the impact of the factor loadings $\boldsymbol{\lambda}_i$ if the same common factors drive both inputs and output.

(iv) Our own AMG estimator yields results much closer to the standard MG than the Pesaran (2006) CCEMG estimator, often failing specification tests in a similar fashion to the MG — Monte Carlo simulations pressented in Chapter 4 will indicate a set of circumstances in which the AMG yields biased estimates but the CCEMG remains unbiased. We would

speculate that the explicit, common TFP estimate $\hat{\mu}_t^{\bullet}$ applied in the AMG estimation may be unable to capture the heterogeneity in the underlying factor structure of the agriculture data: while common TFP across countries prevails in manufacturing, as was shown in Chapter 2, it is not a sound concept to represent production technology in agriculture, where heterogeneous factors play a much more prominent role (see next point).

(v) Our extension to the CCEMG estimator of applying (exogenous) weight matrices before computing the cross-section averages and in effect imposing more structure on the nature of cross-section correlation in the data has provided very interesting insights. We suggest that given the close match between the standard and agro-climate CCEMG that the common correlated effects driving the cross-section dependence are closely proxied by our measure for agro-climatic distance. In contrast, neighbourhood or geographical distance measures do not seem to play a major role. The implication of this finding is that agricultural TFP is affected by different factors and has different levels of responsiveness across geographic regions of the world due to agro-climatic diversity. Furthermore, technology transfer between countries is limited by the adaptiveness of technology to the local environment — both of these statements are widely accepted in the literature but prior to this study were ignored in empirical analyses of cross-country productivity.

(vi) While the preferred standard and agro-climate CCEMG estimators yielded very similar results across static and dynamic specifications, the other heterogeneous parameter estimators collapse to a larger or smaller extent when we introduce a dynamic specification. In our mind this is strong evidence for the robustness of the standard and agro-climate CCEMG.

Appendix to Chapter 3

3.A Data construction and descriptives

The principal data source for our empirical analysis is the Food and Agriculture Organisation's *FAOSTAT* panel database (FAO, 2007), from which we obtain annual observations for agricultural net output, economically active labour force in agriculture, number of tractors used in agriculture, arable land and permanent crop land and fertilizer use in 128 countries from 1961 to 2002. The specific sources and definitions are detailed below. Descriptive statistics for all variables are presented in Table 3.7. The countries in our sample (ordered by country iso-code) and the number of country observations for the regressions in levels and first differences respectively are presented in Table 3.8. The total number of observations is 5,162 in the levels and 5,013 in the first difference specification, minimum T is 19 and 17 respectively, and average T 40.3 and 39.2 respectively.

Real agricultural net output (in thousand International $), taken from *ProdSTAT* is based on all crops and livestock products originating in each country: "practically all products are covered, with the main exception of fodder crops." (FAO, 2007). Intermediate primary inputs of agricultural origin are deducted, including fodder and seed. The quantities for each commodity are weighted by the respective 1999-2001 average international commodity prices and then summed for each year by country. The prices are in international dollars, which are derived using a Geary-Khamis formula for the agricultural sector. The labour variable is taken from *PopSTAT* and represents the annual time series for **total economically active population in agriculture**. For capital stock in agriculture we follow a common convention and use **total number of agricultural tractors in use** as a proxy. This variable can be found in the 'Machinery' section of *ResourceSTAT*. The **livestock** variable is constructed from the data for asses (donkeys), buffalos, camels, cattle, chickens, ducks, horses, mules, pigs, sheep & goats and turkeys in the 'Live animals' section of *ProdStat*. Following convention we use the below formula to convert the numbers for individual animal species into the livestock variable:

$$\begin{aligned}\text{livestock} = {} & 1.1^{*}\text{camels} + \text{buffalos} + \text{horses} + \text{mules} + 0.8^{*}\text{cattle} + 0.8^{*}\text{asses} \\ & + 0.2^{*}\text{pigs} + 0.1^{*}(\text{sheep}+\text{goats}) + 0.01^{*}(\text{chickens}+\text{ducks}+\text{turkeys})\end{aligned}$$

The fertilizer variable is taken from the 'Fertilizers archive' of *ResourceSTAT* and represents **agricultural fertilizer consumed in metric tons**, which includes 'crude' and 'manufactured' fertilizers. The land variable is taken from *ResourceSTAT* and represents **arable and permanent crop land (in 1000 hectare)**.

Additional time-invariant data on geographical distance between countries and contiguity (neighbourhood) is taken from CEPII (2006), and data on the share of agricultural land by climatic zone from Matthews (1983), available in Gallup et al. (1999). For a number of diagnostics we group the data by region or climate zone — Table 3.9 details these groupings using country iso-codes.

Finally, Table 3.10 provides an overview of the climatic zones referred to in the Köppen-Geiger classification, taken from Kottek, Grieser, Beck, Rudolf, and Rubel (2006).

Table 3.7: Descriptive statistics

Variables in untransformed level terms					
Variable	**mean**	**median**	**std. dev.**	**min.**	**max.**
levels					
output	6,659,293	1,523,753	19,900,000	3,196	317,000,000
labour	8,528,396	1,317,000	42,400,000	3,000	511,000,000
tractors	149,528	7,124	512,278	2	5,470,000
livestock	11,800,000	3,031,056	31,900,000	6,644	298,000,000
fertilizer	721,176	58,000	2,680,786	5	39,600,000
land	9,799,731	2,628,500	25,500,000	1,000	191,000,000
logs					
output	14.24	14.24	1.71	8.07	19.57
labour	14.01	14.09	1.84	8.01	20.05
tractors	9.01	8.87	2.79	0.69	15.51
livestock	14.90	14.92	1.71	8.80	19.51
fertilizer	10.82	10.97	2.69	1.61	17.49
land	14.69	14.78	1.80	6.91	19.07
annual growth rate					
output	2.3%	2.4%	8.8%	-83.0%	87.6%
labour	0.3%	0.8%	2.6%	-28.8%	28.8%
tractors	4.4%	2.0%	9.9%	-121.8%	138.6%
livestock	1.4%	1.6%	6.4%	-93.3%	182.9%
fertilizer	5.6%	3.5%	40.1%	-626.3%	393.2%
land	0.8%	0.1%	3.6%	-41.8%	79.0%

Variables in per worker terms					
variable	**mean**	**median**	**std. dev.**	**min.**	**max.**
levels					
output	3.86	0.97	7.40	0.11	54.85
tractors	0.13	0.01	0.29	0.00	1.97
livestock	6.60	2.24	12.76	0.06	102.28
fertilizer	0.43	0.06	0.94	0.00	7.05
land	4.88	1.96	12.87	0.11	140.47
logs					
output	0.23	-0.03	1.42	-2.22	4.00
tractors	-5.00	-4.97	3.01	-13.67	0.68
livestock	0.89	0.81	1.38	-2.77	4.63
fertilizer	-3.19	-2.87	2.67	-11.56	1.95
land	0.68	0.67	1.15	-2.20	4.95
annual growth rate					
output	2.0%	2.0%	9.0%	-80.3%	109.9%
tractors	4.1%	2.1%	10.1%	-120.2%	136.5%
livestock	1.2%	1.2%	6.6%	-93.5%	182.9%
fertilizer	5.4%	4.2%	40.0%	-627.8%	390.8%
land	0.5%	0.0%	4.1%	-43.0%	81.6%

Notes: We report the descriptive statistics for output (in I$1,000), labour (headcount), tractors (number), livestock (cattle-equivalent numbers), fertilizer (in metric tonnes) and land (in hectare) for the full regression sample ($n = 5,162$; $N = 128$).

Table 3.8: Sample of countries and number of observations

Country	Code	levels	FD		Country	Code	levels	FD
Afghanistan	AFG	40	38		Cambodia	KHM	33	30
Angola	AGO	40	39		South Korea	KOR	42	41
Albania	ALB	42	41		Kuwait	KWT	24	22
United Arab Emirates	ARE	31	30		Lao PDR	LAO	38	35
Argentina	ARG	42	41		Lebanon	LBN	42	41
Australia	AUS	42	41		Liberia	LBR	30	29
Austria	AUT	42	41		Libya	LBY	42	41
Burundi	BDI	37	36		Sri Lanka	LKA	42	41
Benin	BEN	42	41		Lesotho	LSO	42	41
Burkina Faso	BFA	42	41		Morocco	MAR	42	41
Bangladesh	BGD	42	41		Madagascar	MDG	42	41
Bulgaria	BGR	42	41		Mexico	MEX	42	41
Belgium-Luxembourg	BLX	39	38		Mali	MLI	42	41
Belize	BLZ	42	41		Myanmar	MMR	42	41
Bolivia	BOL	42	41		Mongolia	MNG	34	33
Brazil	BRA	42	41		Mozambique	MOZ	42	41
Botswana	BWA	42	41		Mauritania	MRT	33	29
Central African Republic	CAF	42	41		Malawi	MWI	42	41
Canada	CAN	42	41		Malaysia	MYS	42	41
Switzerland	CHE	42	41		Niger	NER	34	33
Chile	CHL	42	41		Nigeria	NGA	42	41
China	CHN	42	41		Nicaragua	NIC	42	41
Côte d'Ivoire	CIV	42	41		Netherlands	NLD	42	41
Cameroon	CMR	42	41		Norway	NOR	42	41
Congo, Republic	COG	41	39		Nepal	NPL	42	41
Colombia	COL	42	41		New Zealand	NZL	42	41
Costa Rica	CRI	42	41		Oman	OMN	30	29
Cuba	CUB	42	41		Pakistan	PAK	42	41
Cyprus	CYP	42	41		Panama	PAN	42	41
Germany	DEU	42	41		Philippines	PHL	42	41
Denmark	DNK	42	41		Papua New Guinea	PNG	42	41
Dominican Republic	DOM	42	41		Poland	POL	42	41
Algeria	DZA	42	41		Korea, DPR	PRK	42	41
Ecuador	ECU	42	41		Portugal	PRT	42	41
Egypt	EGY	42	41		Paraguay	PRY	42	41
Spain	ESP	42	41		Qatar	QAT	27	26
Ethiopia	ETH	42	41		Romania	ROM	42	41
Finland	FIN	30	29		Rwanda	RWA	34	32
France	FRA	42	41		Saudi Arabia	SAU	42	41
Gabon	GAB	31	30		Sudan	SDN	42	41
United Kingdom	GBR	42	41		Senegal	SEN	42	41
Ghana	GHA	42	41		Sierra Leone	SLE	42	41
Guinea	GIN	41	39		El Salvador	SLV	42	41
Gambia	GMB	39	38		Somalia	SOM	36	33
Guinea-Bissau	GNB	26	23		Suriname	SUR	42	41
Equatorial Guinea	GNQ	19	17		Sweden	SWE	42	41
Greece	GRC	42	41		Swaziland	SWZ	42	41
Guatemala	GTM	42	41		Syria	SYR	42	41
Guyana	GUY	42	41		Chad	TCD	41	39
Honduras	HND	42	41		Togo	TGO	37	36
Haiti	HTI	42	41		Thailand	THA	42	41
Hungary	HUN	42	41		Trinidad & Tobago	TTO	42	41
Indonesia	IDN	42	41		Tunisia	TUN	42	41
India	IND	42	41		Turkey	TUR	42	41
Ireland	IRL	42	41		Tanzania	TZA	42	41
Iran	IRN	42	41		Uganda	UGA	39	36
Iraq	IRQ	42	41		Uruguay	URY	42	41
Iceland	ISL	42	41		United States	USA	42	41
Israel	ISR	42	41		Venezuela	VEN	42	41
Italy	ITA	42	41		Vietnam	VNM	42	41
Jamaica	JAM	42	41		Yemen, Republic	YEM	37	36
Jordan	JOR	42	41		South Africa	ZAF	42	41
Japan	JPN	42	41		Congo, DR	ZAR	41	39
Kenya	KEN	42	41		Zimbabwe	ZWE	42	41

Notes: The full levels sample contains n=5,162 observations, the sample with variables in first differences (FD) contains n=5,013 observations. The sample period is from 1961 to 2002.

Table 3.9: Country groupings

Regional country groups		
Grouping	N	Country Codes
East Asia	14	CHN, IDN, JPN, KHM, KOR, LAO, MMR, MNG, MYS, PHL, PNG, PRK, THA, VNM
Europe & other WIC	28	ALB, AUS, AUT, BGR, BLX, CAN, CHE, CYP, DEU, DNK, ESP, FIN, FRA, GBR, GRC, HUN, IRL, ISL, ITA, NLD, NOR, NZL, POL, PRT, ROM, SWE, TUR, USA
Latin America & Caribbean	24	ARG, BLZ, BOL, BRA, CHL, COL, CRI, CUB, DOM, ECU, GTM, GUY, HND, HTI, JAM, MEX, NIC, PAN, PRY, SLV, SUR, TTO, URY, VEN
South America	11	ARG, BOL, BRA, CHL, COL, ECU, GUY, PRY, SUR, URY, VEN
C. America & Caribbean	13	BLZ, CRI, CUB, DOM, GTM, HND, HTI, JAM, MEX, NIC, PAN, SLV, TTO
Middle East & North Africa	17	ARE, DZA, EGY, IRN, IRQ, ISR, JOR, KWT, LBN, LBY, MAR, OMN, QAT, SAU, SYR, TUN, YEM
North Africa	5	DZA, EGY, LBY, MAR, TUN
South Asia	6	AFG, BGD, IND, LKA, NPL, PAK
Sub-Saharan Africa	39	AGO, BDI, BEN, BFA, BWA, CAF, CIV, CMR, COG, ETH, GAB, GHA, GIN, GMB, GNB, GNQ, KEN, LBR, LSO, MDG, MLI, MOZ, MRT, MWI, NER, NGA, RWA, SDN, SEN, SLE, SOM, SWZ, TCD, TGO, TZA, UGA, ZAF, ZAR, ZWE
Eastern Africa	9	BDI, ETH, KEN, MOZ, MWI, RWA, SOM, TZA, UGA
Western Africa	16	BEN, CAF, CIV, CMR, COG, GAB, GHA, GIN, GMB, GNB, GNQ, LBR, NGA, SEN, SLE, TGO
Southern Africa	7	AGO, BWA, LSO, MDG, SWZ, ZAR, ZWE (does not include ZAF)
Sahel	6	BFA, MLI, MRT, NER, SDN, TCD

Climate-zone country groups		
Grouping	N	Country Codes
Equatorial (EQ)	52	AGO, BDI, BEN, BGD, BLZ, BOL, BRA, CAF, CIV, CMR, COG, COL, CRI, CUB, DOM, GAB, GHA, GIN, GMB, GNB, GNQ, GTM, GUY, HND, HTI, IDN, JAM, KHM, LAO, LBR, LKA, MDG, MMR, MOZ, MYS, NGA, NIC, PAN, PHL, PNG, RWA, SLE, SLV, SUR, TGO, THA, TTO, TZA, UGA, VEN, VNM, ZAR
Arid (AR)	19	AUS, BEN, BFA, BWA, EGY, IRQ, LBY, MLI, MNG, MRT, NER, PAK, SAU, SDN, SEN, SOM, TCD, ZAF, ZWE
Temperate & Cold (TEMP)	51	ALB, ARG, AUS, AUT, BGR, BLX, CAN, CHE, CHL, CHN, CYP, DEU, DNK, DZA, ESP, FIN, FRA, GBR, GRC, HUN, IRL, ISL, ISR, ITA, JOR, JPN, KOR, LAO, LBN, LSO, MAR, MWI, NLD, NOR, NPL, NZL, POL, PRK, PRT, PRY, ROM, SWE, SWZ, SYR, TUN, TUR, URY, USA, VNM, ZAF, ZWE
Highland	12	AFG, CHE, CHL, ECU, ETH, GTM, HND, IRN, MEX, NIC, SAU, YEM
Inbetween categories[1]	2	IND, KEN
Multiple categories	11	AUS, BEN, CHL, GTM, HND, LAO, NIC, SAU, VNM, ZAF, ZWE
Not classified	4	ARE, OMN, KWT, QAT

Notes: [1] India has around 30-35% of arable land in each of the equatorial, arid and temperate & cold zones; Kenya has around 39% in the equatorial, 36% in the arid and 25% in the highland zones respectively.

Table 3.10: Climate Zones following Köppen-Geiger

Köppen-Geiger Classification			
A	Equatorial climates	Af	Equatorial rainforest, fully humid
		Am	Equatorial monsoon
		As	Equatorial savannah with dry summer
		Aw	Equatorial savannah with dry winter
B	Arid climates	BS	Steppe climate
		BW	Desert climate
C	Warm temperate climates	Cs	Warm temperate climate with dry summer
		Cw	Warm temperate climate with dry winter
		Cf	Warm temperate climate, fully humid
D	Snow climates	Ds	Snow climate with dry summer
		Dw	Snow climate with dry winter
		Df	Snow climate, fully humid
E	Polar climates	ET	Tundra climate
		EF	Frost climate
H	Highland climate		above 2,500m elevation

Notes: This classification is taken from Kottek et al. (2006). The H category for Highland climate was added after the creation of the Köppen-Geiger classification, however it was impossible to establish the elevation cut-off definitively. The 2,500m mark is implicitly suggested in a number of online databases.

3.B Time-series properties of the data

In this section we report results relating to the time-series properties of the data. Since the time dimension of the panel is sizeable (T ranges from 19 to 42, average $T = 40.3$), we first carry out Augmented Dickey-Fuller (Dickey & Fuller, 1979) and KPSS (Kwiatkowski et al., 1992) tests for the variable series within each individual country.[15] We use this combination of tests since the ADF test has the null of nonstationary variable series, whereas the KPSS test has the null of stationary variable series. The **time-series unit root test** rejection frequencies for variables in *levels* and in *first differences* are shown in Table 3.11: we report the share of countries (in %) for which the null hypothesis (stationarity or nonstationarity as indicated) is rejected. The theoretical rejection frequencies at our sample size are 12.8% (H_0 : nonstationarity) and 87.2% (H_0 : stationarity) for the level of statistical significance we adopted (10%).

Table 3.11: Time-series unit root tests — rejection frequency

Unit root tests
share of country tests which reject H_0 (stationarity or nonstationarity as indicated);
no adjustment for cross-sectional dependence; all variables in logs

Testing for levels-stationarity								
Test			log output pw	labour	tractors pw	livestock pw	fertilizer pw	land pw
ADF without trend	H_0: nonstationary	H_1: levels-stationary	9%	9%	48%	16%	41%	10%
KPSS without trend	H_0: levels-stationary	H_1: nonstationary	82%	91%	81%	85%	70%	82%

Testing for trend-stationarity								
Test			output pw	labour	tractors pw	livestock pw	fertilizer pw	land pw
ADF with trend	H_0: nonstationary	H_1: trend-stationary	16%	15%	24%	12%	21%	11%
KPSS with trend	H_0: trend-stationary	H_1: nonstationary	65%	71%	88%	65%	74%	66%

Testing for difference-stationarity								
Test			Δoutput pw	Δlabour	Δtractors pw	Δlivestock pw	Δfertilizer pw	Δland pw
ADF with drift	H_0: nonstationary	H_1: stationary	94%	16%	48%	88%	78%	67%
KPSS with drift	H_0: stationary	H_1: nonstationary	13%	38%	81%	81%	70%	82%

Notes: All variables are in logs. We report the share of countries (out of $N = 128$) for which the respective unit root test is rejected at the 10% level of significance. All unit root tests for variables in levels contain an intercept term in the estimating equation. ADF refers to the augmented Dickey-Fuller test, which has the null of nonstationarity. KPSS refers to the Kwiatkowski et al. (1992) unit root test, which has the null of (trend-)stationarity. Lag-augmentation or bandwidth selection in these tests to account for serial correlation in the variables is allowed to vary by country. For the ADF test we determined 'ideal' lag-augmentation using the Akaike Information Criterion (AIC). For the KPSS tests an automated bandwidth selection following Newey and West (1994) and discussed in Hobijn et al. (2004) is used. For KPSS we use the `kpss` command in `Stata` written by Kit Baum.

For the majority of countries the ADF tests for the variables in levels cannot reject nonstationarity, with the notable exceptions of tractors per worker and fertilizer per worker. Consistently with this finding the majority of country KPSS tests reject the null of level stationarity. The tests for trend stationarity reveal a similar pattern. The difference stationarity tests show considerable differences across variables: in the ADF tests the labour and tractors per worker variables reject the nonstationarity null in far less countries than we would expect (87.2%) and

[15] Whereas the `Stata` command for ADF allows us to run country regressions with gaps in the data, this is not possible for the KPSS tests. We interpolate data in order to run the KPSS for a balanced panel.

the KPSS tests reject stationarity in the vast majority of countries for the tractors, livestock, fertilizer and land (all in per worker terms) variables.

Our analysis based on standard time-series (non)stationarity tests therefore has no clearcut message regarding variable properties. It needs to be emphasised that country-specific unit root tests suffer from low power, in particular in the case where the persistence in the variable is high — i.e. in the case when the test matters most (Harris, 1994).

Next we apply **'first generation' panel unit root tests** to the data. These were developed due to the desirable property of increased power from pooling the results from many low-powered country unit root tests. It is important to stress that rejection of the unit root null hypothesis does not imply that the panel is stationary, but rather that the variable series does not follow a unit root process *in all countries*. Table 3.12 presents the results for the Maddala and Wu (1999) (MW) panel unit root test and a panel version of Phillips and Perron (1988) (PP) test, for which serial correlation is accounted for using nonparametric methods rather than lagged differences. Following Fisher's suggestion the MW statistic is constructed as $P = -2\sum_i log(p_i)$, where p_i is the p-value for the individual country ADF statistics. The PP test is constructed in analogy. For both tests the theoretical distribution of the statistic is $\chi^2(2N)$, s.t. critical values are 97.35 for 5% and 92.16 for 10%.

For both tests tractors per worker and fertilizer per worker in levels reject nonstationarity in both the standard ADF equation and the ADF equation with a trend. All other variables in levels seemingly cannot reject the null of nonstationarity once augmented with sufficient lags or once a trend is added to the ADF equation. For the variables in first differences the tests unanimously reject nonstationarity.

Table 3.12: First generation panel unit root tests: Fisher tests

Maddala & Wu (1999) unit root test												
Variables in levels: ADF equation contains intercept												
variable †	**ly**		**lL**		**ltr**		**llive**		**lf**		**ln**	
lags	χ^2	p	χ^2	p	χ^2	p	χ^2	p	χ^2	p	χ^2	p
0	264.33	.35	836.91	.00	2776.92	.00	318.88	.00	713.80	.00	202.25	.99
1	248.31	.62	212.39	.98	897.70	.00	269.52	.27	672.69	.00	174.36	1.00
2	216.15	.97	161.97	1.00	716.62	.00	298.71	.03	593.96	.00	185.48	1.00
3	205.40	.99	137.31	1.00	707.83	.00	274.18	.21	561.27	.00	172.74	1.00
4	198.13	1.00	139.56	1.00	614.63	.00	236.64	.80	586.15	.00	252.63	.55
5	217.60	.96	138.40	1.00	663.07	.00	236.17	.81	551.27	.00	189.16	1.00
6	180.78	1.00	140.82	1.00	532.05	.00	219.06	.95	399.06	.00	159.21	1.00
Variables in levels: ADF equation contains intercept & trend												
variable †	**ly**		**lL**		**ltr**		**llive**		**lf**		**ln**	
lags	χ^2	p	χ^2	p	χ^2	p	χ^2	p	χ^2	p	χ^2	p
0	473.86	.00	205.02	.99	916.00	.00	272.79	.22	411.14	.00	160.72	1.00
1	322.37	.00	676.49	.00	499.77	.00	326.65	.00	319.24	.00	340.40	.00
2	241.67	.73	293.71	.05	423.53	.00	318.07	.00	256.12	.49	255.62	.50
3	230.43	.87	266.99	.31	335.91	.00	319.78	.00	235.53	.82	269.17	.27
4	224.20	.92	249.02	.61	449.23	.00	286.63	.09	256.61	.48	305.17	.02
5	205.05	.99	254.56	.51	479.71	.00	286.74	.09	313.58	.01	290.07	.07
6	184.29	1.00	230.35	.87	436.98	.00	248.59	.62	190.53	1.00	275.99	.19
Variables in first differences: ADF equation contains drift												
variable ‡	**Δly**		**ΔlL**		**Δltr**		**Δllive**		**Δlf**		**Δln**	
lags	χ^2	p	χ^2	p	χ^2	p	χ^2	p	χ^2	p	χ^2	p
0	5893.19	.00	513.12	.00	2092.88	.00	2780.22	.00	5138.24	.00	2408.11	.00
1	2875.38	.00	538.34	.00	1206.83	.00	1561.93	.00	2475.73	.00	1153.82	.00
2	1571.83	.00	414.45	.00	773.43	.00	1052.88	.00	1437.93	.00	806.68	.00
3	1048.43	.00	338.69	.00	603.03	.00	867.15	.00	956.45	.00	616.82	.00
4	779.00	.00	289.06	.08	534.76	.00	669.12	.00	662.53	.00	473.58	.00
5	564.99	.00	312.39	.01	627.47	.00	601.37	.00	517.31	.00	441.97	.00
6	425.31	.00	341.73	.00	534.53	.00	504.60	.00	420.91	.00	426.10	.00

Philipps & Perron (1988)-type unit root test (panel version)												
Variables in levels: ADF equation contains intercept												
variable †	**ly**		**lL**		**ltr**		**llive**		**lf**		**ln**	
lags	χ^2	p	χ^2	p	χ^2	p	χ^2	p	χ^2	p	χ^2	p
0	264.33	.35	836.91	.00	2776.92	.00	318.88	.00	713.80	.00	202.25	.99
1	264.64	.34	523.52	.00	2373.71	.00	289.99	.07	764.06	.00	187.07	1.00
2	274.14	.21	403.42	.00	2180.19	.00	280.98	.14	815.30	.00	179.45	1.00
3	281.71	.13	344.04	.00	2100.88	.00	279.59	.15	858.99	.00	181.16	1.00
4	288.05	.08	311.81	.01	2063.62	.00	279.96	.15	899.75	.00	180.85	1.00
5	298.86	.03	294.77	.05	2089.83	.00	283.25	.12	946.55	.00	183.07	1.00
6	308.91	.01	287.25	.09	2013.23	.00	285.65	.10	983.57	.00	184.63	1.00
Variables in levels: ADF equation contains intercept & trend												
variable †	**ly**		**lL**		**ltr**		**llive**		**lf**		**ln**	
lags	χ^2	p	χ^2	p	χ^2	p	χ^2	p	χ^2	p	χ^2	p
0	473.86	.00	205.02	.99	916.00	.00	272.79	.22	411.14	.00	160.72	1.00
1	459.81	.00	136.91	1.00	803.17	.00	267.51	.30	408.40	.00	167.54	1.00
2	463.28	.00	120.88	1.00	759.12	.00	260.85	.40	411.18	.00	174.00	1.00
3	468.30	.00	117.74	1.00	762.21	.00	258.26	.45	412.02	.00	180.44	1.00
4	472.24	.00	119.26	1.00	775.59	.00	254.96	.51	411.84	.00	183.58	1.00
5	475.65	.00	121.56	1.00	819.31	.00	251.47	.57	411.58	.00	184.72	1.00
6	478.72	.00	123.13	1.00	883.30	.00	246.73	.65	410.36	.00	184.77	1.00
Variables in first differences: ADF equation contains drift												
variable ‡	**Δly**		**ΔlL**		**Δltr**		**Δllive**		**Δlf**		**Δln**	
lags	χ^2	p	χ^2	p	χ^2	p	χ^2	p	χ^2	p	χ^2	p
0	5893.19	.00	513.12	.00	2092.88	.00	2780.22	.00	5138.24	.00	2408.11	.00
1	5958.33	.00	517.81	.00	2103.97	.00	2763.12	.00	5179.80	.00	2390.08	.00
2	6140.51	.00	548.19	.00	2108.97	.00	2759.40	.00	5284.52	.00	2410.07	.00
3	6261.04	.00	571.63	.00	2106.40	.00	2768.50	.00	5357.08	.00	2438.08	.00
4	6369.17	.00	571.18	.00	2111.24	.00	2774.13	.00	5418.36	.00	2463.51	.00
5	6487.63	.00	566.35	.00	2121.16	.00	2782.00	.00	5469.27	.00	2484.50	.00
6	6592.92	.00	565.60	.00	2135.74	.00	2798.47	.00	5520.19	.00	2499.74	.00

Notes: †Output per worker (lo), labour (lL), tractors per worker (ltr), livestock per worker (llive), fertilizer per worker (lf) and land per worker (ln) — all in logs. ‡The Δ symbolise the growth rates for the above variables (first differences of the variables in logs).
The null is nonstationarity in all countries' variable series, the alternative stationarity in *all* countries' variable series.

