Aggregation versus Heterogeneity in Cross-Country Growth Empirics*

Markus Eberhardt[a,b]
markus.eberhardt@economics.ox.ac.uk

Francis Teal[b,c]
francis.teal@economics.ox.ac.uk

[a] *St Catherine's College, Oxford OX1 3UJ, England*

[b] *Centre for the Study of African Economies,*
Department of Economics, University of Oxford, Manor Road, Oxford OX1 3UQ, England

[c] *Institute for the Study of Labor (IZA),*
Schaumburg-Lippe-Str. 5-9, 53113 Bonn, Germany

Incomplete first draft: 24th March 2010
submitted to the *XIIIth Applied Economics Meeting*, Seville

Abstract:

The empirical literature on cross-country growth and development commonly employs aggregate economy data such as the Penn World Table Heston, Summers, and Aten (2009) to estimate homogeneous production function or convergence regression models. Against the background of a dual economy framework this paper investigates the potential sources of bias when aggregate economy data is adopted instead of sectoral data. A general model is developed, encompassing the different approaches in the dual economy literature. Following appropriate empirical specification and testing we estimate production functions in agriculture and manufacturing for a large panel of developing and developed countries (1963-1992). Our focus here is on novel panel time-series methods which can accommodate technology heterogeneity, variable time-series properties and the breakdown of the standard cross-section independence assumption in panels. We investigate the potential for bias in the production parameter coefficients due to aggregation of sectors and empirical misspecification, employing both theory and Monte Carlo simulations. Our empirical results identify substantial bias in the technology coefficient in a stylised aggregate economy made up of agricultural and manufacturing sectors. Further analysis to attribute the source of this effect is carried out.

Keywords: dual economy model, production function, common factor model
JEL classification: C33, O13

*The first author gratefully acknowledges financial support from the ESRC [grant numbers PTA-031-2004-00345 and PTA-026-27-2048]. For correspondence: Markus Eberhardt, Centre for the Study of African Economies (CSAE), Department of Economics, Manor Road Building, Oxford OX1 3UQ, UK; Email: markus.eberhardt@economics.ox.ac.uk; phone: +44-(0)1865 271084; fax: +44-(0)1865 281447.

1. INTRODUCTION

"The reason why savings are low in an undeveloped economy relatively to national income is not that the people are poor, but that capitalist profits are low relatively to national income. As the capitalist sector expands, profits grow relatively, and an increasing proportion of national income is re-invested." Lewis (1954, p.190)

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.[1] 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).

A common feature across these literatures on both economic development and growth was the use of closed economy models. The basic models put forward by Lewis and by Solow-Swan were closed economy models in which structural transformation and growth occurred within economies.[2] However it was soon realised that these were not the most appropriate models for economies which were small in geographical area and open to the world economy in the sense that their influence on the prices of their products was minimal. As noted by Lucas (1988), the theory of trade as developed by Ricardo and Heckscher-Ohlin implies that trade can have "a level effect, analogous to the one-time shifting upward in production possibilities, [but] not a growth effect" (12) on income. The strong correlation apparent in the data between income growth and trade led to much new work on the theory of how trade may impact growth (e.g. Grossman & Helpman, 1991; Rivera-Batiz & Romer, 1991; Aghion & Howitt, 1992; Matsuyama, 1992), one key mechanism being via improvement in technical progress, another being externalities. In fact much of the empirical work on this topic (Coe, Helpman, & Hoffmaister, 1997; Frankel & Romer, 1999; Rodriguez & Rodrik, 2001; Greenaway, Morgan, & Wright, 2002; Dollar & Kraay, 2002, 2004) used reduced form models and side-stepped the theoretical issues as to exactly why more open economies might grow faster.

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

[1] A note on nomenclature: we refer to 'duality' or 'dual economy models' as representing economies with two stylised sectors of production (agriculture and manufacturing), while 'dualism' refers to wage or marginal labour product differences between sectors. Total Factor Productivity (TFP) is referred to as technology or technology levels, TFP growth as technical/technological progress. We use productivity to refer to income/output per worker, and commonly make this clear by referring to 'labour productivity' in contrast to productivity referring to TFP levels.

[2] This is in one sense not surprising as the major 'development' question of the 1930s still influencing these authors' thinking was the experiment in the Soviet Union to industrialise in autarky in the space of a decade. This aside, the largest economy in the world – the United States – occupying as it does an entire continent was one in which external trade certainly did not seem the major agent in the growth process.

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, Romer, and Weil (1992), which initiated a major revival in the Solow-Swan model and effectively merged the concerns of economic development with those of growth.[3]

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; Easterly, 2002; Durlauf, Johnson, & Temple, 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 (Young, 1995; Chen, 1997; Klenow & Rodriguez-Clare, 1997a, 1997b; Easterly & Levine, 2001; Lipsey & Carlaw, 2001; Baier, Dwyer, & Tamura, 2006). While there is some new theoretical and empirical work using a dual economy model (e.g. Vollrath, 2009c, 2009a, 2009b), this is largely absent from textbooks on economic growth and has not been the central focus of attention for most of the empirical analyses (Temple, 2005). A primary reason for the focus has been the availability of data. The Penn World Table (PWT) dataset — most recently (Heston et al., 2009) — and the Barro-Lee data on human capital (Barro & Lee, 1993, 2001) have supplied macro-data which ensure that the aggregate Solow-Swan model can be readily estimated. In recent years there has however been a development of datasets that allow a closer matching between the dual economy models and the data (Larson, Butzer, Mundlak, & Crego, 2000), which this paper will exploit to throw light on several of the empirical issues that have been central to the analysis of the sources of growth.