Similarly to the above analysis we cannot definitely reject nonstationarity in all variables. However, as Baltagi et al. (2007) point out the first generation panel unit root tests which do not account for cross-section dependence can be subject to considerable size distortions, such that the test tends to overreject. This issue led to the development of **'second generation' panel unit root tests**, namely the Im et al. (2003) and the Pesaran (2007) tests, results for which are presented in Table 3.13. These tests explicitly allow for cross-sectional dependence in the data and therefore have better performance than the 'classic' panel unit root tests that assume cross-sectional independence.

Our results now seem to provide a more consistent theme across the different variables and specifications: following augmentation with lags or a linear trend term the levels variables cannot reject the null of nonstationarity. For the variables in first differences all variables reject nonstationarity with the exception of labour.

It should be noted that our preferred panel unit root test by Pesaran (2007) can only account for a single unobserved common factor as the cause for cross-sectional dependence in the data. More recent extensions of this test (Pesaran et al., 2008) can accommodate multiple unobserved common factors, but this development was too recent for the test to be included in this thesis. Further issues which have received attention in the recent panel unit root literature such as structural breaks in data series (Westerlund, 2006; Kao, Trapani, & Urga, 2007) and changes in the volatility of the innovations (Hanck, 2008) could also not be accommodated.

Table 3.13: Second generation panel unit root tests

Im, Pesaran & Shin (2003) unit root test for heterogeneous panels with cross-section dependence												
Variables in levels: ADF equation contains intercept												
variable †	**ly**		**lL**		**ltr**		**llive**		**lf**		**ln**	
lags	Wtbar	p	Wtbar	p	Wtbar	p	Wtbar	p	Wtbar	p	Wtbar	p
0	0.70	.76	28.08	1.00	-20.32	.00	9.65	1.00	-7.03	.00	15.85	1.00
1	2.71	1.00	13.00	1.00	-3.05	.00	6.37	1.00	-3.84	.00	11.00	1.00
2	3.37	1.00	11.58	1.00	-1.32	.09	5.11	1.00	-3.21	.00	9.28	1.00
3	2.77	1.00	11.43	1.00	-2.11	.02	4.51	1.00	-1.68	.05	8.35	1.00
4	2.35	.99	12.44	1.00	-3.19	.00	4.47	1.00	-2.00	.02	7.16	1.00
5	2.70	1.00	13.47	1.00	-4.60	.00	4.33	1.00	-1.50	.07	7.10	1.00
6	2.80	1.00	14.24	1.00	-1.71	.04	3.74	1.00	-3.11	.00	7.89	1.00
Variables in levels: ADF equation contains intercept & trend												
variable †	**ly**		**lL**		**ltr**		**llive**		**lf**		**ln**	
lags	Wtbar	p	Wtbar	p	Wtbar	p	Wtbar	p	Wtbar	p	Wtbar	p
0	-4.59	.00	16.98	1.00	-9.05	.00	6.63	1.00	-9.30	.00	10.28	1.00
1	0.13	.55	-0.02	.49	-0.53	.30	1.87	.97	-3.66	.00	1.69	.95
2	2.53	.99	0.08	.53	-0.61	.27	0.52	.70	-2.54	.01	1.65	.95
3	2.46	.99	2.02	.98	-1.99	.02	0.75	.77	-0.27	.39	0.30	.62
4	1.56	.94	1.46	.93	-1.07	.14	1.19	.88	-1.00	.16	-0.28	.39
5	2.50	.99	1.75	.96	-2.96	.00	0.45	.67	-0.01	.50	-1.01	.16
6	2.27	.99	3.53	1.00	0.15	.56	1.39	.92	-3.77	.00	-0.85	.20
Variables in first differences: ADF equation contains drift												
variable ‡	**Δly**		**ΔlL**		**Δltr**		**Δllive**		**Δlf**		**Δln**	
lags	Wtbar	p	Wtbar	p	Wtbar	p	Wtbar	p	Wtbar	p	Wtbar	p
0	-77.80	.00	-10.40	.00	-42.24	.00	-46.39	.00	-69.35	.00	-39.20	.00
1	-47.05	.00	-6.55	.00	-28.17	.00	-28.74	.00	-39.86	.00	-23.90	.00
2	-29.98	.00	-6.85	.00	-18.86	.00	-20.97	.00	-29.20	.00	-15.43	.00
3	-21.77	.00	-5.17	.00	-15.16	.00	-17.02	.00	-21.12	.00	-11.75	.00
4	-16.67	.00	-4.88	.00	-11.55	.00	-12.87	.00	-16.44	.00	-8.77	.00
5	-12.57	.00	-5.75	.00	-10.40	.00	-10.83	.00	-13.12	.00	-7.89	.00
6	-9.57	.00	-3.90	.00	-8.52	.00	-7.83	.00	-12.38	.00	-7.58	.00

Pesaran (2007) unit root test (CIPS) for heterogeneous panels with cross-section dependence												
Variables in levels: CADF equation contains intercept												
variable †	**ly**		**lL**		**ltr**		**llive**		**lf**		**ln**	
lags	Ztbar	p	Ztbar	p	Ztbar	p	Ztbar	p	Ztbar	p	Ztbar	p
0	-7.19	.00	14.77	1.00	-0.38	.35	2.56	.99	-9.53	.00	12.13	1.00
1	-2.55	.01	11.70	1.00	-3.17	.00	-1.37	.08	-5.23	.00	7.95	1.00
2	-0.78	.22	13.78	1.00	-2.16	.02	-0.96	.17	-2.51	.01	7.94	1.00
3	-0.34	.37	16.22	1.00	-3.37	.00	0.14	.55	0.54	.71	6.33	1.00
4	0.29	.61	17.63	1.00	0.08	.53	3.06	1.00	1.54	.94	7.89	1.00
Variables in levels: CADF equation contains intercept & trend												
variable †	**ly**		**lL**		**ltr**		**llive**		**lf**		**ln**	
lags	Ztbar	p	Ztbar	p	Ztbar	p	Ztbar	p	Ztbar	p	Ztbar	p
0	-2.31	.01	7.90	1.00	3.26	1.00	6.68	1.00	-9.31	.00	9.69	1.00
1	2.94	1.00	-0.82	.20	-1.09	.14	1.80	.96	-5.61	.00	1.70	.96
2	5.85	1.00	5.24	1.00	-0.59	.28	2.74	1.00	-2.33	.01	0.71	.76
3	7.15	1.00	8.98	1.00	-0.17	.43	4.51	1.00	0.63	.74	-0.07	.47
4	7.38	1.00	10.17	1.00	0.42	.66	6.97	1.00	1.90	.97	0.59	.72
Variables in first differences: CADF equation contains drift												
variable ‡	**Δly**		**ΔlL**		**Δltr**		**Δllive**		**Δlf**		**Δln**	
lags	Ztbar	p	Ztbar	p	Ztbar	p	Ztbar	p	Ztbar	p	Ztbar	p
0	-48.44	.00	-1.37	.09	-32.95	.00	-33.72	.00	-48.55	.00	-28.58	.00
1	-33.63	.00	-0.84	.20	-20.81	.00	-21.62	.00	-35.92	.00	-15.30	.00
2	-20.34	.00	-0.52	.30	-14.15	.00	-13.72	.00	-25.44	.00	-7.51	.00
3	-14.30	.00	1.84	.97	-8.11	.00	-8.41	.00	-16.36	.00	-4.64	.00
4	-7.54	.00	3.41	1.00	-4.26	.00	-4.01	.00	-10.44	.00	-0.92	.18

Notes: †Output per worker (ly), labour (lL), tractors per worker (ltr), livestock per worker (llive), fertilizer per worker (lf) and land per worker (ln) — all in logs. ‡The Δ symbolise the growth rates for the above variables (first differences of the variables in logs).
The null is nonstationarity in all countries' variable series, the alternative stationarity in *some* countries' variable series.

3.C Cross-section dependence in the data

In this section we investigate the potential for cross-section dependence in the data. We initially focus on the full sample ('global') data ($N = 128$, average $T = 40.3$ for the variables in levels) and later on investigate geographic 'regions' and other subsamples.

Table 3.14 details the share of the variance accounted for by the first two principal components (PCs) to indicate the factor structure of the data, as suggested by Coakley et al. (2006).[16] In Panel [a] we apply this method to the variables in (log) levels and find that the first two principal components account for 76-93% of the variance. We also investigate whether the residuals from pooled OLS ($\hat{e}^{POLS^*}$) and 2-way Fixed Effects ($\hat{e}^{2FE^*}$) regressions[17] show signs of factor structures. Again the share of the first two PCs is very high, around 65% in both cases. In Panel [b] we carry out the same analysis for the variables in first differences (for the residuals the production function OLS and 2FE regressions are run with variables in first differences); the explained variance is now considerably lower, although labour force growth (dlL) and growth in tractors per worker (dltr) still exhibit strong underlying factor structures.

In Table 3.15 we report the means for the $N(N-1)$ correlation coefficients for variable series or regression residuals, as well as the Pesaran (2004) Cross-Section Dependence (CD) test statistics. The former represents the simple average of the pairwise correlation coefficients between all country series ($\hat{\rho}_{ij}$) or the average of their absolute values ($|\hat{\rho}_{ij}|$). The CD test statistic is also based on the mean pairwise correlation coefficients. In the unbalanced panel case it is defined as

$$CD = \sqrt{\left(\frac{2}{N(N-1)}\right)}\left(\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\sqrt{T_{ij}}\hat{\rho}_{ij}\right)$$

where T_{ij} is the number of observations used to estimate the correlation coefficient between the series in country i and j and CD is distributed $N(0,1)$ under the null of cross-section independence for $T_{ij} > 3$ and sufficiently large N. This test is robust to the presence of non-stationary processes, parameter heterogeneity or structural breaks, and was shown to perform well even in small samples.

[16] In principal component analysis (PCA) the eigenvalues (ordered by magnitude) over the cumulated eigenvalues give an indication of the variance in the standardized data explained by the different 'principal components'. The latter are linear combinations of the $N(N-1)$ data time-series to account for the maximum variation in the overall dataset.

[17] These are production function regressions as outlined in the main section of this chapter.

Table 3.14: 'Global' Cross-section Dependence (i)

Principal Component Analysis Share of variance (in %) accounted for by the first two Principal Components								
Panel [a]	**Variables in levels**†						**Residuals**‡	
	ly	lL	ltr	llive	lf*	ln	$\hat{e}^{POLS*}$	$\hat{e}^{2FE*}$
%V Comp1	66.8	83.2	75.0	59.4	62.8	67.5	48.3	44.5
%V Comp2	11.5	9.5	15.0	17.0	13.8	16.0	16.8	20.3
sum	78.4	92.7	90.0	76.3	76.7	83.5	65.1	64.8
Panel [b]	**Variables in first differences**†						**Residuals**‡	
	dly*	dlL*	dltr*	dllive*	dlf*	dln*	$\hat{e}^{FD-OLS*}$	$\hat{e}^{FD2FE*}$
%V Comp1	12.03	28.6	29.7	10.7	15.4	13.8	12.5	12.3
%V Comp2	5.97	22.3	11.9	7.4	8.9	10.5	6.1	7.6
sum	18.0	50.9	41.6	18.0	24.3	24.2	18.6	19.9
N	127	127	127	126	128	126	128	128
excluded	FIN	FIN	FIN	BLX,FIN	-	BLX,FIN	-	-

Notes: †Output per worker (ly), labour (lL), tractors per worker (ltr), livestock per worker (llive), fertilizer per worker (lf) and land per worker (ln) — all in logs. The prefix d- identifies data in first differences.
‡These are the residuals from a pooled OLS regression with $T-1$ year dummies (POLS), the 2-way Fixed Effects regression (2FE), the pooled OLS regression with variables (and year dummies) in first differences (FD-OLS) and the 2-way Fixed Effects regression with variables in first differences (FD2FE).
*This indicates that the variable had to be interpolated since there were not enough common years of data across all countries to carry out PCA. Otherwise we excluded some countries from the analysis as indicated.

Panel [a] of Table 3.15 again investigates the variable series in levels and residuals from the pooled OLS and 2FE regressions. Average correlation varies considerably across the variables, from .53 in the case of tractors per worker (ltr) to virtually no correlation in land per worker (ln). Average correlation is low for the regression residuals, however the CD statistic rejects the null of cross-section independence at $p < .01$ in both cases. This result emphasises the importance of parameter heterogeneity (in the presence of nonstationarity): if production function parameters and the influence of the unobserved common factor(s) were *identical* across countries the $2FE$ transformation should be able to eliminate all the cross-section dependence in the data (Coakley et al., 2006). This is seemingly not the case here.

Panel [b] shows the average correlations for the data in first differences and the CD statistic for residuals from OLS and 2FE regressions with the data in first differences. A similar pattern to the PCA results emerges, in that the average correlations are considerably lower than in the levels case in Panel [a]. The CD test cannot reject cross-section independence for the FD-OLS residuals (CD=0.04) — recall that this regression includes $(T-1)$ year dummies which seem to capture the average impact of the unobserved common factor(s) across countries — we discuss this finding further in Chapter 4. In contrast the residuals from the 2FE regression with data in first differences (*after* the 2FE transformation) display cross-section dependence.

Table 3.15: 'Global' Cross-section Dependence (ii)

Mean Correlation Coefficients and Pesaran (2004) CD test$^\sharp$ Global cross-section dependence in the data; $N = 128$, $n = 16,256$								
Panel [a]	**Variables in levels**†						**Residuals**‡	
	ly	lL	ltr	llive	lf	ln	$\hat{e}^{POLS}$	$\hat{e}^{2FE}$
$(N(N-1))^{-1}\sum_i\sum_j \hat{\rho}_{ij}$	0.372	0.100	0.532	0.125	0.458	-0.007	-0.005	0.015
$(N(N-1))^{-1}\sum_i\sum_j \lvert\hat{\rho}_{ij}\rvert$	0.611	0.799	0.699	0.561	0.544	0.645	0.424	0.408
Pesaran CD Statistic							-2.49	9.64
Panel [b]	**Variables in first differences**†						**Residuals**‡	
	dly	dlL	dltr	dllive	dlf	dln	$\hat{e}^{FD-OLS}$	$\hat{e}^{FD2FE}$
$(N(N-1))^{-1}\sum_i\sum_j \hat{\rho}_{ij}$	0.021	0.053	0.210	0.020	0.051	0.008	0.000	0.023
$(N(N-1))^{-1}\sum_i\sum_j \lvert\hat{\rho}_{ij}\rvert$	0.141	0.351	0.276	0.144	0.147	0.172	0.145	0.149
Pesaran CD Statistic							0.04	12.84
Panel [c]	**AR regression residuals**$^\natural$						**Residuals**‡	
	$\hat{e}^{ly}$	$\hat{e}^{lL}$	$\hat{e}^{ltr}$	$\hat{e}^{llive}$	$\hat{e}^{lf}$	$\hat{e}^{ln}$	$\hat{e}^{MG}$	$\hat{e}^{CCEMG}$
$(N(N-1))^{-1}\sum_i\sum_j \hat{\rho}_{ij}$	0.020	0.075	0.012	0.013	0.025	0.008	0.017	0.001
$(N(N-1))^{-1}\sum_i\sum_j \lvert\hat{\rho}_{ij}\rvert$	0.137	0.301	0.143	0.134	0.143	0.141	0.147	0.150
Pesaran CD Statistic	11.52	42.23	7.02	7.47	14.63	4.36	9.16	0.21

Notes: †Variables as defined in Table 3.15. ‡These are the residuals from a pooled OLS regression with $T-1$ year dummies (POLS), the 2-way Fixed Effects regression (2FE), individual country regressions with intercept and linear trend (MG) and from the Pesaran (2006) Common Correlated Effects MG estimator (CCEMG).
♮ Each of the variables in levels is entered into a regression $z_{it} = \pi_{1,i}z_{i,t-1} + \pi_{2,i}z_{i,t-2} + \pi_{t,i}t + \pi_{0,i}$, conducted separately for each country i. The correlations and cross-section dependence statistic are then based on the residuals from these AR regressions.
♯ $\hat{\rho}_{ij}$ where $i \neq j$ refers to the correlation coefficient for the variable/residuals in question between countries i and j. $|\hat{\rho}_{ij}|$ is the absolute value of the same statistic. The CD Test Statistic (for unbalanced panels) is

$$CD = \sqrt{\left(\frac{2}{N(N-1)}\right)}\left(\sum_i\sum_j \sqrt{T_{ij}}\hat{\rho}_{ij}\right)$$

where $i = 1, \ldots, N-1$ and $j = i+1, \ldots, N$ and T_{ij} is the number of observations for each pairwise correlation. In our case $N = 128$ and average $T = 38.8$. Note that we adjusted the residual series for each i by subtracting their mean for the period T_{ij} since they may not sum to zero otherwise (Pesaran, 2004, p.17). For $N \to \infty$ the CD statistic is distributed standard normal under the null of cross-section independence.

Finally, we follow Pesaran (2004) and run autoregressions for each variable in each country. Panel [c] reports CD statistics and the mean correlations across countries for the residuals ($\hat{e}_{it}$) from an AR(2) regression defined $z_{it} = \pi_{1,i}z_{i,t-1} + \pi_{2,i}z_{i,t-2} + \pi_{t,i}t + \pi_{0,i} + e_{it}$. We also report these statistics for the residuals from individual country regressions ($\hat{e}^{MG}$) and the Pesaran (2006) Common Correlated Effects country regressions ($\hat{e}^{CCEMG}$) — again, these are the production functions discussed in the main section. The AR regression and country regression residual series fail the test for cross-section independence since these regressions do not account for the impact of unobserved common factors. In contrast we cannot reject cross-section independence for the CCEMG residuals (CD=0.21).

In summary, the investigation of the full sample ('global' data) offers strong evidence of cross-section dependence in the variable series studied. The basic assumption of the standard panel estimators that data is cross-sectionally independent is therefore violated. We can see this in our analysis of the regression residuals from the pooled OLS and 2-way Fixed Effects estimators, as well as the individual country regressions (MG). Crucially, residuals from the pooled

OLS estimator with variables in first differences (FD-OLS) and the Pesaran (2006) CCEMG estimator, the latter specifically developed to deal with heterogeneous and cross-sectionally dependent panels, both pass the test for cross-section independence.

Having investigated the full sample case, we now turn to analyse unobserved common factors in various spatial groupings. We begin with an investigation of the within-group correlations in a number of 'geo-historical' categories, where we distinguish countries from East Asia (EA), Latin America & the Caribbean (LAC), Middle East & North Africa (MENA), South Asia (SA), Sub-Saharan Africa (SSA) and Europe & other Western Industrialised Countries (EU). The latter category includes the United States, Canada, New Zealand and Australia, pointing to the general makeup of each category not just based on geography, but also cultural characteristics and historical grounds.

Table 3.16 presents variable (and residual) mean absolute correlations (we omit the mean standard correlations to save space) within each of these country groupings. The number of observations ranges from 30 in South Asia ($N = 6$) to 1,482 in Sub-Saharan Africa ($N = 39$). In Panel [a] we report results for the variables in levels and the regression residuals from OLS and 2FE regressions. All variables are uniformly highly correlated within country groupings. For the regression residuals the members of the Europe & WIC and East Asia groups in particular show high levels of correlation.[18] In Panel [b] we investigate residuals from the autoregressions outlined above as well as the country regressions and CCE country regressions. All of these show significantly lower levels of correlation. Average correlation (not reported) shows a much higher range of correlation coefficients across the groups, in the case of land from -.06 in East Asia to .58 in Europe & WIC and .87 in South Asia. In the latter analysis regression residuals from POLS and 2FE remain high in the Europe & WIC case (.41 and .28 respectively), but are virtually insignificant in the MG and CCEMG case (.05 and .02).

In Table 3.17 we present the results for the CD test statistic within and across regions applied to various regression residuals. Panels [a] to [e] are for the pooled OLS (POLS), 2-way Fixed Effects (2FE), OLS in first differences (FD-OLS), country regression and CCE country regres-

[18]Note that the South Asian group is made up of 6 countries only, so that the CD test statistics are likely to be highly distorted. We nevertheless kept these countries separate due to their distinct geographical and climatic characteristics.

Table 3.16: Regional Cross-section Dependence (i)

Mean Absolute Correlation Coefficients Regional cross-section dependence									
Panel [a]		**Variables**[†]						**Residuals**[‡]	
	obs	ly	lL	ltr	llive	lf	ln	$\hat{e}^{POLS}$	$\hat{e}^{2FE}$
Mean absolute correlations: $(N(N-1))^{-1}\sum_i\sum_j\lvert\hat{\rho}_{ij}\rvert$									
East Asia	182	0.727	0.740	0.617	0.461	0.600	0.554	0.407	0.516
Europe & other WIC	756	0.893	0.870	0.848	0.732	0.726	0.843	0.682	0.479
Latin America & Caribbean	552	0.610	0.661	0.657	0.465	0.566	0.544	0.365	0.445
MENA	272	0.697	0.643	0.871	0.640	0.725	0.531	0.320	0.331
South Asia	30	0.497	0.914	0.740	0.680	0.733	0.871	0.781	0.313
SSA	1482	0.427	0.914	0.631	0.520	0.385	0.634	0.367	0.350
Panel [b]		**AR regression residuals**[♮]						**Residuals**[‡]	
	obs	$\hat{e}^{ly}$	$\hat{e}^{lL}$	$\hat{e}^{ltr}$	$\hat{e}^{llive}$	$\hat{e}^{lf}$	$\hat{e}^{ln}$	$\hat{e}^{MG}$	$\hat{e}^{CCEMG}$
Mean absolute correlations: $(N(N-1))^{-1}\sum_i\sum_j\lvert\hat{\rho}_{ij}\rvert$									
East Asia	182	0.132	0.409	0.170	0.142	0.164	0.136	0.139	0.159
Europe & other WIC	756	0.155	0.420	0.163	0.154	0.176	0.155	0.149	0.156
Latin America & Caribbean	552	0.139	0.347	0.146	0.131	0.135	0.140	0.154	0.154
MENA	272	0.156	0.321	0.131	0.133	0.156	0.151	0.154	0.171
South Asia	30	0.108	0.388	0.125	0.138	0.169	0.142	0.137	0.126
SSA	1482	0.131	0.232	0.151	0.133	0.149	0.135	0.152	0.157

Notes: See Table 3.15 for notes and descriptions. Country groupings are described in Table 3.9 ("other WIC" refers to other Western industrialised countries, namely the US, Canada, New Zealand and Australia).

sion residuals respectively. Panel [f] shows the number of observations (pairwise correlations) for each test. We marked the test statistics which indicate rejection of the null of cross-section independence (5% level) in bold.

Focusing first on the *within-group* CD test statistics (on the diagonals), we can see that these commonly reject cross-section independence across all estimation strategies, although less so in the country and CCEMG regression cases. With regard to *cross-group* dependence tests, it can be seen that residuals from the pooled regression in levels and the pooled 2FE regression display a high level of common factor structure, leading to the rejection of the null of cross-section independence in almost all cases. For the pooled regression with variables in first differences 10 out of 15 cross-group tests cannot reject the null of cross-section independence, whereas for country regressions 12 tests cannot reject. For CCEMG all 15 tests cannot reject independence across groups. The results in Panel [e] for CCEMG residuals suggest that this estimation approach cannot entirely wipe out the cross-section dependence in the data if viewed from a group-perspective. It is no longer the case that the number of observations is systematically linked to the test statistic, suggesting that the results are not driven by small sample size.

Table 3.17: Regional Cross-section Dependence (ii)

Pesaran CD statistic — within and across regions Regional cross-section dependence						
Panel [a]	***Residuals from pooled OLS regression***					
	EA	EU	LAC	MENA	SA	SSA
East Asia (EA)	0.42					
Europe & other WIC (EU)	**-11.31**	**50.77**				
Latin America & Caribbean (LAC)	0.64	**7.11**	**3.45**			
Middle East & N Africa (MENA)	**-4.33**	**9.72**	**-2.37**	1.06		
South Asia (SA)	**4.84**	**-21.47**	**-4.20**	**-2.85**	**7.07**	
Sub-Saharan Africa (SSA)	**5.41**	**-26.89**	**-4.85**	**-5.97**	**11.73**	**10.42**
Panel [b]	***Residuals from 2-way FE regression***					
	EA	EU	LAC	MENA	SA	SSA
East Asia (EA)	0.59					
Europe & other WIC (EU)	**10.00**	**34.51**				
Latin America & Caribbean (LAC)	**4.00**	**13.46**	**3.69**			
Middle East & N Africa (MENA)	**3.54**	**12.20**	**2.07**	**11.62**		
South Asia (SA)	0.36	**-3.44**	**-2.11**	**6.14**	**3.80**	
Sub-Saharan Africa (SSA)	**-5.14**	**-15.29**	**-4.93**	**-13.44**	**-5.07**	**10.82**
Panel [c]	***Residuals from FD-OLS regression***					
	EA	EU	LAC	MENA	SA	SSA
East Asia (EA)	**2.15**					
Europe & other WIC (EU)	**2.04**	**3.17**				
Latin America & Caribbean (LAC)	**2.32**	0.60	**2.45**			
Middle East & N Africa (MENA)	-1.82	0.16	**-2.04**	0.11		
South Asia (SA)	1.63	-0.04	0.83	-0.99	**2.71**	
Sub-Saharan Africa (SSA)	-0.10	**-2.20**	-1.11	**-2.07**	0.24	0.96
Panel [d]	***Residuals from country regressions***					
	EA	EU	LAC	MENA	SA	SSA
East Asia (EA)	0.83					
Europe & other WIC (EU)	**3.00**	**6.21**				
Latin America & Caribbean (LAC)	0.58	0.02	1.86			
Middle East & N Africa (MENA)	0.29	**2.77**	0.18	0.32		
South Asia (SA)	0.87	1.65	0.73	-0.15	1.87	
Sub-Saharan Africa (SSA)	0.05	1.92	**2.04**	1.03	1.31	**5.92**
Panel [e]	***Residuals from CCE country regressions***					
	EA	EU	LAC	MENA	SA	SSA
East Asia (EA)	**2.14**					
Europe & other WIC (EU)	1.35	**2.36**				
Latin America & Caribbean (LAC)	0.13	-1.92	1.58			
Middle East & N Africa (MENA)	0.27	0.40	-0.94	1.29		
South Asia (SA)	0.94	-0.26	0.34	-0.10	**2.43**	
Sub-Saharan Africa (SSA)	-0.47	-1.91	0.53	-0.48	-0.37	1.29

Panel [f]	***Number of observations***					
	EA	EU	LAC	MENA	SA	SSA
East Asia (EA)	182					
Europe & other WIC (EU)	392	756				
Latin America & Caribbean (LAC)	336	672	552			
Middle East & N Africa (MENA)	238	476	408	272		
South Asia (SA)	84	168	144	102	30	
Sub-Saharan Africa (SSA)	546	1092	936	663	234	1482

Notes: See Table 3.15 for notes and descriptions. Country groupings are described in Table 3.9 ("other WIC" refers to other Western industrialised countries, namely the US, Canada, New Zealand and Australia).
The test statistics in bold are statistically significant at the 5% level and thus reject the null of cross-section independence.
Panel [f] reports the number of individual correlation estimates ($N(N-1)$) used in the construction of the CD statistic.

Table 3.18: Climate-zone Cross-section Dependence

Pesaran CD statistics — within and across climate zones
We take 40% of arable land in each climate-zone as cut-off

Panel [a] **POLS residuals**	EQ	AR	TEMP	**Panel [b]** **2FE residuals**	EQ	AR	TEMP
Equatorial (EQ)	**2.29**			Equatorial (EQ)	-0.09		
Arid (AR)	2.00	-0.59		Arid (AR)	-1.18	-1.78	
Temperate & Cold (TEMP)	**-11.71**	**-9.36**	**31.28**	Temperate & Cold (TEMP)	-0.89	1.38	**32.79**

Panel [c] **FD-OLS**	EQ	AR	TEMP	**Panel [d]** **MG residuals**	EQ	AR	TEMP
quatorial (EQ)	**4.10**			Equatorial (EQ)	**2.50**		
Arid (AR)	**-3.02**	0.68		Arid (AR)	**3.47**	**3.90**	
Temperate & Cold (TEMP)	0.39	**-2.74**	0.20	Temperate & Cold (TEMP)	**2.64**	**3.47**	**6.67**

Panel [e] **CCEMG residuals**	EQ	AR	TEMP	**Panel [f]** **Observations**	EQ	AR	TEMP
Equatorial (EQ)	1.93			Equatorial (EQ)	2652		
Arid (AR)	0.29	1.31		Arid (AR)	987	342	
Temperate & Cold (TEMP)	-0.78	-1.78	1.39	Temperate & Cold (TEMP)	2650	966	2550

Notes: We divide countries into climatic zones based on the share of their arable land within the (a) equatorial, (b) arid, or (c) temperate & cold climate zones, based on the data provided in Gallup et al. (1999). The cut-off is 40%; for instance a country is grouped as equatorial if 40% or more of its arable land is in the equatorial climate zone. The following 12 countries are excluded from the analysis since they do not have 40% of more in either of the three categories: AFG, ARE, ECU, ETH, IND, IRN, KEN, KWT, MEX, OMN, QAT, YEM. Note that half of these 12 countries have 59% or more land in the 'highland' climatic zone. Further, the following 6 countries are contained in more than one climatic zone: AUS, BEN, LAO, VNM, ZAF, ZWE; since the own-correlation coefficients ($\hat{\rho}_{ii} = 1$) are excluded from the analysis, this does not affect the CD statistics.
The test statistics in bold are statistically significant at the 5% level and thus reject cross-section independence.

In Table 3.18 we apply an alternative 'spatial' grouping, whereby the share of arable land in one of three eco-climatic zones determines group membership. We use data from Matthews (1983), reported in Gallup et al. (1999), on this variable and group countries into equatorial (EQ), arid (AR) or temperate & cold (TEMP) climate zones if they have 40% or more arable land share in the respective climatic zone. We provide details on the countries that drop out of the sample as a result of this categorisation and the treatment of those countries that enter more than one category in the table footnote. Panels [a] to [e] again provide CD test statistics for regression residuals, Panel [f] shows the number of observations for each climate-zone combination. Again the pooled and standard country regression estimators fail most or all tests for residual independence within and across groups, although for 2FE only the temperate & cold zone shows high levels of dependence. The CCEMG residuals now cannot reject independence in any of the tests.

3.D Additional tables and figures

Table 3.19: Dynamic specification — Pooled regressions (CRS imposed)

Pooled regressions dependent variable: [1],[2] & [4]-[7] log output per worker, [3] dto. in 2FE transformation♮, [8] Δlog output per worker, [9] dto. in 2FE transformation♮						
	[1]	[2]	[3]	[4]	[5]	[6]
	POLS◇	**2FE**♮	**CCEP**	**CCEP**	**CCEP**	**CCEP**
weight matrix‡			none	neighbour	distance	agro-climate
lagged levels						
output pw (t − 1)	-0.0189 [5.71]**	-0.1764 [22.80]**	-0.6072 [22.54]**	-0.4024 [19.87]**	-0.4959 [21.44]**	-0.5699 [22.80]**
tractors pw (t − 1)	0.0022 [1.85]	0.0134 [5.47]**	0.0678 [7.02]**	0.0454 [7.51]**	0.0429 [5.90]**	0.0493 [5.76]**
livestock pw (t − 1)	0.0042 [2.49]*	0.0709 [10.07]**	0.2045 [9.67]**	0.1064 [5.89]**	0.1850 [9.31]**	0.2298 [10.32]**
fertilizer pw (t − 1)	0.0061 [5.86]**	0.0147 [7.76]**	0.0258 [5.79]**	0.0360 [8.45]**	0.0243 [4.10]**	0.0237 [5.64]**
land pw (t − 1)	0.0045 [2.28]*	0.0554 [7.28]**	0.1959 [5.35]**	0.1097 [4.46]**	0.1980 [6.38]**	0.1837 [5.77]**
growth rates						
Δtractors pw	0.0555 [4.12]**	0.0970 [8.49]**	0.0828 [4.37]**	0.0591 [3.55]**	0.0558 [4.61]**	0.0771 [4.82]**
Δlivestock pw	0.3153 [7.57]**	0.3307 [17.31]**	0.3097 [9.43]**	0.2899 [8.82]**	0.3175 [10.65]**	0.2726 [8.21]**
Δfertilizer pw	0.0123 [3.48]**	0.0172 [5.63]**	0.0132 [3.27]**	0.0196 [4.79]**	0.0171 [3.67]**	0.0152 [3.99]**
Δland pw	0.3010 [4.38]**	0.3290 [11.67]**	0.3249 [4.86]**	0.3272 [5.88]**	0.3404 [7.42]**	0.3170 [6.42]**
Constant	0.0485 [5.34]**		2.6725 [0.35]	1.7939 [1.75]	0.0141 [0.01]	3.5301 [1.85]
Observations	5,013	5,013	5,013	5,013	5,013	5,013
R-squared	0.12	0.20	0.54	0.54	0.81	0.60
RMSE	.084	.081	.061	.061	.039	.057
Estimated long-run coefficients♭‡						
tractors pw	0.118 [2.09]*	0.076 [5.58]**	0.112 [7.21]**	0.113 [7.73]**	0.087 [6.33]**	0.087 [5.80]**
livestock pw	0.224 [2.77]**	0.402 [11.30]**	0.337 [10.51]**	0.264 [6.30]**	0.373 [10.40]**	0.403 [11.27]**
fertilizer pw	0.322 [5.07]**	0.083 [8.09]**	0.042 [5.85]**	0.089 [8.67]**	0.049 [4.16]**	0.042 [5.80]**
land pw	0.238 [2.56]*	0.314 [7.95]**	0.323 [5.34]**	0.273 [4.50]**	0.399 [6.69]**	0.322 [5.93]**
Implied β_L	0.098	0.125	0.186	0.261	0.092	0.146
Arellano-Bond Serial Correlation Test — H_0: no serial correlation in the residuals						
AR(1) (p)	-9.69 (.00)	-14.95 (.00)	-3.30 (.00)	-6.46 (.00)	-4.71 (.00)	-3.88 (.00)
AR(2) (p)	-0.41 (.68)	3.15 (.00)	-1.28 (.20)	1.43 (.15)	-3.36 (.00)	0.06 (.95)
AR(3) (p)	0.60 (.55)	4.12 (.00)	-1.68 (.09)	0.39 (.70)	-0.82 (.41)	-1.98 (.05)
Nonstationarity: Pesaran (2007) CIPS test applied to residuals♯ — H_0: residuals are I(1)						
lags	Ztbar (p)	Ztbar (p)	Ztbar (p)	Ztbar (p)	Ztbar (p)	Ztbar (p)
none	-48.93 (.00)	-44.51 (.00)	-45.04 (.00)	-47.06 (.00)	-45.84 (.00)	-44.61 (.00)
1 lag	-38.33 (.00)	-29.68 (.00)	-32.18 (.00)	-32.43 (.00)	-32.08 (.00)	-31.67 (.00)
2 lags	-23.23 (.00)	-16.08 (.00)	-20.66 (.00)	-21.58 (.00)	-22.27 (.00)	-20.41 (.00)
3 lags	-15.77 (.00)	-8.30 (.00)	-16.05 (.00)	-14.15 (.00)	-16.51 (.00)	-15.44 (.00)
CSD: Mean Correlation coefficient & Pesaran (2004) CD test — H_0: no CSD						
Mean ρ_{ij}	0.00	0.01	0.00	0.00	0.10	0.00
Mean $\lvert\rho_{ij}\rvert$	0.14	0.15	0.17	0.14	0.16	0.15
CD (p)	-0.08 (.94)	2.66 (.00)	-0.47 (.64)	1.31 (.19)	26.34 (.00)	-0.62 (.53)
Further residual diagnostic tests⋆ — H_0: no heteroskedasticity, regular skewness/kurtosis						
(A) Joint test χ^2 (p)	1019.7 (.00)	943.6 (.00)	623.4 (.00)	551.0 (.00)	559.4 (.00)	566.1 (.00)
(B) Joint test χ^2 (p)	1463.7 (.00)	1166.1 (.00)	-	-	-	-

Notes: See Table 3.6 for details.
◇ We include $T-1$ year dummies in equation [1].
‡ We compute implied labour coefficients by subtracting the averaged long-run coefficients on the factor inputs in this model from unity (given that CRS is imposed).

Chapter 4

Estimation and Inference in Nonstationary Heterogeneous Panels with Multi-Factor Error Structure: New Insights

Abstract:

This chapter uses Monte Carlo simulations to investigate the impact of nonstationarity, parameter heterogeneity and cross-section dependence on estimation and inference in macro panel data. We compare the performance of standard panel estimators with that of our own two-step method (the AMG) and the Pesaran Common Correlated Effects (CCE) estimators in time-series panels with arguably similar characteristics to those encountered in the empirical chapters of this thesis. We first replicate the design and results from two recent studies on the topic (Coakley et al., 2006; Kapetanios et al., 2009) and show that simple augmentation with year dummies can do away with most of the bias in standard pooled estimators reported. Our own Monte Carlo design is able to replicate the patterns we observe for different estimators in the analysis of manufacturing and agriculture data somewhat more closely, namely identifying scenarios under which the CCEMG clearly outperforms the AMG. As was highlighted in previous chapters the empirical model we adopted leads to an identification problem in standard estimation approaches in the case where the same unobserved common factors drive the evolution of both dependent and independent variables. Our simulation results confirm that the Pesaran (2006) CCE approach as well as our AMG estimator solve this identification problem by implicitly accounting for the unobserved common factors in the regression equation. We further highlight the performance of the pooled OLS estimator with variables in first differences and speculate about the reasons for its favourable results.

> *"Monte Carlo results should be treated with some caution since the DGP [Data Generating Process] used in the simulation may not reflect factors typical in real applications. This may be the case here [results for the analysis published in Coakley et al. (2006)], since some of the applications discussed ... do not show the same features as the Monte Carlo results."*
> Smith and Fuertes (2007, p.76)

In this chapter we consider the estimation of the mean slope coefficient in a linear heterogeneous panel model where unobserved common factors lead to correlations in the disturbances across units and to correlations between the disturbances and the regressors. We employ Monte Carlo experiments to investigate the small sample properties of a number of standard pooled and heterogeneous parameter estimators (POLS, FE, FD-OLS, MG) as well as the recently developed Pesaran (2006) CCE estimators and our own contribution, the Augmented Mean Group (AMG) estimator in various contexts, including variable and factor nonstationarity.

Our investigation is motivated by the range of parameter estimates obtained from the production function regressions for manufacturing and agriculture in the previous chapters. We argued that the combination of technology heterogeneity, variable nonstationarity and cross-section dependence (CSD) led to severe distortions in standard panel estimators, and pointed to various diagnostic test results to support our claims. In the present chapter we now seek to provide some further evidence as to the importance of accounting for these matters in empirical analysis.

In our simulation exercises we first build on the setups presented in two recent papers from the literature on cross-section dependence in nonstationary panels, Coakley et al. (2006) and Kapetanios et al. (2009). We also present our own simulation setup and investigate estimation and inference for panel estimators in a number of scenarios.

The chapter makes the following contributions to this thesis and the wider literature:

(i) We compare the small sample properties of the AMG, standard (micro-)panel and Common Correlated Effects estimators in cross-sectionally dependent macro-panels. We are able to identify scenarios in which the performance of the AMG estimators matches that of the Pesaran (2006) CCEMG and CCEP, as well as circumstances under which there is

evidence of bias in the former but not the latter. We can thus provide an explanation for the differential pattern of results observed in the two empirical chapters of this thesis.

(ii) Our investigation highlights that in the recent simulation studies by Coakley et al. (2006) and Kapetanios et al. (2009) the 'naïve' pooled estimators (POLS, FE), which ignore cross-section dependence, are not specified in a manner that reflects practical empirical convention. Namely they do not contain year dummies, a standard augmentation in the applied literature to account for 'common shocks' to all units in the panel. Our Monte Carlo results reveal that once augmented with a set of year dummies the bias in these estimators drops dramatically in *all* of the scenarios studied, even if factor loadings on the unobserved common factors are heterogeneous across countries.

(iii) The simulations further highlight the performance of the pooled first difference estimator (FD-OLS). For cases where FD-OLS is biased (which also leads to bias in the AMG estimators) we suggest a simple IV version of this estimator and investigate its resulting performance (and that of the AMG based on first-stage IV estimates). We further briefly speculate on the origins of the good performance of FD-OLS in standard setups and provide some tentative evidence for our suggestion.

The remainder of the chapter is organised as follows: Section 4.1 introduces the Monte Carlo setups by Coakley et al. (2006) and Kapetanios et al. (2009). We replicate their results and show that introduction of year dummies in the 'naïve estimators' wipes out the most serious bias from cross-section dependence.[1] Section 4.2 then introduces our own simulation setup and the four scenarios for analysis. For the simulation setups which induce biased FD-OLS estimates *by construction* we further investigate an IV version of this estimator. The final section summarises and interprets the results of the chapter and concludes.

[1] Simulations were carried out in `GAUSS` version 9.

4.1 Two recent Monte Carlo studies on multifactor models

4.1.1 Coakley, Fuertes and Smith (2006)

The authors introduce the following Data Generating Process (DGP):

$$\begin{aligned} y_{it} &= \alpha_i + \beta x_{it} + u_{it} \qquad u_{it} = \rho_{ui} u_{i,t-1} + \lambda_i f_t + \varepsilon_{u,it} \\ &\varepsilon_{u,it} \sim \text{i.i.d. } N(0, \sigma_{ui}^2), \text{ where } \sigma_{ui}^2 = 1 \end{aligned} \tag{4.1}$$

for $i = 1, \ldots, N$ and $t = 1, \ldots, T$, where we adjust the notation to concentrate on the non-stationary observables settings with homogeneous β (Cases A-G). Coakley et al. (2006) do not report any simulation results for heterogeneous β but suggest that findings were rather similar to those for the homogeneous setup. The single regressor is defined as

$$\begin{aligned} x_{it} &= \rho_{xi} x_{i,t-1} + \phi_i f_t + \psi_i \chi_t + \varepsilon_{x,it} \\ &\varepsilon_{x,it} \sim \text{i.i.d. } N(0, \sigma_{xi}^2), \text{ where } \sigma_{xi} = \text{i.i.d. } U[0.5, 1.5] \end{aligned} \tag{4.2}$$

The unobserved common factors are generated as

$$f_t = \rho_f f_{t-1} + \varepsilon_{ft} \qquad \varepsilon_{ft} \sim \text{iid } N(0, \sigma_f^2), \text{ where } \sigma_f^2 = 1 \tag{4.3}$$

$$\chi_t = \rho_\chi \chi_{t-1} + \varepsilon_{\chi t} \qquad \varepsilon_{\chi t} \sim \text{iid } N(0, \sigma_\chi^2), \text{ where } \sigma_\chi^2 = 1 \tag{4.4}$$

Heterogeneous intercepts are distributed $\alpha_i \sim \text{iid } U[-0.5, 0.5]$ s.t. $\bar{\alpha} = 0$. Unless indicated the independently drawn factor loadings are heterogeneous across countries: $\lambda_i \sim \text{iid } U[0.5, 1.5]$, $\phi_i \sim \text{iid } U[0.5, 1.5]$ and $\psi_i \sim \text{iid } U[0.5, 1.5]$. Regressors are nonstationary ($\rho_{xi} = 1$) in all the cases presented here,[2] and unless indicated $\rho_f = \rho_\chi = 0$ (stationary common factors). The variation in the regressors (σ_{xi}) differs uniformly across countries. The slope coefficient is common and set to unity ($\beta = 1$).

With reference to our own empirical model applied in Chapters 2 and 3, we can highlight the following points of departure: *firstly*, in equation (4.1) Coakley et al. (2006) allow for serially correlated errors from *other* sources than the presence of unobserved common factors, which includes nonstationary u_{it} (noncointegration) regardless of the nature of the unobserved com-

[2] Further investigation in the Coakley et al. (2006) paper relates to stationary variables y and x, cases (a-f).

mon factors f_t. *Secondly*, in equation (4.2) the single regressor x is nonstationary for reasons *other than* the presence of I(1) common factors: this allows Coakley et al. (2006) to focus their investigation on the impact of *stationary common factors* f_t and χ_t on estimation and inference in a model with two *nonstationary observables* which do or do not cointegrate. *Thirdly*, the authors only allow for cointegration between y and x, but not between these observables and the unobservable common factors f — the presence of the latter is treated as a nuisance in the consistent estimation of the slope coefficient β.

As our later analysis shows, none of these issues lead to fundamental differences in the simulation results. With empirical cross-country production functions in mind we have highlighted the desirability of modelling unobservables (TFP) as a unit root process, as well as the heterogeneous nature of production technology (β_i) across countries, which will both be addressed in our own simulations as well as those by Kapetanios et al. (2009).

In detail, Coakley et al. (2006) consider the following scenarios:

Case *A*: $\rho_{ui} = 0$, $\lambda_i = \phi_i = \psi_i = 0$: Cointegration between y and x. No common factors and thus no cross-section dependence (CSD).

Case *B*: $\rho_{ui} = 1$, $\lambda_i = \phi_i = \psi_i = 0$: No cointegration between y and x. No CSD.

Case *C*: $\rho_{ui} = 1$, $\phi_i = 0$: No cointegration between y and x. An I(0) factor f_t drives the errors, a different I(0) factor χ_t drives the regressors.

Case *D*: $\rho_{ui} = 1$, $\psi_i = 0$: No cointegration between y and x. An I(0) factor f_t drives both the errors and the regressors.

Case $\tilde{D}$: Like Case D, but $\lambda_i = \phi_i$ for all i — we refer to this as factor loading dependence.

Case *E*: $\rho_{ui} = 0$, $\psi_i = 0$: Cointegration between y and x. An I(0) factor f_t drives both the errors and the regressors.

Case *F*: $\rho_{ui} = 1$: No cointegration between y and x. An I(0) factor f_t drives both the errors and the regressors, a different I(0) factor χ_t drives the regressors.

Case *G*: $\rho_f = \rho_\chi = 1$, $\rho_{ui} = 0$: No cointegration between y and x. An I(1) factor f_t drives both the errors and the regressors, a different I(1) factor χ_t drives the regressors.

By construction the simulations are primarily interested in the cointegrating relationship (or lack thereof) between y and x, and exclude the possibility of a three-way cointegrating relation (y, x, f). Furthermore, in most of the scenarios the unobserved common factors are stationary.

In the present and all the following Monte Carlo simulations we compare the small sample performance of the following estimators:[3]

Pooled estimators: POLS — pooled OLS, FE — pooled OLS with Fixed Effects, CCEP — pooled version of the Pesaran (2006) Common Correlated Effects estimator, FD-OLS — pooled OLS with variables in first differences. The estimation equations are augmented with year dummies as indicated in the results tables.

MG-type estimators: CCEMG — Mean Groups version of the Pesaran (2006) Common Correlated Effects estimator, AMG(i) — Augmented Mean Groups estimator with 'common dynamic process' imposed with unit coefficient, AMG(ii) with 'common dynamic process' included as additional regressor, MG — Pesaran and Smith (1995) Mean Groups estimator. All of these are based on averaged country-regression estimates, and we include linear trends in all but the CCEMG.

We present the simulation results across the sample of 5,000 replications for the panel dimensions $N = 30$, $T = 20$ in Table 4.1 in the Appendix. For each estimator we provide the mean, median and ('empirical') standard error of the 5,000 estimates, as well as the sample mean of the standard errors. This replicates the results in Table 3(II) of Coakley et al. (2006).

- In the baseline **Case A** with cointegration and cross-section independence all estimators are unbiased and due to the large variance in the I(1) regressors rather precise.

- The setup with nonstationary errors (**Case B**) represents a 'spurious panel regression' — as established by Phillips and Moon (1999) the pooled estimators in effect average across spurious regressions and provide unbiased estimates, although the empirical standard errors are much larger now, e.g. .1351 instead of .0182 for pooled FE without year dummies ('one-way FE', marked FE†).

- If we introduce cross-section dependence to the non-cointegration scenario (**Case C**) nothing much changes. This is because the omitted factors in the errors and the regres-

[3]Details on each of the regression models can be found in Section 2.1.2 of Chapter 2.

sors are independent. The exceptions are the FE estimator without year dummies (FE†) and the MG estimator, for which the factor f_t in the errors leads to a doubling of the empirical standard errors.

- In **Case D** the correlation between the regressors and the errors via the common factor f_t leads to serious bias in the pooled OLS and FE without year dummies (POLS†, FE†) and the MG estimator. POLS is much less biased at .0766 than FE at .4157. In either case the bias virtually disappears once year dummies are included in the estimation equation (POLS‡, FE‡) — we will speculate about the source of this benign correction in the conclusion of this chapter. The CCE and AMG estimators are unbiased and remain comparatively precise, though not dramatically more so than the POLS‡ or FE‡.

- For factor loading dependence between the errors and regressors (**Case $\tilde{\boldsymbol{D}}$**), whereby the impact of the common factor on the dependent variable and the covariate is *identical* in all countries, we observe a similar pattern of results across estimators, with the bias in POLS† and FE† slightly elevated. FD-OLS is biased for the first time and this bias naturally carries over to our AMG estimates, although the latter display only mild distortion.

- If y and x are cointegrated any correlation between the regressors and the errors via the common factor f_t leads to only modest bias in FE† and MG (**Case E**), since the correlation between the I(1) regressors and I(0) errors goes to zero with T.

- If several, rather than a single factor drive the regressors in the case of no cointegration between y and x and correlation between regressors and the errors (**Case F**) nothing much changes compared to the single factor scenario in Case D, except that the higher variation in the x leads to more precise estimates.

- Finally, the scenario where the unobserved factors are I(1), residuals are nonstationary and a common factor drives both y and x (**Case G**) we can observe the most serious bias of all cases considered here. The POLS† and FE† are biased by .2273 and .4374 respectively, while the bias for the MG is .5110 — all of these estimators are further very imprecise. Once we use year dummies for the pooled estimators, however, their bias goes to zero (POLS‡, FE‡) and the estimators are highly efficient. The CCE estimators are unbiased with relative precision, while the bias in the FD-OLS leads to bias in the AMG estimators — this time of similar magnitude.

In summary, our replication of the Monte Carlo results by Coakley et al. (2006) with alternative POLS‡ and FE‡ estimators, as well as our own AMG-type estimators for the cases considered cannot reveal any serious bias in the standard pooled estimators, provided year dummies are added to the estimation equation. The AMG estimators commonly perform similarly well to the Pesaran (2006) CCE estimators, with the notable exception of Case G (noncointegration even after nonstationary factors are accounted for).

4.1.2 Kapetanios, Pesaran and Yamagata (2008)

The authors introduce the following DGP:

$$y_{it} = \beta_i x_{it} + u_{it} \qquad u_{it} = \alpha_i + \lambda_{i1}^{y} f_{1t} + \lambda_{i2}^{y} f_{2t} + \varepsilon_{it} \tag{4.5}$$

$$x_{it} = a_{i1} + a_{i1} d_t + \lambda_{i1}^{x} f_{1t} + \lambda_{i3}^{x} f_{3t} + v_{it} \tag{4.6}$$

for $i = 1, \ldots, N$ unless indicated below and $t = 1, \ldots, T$, where we adjust the notation by Kapetanios et al. (2009) since we limit our analysis to the case with a single regressor (x).