Cross-country growth regressions represent one of the most active fields of empirical analysis within applied development economics, however the viability of this empirical approach has been seriously questioned over the past decade and at present these methods are deeply unfashionable. We have argued elsewhere that much can be learned from cross-country empirics provided the empirical setup allows for greater flexibility in the estimation equation and recognises the salient data properties of macro panel datasets (Eberhardt & Teal, 2010). Methods developed in the emerging panel time series literature (Bai & Ng, 2002, 2004; Coakley, Fuertes, & Smith, 2006; Pesaran, 2006; Bai, 2009) can go further in providing robust estimation and inference for nonstationary panel data where variable series may be correlated across countries and where common shocks are likely to impact all countries in the sample, albeit to a different extent.

This paper, providing empirical analysis of panel data for developing and developed economies, sets out to address three main objectives: (i) rather than using a calibrated dual economy model for quantitative analysis we provide empirical estimates for technology coefficients in sectoral production functions. This allows for the integration of recent developments in the literature on applied panel data econometrics, including the insights of the emerging panel time series literature. (ii) We estimate a stylised aggregate production function model from agriculture and manufacturing data, and compare results with those from disaggregated regressions. This will allow us to judge whether neglecting a dual economy structure leads to bias in the empirical technology coefficients. (iii) In the light of the results from our sectoral production function estimations we assess the relative sources of growth in a dual economy model: TFP growth and level differences across sectors, and marginal factor dualism.

[3]The addition of human capital to the Solow model in Mankiw et al. (1992) "leads to quantitative predictions that look consistent with the data" (Temple, 2005, p.436).

The remainder of this paper is organised as follows: Section 2 provides an encompassing conceptual framework for the analysis of dual economy effects at the macro level and discusses technology heterogeneity. In the following section we then introduce an empirical specification of our dual economy framework, discuss the data and briefly review the empirical methods and estimators employed. Section 4 reports and discusses empirical findings at the sector-level. 5 then investigates the potential sources of bias in aggregate economy data, employs Monte Carlo simulations to provide support and presents empirical findings from stylised and PWT aggregate data. Section 6 summarizes and concludes.

2. DUALITY AND AGGREGATE EMPIRICS

The literature on dual economy models is surprisingly large, given the relatively limited impact this approach has had in entering textbooks on economic growth theory and analysis, and economics 'orthodoxy' in general. With the availability of sectoral data for a cross-section of countries limited until recently, some of the existing work in this area is built on models relatively disjoint from the formulation of empirically testable questions, while other studies have focused on very specific details of the growth and development process which are then 'tested' using simulation or calibrated models. As a result many of the dual economy models, given their complexity and data requirements, do not suit themselves for empirical testing. In this section we present a theoretical dual economy model based on the existing literature and motivate our emphasis on technology heterogeneity across sectors and economies.

2.1 A model of an open dual economy

The early literature on structural change did not pursue formal modelling of the small open dual economy setup, but limited itself to a conceptual understanding of the link between structural change and potential growth in a closed economy. Lewis (1954), Kaldor (1966), Kindleberger (1967) and Ranis and Fei (1961), for instance, all emphasize the potential for surplus labour in agriculture to act as a major driver for structural change via the migration of labour into the emerging manufacturing sector. In their analyses elastic labour supply enables economic growth by keeping wages in the modern sector low and preserving industrial peace (Temple, 2001; Temin, 2002; Barbier & Rauscher, 2007). A somewhat more complex analysis suggests that agricultural income and food supply constraints should be the focus of analysis, since they represent barriers to structural change and thus development (Jorgensen, 1961). Openness to trade, however, somewhat relaxes these constraints. As noted in the introduction, modelling structural change and growth in a closed economy model is not deemed appropriate to model the development process in small open economies.

The supply side of a small, open dual economy model can be represented by two sectors, assumed to be agriculture ('traditional sector') and manufacturing ('modern sector'), producing distinct goods. It is posited that these two types of production are geographically distinct, the former present in rural areas and the latter in urban areas. Their respective technologies are assumed Cobb-Douglas but unrestricted with regard to returns to scale[4]

$$
\begin{aligned}
Y_{a,t} &= A_{a,t}\, F(K_{a,t},\, L_{a,t},\, N_{a,t}) = A_{a,t} K^{\alpha}_{a,t},\, L^{\beta}_{a,t},\, N^{\gamma}_{a,t} \qquad \alpha, \beta, \gamma < 1 && (1)\\
Y_{m,t} &= A_{m,t}\, G(K_{m,t},\, L_{m,t}) = A_{m,t} K^{\phi}_{m,t},\, L^{\psi}_{m,t} \qquad \phi, \psi < 1 && (2)\\
A_{j,t} &= A_{j,0}\, exp(\lambda_j t) \qquad \text{for } j = a, m && (3)
\end{aligned}
$$

[4]Our model specification is guided by Temple (2005) and Corden and Findlay (1975).

where A represents disembodied technical efficiency of production (TFP),[5] K is physical, reproducible capital, and L is labour (either raw labour or adjusted for human capital differences) for both agricultural and manufacturing sectors a and m.[6] Capital and labour are stock variables which can be accumulated infinitely, but are subject to diminishing returns. N is non-reproducible capital (assumed to be arable land and other forms of capital), and only enters the agricultural production function. We drop the time subscript for ease of exposition.