The common deterministic trend term (d_t) and individual-specific errors for the x-equation are zero-mean independent AR(1) processes defined as

$$d_t = 0.5 d_{t-1} + \upsilon_{dt} \qquad \upsilon_{dt} \sim N(0, 0.75) \qquad t = -48, \ldots, 1, \ldots, T \qquad d_{-49} = 0$$

$$v_{it} = \rho_{vi} v_{i,t-1} + \upsilon_{it} \qquad \upsilon_{it} \sim N(0, (1 - \rho_{vi}^2)) \qquad t = -48, \ldots, 1, \ldots, T \qquad v_{i,-49} = 0$$

where $\rho_{vi} \sim U[0.05, 0.95]$. The three common factors are nonstationary processes

$$f_{jt} = f_{j,t-1} + \upsilon_{ft} \qquad j = 1, 2, 3 \qquad \upsilon_{ft} \sim N(0, 1) \tag{4.7}$$
$$t = -49, \ldots, 1, \ldots, T \qquad f_{j,-50} = 0$$

The authors generate innovations to y as a mix of heterogeneous AR(1) and MA(1) errors

$$\varepsilon_{it} = \rho_{i\varepsilon} \varepsilon_{i,t-1} + \sigma_i \sqrt{1 - \rho_{i\varepsilon}^2}\, \omega_{it} \qquad i = 1, \ldots, N_1 \qquad t = -48, \ldots, 0, \ldots, T$$

$$\varepsilon_{it} = \frac{\sigma_i}{\sqrt{1 + \theta_{i\varepsilon}^2}} (\omega_{it} + \theta_{i\varepsilon} \omega_{i,t-1}) \qquad i = N_1 + 1, \ldots, N \qquad t = -48, \ldots, 0, \ldots, T$$

where N_1 is the nearest integer to $N/2$ and $\omega_{it} \sim N(0, 1)$, $\sigma_i^2 \sim U[0.5, 1.5]$, $\rho_{i\varepsilon} \sim U[0.05, 0.95]$,

and $\theta_{i\varepsilon} \sim U[0,1]$. ρ_{vi}, $\rho_{i\varepsilon}$, $\theta_{i\varepsilon}$ and σ_i do not change across replications. Initial values are set to zero and the first 50 observations are discarded for all of the above.

Regarding parameter values, $\alpha_i \sim N(0,1)$ and a_{i1}, $a_{i2} \sim \text{iid}N(0.5, 0.5)$ do not change across replications. We limit ourselves to *'Experiment 1'* in Kapetanios et al. (2009), where $\beta_i = \beta + \eta_i$ with $\beta = 1$ and $\eta_i \sim N(0, 0.04)$. For the factor loadings the authors consider

$$\lambda^x_{i1} \sim N(0.5, 0.5) \qquad \text{and} \qquad \lambda^x_{i3} \sim N(0.5, 0.5) \tag{4.8}$$

$$\text{with either} \quad \mathcal{A}: \quad \lambda^y_{i1} \sim N(1, 0.2) \qquad \text{and} \qquad \lambda^y_{i2\mathcal{A}} \sim N(1, 0.2) \tag{4.9}$$

$$\text{or} \quad \mathcal{B}: \quad \lambda^y_{i1} \sim N(1, 0.2) \qquad \text{and} \qquad \lambda^y_{i2\mathcal{B}} \sim N(0, 1) \tag{4.10}$$

Since we are interested in consistent estimation of the mean parameter estimate ($\mathbb{E}[\beta_i]$) and therefore did not find considerable differences in the patterns of the results in setup $\mathcal{A}$ and $\mathcal{B}$ we only present the former to save space.[4]

With reference to our own empirical model we can state that the points of departure (e.g. the complex structure of innovations in y) are not substantial by any measure and were introduced by the authors to highlight the robustness of their results to a range of alternative sources of heterogeneity.

We investigate combinations of T and N for $T, N = \{20, 30, 50, 100\}$, but with 1,000 instead of the 2,000 replications in Kapetanios et al. (2009) for each case. Our results in Table 4.2 in the Appendix replicate those in Table 1 of Kapetanios et al. (2009). In addition to the mean, median, empirical standard errors and mean estimated standard errors we also report the average bias and the root mean squared error (RMSE), in line with the presentation in Kapetanios et al. (2009).[5] We also introduce 'infeasible' estimators, namely for fixed effects and MG — these represent estimators where the unobserved common factors in y are included in the estimation equation to provide a benchmark.

[4] In setup $\mathcal{B}$ the mean $\mathbb{E}[\beta_i]$ can be estimated consistently but not the individual β_i — see Kapetanios et al. (2009, p.6) for details.

[5] The bias is computed as $M^{-1}\sum_{m=1}^{M}\hat{\beta}_m - 1$, the average deviation across replications (here $M = 1,000$) of the estimate from the true mean parameter $\beta = 1$. The RMSE is computed as $\sqrt{M^{-1}\sum_{m=1}^{M}(\hat{\beta}_m - 1)^2}$, the average squared deviation across replications of the estimate from the true mean parameter. In case of both statistics we multiplied the results by 100.

The POLS and FE estimators without year dummies (marked †) indicate serious bias which increases in T but is stable as N increases. In all cases the bias in the one-way FE estimator (marked †) is larger. The standard MG estimator (with linear trend) similarly performs quite poorly, in general no better (or worse) than the FE estimator. In contrast the CCEP and FD-OLS (with $T-1$ year dummies) for the pooled case and the augmented MG-estimators display no bias. In data dimensions investigated the FD-OLS estimator has RMSE closest to the infeasible estimators.

The significant bias in the POLS and FE estimator however is almost entirely absent once these are augmented with $(T-1)$ year dummies (again marked ‡). RMSE are still slightly elevated for the latter two estimators, but on the whole the year dummies in the POLS and FE estimators can accommodate the cross-section dependence (as well as the other data properties) introduced in this setup quite well.

4.2 New simulations

Our analysis in the previous section has indicated that under the setups considered the use of 'naïve estimators' such as the pooled OLS or Fixed Effects estimators does not create considerable bias in the estimates provided the estimation equation is augmented with $(T-1)$ year dummies. This is surprising, since these estimators impose *common* factor loadings across countries on all unobserved common factors. Although the CCE and to a lesser extent the AMG estimators provide unbiased and more efficient estimates, the analysis seems to suggest that using estimators which account for cross-section dependence does not yield a dramatically different result from the use of standard pooled estimators that ignore it. This finding is somewhat at odds with our experience from the agriculture and manufacturing data, where we saw vast differences between the pooled OLS, Fixed Effects on the one hand (both augmented with year dummies), and the CCE and AMG-type estimators on the other. In the following we therefore present a new simulation setup and consider a number of alternative scenarios which create results somewhat more in line with those observed in our regressions with real data.

4.2.1 Our Monte Carlo setup

We define our dependent variable and regressor as

$$y_{it} = \beta_i x_{it} + u_{it} \qquad u_{it} = \alpha_i + \lambda^y_{i1} f_{1t} + \lambda^y_{i2} f_{2t} + \varepsilon_{it} \tag{4.11}$$

$$x_{it} = a_i + \lambda^x_{i1} f_{1t} + \lambda^x_{i3} f_{3t} + \epsilon_{it} \qquad \epsilon_{it} = \rho \epsilon_{i,t-1} + e_{it} \tag{4.12}$$

The serially-correlated x-variable is in practice constructed using a dynamic equation

$$x_{it} = (1-\rho)a_i + \lambda_i^{x_1} f_{1,t} - \rho \lambda_i^{x_1} f_{1,t-1} + \lambda_i^{x_3} f_{3,t} - \rho \lambda_i^{x_3} f_{3,t-1} + \rho x_{i,t-1} + e_{it}$$

which we begin with $x_{i,-49} = a_i$ and then accumulate for $t = -48, \ldots, 0, 1, \ldots, T$, discarding the first 50 time-series observations for all i. The common AR-coefficient is $\rho = .25$.

The unobserved common factors are nonstationary processes with individual drifts so as to ensure upward evolution over time, as observed in macro production data.

$$f_{j,t} = \mu_j + f_{j,t-1} + \upsilon_{fjt} \qquad t = -48, \ldots, 0, 1, \ldots, T \qquad f_{j,-49} = 0 \tag{4.13}$$

$$\upsilon_{fjt} \sim N\left(0, \sigma^2_{fj}\right) \qquad \sigma^2_{fj} = .00125 \qquad \mu_j = \{0.015, 0.012, 0.01\} \qquad j = 1, 2, 3$$

The error terms for the y and x equations are defined as

$$e_{it} \sim iid\, N(0, \sigma^2_{e,i}) \qquad \text{where } \sigma^2_{e,i} \sim U[.001, .003]$$

$$\varepsilon_{it} \sim iid\, N(0, \sigma^2_{\varepsilon}) \qquad \sigma^2_{\varepsilon} = .00125$$

The slope coefficient on x is set to $\beta_i = 1 + e^{\beta}_i$ where $e^{\beta}_i \sim U[-.25, +.25]$. The factor loadings are uniformly distributed, with λ^x_{i1} and λ^y_{i1} iid $U[0,1]$ respectively, and λ^x_{i3} and λ^y_{i2} iid $U[.25, 1.25]$ respectively.

We consider the following cases (as well as combinations of these)

(i) baseline (as above).

(ii) baseline with additional country-specific linear trends.

(iii) feedbacks: an idiosyncratic shock to y feeds back into x with one period lag.

(iv) two 'clubs' of countries with the same β coefficient.

For the linear country trend case, the trends are distributed $U[-.02, +.03]$, s.t. that the mean annual growth rate across countries is non-zero. For the feedback case, the lagged error ε_{t-1} from the y-equation in (4.11) is included in the x-equation in (4.12) with coefficient .25 (in practice we enter this term in the same way as the other terms in the dynamic equation as described above to allow for serial correlation in the x-equation). Finally, for the 'two clubs' case 20% of countries have $\beta = 2$, while 80% have $\beta = .75$, s.t. the mean β across all countries is still unity.

4.2.2 Results

Results for our benchmark specification — **Case (i)** — are presented in Table 4.3. For the remainder of this chapter the POLS and FE estimators always contain $(T-1)$ year dummies in their estimation equations. As can be seen both of these estimators are biased (increasing in T, decreasing in N) and in particular the POLS is very imprecise. For $T = 30$, $N = 50$ FE has only limited bias of .0324 with empirical standard error of .0876, compared to .0271 for the infeasible FE estimator. Similarly for the MG estimator, although the magnitude of the bias is higher. For the CCE and AMG estimators, all of which are unbiased, the AMG(ii) commonly is most efficient and very close to the standard errors for the infeasible estimators. The FD-OLS similarly performs very well.

Once we add the country-specific trend terms — **Case (ii)** — as presented in Table 4.4 the bias in the standard pooled estimators does not change by any significant margin, however their imprecision now increases with T. Again for $T = 30$, $N = 50$ FE has a bias of .0277, but a very substantial empirical standard error of .1973 (over twice that of the benchmark case), compared with .0280 for the infeasible FE estimator in the same setup. In contrast the unbiased CCE and AMG estimators are still efficient, although the FD-OLS estimator now emerges as the most efficient feasible estimator. The MG estimator is seriously biased and inefficient.

By construction, the feedback setup — **Case (iii)** — leads to bias in the FD-OLS, which carries over to the AMG estimators, as can be seen in Table 4.5 (AMG estimates marked †): due to first differencing the lagged error $\varepsilon_{i,t-1}$ is contained in both the errors and the regressors of the FD-OLS estimation equation, whereas this is not the case in the other (levels-based)

estimators which account for common factors. We therefore also present the results for an IV-version of the FD-OLS estimator, where we use growth rates at time $(t-1)$ as instruments for the endogenous growth rates at time t (FD-IV), and AMG estimators which are based on the year dummies from the instrumented regression (AMG estimates marked ‡). The pooled OLS, FE and MG results are virtually unchanged from the baseline results: for $T = 30$, $N = 50$ FE has a bias of .0299 with empirical standard error of .0865 compared with .0271 for the infeasible FE estimator. The augmented estimators all display small bias, albeit very modest in case of the CCE estimators, while the new AMG estimates based on the FD-IV results are unbiased. The latter is unbiased, but inefficient compared with the new AMG estimators.

If we combine the feedback setup with linear country trends (results presented in Table 4.6) the pattern of bias does not change considerably, although the empirical standard errors for POLS and FE increase by around a third: for $T = 30$, $N = 50$ FE has a bias of .0253 with empirical standard error of .1943 — this is driven by the presence of the additional trend terms, since results for Case (ii) are very similar.

In the setup where β is heterogeneous but only takes two values for different 'clubs' of countries — **Case (iv)** — the results in Table 4.7 show considerable bias for the POLS estimator, while the results for all other estimators remain relatively similar to those presented before. For $T = 30$, $N = 50$ the FE estimator has a bias of .0224 and an empirical standard error of .1375 compared with .0357 for the infeasible FE.

We also ran the simulations for this setup with additional linear country trends, with results in Table 4.8. This specification leads to some bias and very serious imprecision in the standard pooled estimators, but hardly has any impact on the CCE and AMG estimators vis-à-vis the setup without country-specific linear trends. For $T = 30$, $N = 50$ the FE estimator has a bias of .0202 and an empirical standard error of .2275 compared with .0364 for the infeasible FE.

4.3 Interpretation and concluding remarks

We carried out a range of Monte Carlo simulations to investigate the properties of various standard and novel panel estimators in the face on nonstationary data series and unobserved common factors. Our review of earlier studies with the same aim (Coakley et al., 2006; Kapetanios et al., 2009) and replication of their results shows that the supposed serious bias in standard pooled estimators under the Monte Carlo setup specified by these authors can be reduced substantially by the use of year dummies in the pooled OLS and Fixed Effects estimators — a standard practice in the applied literature. Since our empirical results in previous chapters however showed very significant differences between pooled and heterogeneous CCE and AMG-type estimators, we specified our own simulations so as to identify under which scenario we can observe the same patterns in the simulation results as in the empirical results.

In particular the specification of heterogeneous country trends led to some bias but primarily imprecision in the POLS, FE and MG estimators, while the FD-OLS, AMG and CCE estimators remained unbiased and comparatively precise. Similarly if in addition to heterogeneous country trends we specify two distinct 'clubs' of countries with differing technology parameter β. Overall, the issues of nonstationarity in observables and unobservables as well as the heterogeneity in their impact considered in this Chapter seem to lead to inefficiency, rather than serious bias in the standard pooled and Mean Group panel estimators.

Note however that from the perspective of an applied economist this finding of serious inefficiency is observationally equivalent to severe bias in the estimator: a single 'draw' from the 1,000 replications carried out during our Monte Carlo simulations would yield a highly distorted parameter estimate in the pooled OLS, Fixed Effects (regardless of year dummy augmentation) and standard MG regression case, but a rather precise estimate in case of the estimators accounting for cross-section dependence. As an illustrative example we compute the 95% confidence interval (CI) of the mean for each estimator in the baseline scenario presented in Table 4.3 ($T = 30$, $N = 50$).[6]

[6] As discussed in Hendry and Nielsen (2007, p.272/3) the mean of the simulated statistic (here: regression coefficient on β) provides an unbiased estimate of the unknown mean of the statistic, and the simulation sample variance is an unbiased estimator of the unknown population variance of the statistic. Note that these authors refer to the 'empirical standard error' as the Monte Carlo standard deviation ($\bar{\omega}$) and use the term Monte Carlo standard error $M^{-1/2}\bar{\omega}$ to refer to a measure of precision of the simulation exercise. See footnote 7 below.

Pooled Estimators					
	mean	*median*	*emp.ste.*	*mean ste.*	*95% CI (mean)*
POLS	0.9754	0.9815	0.2139	0.0413	[0.5562 , 1.3946]
FE	1.0324	1.0312	0.0876	0.0269	[0.8607 , 1.2041]
CCEP	0.9995	0.9975	0.0333	0.0222	[0.9342 , 1.0648]
FD-OLS	1.0021	1.0015	0.0342	0.0237	[0.9351 , 1.0691]
FE (inf)	1.0000	0.9996	0.0271	0.0161	[0.9469 , 1.0531]
MG-type Estimators					
	mean	*median*	*emp.ste.*	*mean ste.*	*95% CI (mean)*
CCEMG	0.9992	0.9975	0.0338	0.0327	[0.9330 , 1.0654]
AMG(i)	1.0026	1.0008	0.0323	0.0319	[0.9393 , 1.0659]
AMG(ii)	1.0018	1.0004	0.0326	0.0304	[0.9379 , 1.0657]
MG	1.1259	1.1143	0.1825	0.0388	[0.7682 , 1.4836]
MG (inf)	0.9999	0.9989	0.0267	0.0267	[0.9476 , 1.0522]

Thus if we ran a single regression with real-world data following this DGP, then in 95% of the cases with a true parameter average of unity, the POLS estimate we would obtain is as likely to be .56 as 1.39, the FE estimate is as likely .87 as 1.20, the MG estimate is as likely .77 as 1.48... while the AMG estimate is as likely to be .94 as 1.06. In each case the t-statistic associated with the obtained estimate would indicate the parameter to be highly significant (due to the uniformly low standard errors across all models, as indicated by their means in our simulation studies) and thus very precisely estimated.[7] The invalidity of t-statistics and their distortion away from zero in the case where regression residuals are nonstationary has been highlighted in the literature by Kao (1999). Our numerical example indicates that while the unobserved common factor structures we considered here may not lead to biased estimates, inference is severely affected in the cases where these common factors are not accounted for.

We furthermore provided evidence for the differential performance of the AMG and CCEMG estimators as was found in the analysis for agriculture in Chapter 3.

The estimators accounting for cross-section dependence aside, these simulations particularly highlight the performance of the FD-OLS estimator: with the exception of the feedback case, this estimator is unbiased and has very similar empirical standard errors as the CCE and AMG

[7] This result is not due to the imprecision of the Monte Carlo experiment: based on the empirical standard error (the standard deviation of the $m = 1,000$ estimates $\hat{b}_i$) we can compute the standard error of the experiment as $m^{-1/2}$ times this value to yield (in case of pooled FE) $2.77 * 10^{-3}$. Thus the *precision of the experiment* is such that the true mean estimate lies between 1.0269 and 1.0379 with 95% confidence.

estimators. This is somewhat puzzling, since the FD-OLS model does not account for unobserved common effects with *heterogeneous* factor loadings: with reference to our own DGP in equation (4.11) we have $\lambda_{i1}^{y}\Delta f_{1t}$ and $\lambda_{i2}^{y}\Delta f_{2t}$ as part of the error term, where $\bar{\lambda}_1^y = .5$ and $\bar{\lambda}_2^y = .75$ and Δf_{1t}, Δf_{2t} have non-zero means (due to the presence of the drift terms). This setup thus specifies *non-zero average impact* for the unobserved factors, but in the FD-OLS regressions augmented with year dummies the correlation this induces with the regressors does not seem to create any bias. Thus even though $\boldsymbol{\beta}$ or the mean of the $\boldsymbol{\beta}_i$ is in theory not identified in this case (separately from $\boldsymbol{\lambda}_i$ — see our discussion in Section 1.3 of Chapter 1) the FD-OLS estimator yields an unbiased and efficient result.

A possible explanation for this phenomenon is that the year dummies lead to a 're-centering' of these factor-loadings around zero, such that the presence of the same factors in the errors and regressors (endogeneity) does not create any problems for consistent estimation of $\mathbb{E}[\beta_i]$, since the correlation between regressor and error terms is on average zero. We can provide the following, tentative evidence for this: adopting the baseline specification — Case (i) — from our own Monte Carlo setup we ran simulations where we dropped the year dummies in the FD-OLS estimators. The resulting bias in the estimates for the case of $\bar{\lambda}_1^y = .5$ and $\bar{\lambda}_2^y = .75$ is considerable, e.g. for $T = 30$, $N = 50$ we get a bias of .1159 and an empirical standard error of .0565 (1,000 replications). Recall that the results for the estimator with year dummies were unbiased with an empirical standard error of .0342. Still keeping FD-OLS without year dummies, we then re-center the factor-loadings around zero, which yields unbiased estimates and an empirical standard error of .0314. This pattern where FD-OLS is seriously biased if mean factor loadings are non-zero and year dummies are missing, but unbiased if mean factor loadings are zero, points to the year dummies as fulfilling the role described above. Future research may provide further evidence to support this speculative explanation.

Appendix to Chapter 4: Monte Carlo simulations

The estimators investigated are: **POLS** — pooled OLS, **FE** — Fixed Effects, **CCEP** — Pesaran (2006) pooled CCE estimator, **FD-OLS** — pooled OLS with variables in first difference, **CCEMG** — Pesaran (2006) Mean Group CCE estimator, **AMG(i)** — Augmented Mean Group estimator, $\hat{\mu}_t^{\bullet}$ imposed with unit coefficient, **AMG(ii)** — AMG estimator with $\hat{\mu}_t^{\bullet}$ included as regressor, **MG** — Pesaran and Smith (1995) Mean Group estimator. Two 'infeasible estimators' (for FE and MG) contain the unobserved common factors as regressors.

Table 4.1: Coakley, Fuertes and Smith (2006)

Monte Carlo Results — replicating Coakley, Fuertes and Smith (2006)
5,000 replications; $N = 30$, $T = 20$; year dummies in the POLS or FE estimation equations: † — no, ‡ — yes; AMG-estimators are constructed from FD-OLS year dummy coefficients

	Case *A* Cointegration, no CSD				**Case *B*** No cointegration, no CSD				**Case *C*** No cointegration, CSD				**Case *D*** No cointegration, CSD, endogenous x			
Pooled Estimators																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
POLS†	1.0001	1.0001	0.0109	0.0055	0.9986	1.0005	0.2155	0.0397	0.9971	1.0013	0.2071	0.0399	1.0766	1.0774	0.2099	0.0397
POLS‡	1.0002	1.0002	0.0088	0.0054	0.9987	0.9995	0.2165	0.0407	0.9981	0.9983	0.2182	0.0409	1.0169	1.0170	0.2185	0.0409
FE†	1.0003	1.0004	0.0182	0.0180	1.0037	1.0038	0.1351	0.0404	0.9973	1.0038	0.2808	0.0410	1.4157	1.4065	0.2012	0.0363
FE‡	1.0004	1.0005	0.0186	0.0185	1.0041	1.0026	0.1381	0.0414	1.0034	1.0009	0.1389	0.0415	1.0208	1.0182	0.1420	0.0416
CCEP	1.0003	1.0003	0.0232	0.0226	1.0049	1.0039	0.1154	0.0421	1.0029	1.0043	0.1137	0.0418	1.0034	1.0041	0.1148	0.0420
FD-OLS	1.0014	1.0006	0.0574	0.0573	1.0010	1.0013	0.0413	0.0405	1.0006	1.0003	0.0413	0.0406	1.0120	1.0124	0.0442	0.0406
MG-type Estimators																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
CCEMG	0.9998	1.0002	0.0335	0.0321	1.0045	1.0062	0.1314	0.1260	1.0027	1.0011	0.1281	0.1239	1.0035	1.0022	0.1303	0.1255
AMG(i)	1.0001	0.9997	0.0392	0.0373	1.0022	1.0031	0.1104	0.1072	1.0009	1.0018	0.0689	0.0727	1.0069	1.0071	0.0695	0.0788
AMG(ii)	1.0000	0.9995	0.0289	0.0273	1.0051	1.0052	0.1410	0.1379	1.0025	1.0022	0.0828	0.0897	1.0056	1.0084	0.1327	0.1319
MG	1.0001	0.9998	0.0283	0.0274	1.0047	1.0039	0.1626	0.1595	0.9985	1.0054	0.3017	0.1338	1.5059	1.4880	0.2196	0.1306

	Case $\tilde{D}$ like Case D, factor loading dependence				**Case *E*** Cointegration, CSD, endogenous x				**Case *F*** like Case D, additional I(0) factor in x				**Case *G*** No cointegration, I(1) factors			
Pooled Estimators																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
POLS†	1.1317	1.1288	0.2181	0.0394	1.0072	1.0066	0.0109	0.0071	1.0648	1.0679	0.1938	0.0374	1.2273	1.2128	0.2439	0.0251
POLS‡	1.0088	1.0088	0.2174	0.0408	1.0003	1.0003	0.0106	0.0055	1.0078	1.0098	0.2101	0.0394	1.0016	1.0010	0.0049	0.0010
FE†	1.4437	1.4352	0.2096	0.0356	1.0516	1.0510	0.0345	0.0188	1.2775	1.2685	0.2069	0.0319	1.4374	1.4640	0.5928	0.0217
FE‡	1.0133	1.0123	0.1398	0.0415	1.0015	1.0014	0.0185	0.0185	1.0124	1.0105	0.1352	0.0401	1.0006	1.0004	0.0077	0.0028
CCEP	1.0051	1.0051	0.1147	0.0420	1.0004	1.0006	0.0234	0.0228	1.0035	1.0046	0.1135	0.0416	1.0031	1.0037	0.0934	0.0416
FD-OLS	1.0725	1.0707	0.0477	0.0405	1.0125	1.0122	0.0599	0.0574	1.0111	1.0102	0.0425	0.0392	1.0647	1.0456	0.1323	0.0303
MG-type Estimators																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
CCEMG	1.0056	1.0035	0.1300	0.1256	0.9998	0.9995	0.0336	0.0322	1.0034	1.0029	0.1271	0.1234	1.0021	1.0030	0.1037	0.0988
AMG(i)	1.0381	1.0394	0.0690	0.0782	1.0068	1.0064	0.0447	0.0272	1.0090	1.0101	0.0564	0.0608	1.0627	1.0444	0.1308	0.0490
AMG(ii)	1.0170	1.0180	0.1314	0.1309	1.0071	1.0059	0.0396	0.0186	1.0088	1.0104	0.0806	0.0877	1.0654	1.0479	0.1341	0.0252
MG	1.5057	1.4972	0.2060	0.1238	1.0812	1.0796	0.0456	0.0259	1.3266	1.3134	0.2112	0.1018	1.5110	1.4921	0.7386	0.1585

Notes: For each estimator we report the mean and median for the 5,000 estimates of β. $*$ emp. ste refers to the empirical standard error, the standard deviation of the 5,000 estimates of β; mean ste refers to the sample mean of the estimated standard errors in the 5,000 estimations of β. See main text for simulation setup and detailed description of the cases.
The estimators are: POLS — pooled OLS, FE — Fixed Effects, CCEP — Pesaran (2006) pooled CCE estimator, FD-OLS — pooled OLS with variables in first difference, CCEMG — Pesaran (2006) Mean Group CCE estimator, AMG(i) — Augmented Mean Group estimator, $\hat{\mu}_t^{\bullet}$ imposed with unit coefficient, AMG(ii) — Augmented Mean Group estimator, $\hat{\mu}_t^{\bullet}$ included as regressor, MG — Pesaran and Smith (1995) Mean Group estimator. ‡ (†) We do (not) include $T - 1$ year dummies.