In the most general specification TFP growth rates λ_j and TFP levels $A_{j,0}$ are allowed to differ across sectors, countries, and in case of TFP growth across time. When a country's manufacturing sector enjoys higher TFP growth than its agricultural sector, this implies *ceteris paribus* higher output growth in manufactured goods, and (deflated by sector share in total output s_a, s_m) higher aggregate output growth g.

$$g = \dot{Y}/Y = \dot{Z}/Z \left[+\eta\dot{L}/L + \mu\dot{K}/K\right] = \dot{Z}/Z = s_a\dot{A}_a/A_a + s_m\dot{A}_m/A_m \tag{4}$$

Allowing TFP growth λ_j to vary over time allows for a more realistic dynamic evolvement of the sectoral technology level than a constant TFP growth rate. Given differential TFP levels between sectors, say $A_a < A_m$, structural transformation in the form of labour migration to manufacturing would result in a temporary level effect on output. Unlike in the TFP growth case this would not change the *perpetual* growth trajectory of the economy.[7] Persistent and significant TFP level differences between sectors signal the presence of barriers to technology acquisition or some other form of friction in the low-TFP sector, while TFP level differences across countries signal frictions on the country-level (Caselli, 2005; Restuccia, Yang, & Zhu, 2008). We assume that the economy is open to trade in products but closed to cross-country factor migration such that

$$Y = Y_a + pY_m \tag{5}$$

where the price of the agricultural good Y_a is the numeraire and p provides the relative price of manufactures, exogenously determined by the world price. We restrict discussion to incompletely specialised economies, i.e. both sectors have positive output. We assume full capital employment

$$K = K_a + K_m \tag{6}$$

with capital perfectly mobile between the two sectors leading to rental rate equalisation

$$r_a = \mathrm{MPK}_a = A_a\partial F/\partial K_a = \alpha\frac{Y_a}{K_a} \qquad r_m = \mathrm{MPK}_m = pA_m\partial G/\partial K_m = p\phi\frac{Y_m}{K_m} \qquad r_m = r_a \tag{7}$$

The first-best equilibrium for the economy is defined by equations (1)-(3) and (5)-(7), in addition to equilibrium conditions in the labour market: under full employment and with wages equal to marginal products, workers will (freely) migrate between sectors until wages (deemed to equate marginal labour products) are equalised. However, in order to provide a specification as general as possible, we do not impose wage equalisation (and first-best solution), but assume labour market disequilibrium in form of some exogenously-determined wedge $0 < k < 1$ which drives manufacturing wages *above* those in

[5] We see this as a 'catch-all' for disembodied levels of productive efficiency and technology as well as characteristics such as taxation, regulation, climate, soil conditions etc., following Gollin, Parente, and Rogerson (2002).

[6] Note that agricultural and rural labour should not be taken as homogeneous, but we can assume a setup that allows us to keep the model as it is laid out above, without losing the appeal of this notion (Temple, 2005): using our specification, we assume human capital to be embodied partly within capital and partly within the technical progress term . Human capital data in Timmer (2000) taken from Chai (1995) would halve the number of countries in our manufacturing dataset since only developing nations are discussed. A UNESCO dataset discussed in Córdoba and Ripoll (2009) contains only a handful of observations across time and countries. Due to reasons of limited space in this paper the option to experiment with these datasets was not pursued.

[7] Nevertheless, TFP levels "capture the differences in long-run economic performance that are most directly relevant to welfare" (Hall & Jones, 1999, p.85).

agriculture:[8]

$$w_a = \mathrm{MPL}_a = A_a \partial F/\partial L_a = \beta \frac{Y_a}{L_a} \qquad w_m = \mathrm{MPL}_m = pA_m \partial G/\partial L_m = p\psi \frac{Y_m}{L_m} \qquad w_a = kw_m \tag{8}$$

where $k > 0$. We know that wage equalisation across sectors would provide the optimal output solution and can deduce that a wage differential between sectors leads to an equilibrium characterised by lower output. Adopting the Harris and Todaro (1970) approach to inter-sectoral labour market equilibrium, we assume unemployment in the urban labour market (L_u), such that

$$L = L_a + L_m + L_u \tag{9}$$

The key assumption in this approach is that in the presence of wage differentials and urban unemployment, rural (agricultural) migrants discount the urban wage, such that migration occurs until actual rural wage is equal to expected urban wage:[9]

$$w_a = \mathbb{E}[w_m] = (1-u)w_m \tag{10}$$

The expectation of the urban wage is simply the probability of obtaining a job $(1-u)$, which is determined by the urban unemployment rate

$$u = L_u/(L_m + L_u) \tag{11}$$

In analogy to the wage dualism developed here we can relax the assumption of rental rate equalisation across sectors, replacing the parity condition in equation (7) with

$$r_a = hr_m \qquad h > 0 \tag{12}$$

In the presence of rental rate dualism the equilibrium capital allocation will result in lower output than in the first-best solution. The resulting open economy Harris-Todaro model is represented by equations (1)-(3), (5)-(6), the labour market conditions (9)-(11), the rental rate condition (12) and the assumption that the manufacturing wage is exogenously fixed above the agricultural wage, while returns to capital can differ freely across sectors. Since we are developing a small open economy and thus prices are fixed exogenously, the demand side and preferences need not enter our study of equilibrium in the economy (Temple, 2005; Córdoba & Ripoll, 2009). As will become clear, the above model encompasses the various modelling approaches taken in the existing literature on dual economy models.

2.2 Technology heterogeneity

2.2.1 Heterogeneity across sectors

From a technical point of view, an aggregate production function only offers an appropriate construct in cross-country analysis if the economies investigated do not display large differences in sectoral structure (Temple, 2005), since a single production function framework assumes common production technology across all firms facing the same factor prices. Take two distinct sectors within this economy, assuming marginal labour product equalisation and capital homogeneity across sectors, and

[8] Temple (2005) suggests migration restrictions, or institutional reasons such as minimum wage legislation, trade unions, or an efficiency wage system in manufacturing as possible sources of this wage gap. Further, migration costs between sectors should be regarded as non-negligible. Additional considerations relate to the family organisation of asset returns (Ranis & Fei, 1961) whereby the wage in agriculture is equal to the average, rather than the marginal labour product which results in too little employment in the modern sector (Robertson, 1999).

[9] Assuming risk-neutral agents who obtain no wage at all if unemployed.

Cobb-Douglas-type production technology. Then if technology parameters differ between sectors, aggregated production technology cannot be of the Cobb-Douglas form (Temple & Wößmann, 2006; Córdoba & Ripoll, 2009). Finding differential technology parameters in sectoral production function estimation thus is potentially a serious challenge to treating production in form of an aggregated function.

An alternative motivation for a focus on sector-level rather than aggregate growth across countries is as follows: it is common practice to exclude oil-producing countries from any aggregate growth analysis, since "the bulk of recorded GDP for these countries represents the extraction of existing resources, not value added" (Mankiw et al., 1992, p.413). The underlying argument is that sectoral 'distortions', such as resource wealth, justify the exclusion of the country observations. By extension of the same argument, we could suggest that given the large share of agriculture in GDP for countries such as Malawi (25-50%), India (25-46%) or Malaysia (8-30%) over the period 1970-2000, these countries should be excluded from any *aggregate* growth analysis since a significant share of their *aggregate* GDP derives from a single resource, namely land.[10] Sector-level analysis, in contrast, does not face these difficulties, since sectors such as manufacturing or agriculture are defined closely enough to represent a reasonably homogeneous conceptual construct.

`[Table I about here]`

Having already indicated the importance of agriculture for GDP for a number of countries, we complete this section by providing some more data to highlight the importance and dynamics of agriculture in a wider set of countries. As can be seen in Table I the shift away from agriculture has been most dramatic in the East Asia group, whereas the Sub-Saharan Africa has seen virtually no change over the same period.

2.2.2 Heterogeneity across countries

A theoretical justification for heterogeneous technology parameters *across countries* 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, Kourtellos, & Minkin, 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.

[10]The quoted shares are from the World Bank World Development Indicators database (World Bank, 2008). For comparison, maximum share of oil revenue in GDP, computed as the difference between 'industry share in GDP' and 'manufacturing share in GDP' from the same database yields the following ranges for some of the countries omitted in Mankiw et al. (1992): Iran (12-51%), Kuwait (15-81%), Gabon (28-60%), Saudi Arabia (29-67%).

3. AN EMPIRICAL MODEL OF A DUAL ECONOMY

In seeking to understand processes of growth at the macro-level, empirical work has focused primarily on an aggregate production specification (see surveys in Barro & Sala-i-Martin, 1995; Aghion & Howitt, 1998; Temple, 1999; Aghion & Durlauf, 2005). While duality has featured prominently in theoretical developments there has been only a very limited matching of this theory to empirical models. This disjunction between theory and testing has reflected in large part the availability of data. In this paper we employ a large-scale cross-country dataset made publicly available by the World Bank in 2003 (henceforth Crego et al (1998), although the data is also described in detail in Larson et al., 2000) which allows us to specify manufacturing and agricultural production functions and thus provides a macro-model of a dual economy that can be compared with the single sector models dominating the empirical literature. In the following we first present a general empirical specification for our sector-specific analysis of agriculture and manufacturing. Next we review a number of empirical estimators, focusing in particular on those arising from the recent panel time series literature, before we briefly discuss the data.

3.1 Empirical specification

The analysis of growth and development using cross-country data is still dominated by variants on the 'convergence equation' introduced by Mankiw et al. (1992), where variables are averaged over the entire time-horizon and estimation is carried out in a single cross-country regression (Durlauf et al., 2005). The multiple short-comings of this approach have been discussed elsewhere in great detail, most recently in Eberhardt and Teal (2010). The latter also point to a number of modelling concerns we will address in our empirical analysis, namely parameter heterogeneity, cross-section dependence and variable time-series properties. Briefly, the notion that equilibrium relationships may differ fundamentally across countries (perhaps at different stages of development) is a familiar one, both in the theoretical (Murphy et al., 1989; Azariadis & Drazen, 1990; Durlauf, 1993; Banerjee & Newman, 1993) and empirical literatures (Durlauf et al., 2001; Basturk, Paap, & Dijk, 2008; Kourtellos, Stengos, & Tan, 2008; Cavalcanti, Mohaddes, & Raissi, 2009) on cross-country growth. In contrast, the notion of cross-section correlation, hypothesised to arise from common global shocks and/or local spillover effects, and concerns about variable nonstationarity have received little attention in the mainstream growth empirics literature. This is despite the rapid developments in econometrics theory over recent years (Bai & Ng, 2004; Andrews, 2005; Pesaran, 2006, 2007; Kapetanios, Pesaran, & Yamagata, 2009; Bai, 2009; Bai, Kao, & Ng, 2009), particularly in panel time series econometrics. In the context of cross-country growth and development analysis, the potential for cross-section dependency is particularly salient, given the interconnectedness of countries through history, geography and trade relations. Besides a number of spatial econometric approaches, where the nature of the spatial association is imposed by the econometrician (Conley & Ligon, 2002; Ertur & Koch, 2007), only a limited number of applied papers have concerned themselves with these matters (Eberhardt & Teal, 2008; Cavalcanti et al., 2009; Costantini & Destefanis, 2009).

Our empirical setup will follow the general model laid out in Eberhardt and Teal (2010), adopting a common factor representation for a standard log-linearised Cobb-Douglas production function model. Each sector/level of aggregation (agriculture, manufacturing, aggregated data, PWT data) is modelled separately — for ease of notation we do not identify this multiplicity in our general model. Let

$$y_{it} = \boldsymbol{\beta}_i' \boldsymbol{x}_{it} + u_{it} \qquad u_{it} = \alpha_i + \boldsymbol{\lambda}_i' \boldsymbol{f}_t + \varepsilon_{it} \tag{13}$$

$$x_{mit} = \pi_{mi} + \boldsymbol{\delta}_{mi}' \boldsymbol{g}_{mt} + \rho_{1mi} f_{1mt} + \ldots + \rho_{nmi} f_{nmt} + v_{mit} \tag{14}$$