Table 4.2: Kapetanios, Pesaran and Yamagata (2008)

Monte Carlo Results — replicating Kapetanios, Pesaran and Yamagata (2008)
1,000 replications; year dummies in the POLS or FE estimation equations: † — no, ‡ — yes; AMG-estimators are constructed from FD-OLS year dummy coefficients

$T = 20$	$N = 20$						$N = 30$						$N = 50$						$N = 100$					
	mean	median	emp. ste.*	mean ste.*	Bias x 100	RMSE x 100	mean	median	emp. ste.*	mean ste.*	Bias x 100	RMSE x 100	mean	median	emp. ste.*	mean ste.*	Bias x 100	RMSE x 100	mean	median	emp. ste.*	mean ste.*	Bias x 100	RMSE x 100
Pooled Estimators																								
POLS†	1.028	1.021	0.197	0.046	2.78	19.85	1.038	1.026	0.163	0.037	3.80	16.69	1.050	1.038	0.144	0.029	5.01	15.26	1.047	1.032	0.115	0.020	4.69	12.42
POLS‡	0.989	0.992	0.181	0.040	-1.09	18.15	0.986	0.992	0.142	0.033	-1.45	14.27	0.996	0.999	0.119	0.026	-0.37	11.85	0.995	0.992	0.081	0.018	-0.47	8.13
FE†	1.224	1.201	0.296	0.062	22.37	37.07	1.213	1.194	0.272	0.048	21.28	34.52	1.231	1.218	0.280	0.039	23.11	36.31	1.228	1.213	0.265	0.027	22.84	35.00
FE‡	0.996	0.994	0.107	0.041	-0.41	10.72	0.999	0.995	0.085	0.031	-0.14	8.54	1.000	0.999	0.070	0.026	-0.02	7.04	1.002	1.003	0.046	0.018	0.21	4.56
CCEP	0.998	1.001	0.089	0.044	-0.17	8.89	0.999	0.995	0.073	0.034	-0.11	7.32	1.001	1.000	0.061	0.030	0.13	6.10	1.002	1.001	0.041	0.020	0.21	4.05
FD-OLS	0.998	0.998	0.074	0.042	-0.21	7.41	0.999	1.000	0.058	0.031	-0.11	5.77	1.001	1.002	0.050	0.028	0.06	4.96	1.001	1.001	0.033	0.019	0.13	3.29
FE (inf)	1.002	1.001	0.068	0.034	0.16	6.81	1.000	0.999	0.053	0.025	-0.04	5.30	0.999	0.999	0.045	0.023	-0.13	4.48	1.001	1.002	0.030	0.015	0.09	3.03
MG-type Estimators																								
CCEMG	0.998	0.997	0.088	0.084	-0.25	8.75	1.000	0.997	0.074	0.070	-0.02	7.42	1.002	1.001	0.062	0.059	0.16	6.17	1.001	1.001	0.041	0.041	0.10	4.10
AMG(i)	0.997	0.999	0.080	0.075	-0.31	8.00	0.996	0.997	0.065	0.062	-0.37	6.51	1.001	1.003	0.057	0.053	0.05	5.71	1.000	1.001	0.037	0.036	-0.01	3.69
AMG(ii)	0.997	0.997	0.078	0.075	-0.26	7.79	0.998	0.998	0.066	0.063	-0.19	6.55	1.002	1.002	0.057	0.053	0.18	5.74	1.001	1.001	0.037	0.036	0.06	3.71
MG	1.217	1.184	0.286	0.163	21.74	35.91	1.209	1.187	0.261	0.133	20.88	33.45	1.230	1.208	0.270	0.113	22.99	35.45	1.230	1.211	0.251	0.077	23.00	34.05
MG (inf)	1.003	1.004	0.063	0.063	0.25	6.33	0.999	0.999	0.052	0.052	-0.14	5.22	0.999	0.999	0.047	0.045	-0.09	4.71	0.999	0.999	0.030	0.030	-0.08	3.01

$T = 30$	$N = 20$						$N = 30$						$N = 50$						$N = 100$					
	mean	median	emp. ste.*	mean ste.*	Bias x 100	RMSE x 100	mean	median	emp. ste.*	mean ste.*	Bias x 100	RMSE x 100	mean	median	emp. ste.*	mean ste.*	Bias x 100	RMSE x 100	mean	median	emp. ste.*	mean ste.*	Bias x 100	RMSE x 100
Pooled Estimators																								
POLS†	1.064	1.050	0.196	0.038	6.43	20.62	1.066	1.049	0.172	0.030	6.61	18.41	1.054	1.039	0.144	0.023	5.45	15.40	1.061	1.042	0.124	0.016	6.10	13.82
POLS‡	1.015	1.020	0.174	0.032	1.51	17.41	1.006	1.000	0.140	0.026	0.59	13.96	0.994	0.997	0.108	0.020	-0.60	10.82	1.001	1.002	0.074	0.014	0.09	7.37
FE†	1.253	1.240	0.318	0.051	25.34	40.65	1.240	1.216	0.287	0.040	23.96	37.40	1.241	1.224	0.285	0.031	24.14	37.37	1.243	1.226	0.283	0.022	24.26	37.29
FE‡	1.002	1.001	0.113	0.032	0.16	11.24	1.006	1.005	0.087	0.025	0.57	8.68	0.999	1.003	0.069	0.020	-0.10	6.87	1.002	1.000	0.050	0.014	0.16	4.98
CCEP	0.998	1.000	0.093	0.036	-0.17	9.31	1.001	1.001	0.070	0.027	0.07	6.96	1.001	1.004	0.056	0.022	0.10	5.63	1.001	1.002	0.041	0.015	0.08	4.12
FD-OLS	1.001	0.999	0.075	0.038	0.13	7.49	1.003	1.000	0.055	0.026	0.27	5.50	0.998	0.999	0.042	0.021	-0.23	4.21	1.002	1.002	0.032	0.015	0.19	3.17
FE (inf)	1.001	0.998	0.066	0.027	0.11	6.56	1.002	1.003	0.053	0.020	0.16	5.28	0.998	0.997	0.041	0.016	-0.22	4.11	1.001	1.002	0.029	0.011	0.11	2.92
MG-type Estimators																								
CCEMG	0.997	0.997	0.088	0.083	-0.33	8.82	1.000	0.999	0.067	0.065	0.04	6.73	1.001	1.002	0.053	0.052	0.12	5.31	1.003	1.004	0.039	0.038	0.25	3.94
AMG(i)	0.998	1.001	0.084	0.078	-0.22	8.44	1.003	1.004	0.063	0.059	0.34	6.33	0.999	1.000	0.048	0.048	-0.07	4.84	1.002	1.004	0.036	0.034	0.24	3.61
AMG(ii)	0.999	0.998	0.085	0.080	-0.14	8.53	1.002	1.002	0.062	0.061	0.20	6.16	0.999	1.000	0.050	0.049	-0.09	4.97	1.002	1.002	0.036	0.035	0.22	3.60
MG	1.247	1.223	0.320	0.183	24.65	40.36	1.231	1.204	0.275	0.137	23.12	35.95	1.241	1.223	0.270	0.111	24.12	36.17	1.243	1.219	0.263	0.079	24.27	35.77
MG (inf)	0.998	0.997	0.060	0.060	-0.17	6.01	1.001	1.000	0.046	0.046	0.08	4.62	0.999	0.999	0.036	0.037	-0.14	3.60	1.002	1.002	0.026	0.026	0.15	2.58

Continued on the following page.

Kapetanios, Pesaran and Yamagata (2008) — continued

$T = 50$	$N = 20$						$N = 30$						$N = 50$						$N = 100$					
	mean	median	emp. ste.*	mean ste.*	Bias x 100	RMSE x 100	mean	median	emp. ste.*	mean ste.*	Bias x 100	RMSE x 100	mean	median	emp. ste.*	mean ste.*	Bias x 100	RMSE x 100	mean	median	emp. ste.*	mean ste.*	Bias x 100	RMSE x 100
Pooled Estimators																								
POLS†	1.083	1.058	0.207	0.029	8.31	22.25	1.083	1.063	0.183	0.024	8.30	20.10	1.083	1.061	0.165	0.018	8.31	18.48	1.078	1.052	0.147	0.013	7.84	16.69
POLS‡	0.998	0.995	0.167	0.024	-0.25	16.70	1.000	1.003	0.129	0.020	0.02	12.87	1.000	0.998	0.098	0.015	-0.04	9.79	0.997	1.000	0.071	0.011	-0.36	7.10
FE†	1.263	1.239	0.334	0.039	26.27	42.48	1.259	1.240	0.315	0.031	25.85	40.71	1.268	1.257	0.315	0.025	26.84	41.38	1.268	1.253	0.306	0.017	26.79	40.67
FE‡	0.997	0.998	0.114	0.024	-0.28	11.43	0.999	0.998	0.092	0.019	-0.14	9.22	0.999	1.001	0.070	0.015	-0.07	6.96	1.001	0.999	0.051	0.011	0.06	5.08
CCEP	1.006	1.005	0.092	0.025	0.55	9.25	1.000	1.001	0.074	0.019	0.02	7.38	1.002	1.005	0.061	0.016	0.24	6.05	1.001	1.001	0.042	0.011	0.08	4.24
FD-OLS	1.001	0.998	0.067	0.027	0.13	6.70	1.000	0.998	0.053	0.020	-0.01	5.33	1.002	1.004	0.042	0.017	0.22	4.17	1.001	1.002	0.029	0.012	0.08	2.93
FE (inf)	0.999	0.998	0.060	0.017	-0.07	6.02	0.998	1.000	0.050	0.013	-0.17	4.95	1.001	1.002	0.041	0.011	0.08	4.08	1.001	1.001	0.029	0.008	0.10	2.90
MG-type Estimators																								
CCEMG	1.005	1.008	0.087	0.083	0.48	8.71	1.000	1.003	0.070	0.068	-0.03	7.02	1.002	1.003	0.057	0.055	0.23	5.68	1.000	1.000	0.040	0.038	-0.02	3.95
AMG(i)	1.003	0.998	0.077	0.073	0.27	7.70	1.002	1.003	0.063	0.062	0.19	6.27	1.004	1.006	0.053	0.051	0.37	5.30	1.001	1.002	0.036	0.035	0.08	3.58
AMG(ii)	1.005	1.002	0.077	0.075	0.47	7.75	1.003	1.004	0.064	0.063	0.25	6.42	1.003	1.004	0.053	0.052	0.29	5.31	1.001	1.001	0.037	0.036	0.06	3.73
MG	1.263	1.241	0.336	0.180	26.29	42.65	1.266	1.236	0.316	0.148	26.58	41.24	1.277	1.246	0.304	0.123	27.73	41.11	1.277	1.254	0.294	0.083	27.65	40.33
MG (inf)	1.000	1.003	0.051	0.050	0.01	5.06	1.000	1.000	0.042	0.042	0.04	4.15	1.002	1.002	0.035	0.033	0.18	3.46	1.000	1.001	0.024	0.023	0.04	2.41

$T = 100$	$N = 20$						$N = 30$						$N = 50$						$N = 100$					
	mean	median	emp. ste.*	mean ste.*	Bias x 100	RMSE x 100	mean	median	emp. ste.*	mean ste.*	Bias x 100	RMSE x 100	mean	median	emp. ste.*	mean ste.*	Bias x 100	RMSE x 100	mean	median	emp. ste.*	mean ste.*	Bias x 100	RMSE x 100
Pooled Estimators																								
POLS†	1.128	1.106	0.221	0.022	12.77	25.54	1.124	1.100	0.202	0.018	12.41	23.68	1.116	1.091	0.182	0.014	11.61	21.62	1.118	1.089	0.171	0.010	11.80	20.78
POLS‡	1.007	1.004	0.158	0.018	0.72	15.84	1.008	1.003	0.131	0.014	0.81	13.12	0.999	0.999	0.097	0.011	-0.11	9.73	1.001	1.003	0.070	0.008	0.07	7.02
FE†	1.318	1.289	0.346	0.028	31.78	46.97	1.322	1.312	0.337	0.023	32.19	46.61	1.319	1.312	0.324	0.018	31.89	45.47	1.322	1.311	0.320	0.012	32.23	45.40
FE‡	1.001	1.000	0.121	0.017	0.05	12.11	1.002	1.002	0.098	0.014	0.24	9.81	0.998	0.997	0.076	0.010	-0.21	7.63	1.000	1.001	0.053	0.007	-0.01	5.29
CCEP	1.001	0.997	0.103	0.016	0.10	10.27	1.007	1.007	0.088	0.014	0.69	8.79	1.003	1.004	0.065	0.010	0.28	6.50	1.004	1.004	0.047	0.008	0.36	4.72
FD-OLS	1.002	1.002	0.065	0.019	0.22	6.46	1.002	1.002	0.053	0.016	0.16	5.33	1.000	1.003	0.039	0.012	-0.05	3.93	1.001	1.001	0.028	0.008	0.12	2.82
FE (inf)	1.002	1.001	0.064	0.010	0.18	6.41	1.000	1.000	0.057	0.008	0.01	5.68	1.000	1.001	0.042	0.006	0.04	4.21	1.000	1.000	0.029	0.004	0.04	2.91
MG-type Estimators																								
CCEMG	1.003	1.001	0.099	0.090	0.27	9.88	1.007	1.006	0.081	0.076	0.71	8.15	1.004	1.003	0.062	0.059	0.36	6.21	1.003	1.001	0.043	0.043	0.25	4.34
AMG(i)	1.001	1.001	0.081	0.079	0.12	8.04	1.003	1.002	0.067	0.069	0.25	6.73	1.001	1.002	0.053	0.052	0.05	5.32	1.001	1.002	0.037	0.038	0.09	3.72
AMG(ii)	1.003	1.000	0.083	0.080	0.34	8.25	1.007	1.007	0.073	0.070	0.72	7.32	1.002	1.004	0.054	0.054	0.22	5.39	1.002	1.003	0.039	0.039	0.18	3.93
MG	1.334	1.298	0.366	0.207	33.41	49.55	1.351	1.327	0.361	0.181	35.10	50.37	1.338	1.310	0.327	0.137	33.84	47.05	1.343	1.326	0.312	0.099	34.33	46.35
MG (inf)	1.003	1.002	0.049	0.047	0.25	4.90	1.002	1.002	0.042	0.039	0.16	4.16	1.002	1.001	0.031	0.030	0.15	3.08	1.000	1.001	0.022	0.021	0.03	2.18

Notes: See Table 4.1 and main text for details. FE (inf) and MG (inf) are 'infeasible estimators' where the true unobserved common factors are included in the regression. ‡ (†) We do (not) include $T - 1$ year dummies.

Table 4.3: New Simulations — (i) Baseline setup

Monte Carlo Results — Baseline Setup
1,000 replications; POLS, FE and FD-OLS all have $T-1$ year dummies; AMG-estimators are constructed from FD-OLS year dummy coefficients

$T = 20$	$N = 20$				$N = 30$				$N = 50$				$N = 100$			
Pooled Estimators																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
POLS	1.0481	1.0618	0.3660	0.0793	1.0448	1.0364	0.2875	0.0618	0.9689	0.9628	0.2142	0.0508	0.9896	0.9845	0.1384	0.0328
FE	1.0543	1.0483	0.1205	0.0499	1.0188	1.0188	0.0934	0.0402	1.0211	1.0201	0.0703	0.0312	1.0093	1.0086	0.0479	0.0218
CCEP	1.0014	0.9994	0.0584	0.0444	0.9999	1.0018	0.0491	0.0365	1.0006	1.0011	0.0370	0.0282	1.0014	0.9998	0.0268	0.0200
FD-OLS	1.0057	1.0054	0.0648	0.0466	1.0016	1.0008	0.0534	0.0377	1.0029	1.0027	0.0396	0.0291	1.0013	1.0005	0.0292	0.0204
FE (inf)	1.0019	1.0028	0.0474	0.0344	1.0003	0.9991	0.0403	0.0281	1.0008	1.0015	0.0309	0.0219	1.0009	1.0001	0.0221	0.0154
MG-type Estimator																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
CCEMG	1.0013	1.0008	0.0598	0.0586	0.9983	1.0003	0.0498	0.0483	1.0004	1.0007	0.0382	0.0376	1.0016	1.0008	0.0277	0.0269
AMG(i)	1.0059	1.0050	0.0598	0.0530	1.0015	1.0021	0.0500	0.0439	1.0040	1.0044	0.0373	0.0344	1.0021	1.0005	0.0271	0.0246
AMG(ii)	1.0046	1.0028	0.0590	0.0499	1.0013	1.0022	0.0492	0.0421	1.0031	1.0041	0.0376	0.0328	1.0023	1.0011	0.0270	0.0233
MG	1.1076	1.1013	0.1651	0.0656	1.1261	1.1160	0.1725	0.0543	1.1128	1.1002	0.1582	0.0421	1.1205	1.1114	0.1656	0.0299
MG (inf)	1.0007	1.0000	0.0488	0.0493	0.9992	0.9981	0.0408	0.0405	1.0003	1.0014	0.0317	0.0314	1.0007	0.9996	0.0224	0.0222

$T = 30$	$N = 20$				$N = 30$				$N = 50$				$N = 100$			
Pooled Estimators																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
POLS	1.0517	1.0593	0.3582	0.0649	1.0370	1.0269	0.2895	0.0507	0.9754	0.9815	0.2139	0.0413	0.9908	0.9940	0.1406	0.0268
FE	1.0735	1.0703	0.1536	0.0431	1.0258	1.0257	0.1178	0.0346	1.0324	1.0312	0.0876	0.0269	1.0111	1.0069	0.0602	0.0188
CCEP	1.0018	1.0049	0.0514	0.0350	1.0012	1.0007	0.0438	0.0287	0.9995	0.9975	0.0333	0.0222	1.0007	1.0006	0.0241	0.0157
FD-OLS	1.0035	1.0052	0.0552	0.0381	1.0037	1.0045	0.0454	0.0308	1.0021	1.0015	0.0342	0.0237	1.0009	1.0004	0.0248	0.0167
FE (inf)	1.0012	1.0035	0.0438	0.0255	1.0014	1.0023	0.0347	0.0207	1.0000	0.9996	0.0271	0.0161	1.0002	1.0003	0.0197	0.0113
MG-type Estimator																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
CCEMG	1.0007	1.0017	0.0517	0.0497	1.0009	1.0014	0.0436	0.0420	0.9992	0.9975	0.0338	0.0327	1.0003	1.0002	0.0241	0.0237
AMG(i)	1.0064	1.0081	0.0523	0.0488	1.0041	1.0036	0.0435	0.0405	1.0026	1.0008	0.0323	0.0319	1.0024	1.0024	0.0237	0.0229
AMG(ii)	1.0035	1.0036	0.0517	0.0461	1.0043	1.0048	0.0429	0.0386	1.0018	1.0004	0.0326	0.0304	1.0024	1.0023	0.0231	0.0217
MG	1.1284	1.1263	0.1827	0.0604	1.1520	1.1369	0.1864	0.0502	1.1259	1.1143	0.1825	0.0388	1.1378	1.1356	0.1839	0.0278
MG (inf)	1.0012	1.0038	0.0431	0.0419	1.0016	1.0019	0.0336	0.0344	0.9999	0.9989	0.0267	0.0267	1.0002	1.0000	0.0194	0.0190

Continued on the following page.

New Simulations — (i) Baseline setup (continued)

$T = 50$	$N = 20$				$N = 30$				$N = 50$				$N = 100$			
Pooled Estimators																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
POLS	1.0502	1.0698	0.3640	0.0504	1.0342	1.0343	0.2919	0.0388	0.9857	0.9825	0.2057	0.0318	0.9893	0.9971	0.1392	0.0209
FE	1.1156	1.1189	0.2044	0.0357	1.0381	1.0356	0.1529	0.0285	1.0451	1.0468	0.1163	0.0221	1.0165	1.0140	0.0823	0.0155
CCEP	1.0024	1.0023	0.0480	0.0264	0.9993	0.9988	0.0405	0.0218	0.9997	1.0001	0.0317	0.0168	0.9996	1.0003	0.0217	0.0119
FD-OLS	1.0055	1.0031	0.0493	0.0295	1.0006	1.0008	0.0387	0.0239	1.0018	1.0022	0.0312	0.0184	1.0003	1.0004	0.0217	0.0129
FE (inf)	1.0009	1.0000	0.0393	0.0172	0.9995	0.9997	0.0324	0.0141	1.0000	1.0001	0.0257	0.0109	0.9997	0.9995	0.0177	0.0077
MG-type Estimator																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
CCEMG	1.0026	0.9999	0.0475	0.0459	0.9982	0.9993	0.0405	0.0387	0.9994	0.9997	0.0310	0.0300	1.0001	0.9999	0.0213	0.0217
AMG(i)	1.0075	1.0064	0.0474	0.0479	1.0024	1.0027	0.0385	0.0398	1.0040	1.0035	0.0301	0.0312	1.0024	1.0023	0.0211	0.0224
AMG(ii)	1.0048	1.0036	0.0464	0.0444	1.0016	1.0020	0.0375	0.0372	1.0024	1.0022	0.0301	0.0290	1.0018	1.0020	0.0207	0.0209
MG	1.1700	1.1564	0.2160	0.0595	1.1761	1.1669	0.2123	0.0499	1.1613	1.1496	0.2088	0.0384	1.1641	1.1584	0.2148	0.0275
MG (inf)	1.0006	1.0017	0.0368	0.0370	0.9996	1.0001	0.0314	0.0300	0.9998	0.9996	0.0241	0.0236	0.9996	1.0005	0.0170	0.0166

$T = 100$	$N = 20$				$N = 30$				$N = 50$				$N = 100$			
Pooled Estimators																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
POLS	1.0973	1.1043	0.3540	0.0349	1.0422	1.0282	0.2762	0.0273	0.9990	0.9964	0.2148	0.0221	0.9993	1.0001	0.1434	0.0145
FE	1.1469	1.1565	0.2527	0.0266	1.0446	1.0479	0.1911	0.0212	1.0557	1.0553	0.1433	0.0166	1.0233	1.0196	0.1042	0.0115
CCEP	1.0068	1.0045	0.0535	0.0185	1.0017	1.0011	0.0428	0.0153	0.9999	1.0001	0.0343	0.0119	0.9984	0.9987	0.0260	0.0084
FD-OLS	1.0063	1.0051	0.0427	0.0208	1.0034	1.0033	0.0346	0.0169	1.0021	1.0023	0.0266	0.0130	1.0002	1.0007	0.0195	0.0091
FE (inf)	1.0019	1.0012	0.0375	0.0101	1.0018	1.0019	0.0302	0.0082	0.9997	0.9996	0.0226	0.0063	0.9994	0.9991	0.0171	0.0045
MG-type Estimator																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
CCEMG	1.0044	1.0024	0.0509	0.0478	0.9998	0.9992	0.0415	0.0408	0.9987	0.9986	0.0328	0.0323	0.9996	0.9992	0.0249	0.0229
AMG(i)	1.0089	1.0075	0.0461	0.0503	1.0052	1.0055	0.0363	0.0420	1.0054	1.0042	0.0291	0.0336	1.0032	1.0038	0.0212	0.0238
AMG(ii)	1.0056	1.0039	0.0436	0.0459	1.0049	1.0046	0.0358	0.0387	1.0026	1.0027	0.0280	0.0308	1.0022	1.0031	0.0209	0.0219
MG	1.2078	1.1970	0.2549	0.0617	1.2084	1.2021	0.2516	0.0519	1.1932	1.1815	0.2678	0.0412	1.1944	1.1858	0.2601	0.0295
MG (inf)	1.0010	0.9995	0.0349	0.0337	1.0014	1.0020	0.0279	0.0277	0.9993	0.9992	0.0209	0.0216	0.9997	0.9996	0.0155	0.0152

Notes: See main text for details of Monte Carlo setup.

Table 4.4: New Simulations — (ii) Additional country trend

Monte Carlo Results — Baseline Setup
1,000 replications; POLS, FE and FD-OLS all have $T-1$ year dummies; AMG-estimators are constructed from FD-OLS year dummy coefficients

$T=20$	$N=20$				$N=30$				$N=50$				$N=100$			
Pooled Estimators																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
POLS	1.0517	1.0587	0.4729	0.1068	1.0321	1.0318	0.3800	0.0839	0.9679	0.9616	0.2921	0.0675	0.9914	0.9951	0.1949	0.0441
FE	1.0403	1.0364	0.2249	0.0905	1.0201	1.0226	0.1840	0.0732	1.0210	1.0192	0.1429	0.0563	1.0113	1.0123	0.0998	0.0396
CCEP	0.9986	0.9999	0.0726	0.0526	0.9995	0.9990	0.0595	0.0435	1.0025	1.0022	0.0445	0.0333	1.0015	1.0006	0.0307	0.0236
FD-OLS	1.0050	1.0067	0.0670	0.0497	1.0015	0.9995	0.0536	0.0401	1.0027	1.0006	0.0393	0.0309	1.0008	1.0000	0.0287	0.0217
FE (inf)	1.0023	1.0036	0.0520	0.0384	0.9998	0.9991	0.0427	0.0313	1.0012	1.0017	0.0322	0.0243	1.0008	1.0004	0.0234	0.0172
MG-type Estimator																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
CCEMG	0.9990	0.9994	0.0741	0.0692	0.9981	0.9968	0.0622	0.0582	1.0022	1.0018	0.0456	0.0441	1.0017	1.0015	0.0320	0.0317
AMG(i)	1.0050	1.0065	0.0947	0.0537	1.0035	1.0054	0.0745	0.0445	1.0054	1.0049	0.0603	0.0347	1.0041	1.0033	0.0458	0.0250
AMG(ii)	1.0161	1.0074	0.1054	0.0718	1.0155	1.0143	0.0819	0.0586	1.0152	1.0115	0.0672	0.0446	1.0120	1.0109	0.0505	0.0315
MG	1.1092	1.1037	0.1686	0.0656	1.1254	1.1144	0.1742	0.0544	1.1129	1.0965	0.1579	0.0423	1.1203	1.1148	0.1650	0.0300
MG (inf)	1.0032	1.0017	0.0545	0.0537	1.0016	1.0021	0.0451	0.0436	1.0029	1.0027	0.0357	0.0338	1.0029	1.0028	0.0255	0.0241

$T=30$	$N=20$				$N=30$				$N=50$				$N=100$			
Pooled Estimators																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
POLS	1.0654	1.0722	0.4969	0.0893	1.0241	1.0298	0.3888	0.0708	0.9731	0.9688	0.3102	0.0567	0.9948	1.0017	0.2054	0.0371
FE	1.0718	1.0761	0.3278	0.0888	1.0109	1.0138	0.2497	0.0717	1.0277	1.0283	0.1973	0.0552	1.0108	1.0159	0.1400	0.0389
CCEP	1.0028	1.0021	0.0612	0.0431	1.0015	0.9991	0.0553	0.0355	0.9991	1.0003	0.0395	0.0275	1.0003	1.0018	0.0286	0.0194
FD-OLS	1.0036	1.0038	0.0556	0.0402	1.0028	1.0024	0.0473	0.0325	1.0025	1.0031	0.0351	0.0250	1.0009	1.0017	0.0243	0.0176
FE (inf)	1.0009	1.0019	0.0446	0.0282	1.0011	1.0001	0.0372	0.0230	0.9998	0.9995	0.0280	0.0179	1.0004	1.0002	0.0191	0.0126
MG-type Estimator																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
CCEMG	1.0019	1.0017	0.0618	0.0603	1.0007	0.9984	0.0554	0.0508	0.9997	0.9997	0.0404	0.0400	1.0002	1.0008	0.0295	0.0284
AMG(i)	1.0041	1.0044	0.0839	0.0495	1.0058	1.0089	0.0697	0.0405	1.0049	1.0047	0.0506	0.0322	1.0019	1.0014	0.0378	0.0229
AMG(ii)	1.0100	1.0071	0.0894	0.0665	1.0130	1.0144	0.0752	0.0555	1.0090	1.0080	0.0558	0.0432	1.0061	1.0053	0.0412	0.0300
MG	1.1262	1.1248	0.1824	0.0603	1.1506	1.1427	0.1853	0.0502	1.1269	1.1185	0.1848	0.0390	1.1379	1.1361	0.1826	0.0277
MG (inf)	1.0013	1.0005	0.0452	0.0441	1.0028	1.0025	0.0375	0.0361	1.0017	1.0019	0.0286	0.0281	1.0012	1.0011	0.0198	0.0200

Continued on the following page.