$$\boldsymbol{f}_t = \boldsymbol{\varrho}' \boldsymbol{f}_{t-1} + \boldsymbol{\omega}_t \qquad \text{and} \qquad \boldsymbol{g}_t = \boldsymbol{\kappa}' \boldsymbol{g}_{t-1} + \boldsymbol{\epsilon}_t \tag{15}$$

for $i = 1, \dots, N$, $t = 1, \dots, T$ and $m = 1, \dots, k$, where $\boldsymbol{f}_{\cdot mt} \subset \boldsymbol{f}_t$ and the error terms ε_{it}, v_{mit}, $\boldsymbol{\omega}_t$ and $\boldsymbol{\epsilon}_t$ are white noise. Equation (13) represents the production function model, with y as sectoral or aggregated value-added and $\boldsymbol{x}$ as a set of inputs: labour, physical capital stock, and a measure for natural capital stock (arable land and permanent crops) in the agriculture specification (all variables are in logs). We consider additional inputs (human capital, livestock, fertilizer) as robustness checks for our general findings. The output elasticities associated with each input ($\boldsymbol{\beta}_i$) are allowed to differ across countries. 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.[11] Equation (15) provides some structure for the unobserved common factors, which are modelled as simple AR(1) processes, where we do not exclude the possibility of unit root processes ($\boldsymbol{\varrho} = 1$, $\boldsymbol{\kappa} = 1$) leading to nonstationary observables. Note that from this the potential for spurious regression results arises if the empirical equation is misspecified. Equation (14) details the evolution of the set of $m = 1, \dots, k$ regressors; crucially, some of the common factors contained in the covariates are also assumed to be driving the unobservables in the production function equation (u_{it}). This setup leads to endogeneity whereby the regressors are correlated with the unobservables, making it difficult to identify $\boldsymbol{\beta}_i$ separately from $\boldsymbol{\lambda}_i$ and $\boldsymbol{\rho}_i$ (Kapetanios et al., 2009).

Our empirical specification thus allows for a maximum of flexibility with regard to the impact of observables and unobservables on output. Empirical implementation will necessarily lead to different degrees of restrictions on this flexibility, which will then be tested by formal statistical means: the emphasis is on comparison of different empirical estimators allowing for or restricting the heterogeneity in observables and unobservables outlined above. A conceptual justification for the pervasive character of unobserved common factors is provided by the nature of macro-economic variables in a globalised world. In our mind latent forces drive all of the variables in our model, and their presence makes it difficult to argue for the validity of traditional approaches to causal interpretation of cross-country growth analyses. For instance, instrumental variable estimation in standard cross-section growth regressions (Clemens & Bazzi, 2009, p.2) or (Arellano & Bond, 1991)-type lag-instrumentation in pooled panel models (Pesaran & Smith, 1995; Lee, Pesaran, & Smith, 1997) are both invalid in the face of common factors and/or heterogeneous equilibrium relationships. We now introduce a novel estimation approach developed by Pesaran (2006) which allows us to bypass these issues by adopting panel time series methods for estimation and inference.

3.2 Empirical implementation

Our empirical approach emphasises the importance of parameter and factor loading heterogeneity across countries. The following 2×2 matrix indicates how the various estimators implemented below account for these matters.[12]

[11] The parameters $\boldsymbol{\beta}_i$ are unknown random coefficients with fixed means and finite variances. The same applies for the unknown factor loadings, i.e. $\boldsymbol{\lambda}_i = \boldsymbol{\lambda} + \boldsymbol{\eta}_i$ where $\boldsymbol{\eta}_i \sim \text{iid}(0, \Omega_\eta)$, similarly for $\boldsymbol{\delta}_{mi}$ and $\boldsymbol{\rho}_{mi}$. 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).

[12] Abbreviations: POLS — Pooled OLS, 2FE — 2-way Fixed Effects, FD2FE — OLS with variables in first differences and accounting for year fixed-effects, GMM — Arellano and Bond (1991) Difference GMM and Blundell and Bond (1998) System GMM, MG — Pesaran and Smith (1995) Mean Group with linear country trend, FDMG — dto. but with variables in first difference and country drift, PMG — Pesaran, Shin, and Smith (1999) Pooled Mean Group estimator, CPMG — dto. but augmented with cross-section averages following Binder and Offermanns (2007), CCEP/CMG — Pesaran (2006) Common Correlated Effects estimators. Note that like our 2FE estimator the OLS models is augmented with $T-1$ year dummies.

		Factor loadings: homogeneous	heterogeneous
Technology parameters:	homogeneous	POLS, 2FE, FD2FE, GMM, PMG	CCEP, CPMG
	heterogeneous	MG, FDMG	CMG

The Common Correlated Effects estimator developed in Pesaran (2006) and extended to nonstationary variables in Kapetanios et al. (2009) augments the regression equation with cross-section averages of the dependent and independent variables to account for the presence of unobserved common factors. For the Mean Group version (CMG), the individual country regression is specified as

$$y_{it} = a_i + \boldsymbol{b}_i' \boldsymbol{x}_{it} + c_{0i}\bar{y}_t + \sum_{m=1}^{k} \boldsymbol{c}_{mi}\bar{x}_{mt} + e_{it} \tag{16}$$

for $k = 1, \ldots, m$ covariates and e_{it} white noise, whereupon the parameter estimates are averaged across countries akin to the Pesaran and Smith (1995) Mean Group estimator. The pooled version (CCEP) is specified as

$$y_{it} = a_i + \boldsymbol{b}' \boldsymbol{x}_{it} + \sum_{j=1}^{N} c_{0i}(\bar{y}_t D_j) + \sum_{m=1}^{k}\sum_{j=1}^{N} \boldsymbol{c}_{mi}(\bar{x}_{mt} D_j) + e_{it} \tag{17}$$

Thus in the MG version we have individual country regressions with $2k+2$ RHS variables (including the intercept) and in the pooled version we have a single regression equation with $k+(k+2)N$ RHS variables (including N intercepts), where k is the number of observed covariates.