New Simulations — (ii) Additional country trend (continued)

$T = 50$	$N = 20$				$N = 30$				$N = 50$				$N = 100$			
Pooled Estimators																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
POLS	1.0407	1.0480	0.5318	0.0724	1.0288	1.0308	0.4090	0.0572	0.9913	0.9831	0.3194	0.0462	0.9868	0.9830	0.2078	0.0304
FE	1.1113	1.1193	0.4462	0.0821	1.0354	1.0517	0.3627	0.0658	1.0486	1.0570	0.2810	0.0512	1.0137	1.0202	0.2003	0.0361
CCEP	1.0053	1.0051	0.0640	0.0356	0.9999	0.9987	0.0511	0.0290	0.9999	1.0007	0.0409	0.0223	0.9991	0.9984	0.0288	0.0159
FD-OLS	1.0058	1.0038	0.0502	0.0309	1.0006	1.0013	0.0395	0.0250	1.0009	1.0012	0.0318	0.0192	1.0007	1.0005	0.0221	0.0135
FE (inf)	1.0015	1.0004	0.0406	0.0192	0.9999	1.0003	0.0330	0.0157	0.9988	0.9998	0.0255	0.0122	1.0001	1.0002	0.0181	0.0086
MG-type Estimator																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
CCEMG	1.0055	1.0051	0.0628	0.0586	0.9989	0.9983	0.0506	0.0482	0.9997	1.0004	0.0398	0.0375	0.9993	0.9992	0.0285	0.0270
AMG(i)	1.0040	1.0080	0.0767	0.0483	0.9991	0.9996	0.0597	0.0395	1.0011	1.0005	0.0461	0.0311	1.0027	1.0021	0.0339	0.0224
AMG(ii)	1.0061	1.0068	0.0829	0.0694	1.0023	1.0013	0.0655	0.0561	1.0027	1.0034	0.0495	0.0437	1.0038	1.0041	0.0360	0.0308
MG	1.1724	1.1591	0.2178	0.0600	1.1755	1.1653	0.2131	0.0496	1.1606	1.1530	0.2088	0.0383	1.1651	1.1605	0.2149	0.0276
MG (inf)	1.0013	0.9999	0.0400	0.0383	1.0002	1.0008	0.0314	0.0311	1.0000	0.9997	0.0245	0.0243	1.0010	1.0011	0.0177	0.0171

$T = 100$	$N = 20$				$N = 30$				$N = 50$				$N = 100$			
Pooled Estimators																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
POLS	1.1143	1.1221	0.5645	0.0569	1.0373	1.0306	0.4429	0.0454	0.9945	0.9797	0.3600	0.0359	1.0010	1.0050	0.2425	0.0238
FE	1.1824	1.1952	0.6226	0.0652	1.0394	1.0153	0.4984	0.0537	1.0535	1.0629	0.3844	0.0409	1.0211	1.0263	0.2716	0.0289
CCEP	1.0051	1.0048	0.0673	0.0280	0.9999	0.9996	0.0586	0.0231	1.0017	1.0009	0.0457	0.0179	0.9992	1.0002	0.0314	0.0126
FD-OLS	1.0063	1.0070	0.0429	0.0217	1.0018	1.0014	0.0337	0.0176	1.0032	1.0024	0.0270	0.0135	1.0020	1.0019	0.0188	0.0095
FE (inf)	1.0013	1.0009	0.0364	0.0114	0.9997	1.0004	0.0301	0.0093	1.0002	0.9994	0.0241	0.0072	1.0012	1.0010	0.0166	0.0051
MG-type Estimator																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
CCEMG	1.0038	1.0017	0.0639	0.0619	0.9976	0.9975	0.0524	0.0530	0.9998	0.9999	0.0429	0.0417	1.0005	1.0016	0.0294	0.0295
AMG(i)	1.0013	1.0019	0.0683	0.0508	1.0043	1.0062	0.0516	0.0421	1.0043	1.0049	0.0425	0.0336	1.0049	1.0056	0.0284	0.0238
AMG(ii)	1.0003	0.9993	0.0732	0.0733	1.0040	1.0050	0.0542	0.0602	1.0037	1.0016	0.0448	0.0479	1.0042	1.0050	0.0311	0.0337
MG	1.2072	1.1991	0.2528	0.0621	1.2091	1.2019	0.2514	0.0521	1.1948	1.1830	0.2692	0.0413	1.1966	1.1871	0.2593	0.0295
MG (inf)	1.0012	1.0021	0.0343	0.0345	1.0002	0.9999	0.0279	0.0280	1.0006	1.0004	0.0221	0.0218	1.0017	1.0017	0.0156	0.0154

Notes: See main text for details of Monte Carlo setup.

Table 4.5: New Simulations — (iii) Feedback setup

Monte Carlo Results — Setup with Feedbacks from y to x
1,000 replications; POLS, FE and FD-OLS all have $T-1$ year dummies; AMG-estimators are constructed from † FD-OLS or ‡ FD-IV year dummy coefficients

$T = 20$	$N = 20$				$N = 30$				$N = 50$				$N = 100$			
Pooled Estimators																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
POLS	1.0481	1.0618	0.3660	0.0793	1.0448	1.0374	0.2874	0.0618	0.9688	0.9629	0.2141	0.0508	0.9896	0.9845	0.1384	0.0328
FE	1.0485	1.0427	0.1183	0.0493	1.0140	1.0117	0.0923	0.0397	1.0163	1.0133	0.0691	0.0309	1.0048	1.0036	0.0473	0.0216
CCEP	0.9823	0.9814	0.0578	0.0436	0.9805	0.9833	0.0492	0.0358	0.9812	0.9822	0.0361	0.0277	0.9823	0.9806	0.0265	0.0196
FD-OLS	0.9181	0.9195	0.0631	0.0467	0.9142	0.9127	0.0530	0.0377	0.9154	0.9146	0.0391	0.0291	0.9142	0.9140	0.0287	0.0204
FD-IV	0.9951	0.9876	0.1662	0.0474	0.9963	0.9936	0.1282	0.0381	1.0027	1.0033	0.1036	0.0293	0.9978	0.9981	0.0706	0.0205
FE (inf)	0.9892	0.9902	0.0472	0.0340	0.9875	0.9867	0.0406	0.0278	0.9880	0.9881	0.0304	0.0216	0.9883	0.9879	0.0224	0.0152
MG-type Estimator																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
CCEMG	0.9772	0.9758	0.0589	0.0576	0.9740	0.9764	0.0495	0.0473	0.9762	0.9761	0.0370	0.0368	0.9772	0.9765	0.0271	0.0264
AMG(i)†	0.9585	0.9597	0.0596	0.0526	0.9540	0.9541	0.0508	0.0435	0.9580	0.9576	0.0378	0.0340	0.9569	0.9562	0.0288	0.0244
AMG(i)‡	0.9922	0.9890	0.0877	0.0532	0.9913	0.9931	0.0704	0.0438	0.9939	0.9926	0.0571	0.0343	0.9914	0.9894	0.0432	0.0245
AMG(ii)†	0.9528	0.9534	0.0573	0.0500	0.9508	0.9514	0.0494	0.0419	0.9537	0.9528	0.0375	0.0327	0.9539	0.9535	0.0286	0.0234
AMG(ii)‡	1.0055	0.9958	0.1005	0.0507	1.0030	0.9990	0.0773	0.0422	1.0037	0.9981	0.0663	0.0328	0.9991	0.9974	0.0490	0.0233
MG	1.0918	1.0852	0.1627	0.0648	1.1105	1.1012	0.1692	0.0535	1.0970	1.0828	0.1554	0.0415	1.1048	1.0956	0.1625	0.0295
MG (inf)	0.9829	0.9815	0.0483	0.0490	0.9814	0.9819	0.0410	0.0400	0.9826	0.9834	0.0311	0.0311	0.9831	0.9823	0.0225	0.0220

$T = 30$	$N = 20$				$N = 30$				$N = 50$				$N = 100$			
Pooled Estimators																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
POLS	1.0518	1.0588	0.3581	0.0648	1.0370	1.0271	0.2894	0.0507	0.9754	0.9818	0.2138	0.0413	0.9908	0.9938	0.1406	0.0268
FE	1.0697	1.0647	0.1513	0.0428	1.0232	1.0231	0.1163	0.0343	1.0299	1.0285	0.0865	0.0267	1.0088	1.0055	0.0596	0.0186
CCEP	0.9888	0.9915	0.0507	0.0343	0.9883	0.9892	0.0432	0.0282	0.9867	0.9851	0.0330	0.0219	0.9880	0.9882	0.0238	0.0155
FD-OLS	0.9162	0.9177	0.0547	0.0377	0.9162	0.9165	0.0447	0.0305	0.9149	0.9136	0.0341	0.0235	0.9139	0.9138	0.0243	0.0165
FD-IV	0.9948	0.9924	0.1312	0.0381	1.0009	1.0006	0.1052	0.0308	1.0004	0.9993	0.0813	0.0237	0.9973	0.9989	0.0569	0.0166
FE (inf)	0.9934	0.9963	0.0436	0.0252	0.9938	0.9943	0.0345	0.0205	0.9924	0.9923	0.0271	0.0159	0.9926	0.9926	0.0198	0.0112
MG-type Estimator																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
CCEMG	0.9847	0.9873	0.0510	0.0490	0.9845	0.9855	0.0431	0.0413	0.9828	0.9819	0.0333	0.0322	0.9841	0.9843	0.0238	0.0233
AMG(i)†	0.9581	0.9580	0.0529	0.0485	0.9563	0.9569	0.0442	0.0401	0.9552	0.9541	0.0340	0.0316	0.9560	0.9558	0.0258	0.0227
AMG(i)‡	0.9979	0.9985	0.0801	0.0486	0.9978	0.9968	0.0653	0.0403	0.9959	0.9953	0.0486	0.0317	0.9949	0.9954	0.0375	0.0227
AMG(ii)†	0.9516	0.9520	0.0513	0.0460	0.9528	0.9535	0.0429	0.0385	0.9511	0.9503	0.0338	0.0303	0.9527	0.9532	0.0255	0.0217
AMG(ii)‡	1.0033	1.0008	0.0874	0.0464	1.0061	1.0004	0.0722	0.0386	1.0015	0.9995	0.0542	0.0303	0.9997	0.9991	0.0415	0.0216
MG	1.1179	1.1159	0.1801	0.0596	1.1413	1.1239	0.1839	0.0496	1.1157	1.1043	0.1799	0.0384	1.1274	1.1261	0.1810	0.0274
MG (inf)	0.9901	0.9913	0.0430	0.0416	0.9906	0.9907	0.0336	0.0341	0.9888	0.9884	0.0265	0.0265	0.9892	0.9891	0.0195	0.0189

Continued on the following page.

New Simulations — (iii) Feedback setup (continued)

$T = 50$	$N = 20$				$N = 30$				$N = 50$				$N = 100$			
Pooled Estimators																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
POLS	1.0501	1.0700	0.3639	0.0503	1.0342	1.0344	0.2919	0.0388	0.9857	0.9823	0.2056	0.0318	0.9893	0.9969	0.1391	0.0209
FE	1.1136	1.1164	0.2024	0.0355	1.0369	1.0349	0.1515	0.0284	1.0439	1.0458	0.1153	0.0220	1.0156	1.0125	0.0816	0.0154
CCEP	0.9947	0.9935	0.0474	0.0260	0.9917	0.9914	0.0400	0.0214	0.9922	0.9925	0.0314	0.0166	0.9921	0.9924	0.0214	0.0117
FD-OLS	0.9179	0.9165	0.0489	0.0290	0.9131	0.9125	0.0385	0.0235	0.9145	0.9143	0.0307	0.0181	0.9132	0.9137	0.0212	0.0127
FD-IV	0.9973	0.9963	0.1028	0.0293	0.9944	0.9924	0.0836	0.0236	0.9977	1.0005	0.0659	0.0182	0.9972	0.9971	0.0439	0.0128
FE (inf)	0.9970	0.9969	0.0392	0.0171	0.9957	0.9960	0.0323	0.0139	0.9962	0.9960	0.0257	0.0108	0.9960	0.9959	0.0177	0.0076
MG-type Estimator																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
CCEMG	0.9927	0.9899	0.0470	0.0453	0.9886	0.9881	0.0399	0.0381	0.9897	0.9895	0.0306	0.0296	0.9903	0.9901	0.0209	0.0214
AMG(i)†	0.9543	0.9541	0.0491	0.0477	0.9503	0.9509	0.0395	0.0395	0.9528	0.9516	0.0318	0.0310	0.9515	0.9512	0.0238	0.0222
AMG(i)‡	1.0001	1.0000	0.0711	0.0477	0.9960	0.9977	0.0564	0.0395	0.9987	0.9990	0.0444	0.0310	0.9978	0.9977	0.0303	0.0222
AMG(ii)†	0.9485	0.9470	0.0469	0.0443	0.9461	0.9459	0.0380	0.0371	0.9477	0.9485	0.0310	0.0290	0.9475	0.9477	0.0236	0.0209
AMG(ii)‡	1.0021	0.9992	0.0774	0.0446	0.9993	0.9978	0.0612	0.0371	1.0007	0.9999	0.0486	0.0290	0.9994	0.9992	0.0327	0.0208
MG	1.1634	1.1486	0.2130	0.0588	1.1694	1.1597	0.2094	0.0493	1.1548	1.1428	0.2061	0.0380	1.1577	1.1515	0.2118	0.0272
MG (inf)	0.9946	0.9958	0.0368	0.0369	0.9938	0.9943	0.0312	0.0299	0.9939	0.9940	0.0242	0.0235	0.9937	0.9943	0.0170	0.0166

$T = 100$	$N = 20$				$N = 30$				$N = 50$				$N = 100$			
Pooled Estimators																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
POLS	1.0973	1.1041	0.3540	0.0349	1.0423	1.0277	0.2761	0.0273	0.9990	0.9964	0.2148	0.0221	0.9993	1.0000	0.1434	0.0145
FE	1.1463	1.1546	0.2518	0.0266	1.0443	1.0476	0.1904	0.0211	1.0554	1.0557	0.1428	0.0165	1.0231	1.0194	0.1039	0.0115
CCEP	1.0031	1.0007	0.0526	0.0182	0.9981	0.9978	0.0423	0.0150	0.9964	0.9967	0.0338	0.0117	0.9949	0.9952	0.0256	0.0083
FD-OLS	0.9189	0.9181	0.0429	0.0204	0.9162	0.9165	0.0346	0.0165	0.9147	0.9146	0.0264	0.0127	0.9132	0.9139	0.0194	0.0089
FD-IV	0.9966	0.9926	0.0807	0.0205	1.0010	1.0005	0.0601	0.0166	0.9988	0.9977	0.0473	0.0128	0.9989	0.9982	0.0327	0.0090
FE (inf)	1.0005	0.9998	0.0375	0.0101	1.0004	1.0009	0.0302	0.0082	0.9983	0.9985	0.0226	0.0063	0.9981	0.9976	0.0171	0.0045
MG-type Estimator																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
CCEMG	0.9997	0.9978	0.0501	0.0470	0.9952	0.9947	0.0410	0.0401	0.9940	0.9947	0.0323	0.0317	0.9948	0.9944	0.0244	0.0225
AMG(i)†	0.9478	0.9458	0.0473	0.0503	0.9447	0.9460	0.0384	0.0418	0.9448	0.9443	0.0315	0.0334	0.9432	0.9429	0.0248	0.0237
AMG(i)‡	1.0011	0.9987	0.0667	0.0501	1.0030	1.0027	0.0497	0.0418	1.0021	1.0019	0.0397	0.0334	1.0011	1.0013	0.0282	0.0237
AMG(ii)†	0.9422	0.9410	0.0447	0.0461	0.9415	0.9413	0.0374	0.0389	0.9388	0.9395	0.0295	0.0310	0.9392	0.9402	0.0239	0.0220
AMG(ii)‡	1.0001	0.9962	0.0689	0.0460	1.0038	1.0026	0.0519	0.0386	1.0007	1.0008	0.0409	0.0307	1.0008	1.0009	0.0294	0.0218
MG	1.2041	1.1925	0.2523	0.0612	1.2048	1.1968	0.2491	0.0514	1.1897	1.1785	0.2651	0.0408	1.1909	1.1814	0.2575	0.0292
MG (inf)	0.9986	0.9970	0.0348	0.0337	0.9990	0.9997	0.0279	0.0277	0.9969	0.9966	0.0209	0.0215	0.9974	0.9971	0.0155	0.0152

Notes: ‡ These use the year dummy coefficients from FD-IV estimator, rather than the FD-OLS estimator.

Table 4.6: New Simulations — (iii)* Feedback and country trend

Monte Carlo Results — Setup with Feedbacks from y to x
1,000 replications; POLS, FE and FD-OLS all have $T-1$ year dummies; AMG-estimators are constructed from † FD-OLS or ‡ FD-IV year dummy coefficients

$T=20$	$N=20$				$N=30$				$N=50$				$N=100$			
Pooled Estimators																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
POLS	1.0517	1.0595	0.4728	0.1068	1.0321	1.0324	0.3799	0.0839	0.9678	0.9614	0.2920	0.0675	0.9914	0.9953	0.1949	0.0441
FE	1.0345	1.0289	0.2202	0.0895	1.0157	1.0187	0.1807	0.0724	1.0162	1.0150	0.1402	0.0557	1.0068	1.0107	0.0979	0.0392
CCEP	0.9793	0.9802	0.0714	0.0517	0.9804	0.9789	0.0590	0.0427	0.9827	0.9819	0.0434	0.0327	0.9824	0.9807	0.0303	0.0232
FD-OLS	0.9173	0.9163	0.0656	0.0484	0.9142	0.9141	0.0529	0.0391	0.9153	0.9138	0.0385	0.0301	0.9137	0.9130	0.0283	0.0212
FD-IV	0.9968	0.9981	0.1686	0.0490	0.9954	0.9893	0.1293	0.0395	1.0007	0.9996	0.1034	0.0304	0.9973	0.9964	0.0723	0.0213
FE (inf)	0.9834	0.9840	0.0518	0.0379	0.9808	0.9806	0.0427	0.0309	0.9824	0.9828	0.0317	0.0240	0.9821	0.9820	0.0236	0.0169
MG-type Estimator																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
CCEMG	0.9745	0.9730	0.0726	0.0678	0.9738	0.9747	0.0612	0.0569	0.9776	0.9775	0.0445	0.0432	0.9772	0.9777	0.0313	0.0311
AMG(i) IV	0.9934	0.9922	0.0910	0.0532	0.9912	0.9929	0.0721	0.0440	0.9931	0.9922	0.0585	0.0343	0.9914	0.9908	0.0445	0.0247
AMG(ii) IV	1.0047	0.9966	0.1030	0.0707	1.0035	1.0011	0.0800	0.0577	1.0034	0.9996	0.0657	0.0440	0.9998	0.9976	0.0494	0.0311
MG	1.0933	1.0899	0.1658	0.0648	1.1099	1.0978	0.1711	0.0537	1.0972	1.0782	0.1552	0.0417	1.1045	1.0977	0.1618	0.0296
MG (inf)	0.9789	0.9769	0.0540	0.0530	0.9773	0.9798	0.0451	0.0431	0.9788	0.9793	0.0351	0.0334	0.9788	0.9790	0.0254	0.0238

$T=30$	$N=20$				$N=30$				$N=50$				$N=100$			
Pooled Estimators																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
POLS	1.0654	1.0724	0.4969	0.0893	1.0241	1.0300	0.3887	0.0708	0.9731	0.9684	0.3101	0.0567	0.9948	1.0020	0.2053	0.0371
FE	1.0680	1.0746	0.3230	0.0880	1.0082	1.0124	0.2460	0.0711	1.0253	1.0278	0.1943	0.0548	1.0083	1.0121	0.1379	0.0386
CCEP	0.9898	0.9902	0.0605	0.0424	0.9885	0.9878	0.0544	0.0349	0.9861	0.9873	0.0387	0.0270	0.9874	0.9887	0.0282	0.0190
FD-OLS	0.9162	0.9168	0.0547	0.0391	0.9153	0.9142	0.0467	0.0316	0.9153	0.9152	0.0347	0.0244	0.9139	0.9145	0.0238	0.0171
FD-IV	0.9965	0.9975	0.1331	0.0395	1.0022	1.0028	0.1072	0.0319	1.0024	1.0029	0.0820	0.0245	0.9968	0.9961	0.0554	0.0172
FE (inf)	0.9893	0.9897	0.0443	0.0278	0.9895	0.9893	0.0371	0.0227	0.9882	0.9879	0.0280	0.0177	0.9889	0.9886	0.0193	0.0125
MG-type Estimator																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
CCEMG	0.9858	0.9848	0.0609	0.0592	0.9842	0.9840	0.0545	0.0498	0.9833	0.9836	0.0392	0.0393	0.9837	0.9841	0.0289	0.0279
AMG(i) IV	0.9978	0.9971	0.0805	0.0490	0.9983	1.0023	0.0663	0.0401	0.9976	0.9973	0.0487	0.0319	0.9945	0.9932	0.0363	0.0227
AMG(ii) IV	1.0041	1.0003	0.0864	0.0658	1.0054	1.0046	0.0725	0.0547	1.0019	1.0013	0.0538	0.0427	0.9988	0.9981	0.0401	0.0296
MG	1.1158	1.1139	0.1797	0.0596	1.1399	1.1316	0.1827	0.0496	1.1167	1.1089	0.1821	0.0386	1.1274	1.1245	0.1797	0.0273
MG (inf)	0.9861	0.9857	0.0448	0.0438	0.9873	0.9869	0.0374	0.0358	0.9862	0.9867	0.0285	0.0279	0.9859	0.9860	0.0199	0.0198

Continued on the following page.

New Simulations — (iii)* Feedback and country trend (continued)

$T = 50$	$N = 20$				$N = 30$				$N = 50$				$N = 100$			
Pooled Estimators																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
POLS	1.0406	1.0486	0.5317	0.0724	1.0288	1.0311	0.4089	0.0571	0.9913	0.9834	0.3193	0.0462	0.9868	0.9831	0.2077	0.0304
FE	1.1093	1.1158	0.4420	0.0817	1.0341	1.0501	0.3591	0.0655	1.0474	1.0540	0.2786	0.0509	1.0127	1.0212	0.1984	0.0359
CCEP	0.9978	0.9964	0.0630	0.0350	0.9921	0.9903	0.0500	0.0285	0.9923	0.9927	0.0403	0.0220	0.9916	0.9906	0.0282	0.0156
FD-OLS	0.9183	0.9183	0.0501	0.0301	0.9132	0.9144	0.0392	0.0243	0.9136	0.9137	0.0312	0.0187	0.9136	0.9131	0.0219	0.0132
FD-IV	0.9979	0.9997	0.1048	0.0303	0.9946	0.9948	0.0858	0.0245	0.9963	0.9984	0.0655	0.0188	0.9988	0.9997	0.0455	0.0132
FE (inf)	0.9954	0.9944	0.0408	0.0191	0.9939	0.9943	0.0330	0.0155	0.9928	0.9936	0.0255	0.0121	0.9942	0.9945	0.0181	0.0085
MG-type Estimator																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
CCEMG	0.9959	0.9963	0.0618	0.0575	0.9892	0.9875	0.0493	0.0473	0.9899	0.9908	0.0391	0.0369	0.9894	0.9889	0.0277	0.0265
AMG(i) IV	1.0009	1.0022	0.0738	0.0479	0.9957	0.9944	0.0572	0.0392	0.9975	0.9975	0.0443	0.0308	0.9987	0.9976	0.0328	0.0222
AMG(ii) IV	1.0030	1.0020	0.0801	0.0686	0.9989	0.9983	0.0632	0.0555	0.9992	0.9993	0.0481	0.0433	1.0000	1.0002	0.0349	0.0304
MG	1.1657	1.1519	0.2148	0.0594	1.1688	1.1590	0.2103	0.0490	1.1542	1.1451	0.2061	0.0379	1.1587	1.1530	0.2120	0.0272
MG (inf)	0.9928	0.9914	0.0400	0.0381	0.9919	0.9927	0.0314	0.0310	0.9916	0.9917	0.0245	0.0242	0.9926	0.9926	0.0177	0.0171

$T = 100$	$N = 20$				$N = 30$				$N = 50$				$N = 100$			
Pooled Estimators																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
POLS	1.1143	1.1221	0.5644	0.0569	1.0373	1.0301	0.4428	0.0454	0.9945	0.9799	0.3599	0.0359	1.0010	1.0052	0.2425	0.0238
FE	1.1816	1.1930	0.6203	0.0651	1.0390	1.0171	0.4965	0.0536	1.0534	1.0632	0.3832	0.0409	1.0209	1.0255	0.2706	0.0288
CCEP	1.0014	1.0008	0.0657	0.0275	0.9962	0.9957	0.0575	0.0227	0.9982	0.9970	0.0447	0.0176	0.9957	0.9967	0.0307	0.0124
FD-OLS	0.9190	0.9189	0.0430	0.0211	0.9145	0.9140	0.0333	0.0171	0.9158	0.9151	0.0267	0.0132	0.9148	0.9150	0.0186	0.0093
FD-IV	0.9965	0.9988	0.0820	0.0213	1.0003	1.0026	0.0605	0.0172	1.0002	0.9994	0.0481	0.0133	1.0010	1.0020	0.0332	0.0093
FE (inf)	0.9990	0.9986	0.0363	0.0113	0.9975	0.9973	0.0300	0.0092	0.9980	0.9971	0.0240	0.0071	0.9990	0.9989	0.0165	0.0051
MG-type Estimator																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
CCEMG	0.9990	0.9976	0.0624	0.0606	0.9928	0.9924	0.0516	0.0519	0.9951	0.9948	0.0420	0.0408	0.9958	0.9963	0.0286	0.0289
AMG(i) IV	1.0012	1.0012	0.0663	0.0504	1.0030	1.0044	0.0499	0.0418	1.0031	1.0023	0.0406	0.0334	1.0034	1.0041	0.0273	0.0237
AMG(ii) IV	1.0002	0.9984	0.0710	0.0727	1.0026	1.0026	0.0527	0.0596	1.0025	1.0009	0.0432	0.0475	1.0027	1.0023	0.0298	0.0334
MG	1.2035	1.1962	0.2503	0.0615	1.2054	1.1978	0.2490	0.0516	1.1914	1.1790	0.2665	0.0409	1.1931	1.1835	0.2567	0.0292
MG (inf)	0.9978	0.9987	0.0342	0.0344	0.9968	0.9970	0.0279	0.0280	0.9972	0.9966	0.0221	0.0218	0.9983	0.9983	0.0155	0.0154

Notes: ‡ These use the year dummy coefficients from FD-IV estimator, rather than the FD-OLS estimator.