In order to get an insight into the workings of this approach, consider the cross-section average of our common factor model in equation (13): given that $\bar{\varepsilon}_t = 0$

$$\bar{y}_t = \bar{\alpha} + \bar{\boldsymbol{\beta}}'\bar{\mathbf{x}}_t + \bar{\boldsymbol{\lambda}}'\bar{\mathbf{f}}_t \tag{18}$$

which can be expressed as

$$\bar{\mathbf{f}}_t = \bar{\lambda}^{-1}(\bar{y}_t - \bar{\alpha} - \bar{\boldsymbol{\beta}}'\bar{\mathbf{x}}_t) \tag{19}$$

where the D_j represent country dummies. Thus we can see that the unobserved common factors can be captured by the cross-sectional means of y and $\mathbf{x}$ since $\bar{f}_t \xrightarrow{p} f_t$ as $N \to \infty$. Given the assumed heterogeneity in factor loadings across countries ($\boldsymbol{\lambda}_i'$) the estimator is implemented in the fashion detailed above which allows for each country i to have different parameter estimates on $\bar{y}_t$ and the $\bar{\mathbf{x}}_t$. Simulation studies (Pesaran, 2006; Coakley et al., 2006; Kapetanios et al., 2009) have shown that this approach performs well even when the cross-section dimension N is small, when variables are nonstationary, subject to structural breaks and in the presence of weak unobserved factors.

We abstract from discussing the standard panel estimators here in great detail and refer to the overview article by Coakley et al. (2006), as well as the article by Bond (2002) for more details. As a robustness check we also investigate the Pooled Mean Group estimator by Pesaran et al. (1999). For a detailed discussion of the pooled Mean Group estimator in the context of cross-country regressions refer to Arnold, Bassanini, and Scarpetta (2007); we further implement a simple extension to the PMG where we include cross-section averages of the dependent and independent variables, as suggested in Binder and Offermanns (2007).

A number of alternative nonstationary panel estimators for the case of homogeneous factor loadings are available in the literature (Pedroni, 2000, 2001), however given our emphasis on cross-section dependence we do not consider them in this work. Finally, we do not adopt any empirical methods accommodating unobserved factor via a two-step method where the number of significant factors in an equilibrium relationship is determined first (Bai & Ng, 2002) before estimates of the factors, loadings and slope parameters are determined jointly (Bai & Kao, 2006; Bai et al., 2009). The reason for this choice is the failure of these methods to account for cross-section dependence of the 'weak' type, such as that arising from local spillover effects (Chudik, Pesaran, & Tosetti, 2009), whereas the CCE estimators are robust to both cross-section dependence of the 'strong' and 'weak' type (Pesaran & Tosetti, 2007) — in fact, it can be shown that the method is robust to the inclusion of an infinite number of weak factors (Chudik et al., 2009).

3.3 Data description

Descriptive statistics and a more detailed discussion of the data can be found in the Appendix. Briefly, we conduct all empirical analysis for four datasets:

(1) for the agricultural sector, building on the sectoral investment series developed by Crego, Larson, Butzer, and Mundlak (1998) and output from the World Development Indicators WDI World Bank (2008), as well as sectoral labour and land data and FAO (2007);
(2) for the manufacturing sector, building on the sectoral investment series developed by Crego et al. (1998), output data from the WDI and labour data from UNIDO (2004);
(3) for a stylised aggregate economy made up of the summed data for the agriculture and manufacturing sectors;
(4) for the aggregate economy, building on data provided by the Penn World Table (PWT; we use version 6.2, Heston, Summers, & Aten, 2006).

The capital stocks in the agriculture, manufacturing and PWT samples are constructed from investment data following the perpetual inventory method (see Klenow & Rodriguez-Clare, 1997b, for details), for the aggregated sample we simple added up sectoral capital stock for agriculture and manufacturing. Comparison across sectors and with the stylised aggregate sector is possible due to the efforts by Crego et al. (1998) in providing sectoral investment data for agriculture and manufacturing. All monetary values in the sectoral and aggregated datasets are transformed into US$ 1990 values (in the capital stock case this transformation is applied to the investment data before the capital stocks are constructed), following the suggestions in Martin and Mitra (2002). Given concerns that the stylised aggregate economy data may not represent a good proxy for aggregate economy data we have adopted the PWT data, which measures monetary values in International $ PPP, as a benchmark for comparison — despite a number of vocal critics (Johnson, Larson, Papageorgiou, & Subramanian, 2009) the latter is without doubt the most popular macro dataset for cross-country empirical analysis. We are of course aware that the difference in deflation between our sectoral and aggregated data on the one hand and PWT on the other makes them conceptually very different measures of growth and development: the former emphasise tradable goods production whereas the latter puts equal emphasis on tradable and non-tradable goods and services. However, we believe that these differences are comparatively unimportant for estimation and inference in comparison to the distortions introduced by neglecting the sectoral makeup and technology heterogeneity of economies at very different stages of economic development.

Our sample is an unbalanced panel for 1963 to 1992 made up of 41 developing and developed countries with a total of 928 observations (average $T = 22.6$) — our desired aim to compare estimates across the four datasets requires us to we match the same sample, thus reducing the number of observations to the smallest common denominator. A detailed description of the sample is available in Table A-I, descriptive statistics in Table A-II are provided for each of the four samples (both tables can be found in the appendix).

Note that in our production function regressions we adopt a very common trick whereby the output and non-labour input variables are all expressed in 'per worker' terms (all variables in logs). If the labour variable is added to this its estimated parameter coefficient provides a simple test for constant returns to scale: if insignificant the relationship is subject to constant returns, if positive (negative) significant the relationship is suggested to be subject to increasing (decreasing) returns. In addition, this setup allows for easy imposition of constant returns by simply dropping the labour variable from the regression equation. Variable tests for stationarity and cross-section dependence are therefore carried out for the variables entering the regression equations, namely output, capital, land (all in logs of per worker terms) and labour (in logs).

4. EMPIRICAL RESULTS

Preliminary data analysis (unit root and cross-section dependence tests) have been confined to the technical appendix of the paper. We adopt the Pesaran (2007) CIPS panel unit root test which assumes a single unobserved common factor. This is clearly restrictive, however given the data restrictions (unbalanced panel, relatively short) we were unable to implement the newer CIPSM version of this test (Pesaran, Smith, & Yamagata, 2009) which allows for multiple common factors. Results (see Table TA-1) strongly suggest that variables in levels for all four datasets are nonstationary. Additional analysis of variables in first difference further suggests that our measure for agricultural labour may in fact be I(2) — this is almost definitely the outcome of variable construction: FAO (2007) data on economically active population in agriculture (and for that matter for all the other labour-related measures) are not evaluated annually, but at 5- or 10-year intervals.

In the following we discuss the empirical results from sectoral production function regressions for agriculture and manufacturing respectively, first assuming technology parameter homogeneity (Section 4.1) and then allowing for differential technology across countries (Section 4.2).

4.1 Pooled models

Table II presents the empirical results for agriculture and manufacturing, Panel A for unrestricted returns to scale and Panel B for the specification with CRS imposed. Beginning with agriculture, the empirical estimates for the models neglecting cross-section dependence are quite similar, with the capital coefficient around .63 and statistically significant decreasing returns to scale. Diagnostic tests indicate that the residuals in these models are cross-sectionally dependent, and that the levels models (POLS, 2FE) have nonstationary residuals and thus may represent spurious regressions. It is important to point out that in the presence of nonstationary residuals the t-statistics in the levels models are invalid (Kao, 1999). The CCEP model yields cross-sectionally independent and stationary residuals, a capital coefficient of around .5 and insignificant land coefficient. Imposition of CRS does not change these results substantially, with the exception of the 2FE estimates, where the land variable (previously negative and significant) is now insignificant and the capital coefficient has been inflated.

In the manufacturing data the models ignoring cross-section dependence yield increasing returns to scale and capital coefficients in excess of .85 for POLS and 2FE while the FD model yields .7. While residuals for the former two models again display nonstationarity the CD tests now suggest that they are cross-sectionally independent. Surprisingly the CCEP model, with a capital coefficient of around

.5 (like in agriculture) does not pass the cross-section correlation test. Following imposition of CRS all models reject cross-section independence, while parameter estimates are more or less identical to those in the unrestricted models.

`[Table II about here]`

Based on these pooled regression results, the diagnostic tests seem to favour the CCEP results in the agriculture data, whereas in the manufacturing data no estimator seems without concern. For the agriculture sample we conducted a number of robustness checks, including further covariates (livestock per worker, fertilizer per worker) in the pooled regression framework. Results (available on request) did not change from those presented above, with the CCEP estimator emerging as the most reliable empirical model.[13] The CCEP therefore remains our estimator of choice for the pooled agriculture data. For manufacturing, we conducted robustness checks including human capital in the estimation equation (linear & squared terms)[14] — as a result a number of countries drop out of our sample (CRI, IRN, KOR, MDG) which now contains $n = 860$ observations ($N = 37$). Results for unconstrained and CRS regressions are presented in Table III.

`[Table III about here]`

Results for the CCEP are very similar as in the previous specifications: the capital elasticity is around .5, while CRS is not rejected by the data. Returns to eduction follow a concave function (wrt years of schooling) and for the mean education value across countries are quite high in these models, around 8%pa and 11% pa in the unrestricted and restricted models respectively. In either case residuals are stationary and cross-sectionally independent. Our shift to heterogeneous technology models in the next session will allow us to judge whether violation of the homogeneity assumption is at the heart of the problem.

4.2 Averaged country regressions

Table IV presents the robust means for each regressors across N country regressions for the unrestricted (Panel A) and CRS models (Panel B) respectively. We adopt robust means[15] as these are more reliable than unweighted means, which are subject to greater distortion by outliers. The t-statistics reported for each average estimate test whether the average parameter is statistically different from zero, following Pesaran et al. (2009). In addition we also provide test statistics for the 'panel t-statistic' following Pedroni (1999).

`[Table IV about here]`

Beginning with the unrestricted models in Panel (A), we can see that MG and FDMG suffer from high imprecision in both agriculture and manufacturing equations. This aside, in the agriculture model MG yields decreasing returns to scale that are nonsensical in magnitude. Monte Carlo simulations for nonstationary and cross-sectionally dependent data (Coakley et al., 2006; Bond & Eberhardt, 2009) frequently show that MG estimates are commonly severely affected by their failure to account for cross-section dependence. As in the pooled models, the CMG estimator yields an insignificant land coefficient in agriculture and in both sectors results are generally very much in line with the CCEP

[13] In some more detail: The specification including livestock and fertilizer (both in log of per worker terms) could not reject CRS. The CRS specification yielded a capital coefficient of .383 [$t = 5.64$] which is lower than the comparable elasticity presented in Table II (.493) but the two estimates are still contained in each other's 95% confidence intervals; residual diagnostics indicate stationary and cross-sectionally independent resiudals. The livestock coefficient of .097 [$t = 3.70$] seems to capture the difference.

[14] We follow convention and pick the average years of schooling in the population as a proxy for Human Capital stock. We assume that the aggregate economy data for schooling developed by Barro and Lee (2001) which is available in 5-year intervals. Simple interpolation to obtain annual data (as is done here) is not ideal, however the evolution of this variable over time is commonly very stable (linear), s.t. we do not feel that linear interpolation creates additional issues.

[15] We use robust regression to produce a robust estimate of the mean — see Hamilton (1992) for details.

results in Table II.