Table 4.7: New Simulations — (iv) Two 'clubs' for β

Monte Carlo Results — Setup with 2 'clubs' of countries
1,000 replications; POLS, FE and FD-OLS all have $T-1$ year dummies; AMG-estimators are constructed from † FD-OLS or ‡ FD-IV year dummy coefficients

$T = 20$	$N = 20$				$N = 30$				$N = 50$				$N = 100$			
Pooled Estimators																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
POLS	1.8149	1.7937	0.5854	0.2526	0.7037	0.6908	0.4523	0.1948	0.5417	0.5377	0.3662	0.1566	0.6228	0.6079	0.2252	0.1022
FE	1.0311	1.0204	0.1755	0.0691	1.0079	1.0082	0.1450	0.0561	1.0157	1.0144	0.1095	0.0434	1.0079	1.0065	0.0756	0.0308
CCEP	0.9989	0.9977	0.0727	0.0512	0.9974	0.9968	0.0591	0.0421	1.0026	1.0016	0.0450	0.0326	1.0000	0.9995	0.0320	0.0231
FD-OLS	1.0023	1.0036	0.0778	0.0563	1.0006	1.0002	0.0650	0.0455	1.0017	1.0016	0.0481	0.0350	1.0005	0.9997	0.0355	0.0246
FD-IV	0.9953	0.9914	0.2019	0.0569	0.9975	0.9970	0.1587	0.0458	0.9993	0.9996	0.1241	0.0352	0.9971	0.9950	0.0864	0.0247
FE (inf)	1.0004	0.9981	0.0636	0.0431	0.9994	1.0006	0.0515	0.0353	1.0018	1.0010	0.0391	0.0274	1.0003	1.0004	0.0285	0.0193
MG-type Estimator																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
CCEMG	1.0013	1.0017	0.0525	0.1205	0.9980	0.9993	0.0419	0.0993	1.0007	1.0001	0.0329	0.0769	1.0012	1.0006	0.0239	0.0548
AMG(i)	1.0058	1.0104	0.1008	0.1195	1.0051	1.0036	0.0820	0.0980	1.0065	1.0048	0.0647	0.0763	1.0036	1.0037	0.0470	0.0541
AMG(ii)	1.0268	1.0199	0.1172	0.1193	1.0228	1.0140	0.0965	0.0980	1.0182	1.0093	0.0768	0.0755	1.0123	1.0105	0.0545	0.0535
MG	1.1070	1.1008	0.1631	0.1283	1.1257	1.1134	0.1705	0.1047	1.1130	1.1011	0.1571	0.0805	1.1202	1.1128	0.1650	0.0571
MG (inf)	1.0000	1.0013	0.0377	0.1205	0.9988	0.9993	0.0310	0.0978	1.0004	1.0012	0.0248	0.0753	1.0003	1.0002	0.0175	0.0530

$T = 30$	$N = 20$				$N = 30$				$N = 50$				$N = 100$			
Pooled Estimators																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
POLS	1.7897	1.7979	0.6062	0.2051	0.7090	0.7064	0.4695	0.1579	0.5539	0.5537	0.3778	0.1267	0.6332	0.6166	0.2479	0.0830
FE	1.0400	1.0279	0.2265	0.0619	1.0098	1.0028	0.1799	0.0505	1.0224	1.0171	0.1375	0.0388	1.0114	1.0051	0.1015	0.0277
CCEP	1.0008	0.9985	0.0601	0.0405	0.9978	0.9978	0.0505	0.0332	1.0017	1.0000	0.0384	0.0257	1.0007	0.9991	0.0279	0.0182
FD-OLS	1.0008	0.9976	0.0654	0.0455	1.0020	1.0019	0.0541	0.0368	1.0023	1.0009	0.0401	0.0283	1.0015	1.0017	0.0297	0.0200
FD-IV	0.9908	0.9908	0.1610	0.0458	1.0044	1.0068	0.1324	0.0370	1.0011	0.9960	0.0968	0.0284	0.9977	0.9988	0.0692	0.0200
FE (inf)	1.0028	1.0004	0.0578	0.0329	0.9983	0.9979	0.0466	0.0266	0.9999	0.9994	0.0357	0.0207	1.0001	0.9993	0.0259	0.0147
MG-type Estimator																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
CCEMG	1.0002	0.9988	0.0406	0.1177	1.0005	1.0005	0.0354	0.0966	0.9996	0.9991	0.0260	0.0747	1.0002	0.9996	0.0194	0.0533
AMG(i)	1.0033	1.0035	0.0919	0.1185	1.0077	1.0087	0.0766	0.0967	1.0045	1.0018	0.0541	0.0750	1.0019	1.0048	0.0417	0.0533
AMG(ii)	1.0162	1.0082	0.1045	0.1183	1.0220	1.0155	0.0879	0.0966	1.0132	1.0108	0.0625	0.0746	1.0075	1.0075	0.0473	0.0528
MG	1.1273	1.1158	0.1797	0.1262	1.1518	1.1404	0.1830	0.1032	1.1260	1.1166	0.1827	0.0788	1.1377	1.1358	0.1834	0.0559
MG (inf)	1.0000	0.9990	0.0277	0.1181	1.0014	1.0015	0.0232	0.0954	1.0000	1.0003	0.0174	0.0734	1.0001	1.0002	0.0125	0.0518

Continued on the following page.

New Simulations — (iv) Two 'clubs' for β (continued)

$T = 50$	$N = 20$				$N = 30$				$N = 50$				$N = 100$			
Pooled Estimators																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
POLS	1.7581	1.7304	0.6804	0.1555	0.7469	0.7299	0.4968	0.1198	0.5878	0.5711	0.4052	0.0958	0.6544	0.6465	0.2592	0.0632
FE	1.0730	1.0577	0.3102	0.0530	1.0259	1.0188	0.2327	0.0436	1.0308	1.0245	0.1809	0.0331	1.0147	1.0109	0.1342	0.0239
CCEP	1.0003	1.0010	0.0541	0.0306	0.9962	0.9955	0.0432	0.0252	1.0012	1.0000	0.0343	0.0195	0.9997	0.9988	0.0240	0.0138
FD-OLS	1.0036	1.0026	0.0562	0.0350	1.0000	1.0002	0.0441	0.0283	1.0019	0.9999	0.0355	0.0218	0.9999	0.9990	0.0251	0.0154
FD-IV	0.9980	0.9944	0.1261	0.0352	0.9957	0.9896	0.1014	0.0284	0.9988	0.9993	0.0760	0.0218	0.9972	0.9983	0.0521	0.0154
FE (inf)	1.0017	1.0001	0.0539	0.0233	1.0001	0.9993	0.0432	0.0190	1.0013	1.0012	0.0345	0.0148	0.9990	0.9980	0.0244	0.0104
MG-type Estimator																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
CCEMG	1.0030	1.0020	0.0361	0.1160	0.9987	0.9984	0.0295	0.0955	0.9992	0.9991	0.0226	0.0739	1.0004	1.0007	0.0159	0.0525
AMG(i)	1.0061	1.0043	0.0834	0.1179	1.0010	0.9992	0.0669	0.0968	1.0040	1.0040	0.0491	0.0749	1.0017	1.0012	0.0339	0.0532
AMG(ii)	1.0142	1.0067	0.0936	0.1179	1.0080	1.0025	0.0761	0.0962	1.0072	1.0063	0.0547	0.0741	1.0038	1.0031	0.0374	0.0524
MG	1.1699	1.1553	0.2150	0.1256	1.1766	1.1668	0.2104	0.1031	1.1608	1.1480	0.2079	0.0785	1.1644	1.1598	0.2147	0.0560
MG (inf)	1.0005	0.9999	0.0196	0.1162	1.0001	0.9997	0.0152	0.0940	0.9993	0.9993	0.0118	0.0725	0.9999	0.9999	0.0084	0.0509

$T = 100$	$N = 20$				$N = 30$				$N = 50$				$N = 100$			
Pooled Estimators																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
POLS	1.6150	1.5904	0.7539	0.1035	0.8114	0.8097	0.5409	0.0804	0.6528	0.6305	0.4439	0.0634	0.7408	0.7289	0.2888	0.0424
FE	1.0786	1.0545	0.3821	0.0408	1.0371	1.0189	0.2999	0.0335	1.0348	1.0376	0.2320	0.0255	1.0310	1.0327	0.1667	0.0184
CCEP	1.0023	1.0023	0.0563	0.0214	0.9965	0.9964	0.0466	0.0177	1.0015	1.0013	0.0362	0.0137	0.9990	0.9990	0.0274	0.0097
FD-OLS	1.0018	1.0033	0.0453	0.0246	1.0022	1.0010	0.0372	0.0199	1.0023	1.0032	0.0282	0.0153	1.0017	1.0015	0.0217	0.0108
FD-IV	0.9934	0.9927	0.0907	0.0246	1.0008	1.0021	0.0716	0.0200	0.9996	1.0001	0.0554	0.0154	0.9994	1.0006	0.0377	0.0108
FE (inf)	1.0008	0.9949	0.0579	0.0149	1.0034	1.0034	0.0448	0.0122	1.0013	1.0015	0.0383	0.0094	1.0010	1.0002	0.0258	0.0067
MG-type Estimator																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
CCEMG	1.0046	1.0052	0.0379	0.1176	0.9989	0.9989	0.0312	0.0971	0.9994	0.9998	0.0253	0.0751	0.9996	1.0000	0.0195	0.0531
AMG(i)	1.0027	1.0011	0.0748	0.1197	1.0033	1.0018	0.0586	0.0982	1.0056	1.0069	0.0466	0.0762	1.0017	1.0030	0.0316	0.0539
AMG(ii)	1.0055	1.0013	0.0798	0.1191	1.0064	1.0053	0.0627	0.0972	1.0057	1.0053	0.0492	0.0750	1.0022	1.0030	0.0335	0.0529
MG	1.2074	1.1958	0.2524	0.1270	1.2075	1.1975	0.2499	0.1043	1.1937	1.1839	0.2683	0.0802	1.1945	1.1898	0.2602	0.0570
MG (inf)	1.0006	1.0001	0.0108	0.1152	1.0005	1.0002	0.0092	0.0933	0.9998	0.9999	0.0070	0.0717	0.9998	0.9999	0.0050	0.0505

Notes: ‡ These use the year dummy coefficients from FD-IV estimator, rather than the FD-OLS estimator.

Table 4.8: New Simulations — (iv)* Two 'clubs', country trends

Monte Carlo Results — Setup with 2 'clubs' of countries and country trends
1,000 replications; POLS, FE and FD-OLS all have $T-1$ year dummies; AMG-estimators are constructed from † FD-OLS or ‡ FD-IV year dummy coefficients

$T = 20$	$N = 20$				$N = 30$				$N = 50$				$N = 100$			
Pooled Estimators																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
POLS	1.8069	1.7834	0.6712	0.2612	0.7184	0.7198	0.5146	0.2032	0.5362	0.5287	0.4150	0.1630	0.6198	0.6131	0.2632	0.1064
FE	1.0184	1.0082	0.2557	0.1025	1.0130	1.0061	0.2168	0.0833	1.0175	1.0129	0.1683	0.0641	1.0075	1.0060	0.1161	0.0454
CCEP	0.9959	0.9960	0.0836	0.0586	0.9975	0.9952	0.0684	0.0484	1.0032	1.0032	0.0507	0.0371	1.0002	0.9998	0.0354	0.0263
FD-OLS	1.0016	1.0017	0.0793	0.0577	1.0013	1.0018	0.0654	0.0467	1.0017	1.0016	0.0478	0.0359	1.0001	0.9986	0.0357	0.0253
FE (inf)	1.0005	1.0008	0.0684	0.0455	1.0006	1.0016	0.0524	0.0371	1.0015	1.0000	0.0397	0.0288	0.9996	0.9999	0.0287	0.0203
MG-type Estimator																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
CCEMG	0.9987	0.9984	0.0677	0.1267	0.9979	0.9964	0.0546	0.1048	1.0020	1.0030	0.0403	0.0806	1.0016	1.0009	0.0285	0.0574
AMG(i)	1.0078	1.0105	0.1029	0.1196	1.0038	1.0026	0.0828	0.0981	1.0058	1.0062	0.0644	0.0763	1.0045	1.0034	0.0477	0.0541
AMG(ii)	1.0274	1.0172	0.1174	0.1307	1.0215	1.0172	0.0918	0.1066	1.0181	1.0140	0.0714	0.0818	1.0141	1.0106	0.0522	0.0578
MG	1.1082	1.0993	0.1640	0.1284	1.1254	1.1132	0.1715	0.1048	1.1127	1.1001	0.1563	0.0806	1.1203	1.1120	0.1650	0.0571
MG (inf)	1.0006	0.9999	0.0428	0.1222	0.9993	0.9975	0.0341	0.0989	1.0004	1.0014	0.0271	0.0762	1.0010	1.0002	0.0192	0.0536

$T = 30$	$N = 20$				$N = 30$				$N = 50$				$N = 100$			
Pooled Estimators																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
POLS	1.8009	1.7707	0.6974	0.2129	0.7297	0.7311	0.5478	0.1663	0.5506	0.5589	0.4528	0.1325	0.6312	0.6313	0.2837	0.0868
FE	1.0403	1.0383	0.3698	0.0994	1.0013	1.0133	0.2795	0.0811	1.0202	1.0222	0.2275	0.0622	1.0077	1.0076	0.1592	0.0441
CCEP	1.0013	0.9991	0.0694	0.0478	0.9994	0.9990	0.0593	0.0393	1.0009	1.0014	0.0445	0.0305	1.0006	1.0011	0.0321	0.0215
FD-OLS	1.0003	0.9970	0.0653	0.0467	1.0019	0.9999	0.0544	0.0378	1.0029	1.0021	0.0405	0.0291	1.0012	1.0015	0.0299	0.0205
FE (inf)	1.0022	0.9988	0.0535	0.0346	1.0016	1.0001	0.0465	0.0282	1.0016	0.9998	0.0364	0.0219	1.0001	0.9992	0.0240	0.0155
MG-type Estimator																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
CCEMG	1.0021	1.0010	0.0548	0.1223	1.0008	0.9994	0.0485	0.1010	0.9997	1.0010	0.0347	0.0784	1.0001	1.0006	0.0261	0.0558
AMG(i)	1.0032	1.0033	0.0946	0.1185	1.0070	1.0094	0.0774	0.0967	1.0053	1.0049	0.0539	0.0750	1.0019	1.0033	0.0415	0.0533
AMG(ii)	1.0160	1.0082	0.1035	0.1286	1.0197	1.0152	0.0859	0.1054	1.0127	1.0104	0.0614	0.0816	1.0077	1.0063	0.0453	0.0571
MG	1.1258	1.1189	0.1801	0.1261	1.1508	1.1381	0.1833	0.1032	1.1267	1.1196	0.1833	0.0789	1.1378	1.1365	0.1830	0.0559
MG (inf)	0.9999	0.9994	0.0305	0.1189	1.0015	1.0018	0.0251	0.0960	0.9999	0.9995	0.0190	0.0739	0.9999	0.9997	0.0134	0.0521

Continued on the following page.

New Simulations — (iv)* Two 'clubs', country trends (continued)

$T = 50$	$N = 20$				$N = 30$				$N = 50$				$N = 100$			
Pooled Estimators																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
POLS	1.7315	1.7307	0.7819	0.1637	0.7518	0.7384	0.5836	0.1268	0.5759	0.5795	0.4862	0.1018	0.6663	0.6606	0.3030	0.0668
FE	1.0674	1.0687	0.4922	0.0913	1.0249	1.0182	0.4090	0.0738	1.0331	1.0291	0.3222	0.0570	1.0165	1.0226	0.2254	0.0405
CCEP	1.0035	1.0035	0.0665	0.0389	0.9969	0.9977	0.0543	0.0317	1.0016	1.0004	0.0429	0.0244	0.9985	0.9988	0.0295	0.0174
FD-OLS	1.0028	1.0028	0.0562	0.0359	0.9998	1.0001	0.0437	0.0291	1.0016	1.0001	0.0357	0.0224	1.0002	0.9991	0.0253	0.0158
FE (inf)	0.9999	0.9972	0.0508	0.0246	0.9994	0.9966	0.0410	0.0201	0.9996	1.0004	0.0316	0.0156	0.9999	0.9996	0.0215	0.0110
MG-type Estimator																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
CCEMG	1.0062	1.0057	0.0521	0.1219	0.9992	0.9994	0.0438	0.1000	1.0001	1.0005	0.0347	0.0774	0.9991	0.9996	0.0239	0.0552
AMG(i)	1.0069	1.0053	0.0843	0.1180	1.0008	1.0002	0.0677	0.0967	1.0039	1.0027	0.0501	0.0749	1.0014	0.9997	0.0352	0.0533
AMG(ii)	1.0144	1.0070	0.0917	0.1311	1.0076	1.0042	0.0753	0.1059	1.0069	1.0043	0.0551	0.0819	1.0032	1.0030	0.0379	0.0577
MG	1.1723	1.1552	0.2146	0.1259	1.1758	1.1663	0.2115	0.1030	1.1608	1.1531	0.2079	0.0784	1.1650	1.1580	0.2149	0.0560
MG (inf)	1.0007	1.0010	0.0213	0.1166	0.9999	0.9998	0.0169	0.0944	0.9994	0.9992	0.0130	0.0727	1.0001	1.0000	0.0093	0.0511

$T = 100$	$N = 20$				$N = 30$				$N = 50$				$N = 100$			
Pooled Estimators																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
POLS	1.6429	1.6258	0.8825	0.1122	0.7961	0.7808	0.6623	0.0884	0.6299	0.6197	0.5156	0.0694	0.7340	0.7407	0.3508	0.0464
FE	1.1202	1.1320	0.6789	0.0721	1.0369	1.0160	0.5495	0.0598	1.0245	1.0204	0.4288	0.0453	1.0213	1.0196	0.2998	0.0324
CCEP	1.0014	1.0012	0.0726	0.0301	0.9962	0.9964	0.0598	0.0248	1.0030	1.0012	0.0480	0.0192	0.9985	0.9992	0.0334	0.0135
FD-OLS	1.0023	1.0026	0.0452	0.0252	1.0023	1.0011	0.0382	0.0205	1.0017	1.0019	0.0287	0.0157	1.0012	1.0009	0.0212	0.0111
FE (inf)	0.9999	0.9975	0.0529	0.0157	1.0018	1.0019	0.0435	0.0128	0.9995	0.9990	0.0327	0.0099	1.0000	0.9993	0.0232	0.0070
MG-type Estimator																
	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*	mean	median	emp. ste*	mean ste*
CCEMG	1.0042	1.0038	0.0575	0.1246	0.9981	0.9981	0.0448	0.1036	0.9994	0.9987	0.0381	0.0801	0.9988	0.9995	0.0266	0.0564
AMG(i)	1.0028	1.0048	0.0760	0.1197	1.0052	1.0044	0.0605	0.0982	1.0066	1.0079	0.0471	0.0762	1.0032	1.0046	0.0319	0.0539
AMG(ii)	1.0052	1.0022	0.0818	0.1335	1.0074	1.0060	0.0646	0.1092	1.0070	1.0059	0.0506	0.0845	1.0035	1.0042	0.0345	0.0593
MG	1.2070	1.1939	0.2515	0.1270	1.2097	1.1979	0.2501	0.1043	1.1944	1.1818	0.2689	0.0802	1.1950	1.1825	0.2590	0.0571
MG (inf)	1.0007	1.0006	0.0121	0.1152	1.0004	1.0001	0.0101	0.0934	0.9999	0.9998	0.0075	0.0718	0.9998	0.9997	0.0056	0.0505

Notes: ‡ These use the year dummy coefficients from FD-IV estimator, rather than the FD-OLS estimator.

Summary and Conclusions

Why do we observe such dramatic differences in labour productivity across countries in the macro data? This question has been central to the empirical investigation of development and growth over the past twenty years. This thesis has argued that current empirical approaches to the analysis of growth and development using macro panel data oversimplify the complexity of the production process across countries and neglect data cross-section and time-series properties.

Many authors talk of heterogeneous production technology across countries, but in their empirical implementations they typically only allow for differential TFP levels and evolution, imposing homogeneous slope coefficients on all countries — a choice that is arguably "particularly inappropriate when one is studying complex heterogeneous objects such as countries" (Brock & Durlauf, 2001, p.8/9). Although the *time-series* literature views variable nonstationarity in macro datasets as the rule rather than the exception, the investigation of these data properties in macro *panel* data has thus far been largely ignored in the mainstream literature on growth and other cross-country empirics. Prominent researchers in the panel time-series econometric literature have highlighted the salience of nonstationarity in standard macro datasets such as the Penn World Table and point out that it has "received virtually no attention in traditional panel regression analysis." (Phillips & Moon, 2000, p.264). The standard econometric assumption of cross-section independence has similarly encountered serious challenge in the econometric literature, but with the exception of 'spatial econometric' approaches (e.g. spatial lags) has not been addressed in the cross-country empirical literature. However, the most recent empirical implementations in the panel time-series literature (Bai & Ng, 2004; Kapetanios et al., 2009) have adopted common factor models to address this issue and thereby introduced a new level flexibility to econometric modelling. At present unobserved common factor approaches similar to those in Pesaran (2006) are primarily applied to macro

and financial time-series data, where their use is already widespread in the areas of forecasting and business cycle analysis (Stock & Watson, 2002; Giannone, Reichlin, & Sala, 2006). The discussion in this thesis is primarily intended to show that these factor models models have a wider appeal and to highlight their usefulness for empirical modelling in cross-country analyses using macro panel data.

This thesis makes a number of contributions to the applied econometrics and econometric theory literatures. These are detailed in the following. We first discuss our contributions to 'Empirical Growth and Productivity Analysis':

(i) **Technology Heterogeneity: impact of observable inputs**

Our conceptual and empirical analyses have shown that specification of cross-country heterogeneity in production technology is integral to consistent estimation and inference. Based on solid theoretical arguments, parameter heterogeneity was then highlighted in the empirical results for both agriculture and manufacturing data and furthermore confirmed in formal heterogeneity tests in the latter.

(ii) **TFP level estimates**

A separate point relates to the interpretation of TFP in the context of heterogeneous production technology: whilst in a pooled regression the *ceteris paribus* property allows for the comparison of TFP-levels across countries since all else remains equal, the adoption of a framework where the parameters on inputs can differ across countries prevents us from doing so, since all else *does not* remain equal. We have suggested an adjustment procedure which allows for TFP-level comparison across countries with differing production technologies, and across different empirical specifications.

(iii) **Technology Heterogeneity: impact of unobservables**

Adopting a common factor model with heterogeneous factor loadings for the analysis of agriculture and manufacturing data has enabled us to allow for a great deal of flexibility in the impact of unobservables — the 'measure of our ignorance', TFP — across countries, which has yielded important insights into the production process in these two differential industrial sectors.

In the manufacturing data we found that specifying an element of TFP which is idiosyncratic alongside another element which is common to all countries fitted the data well. The implication of the latter is that while R&D and technical innovation in global manu-

facturing ('technical progress') has a differential impact across countries, this 'common TFP' nevertheless has the ability to 'lift all boats'. Previous work on the spillovers from R&D through trade and exchange are thus supported by this finding (Coe & Helpman, 1995; Coe et al., 1997).

In the agriculture data, in contrast, our analysis showed that while TFP can similarly be represented as a common process across countries, the factors that make up TFP differ considerably across a spatial dimension. We found that imposing structure on the unobserved common factors — countries with similar (different) agro-climatic environment were assumed to have very similar (different) unobservables driving output — yielded the same results as leaving the data to specify the relationship freely, thus vindicating our choice of structure. We could conclude that countries with similar agro-climatic environment share a common TFP process, rather than neighbours or those in close geographical proximity as postulated by spatial econometric approaches. This finding may be interpreted as an explanation for the observed failure of technology transfer from advanced countries in the temperate 'North' to arid or equatorial developing countries of the 'South'.

(iv) **Identification**

By adopting a common factor model specification, our conceptual development in Chapter 1 was able to highlight an identification problem in the standard panel econometric approaches to cross-country analysis: in the case where production inputs are driven by the same unobserved factors as output the common parameter on observable inputs $\boldsymbol{\beta}$ (or mean of the country-specific parameters $\mathbb{E}[\boldsymbol{\beta}_i]$) is unidentified, since it is compounded with the impact of the unobserved common factors on output $\boldsymbol{\lambda}_i$. Standard instrumentation strategies in a pooled empirical model (2SLS or 'own-instrumentation' via lagged values in levels or first difference) cannot address this endogeneity if technology is heterogeneous across countries, since any informative instrument must be invalid due to the misspecification of parameter heterogeneity.

Our conceptual and empirical results suggest that the Pesaran (2006) CCE estimators as well as our own Augmented Mean Group estimators solve this identification problem by implicitly accounting for the heterogeneous impact of unobserved common factors on output. At the same time we suggest that these empirical strategies are robust to reverse causality.

We now summarise our contributions to the 'Nonstationary Panel Econometric' literature.

(i) **New estimation strategy**

We introduced the Augmented Mean Group (AMG) estimator for empirical analysis of macro panel data with potentially nonstationary observables and unobservables as well as cross-section dependence. Our Monte Carlo simulations showed that the performance of the AMG is similar to that of the Pesaran (2006) CCE estimators in most empirical setups, and that a simple instrumentation approach can remedy any bias in those setups where the AMG underperforms. The AMG provides an economically meaningful interpretation of the common correlation, namely the evolution of common or mean Total Factor Productivity across countries. This will allow us to investigate the country-specific parameters on common TFP and makes for a more complete analysis of productivity analysis. While the AMG is primarily intended for macro panel analysis, its intuition makes it an interesting approach to pursue in and adapt for other data contexts, in particular with view to large N, short T micro panel datasets.

(ii) **Standard panel estimators and cross-section dependence**

Our analysis further pinpointed the shortcomings of the present econometric literature on estimation and inference in nonstationary panel data with a multi-factor error structure. We showed that recent studies have overstated the extent of bias in standard (micro-)panel estimators such as pooled OLS or Fixed Effects in the DGP setups studied by failing to augment these estimators with year dummies — a common practice in the applied literature. As a result of our improved understanding gained, it would seem that the impact of nonstationarity and unobserved common factors leads to relatively limited bias (across all Monte Carlo replications), however individual estimates are very far away from the true DGP parameter values (single replication) thus indicating a very large sampling variation. While this is an interesting finding in itself, we highlight that with empirical application in mind the Monte Carlo results seem to back up our findings in the real data that these estimators yield severely distorted results. Prime among the reasons for this has to be the identification problem we summarised above.

A second interesting finding is the performance of the pooled OLS estimator with data in first differences (FD-OLS). Although in theory subject to the same identification problem as its level cousins the FD-OLS nevertheless yields consistent and efficient estimates in a wide variety of setups. We speculated that the use of year dummies in this case

may lead to a re-centering of the heterogeneous factor loadings around zero, such that *on average* no correlation between regression residuals and regressors arises. Further research is necessary to show this relationship more formally, although a simple exercise we completed provided encouraging results. The implication of this finding is that innovative augmentations such as that used by the CCE estimators may be redundant under many scenarios if a much simpler approach can yield identical results.

From our discussion above it is clear that there is ample scope for future research arising from this thesis, both on the econometric theory side and the application to cross-country growth empirics. The study of nonstationary macro panels with multi-factor error structure quite clearly has gained momentum over the past five years, and it is anticipated that further progress in this field will be rapid. At present, it does however seem that much of this research is carried out in isolation from applied work focused on datasets covering both developed *and* developing economies. Therefore the application of the insights and methods from this literature should be pursued first and foremost, so as to gain wider recognition within the field of empirical development economics. Separate to this, using the initial steps made in this thesis, the study of production and development should be allowed to move towards a more 'integrated' treatment of the production technology *in its entirety*, in the sense that "a full understanding of cross-country growth differences will need to explain *why* this parameter heterogeneity exists" (Durlauf et al., 2001, p.935, emphasis added). With specific reference to our empirical results we can highlight the investigation of the distribution and determinants of country-specific technology parameters in agriculture and manufacturing respectively, as well as in their relation across the two sectors as future avenues for research. The latter would represent a formal analysis of the dual economy model ideas we used to motivate our empirical investigations. In the same vein a study of *aggregate* economy data would also be insightful, in particular if coupled with simulations allowing for an underlying dual economy structure or with the analysis of comparable sectoral data. All of these extensions and future avenues for research then highlight the importance of allowing for heterogeneity in many dimensions when investigating economic growth and development across countries and over time.

Bibliography

Abdih, Y., & 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 trend in the United States since 1870. *American Economic Review*, *46*(2), 5-23.

Abreu, M., de Groot, H. L. F., & Florax, R. (2005). Space and Growth: A Survey of Empirical Evidence and Methods. *Région et Développement*, *21*.

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

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

Aiyar, S., & 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.

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

Azariadis, C., & Drazen, A. (1990). Threshold externalities in economic development. *Quarterly Journal of Economics*, *105*(2), 501-26.

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

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

Bai, J., & Ng, S. (2008). Large Dimensional Factor Analysis. *Foundations and Trends in Econometrics*, *3*(2), 89-163.

Baier, S. L., Dwyer, G. P., & Tamura, R. (2006). How important are capital and total factor productivity for economic growth? *Economic Enquiry*, *44*(1), 23-49.

Baily, M. N. (1986). Productivity Growth and Materials Use in U.S. Manufacturing. *Quarterly Journal of Economics*, *101*(1), 185-196.

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

Baltagi, B. H., Bresson, G., Griffin, J. M., & 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., Bresson, G., & Pirotte, A. (2007). Panel unit root tests and spatial dependence. *Journal of Applied Econometrics*, *22*(2), 339-360.

Baltagi, B. H., & 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.

Banerjee, A. V., & Newman, A. F. (1993). Occupational Choice and the Process of Development. *Journal of Political Economy*, *101*(2), 274-98.

Baptist, S., & Teal, F. (2008). *Why do South Korean firms produce so much more output per worker than Ghanaian ones?* (CSAE working paper, WPS/2008-10)

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

Barro, R. J. (1999). Notes on growth accounting. *Journal of Economic Growth*, *4*(2), 119-137.

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

Barro, R. J., & Lee, J. W. (1996). International measures of schooling years and schooling quality. *American Economic Review*, *86*(2), 218-23.

Barro, R. J., & Lee, J.-W. (2001). International data on educational attainment: Updates and implications. *Oxford Economic Papers*, *53*(3), 541-63.

Basile, R., Costantini, M., & Destefanis, S. (2005). *Unit root and cointegration tests for cross-sectionally correlated panels: Estimating regional production functions.* (CELPE Discussion Paper 94)

Basu, S., & Fernald, J. G. (1995). Are apparent productive spillovers a figment of specification error? *Journal of Monetary Economics*, *36*(1), 165-188.

Basu, S., & Fernald, J. G. (1997). Returns to scale in U.S. production: Estimates and implications. *Journal of Political Economy*, *105*(2), 249-83.

Bernard, A. B., & Jones, C. I. (1996a). Comparing apples to oranges: Productivity convergence and measurement across industries and countries. *American Economic Review*, *86*(5), 1216-38.

Bernard, A. B., & Jones, C. I. (1996b). Productivity across industries and countries: Time series theory and evidence. *The Review of Economics and Statistics*, *78*(1), 135-46.

Blomström, M., Lipsey, R. E., & Zejan, M. (1996). Is fixed investment the key to economic growth? *Quarterly Journal of Economics*, *111*(1), 269-76.

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

Blundell, R., & Bond, S. (1999). *GMM estimation with persistent panel data: an application to production functions* (IFS Working Papers No. W99/04).

Bode, E., & Rey, S. J. (2006). The spatial dimension of economic growth and convergence. *Papers in Regional Science*, *85*(2), 171-176.

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

Bond, S., Hoeffler, A., & Temple, J. (2001). *GMM estimation of empirical growth models* (Discussion Paper No. 3048).

Bond, S., Leblebicioglu, A., & Schiantarelli, F. (2007). *Capital Accumulation and Growth: A New Look at the Empirical Evidence.* (unpublished, updated version of 2004 paper)

Bowsher, C. G. (2002). On testing overidentifying restrictions in dynamic panel data models. *Economics Letters*, *77*(2), 211-220.

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

Bravo-Ortega, C., & Lederman, D. (2004). Agricultural Productivity and its Determinants: Revisiting International Experiences. *Estudios de Economia*, *31*(2), 133-163.

Breitung, J., & Pesaran, M. H. (2005). *Unit roots and cointegration in panels* (Discussion Paper Series 1: Economic Studies No. 2005-42).

Breuer, J. B., McNown, R., & Wallace, M. S. (2001). Misleading Inferences from Panel Unit-Root Tests with an Illustration from Purchasing Power Parity. *Review of International Economics*, *9*(3), 482-93.

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

Cameron, A. C., & Trivedi, P. (1990). *The information matrix test and its applied alternative hypotheses.* (Working Paper, University of California, Davis)

Canning, D., & Pedroni, P. (2008). Infrastructure, Long-Run Economic Growth And Causality Tests For Cointegrated Panels. *Manchester School*, *76*(5), 504-527.

Carroll, C. D., & Weil, D. N. (1994). Saving and growth: a reinterpretation. *Carnegie-Rochester Conference Series on Public Policy*, *40*, 133-192.

Caselli, F. (2005). Accounting for Cross-Country Income Differences. In P. Aghion & S. Durlauf (Eds.), *Handbook of Economic Growth* (Vol. 1, p. 679-741). Elsevier.

Caselli, F. (2008a). Growth Accounting. In S. N. Durlauf & L. E. Blume (Eds.), *The New Palgrave Dictionary of Economics Online* (Second edition ed.). Palgrave Macmillan. (Accessed March 2009)

Caselli, F. (2008b). Level Accounting. In S. N. Durlauf & L. E. Blume (Eds.), *The New Palgrave Dictionary of Economics Online* (Second edition ed.). Palgrave Macmillan. (Accessed March 2009)

Caselli, F., Esquivel, G., & Lefort, F. (1996). Reopening the convergence debate: A new look at cross-country growth empirics. *Journal of Economic Growth*, *1*(3), 363-89.

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

Cass, D. (1965). Optimum growth in an aggregative model of capital accumulation. *Review of Economic Studies*, *32*(3), 233-240.

Caves, D. W., Christensen, L. R., & Diewert, W. E. (1982). Multilateral Comparisons of Output, Input, and Productivity Using Superlative Index Numbers. *Economic Journal*, *92*(365), 73-86.

CEPII. (2006). *Datasets for Geodesic Distances.* (Centre D'Etudes Prospectives et D'Information Internationales. See note on CEPII website by Thierry Mayer and Soledad Zignago)

Cermeño, R., Maddala, G., & Trueblood, M. A. (2003). Modelling technology as a dynamic error components process: The case of the inter-country agricultural production function. *Econometric Reviews*, *22*, 289-306.

Chen, E. K. (1997). The Total Factor Productivity debate: Determinants of economic growth in East Asia. *Asian-Pacific Economic Literature*, *11*(1), 18-38.

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

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

Coakley, J., Fuertes, A., & Smith, R. (2001). *Small sample properties of panel time-series estimators with I(1) errors.* (unpublished working paper)

Coakley, J., Fuertes, A., & Smith, R. (2002). *A principle components approach to cross-section dependence in panels.* (Unpublished working paper)

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

Cobbold, T. (2003). *A Comparison of Gross Output and Value-Added Methods of Productivity Estimation* [Research Memorandum].

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

Coe, D. T., Helpman, E., & Hoffmaister, A. W. (1997). North-South R&D spillovers. *Economic Journal*, *107*(440), 134-49.

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

Craig, B. J., Pardey, P. G., & Roseboom, J. (1997). International Productivity Patterns: Accounting for Input Quality, Infrastructure, and Research. *American Journal of Agricultural Economics*, *79*(4), 1064-1076.

Crego, A., Larson, D., Butzer, R., & Mundlak, Y. (1998). *A new database on investment and capital for agriculture and manufacturing* (Policy Research Working Paper Series No. 2013).

D'Agostino, R. B., Balanger, A., & D'Agostino Jr., R. B. (1990). A suggestion for using powerful and informative tests of normality. *American Statistician*, *44*, 316-321.

Dickey, D., & Fuller, W. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, *74*(366), 427-431.

Duffy, J., & 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), 349-66.

Durlauf, S. N. (2001). Manifesto for a growth econometrics. *Journal of Econometrics*, *100*(1), 65-69.

Durlauf, S. N., & Johnson, P. A. (1995). Multiple regimes and cross-country growth behaviour. *Journal of Applied Econometrics*, *10*(4), 365-84.

Durlauf, S. N., & Johnson, P. A. (2008). Convergence. In S. N. Durlauf & L. E. Blume (Eds.), *The New Palgrave Dictionary of Economics Online* (Second Edition ed.). Palgrave Macmillan. (Accessed March 2009)

Durlauf, S. N., Johnson, P. A., & Temple, J. R. (2005). Growth econometrics. In P. Aghion & S. Durlauf (Eds.), *Handbook of Economic Growth* (Vol. 1, p. 555-677). Elsevier.

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

Durlauf, S. N., & Quah, D. T. (1999). The new empirics of economic growth. In J. B. Taylor & M. Woodford (Eds.), *Handbook of Macroeconomics* (Vol. 1, p. 235-308). Elsevier.

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

Easterly, W., & Levine, R. (2001). It's not factor accumulation: Stylised facts and growth models. *World Bank Economic Review*, *15*(2), 177-219.

Enders, W. (2004). *Applied Econometric Time Series* (2nd edition ed.). J Wiley.

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), 831-41.

Engle, R., & Granger, C. (1987). Cointegration and error correction: representations, estimation and testing. *Econometrica*, *55*(2), 252-276.

Ertur, C., & 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)

Fisher, R. A. (1932). *Statistical Methods for Research Workers* (4th Edition. ed.). Oliver & Boyd, Edinburgh.

Ford, G. S., Jackson, J. D., & Kline, A. D. (2006). Misleading Inferences from Panel Unit Root Tests: a Comment. *Review of International Economics*, *14*(3), 508-511.

Frankel, J. A., & Romer, D. (1999). Does trade cause growth? *American Economic Review*, *89*(3), 379-399.

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

Gallup, J. L., Mellinger, A. D., & Sachs, J. D. (1999). *Geography Datasets: Agricultural Measures.* Available online at CID, Harvard University.

Gengenbach, C., Palm, F. C., & Urbain, J.-P. (2008). *Panel Unit Root Tests in the Presence of Cross-Sectional Dependencies: Comparison and Implications for Modelling.* (Department of Quantitative Economics, University of Maastricht, version: 12 March, 2008)

Giannone, D., Reichlin, L., & Sala, L. (2006, May). VARs, common factors and the empirical validation of equilibrium business cycle models. *Journal of Econometrics*, *132*(1), 257-279. Available from `http://ideas.repec.org/a/eee/econom/v132y2006i1p257-279.html`

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

Gollop, F. M. (1979). Accounting for intermediate input: The link between sectoral and aggregate measures of productivity growth. In National Research Council (Ed.), *Measurement and interpretation of productivity* (p. 318-333). Washington: National Academy of Sciences.

Gomme, P., & Rupert, P. (2004). *Measuring labor's share of income* [Federal Reserve Bank

of Cleveland Policy Discussion Paper]. (November)
Graham, B., & Temple, J. (2006). Rich nations, poor nations: how much can multiple equilibria explain? *Journal of Economic Growth*, *11*(1), 5–41.
Granger, C. W. J. (1988). *Aggregation of time series variables: a survey* (Discussion Paper/Institute for Empirical Macroeconomics No. 1).
Granger, C. W. J. (1997). On modelling the long run in applied economics. *Economic Journal*, *107*(440), 169-77.
Granger, C. W. J., & Newbold, P. (1974). Spurious regressions in econometrics. *Journal of Econometrics*, *2*(2), 111-120.
Granger, C. W. J., & 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., & Mairesse, J. (1998). Production Functions: The Search for Identification. In S. Strøm (Ed.), *The Ragnar Frisch Centennial Symposium* (p. 169-203). Cambridge: Cambridge University Press.
Grossman, G. M., & Helpman, E. (1991). *Innovation and growth in the global economy.* Cambridge, Mass: MIT Press.
Gutierrez, L. (2002). Why is agricultural labour productivity higher in some countries than in others? *Agricultural Economics Review*, *3*(1), 58-72.
Gutierrez, L., & Gutierrez, M. M. (2003). International R&D spillovers and productivity growth in the agricultural sector. A panel cointegration approach. *European Review of Agriculture Economics*, *30*(3), 281-303.
Hall, R. E. (1990). Growth/productivity/employment. In P. Diamond (Ed.), (p. 71-112). MIT Press, Cambridge, MA.
Hall, R. E., & 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.
Hamilton, J. D. (1994). *Time-Series Analysis.* Princeton University Press.
Hanck, C. (2008). *Nonstationary-Volatility Robust Panel Unit Root Tests and the Great Moderation* [MPRA Paper, #11,988].
Harberger, A. (1987). Comment. In S. Fischer (Ed.), *NBER Macroeconomics Annual* (p. 255-258). Cambridge/Mass.: MIT Press.
Harrigan, J. (1999). Estimation of cross-country differences in industry production functions. *Journal of International Economics*, *47*(2), 267-293.
Harris, R. (1994). *Using Cointegration Analysis in Econometric Modelling.* London: Prentice Hall.
Hayami, Y., & Ruttan, V. (1970). Agricultural Productivity Differences Among Countries. *American Economic Review*, *60*(5), 895-911.
Hayami, Y., & Ruttan, V. (1985). *Agricultural development: An international perspective.* Baltimore: Johns Hopkins University Press.
Hendry, D. (1995). *Dynamic Econometrics.* Oxford University Press.
Hendry, D., & Nielsen, B. (2007). *Econometric Modeling: A Likelihood Approach.* Princeton University Press.
Heston, A., Summers, R., & Aten, B. (2006). *Penn World Table version 6.2.* (Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania)
Hobijn, B., Franses, P.-H., & Ooms, M. (2004). Generalizations of the KPSS-test for stationarity. *Statistica Neerlandica*, *58*(4), 483-502.
Hulten, C. R. (2000). *Total factor productivity: A short biography* (NBER Working Papers No. 7471).
Im, K., Pesaran, M. H., & Shin, Y. (1997). *Testing for unit roots in heterogeneous panels.* (Discussion Paper, University of Cambridge)
Im, K., Pesaran, M. H., & 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), 1127-70.

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

Jaffe, A. (1986). Technological Opportunity and Spillovers of R&D: Evidence from Firms' Patents, Profits and Market Value [NBER Working Paper Series]. (No. 1815).

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

Jorgensen, D. W. (1961). The development of a dual economy. *Economic Journal*, *72*, 309-334.

Kaldor, N. (1957). A model of economic growth. *Economic Journal*, *67*, 591-624.

Kaldor, N. (1966). *Causes of the Slow Rate of Economic Growth in the United Kingdom.* Cambridge University Press.

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

Kao, C., & 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., & 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.

Kao, C., Trapani, L., & Urga, G. (2007). *Modelling and Testing for Structural Changes in Panel Cointegration Models with Common and Idiosyncratic Stochastic Trends* [Syracuse University, CPR Working Paper 92].

Kapetanios, G., Pesaran, M. H., & Yamagata, T. (2009). *Panels with Nonstationary Multifactor Error Structures.* (unpublished working paper, July 2009, updated version of IZA Discussion Paper #2243)

Kendrick, J. W. (1956). Productivity Trends: Capital and Labor. *The Review of Economics and Statistics*, *38*(3), 248-257.

Kindleberger, C. P. (1967). *Europe's post-war growth: the role of labor supply.* Cambridge, Mass.: Harvard University Press.

Klenow, P. J., & Rodriguez-Clare, A. (1997a). Economic growth: A review essay. *Journal of Monetary Economics*, *40*(3), 597-617.

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

Knight, M., Loayza, N., & Villanueva, D. (1993). Testing the Neoclassical Theory of Economic Growth: A Panel Data Approach. *IMF Staff Papers*, *40*(3), 512-541.

Kottek, M., Grieser, J., Beck, C., Rudolf, B., & Rubel, F. (2006). World Map of the Köppen-Geiger climate classification updated. *Meteorologische Zeitschrift*, *15*(3), 259-263.

Kuznets, S. (1961). Economic growth and the contribution of agriculture: notes on measurement. *International Journal of Agrarian Affairs*, *3*(1), 56-75.

Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics*, *54*(1-3), 159-178.

Larson, D. F., Butzer, R., Mundlak, Y., & Crego, A. (2000). A cross-country database for sector investment and capital. *World Bank Economic Review*, *14*(2), 371-391.

Lee, K., Pesaran, M. H., & Smith, R. (1997). Growth and convergence in a multi-country empirical stochastic Solow model. *Journal of Applied Econometrics*, *12*(4), 357-92.

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

Leibenstein, H. (1957). *Economic backwardness and economic growth.* NY: Wiley.

Levine, R., & Renelt, D. (1991). *Cross-country studies of growth and policy - methodological, conceptual and statistical problems.* (World Bank WPS# 608)

Lewis, W. A. (1954). Economic development with unlimited supplies of labour. *The Manchester School*, *22*, 139-191.

Lipsey, R. G., & Carlaw, K. I. (2001). *What does Total Factor Productivity measure?* (Unpublished manuscript)

Lipsey, R. G., & Carlaw, K. I. (2004). Total factor productivity and the measurement of technological change. *Canadian Journal of Economics*, *37*(4), 1118-1150.

Liu, Z., & Stengos, T. (1999). Non-linearities in Cross-Country Growth Regressions: A Semiparametric Approach. *Journal of Applied Econometrics*, *14*(5), 527-38.

Loayza, N. (1994). *A test of the international convergence hypothesis using panel data* (Policy Research Working Paper Series No. 1333).

Long, J. B. D., & Summers, L. H. (1991). Equipment Investment and Economic Growth. *The Quarterly Journal of Economics*, *106*(2), 445-502.

Lucas, J., Robert E. (1993). Making a Miracle. *Econometrica*, *61*(2), 251-72.

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

Maddala, G. S., & Kim, I.-M. (1998). *Unit roots, cointegration and structural change*. Cambridge University Press.

Maddala, G. S., & 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.

Magrini, S. (2004). Regional (di)convergence. In J. V. Henderson & J. F. Thisse (Eds.), *Handbook of Regional and Urban Economics* (Vol. 4, p. 2741-2796). Elsevier.

Mankiw, G., Romer, D., & 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 & Economic Statistics*, *17*(1), 129-36.

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

Martin, W., & 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.

Moran, P. (1950). Notes on continuous stochastic phenomena. *Biometrika*, *37*(1/2), 17-33.

Mundlak, Y., Larson, D., & Butzer, R. (1997). *The determinants of agricultural production: a cross-country analysis* (Policy Research Working Paper Series No. 1827).

Mundlak, Y., Larson, D., & Butzer, R. (1999). Rethinking within and between regressions: The case of agricultural production functions. *Annales D'Economie et de Statistique*, *55/56*, 475-501.

Mundlak, Y., Larson, D. F., & Butzer, R. (2008). *Heterogeneous technology and panel data: the case of the agricultural production function* (Policy Research Working Paper Series No. 4536).

Murphy, K. M., Shleifer, A., & Vishny, R. W. (1989). Industrialization and the Big Push. *Journal of Political Economy*, *97*(5), 1003-26.

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

Nelson, R. (1968). A 'Diffusion' Model of International Productivity Differences in Manufacturing Industry. *American Economic Review*, *58*(5), 1219-1248.

Newey, W. K., & West, K. D. (1994). Automatic lag selection in covariance matrix estimation. *Review of Economic Studies*, *61*(4), 631-53.

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

Oulton, N., & O'Mahony, M. (1994). *Productivity and growth: A Study of British Industry, 1954-86*. Cambridge University Press.

Paauw, D. S., & Fei, J. C. (1973). *The Transition in Open Dualistic Economies — Theory and Southeast Asian Experience*. New Haven, CN: Yale University Press.

Pagan, A. (1984). Econometric Issues in the Analysis of Regressions with Generated Regressors. *International Economic Review*, *25*(1), 221-247.

Palm, F. C., & Pfann, G. A. (1995). Unraveling Trend and Stationary Components of Total Factor Productivity. *Annales D'Economie et de Statistique*, *39*, 67-92.

Pardey, P., James, J., Alston, J., Wood, S., Koo, B., Binenbaum, E., et al. (2007). *Science, Technology and Skills* [Background Paper for the World Bank 'World Development Report 2008: Agriculture for Development'].

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

Parente, S. L., & 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 No. 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. (2004). *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 2004)

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 No. 1240 and CESifo Working Paper No. 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., & Smith, R. (1995). Estimating long-run relationships from dynamic heterogeneous panels. *Journal of Econometrics*, *68*(1), 79-113.

Pesaran, M. H., Smith, V., & Yamagata, T. (2008). *Panel unit root tests in the presence of a multifactor error structure.* (Paper presented at the 'Unobserved Factor Models' Conference, Birkbeck College, 20-21 November 2008)

Pesaran, M. H., & Tosetti, E. (2007). *Large panels with common factors and spatial correlations.* (IZA Discussion Papers No. 3032)

Pesaran, M. H., & Yamagata, T. (2008). Testing Slope Homogeneity In Large Panels. *Journal of Econometrics*, *142*(1), 50-93.

Phillips, P. C. B. (1995). Fully Modified Least Squares and Vector Autoregression. *Econometrica*, *63*(5), 1023–1078.

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

Phillips, P. C. B., & Moon, H. R. (2000). Nonstationary panel data analysis: An overview of

some recent developments. *Econometric Reviews*, *19*(3), 263-286.
Phillips, P. C. B., & Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, *75*(2), 335-346.
Phillips, P. C. B., & Sul, D. (2003). Dynamic panel estimation and homogeneity testing under cross section dependence. *Econometrics Journal*, *6*(1), 217-259.
Ranis, G., & Fei, J. (1961). A theory of economic development. *American Economic Review*, *51*(4), 533-556.
Rapach, D. E. (2002). Are real GDP levels nonstationary? Evidence from panel data tests. *Southern Economic Journal*, *68*(3), 473-495.
Redding, S., & Venables, A. J. (2004). Economic geography and international inequality. *Journal of International Economics*, *62*(1), 53-82.
Robinson, S. (1971). Sources of growth in less developed countries: A cross-section study. *Quarterly Journal of Economics*, *85*(3), 391-408.
Romer, P. M. (1990). Endogenous technological change. *Journal of Political Economy*, *98*(5), S71-102.
Roodman, D. (2009). A Note on the Theme of Too Many Instruments. *Oxford Bullentin of Economics and Statistics*, *71*(1), 135-158.
Ruttan, V. W. (2002). Productivity Growth in World Agriculture: Sources and Constraints. *Journal of Economic Perspectives*, *16*(4), 161-184.
Ryan, K., & Giles, D. (1998). *Testing for unit roots with missing observations* (Department Discussion Papers No. 9802).
Sachs, J. D. (2003). *Institutions Don't Rule: Direct Effects of Geography on Per Capita Income* (NBER Working Papers No. 9490).
Sachs, J. D., & Warner, A. (1995). Economic reform and the process of global integration. *Brookings Papers on Economic Activity*, *1*, 1-118.
Sachs, J. D., & Warner, A. M. (2001). The curse of natural resources. *European Economic Review*, *45*(4-6), 827-838.
Shepotylo, O. (2008). *Spatial HAC Estimator: Analysis of Convergence of European Regions.* (National University of Kyiv-Mohyla Academy, Kyiv School of Economics (KSE), December 2008)
Smith, R. P., & Fuertes, A.-M. (2004). *Panel Time Series.* (Centre for Microdata Methods and Practice (cemmap) mimeo, April 2004.)
Smith, R. P., & Fuertes, A.-M. (2007). *Panel Time Series.* (Centre for Microdata Methods and Practice (cemmap) mimeo, April 2007.)
Söderbom, M., & 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), 312-20.
Solow, R. M. (1970). *Growth theory: an exposition.* Oxford: Oxford University Press.
Solow, R. M. (1986). Unemployment: Getting the questions right. *Economica*, *53*(210), S23-34.
Stock, J. H., & Watson, M. W. (2002). Macroeconomic Forecasting Using Diffusion Indexes. *Journal of Business & Economic Statistics*, *20*(2), 147-62.
Swamy, P. A. V. B. (1970). Efficient inference in a random coefficient regression model. *Econometrica*, *38*(2), 311-23.
Swan, T. W. (1956). Economic growth and capital accumulation. *Economic Record*, *32*(2), 334-61.
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].

UN. (2005). *UN Common Statistics 2005* [Online database, New York: UN].

UNIDO. (2004). *UNIDO Industrial Statistics 2004* [Online database, Vienna: UNIDO].

Westerlund, J. (2006). Testing for panel cointegration with a level break. *Economics Letters*, *91*(1), 27-33.

Westerlund, J., & Breitung, J. (2009). *Myths and Facts about Panel Unit Root Tests.* (Paper presented at the Panel 2009 Conference in Bonn, 3-5 July 2009)

Westerlund, J., & 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.

World Bank. (2008). *World Development Indicators.* Database.

Young, A. (1994). Lessons from the East Asian NICs: a contrarian view. *European Economic Review*, *38*(3-4), 964-973.

Young, A. (1995). The tyranny of numbers: confronting the statistical realities of the East Asian growth experience. *Quarterly Journal of Economics*, *110*(3), 641-680.