All unrestricted models yield stationary residuals, however the agriculture CMG is the only model that cannot reject cross-sectionally independent residuals. Moving to the models where CRS is imposed in Panel (B), we can see that MG and FD-MG estimates are now somewhat more precise, while the CMG estimates are virtually unchanged. The residual diagnostics are sound in the agriculture CMG, but the manufacturing CMG still suffers from cross-sectionally dependent residuals. We therefore implement an alternative specification for manufacturing which includes a proxy for human capital (average years of schooling in the adult population) as additional covariate.

[Table V about here]

We also estimated the human-capital augmented models for manufacturing allowing for heterogeneous technology parameters. Results for the MG and FDMG in Table V mirror those in the unaugmented models presented above. In the unrestricted models these estimators yield very imprecise estimates, although if CRS is imposed the capital coefficients are again estimated around .3; average estimates on the linear and quadratic eduction terms are insignificant and the implied returns to education are negative albeit insignificant by the robust regression approach we adopted. For the CMG models we find capital coefficients somewhat below those in the unaugmented models, albeit still within each other's 95% confidence intervals. Average education coefficients are significant in both models (marginally so in the CRS version) and indicate rather high returns to eduction: 11% and 12% in the unrestricted and CRS model respectively. This merits two comments: firstly, we might argue for the validity of this high estimate given that the manufacturing sector is arguably the more dynamic sector in comparison to agriculture, which builds on innovation and R&D, thus relying on knowledge-accumulation which is heavily linked to human capital. Secondly, adopting country-wide education data as a proxy for manufacturing may severely distort the results we present here — firm-level data (Baptist & Teal, 2008) on Ghana and Korea in the late 1990s for instance indicates that average worker eduction (crudely measured in years of schooling) does not differ substantially between these countries.

5. AGGREGATION BIAS

In this section we will return to the themes developed in Section **??** and ask what the implications of the dual economy model are for aggregate cross-country growth analysis. First we discuss the econometric concerns arising from aggregation of heterogeneous sectoral data created by separate technologies. Our hypotheses are put to the test in a Monte Carlo simulation exercise of stylised aggregate data constructed from heterogeneous sectors. We then investigate whether the assumption of an aggregate production function yields biased estimation results. To the best of our knowledge this is the first paper to consider this issue empirically, enabled by pioneering work of Crego et al. (1998) in providing comparable investment and capital stock measures for Agriculture and Manufacturing. Finally we test our hypotheses about the sources of aggregation bias in cross-country growth empirics using our data.

5.1 Aggregation bias — Conceptual development

This section provides an insight into the problems for estimation arising from aggregation.

[Add conceptual literature on aggregation and econometrics considerations.]

5.2 Aggregation bias — some Monte Carlo evidence

This section provides simulation results based on a sample of stylised aggregate economies made up of two heterogeneous sectors.

`[Add Monte Carlo simulations.]`

5.3 Aggregation bias — empirical evidence

Our empirical results in Section 4 suggested fairly similar pooled and averaged capital coefficients for manufacturing and agriculture across the various empirical models. This might lead one to suggest that carrying out cross-country growth empirics may best be conducted taking the aggregate economy, and thus the Penn World Table (PWT) data (Heston et al., 2009), as the basic unit of analysis. Our empirical approach emphasised the importance of unobserved heterogeneity across countries, but our analysis refrained from testing technology parameter differences across sectors with any formal methods — our justification is that in our most flexible specification (CMG) the individual country-estimates are not reliable (Pedroni, 2007) and should not be the basis for comparison. In this section we will instead provide practical evidence that the use of an *aggregate* production function will lead to seriously biased estimates of the capital coefficient. We carry out this analysis by creating a stylised 'aggregated economy' from our data on agriculture and manufacturing. Since it might be suggested that results could be severely distorted by the overly simplistic nature of our inquiry, we compare results with those from a matched sample of data from the PWT.

We begin with the pooled models in Table VI. Across all specifications the estimated capital coefficients in the aggregated data far exceed those derived from the respective agriculture and manufacturing samples in Table II. Furthermore, the patterns in the aggregated data are replicated one-to-one in the PWT data, which also yields excessively high capital coefficients across all models. All models suffer from cross-sectional dependence in the residuals, while there are also indications that the residuals in the CCEP model for the aggregated data are nonstationary (those in the two other levels specifications are *always* nonstationary). We also investigated the impact of human capital (proxied via average years of schooling attained in the population over 15 years of age) in these aggregate economy data models, but as Table TA-IV in the Technical Appendix reveals the basic bias remains.

`[Table VI about here]`

In addition we estimated pooled dynamic models (introducing the PMG and CPMG estimators) in Table TA-III in the Technical Appendix — *all* of these results follow the patterns we described in this and the previous sections.

Turning to the results from averaged country regressions in Table VII. The MG and FDMG model point to some differences between the aggregated and PWT data, whereby the capital coefficients in the former are very imprecisely estimated but seem to centre around .3, whereas in the latter they are considerably higher at around .7 to .9. Results for the CMG, however, are again very consistent between the two data samples and across unrestricted and CRS models, with capital coefficients around .7. Residual testing suggests that all specification yield stationary residuals — this is somewhat surprising in the MG case, given the misspecification implicit in this equation. Cross-section correlation tests reject independence in all residual series tests — in case of the aggregated data the CMG rejects marginally.

`[Table VII about here]`

The results from this exercise should perhaps be viewed with caution, given the overly stylised nature of the aggregate economy data we created from manufacturing and agriculture values. However, given the standard finding in aggregate growth regressions of an inflated capital coefficient around .7, we feel vindicated in our claim that empirical estimation at the aggregate economy level hides important structural differences within countries and yields misleading results.

6. CONCLUSIONS

In this paper we have developed a general framework for dual economy models and used some unique panel data for agriculture and manufacturing to estimate sector-level and aggregated production functions. Our conceptual development built on the theoretical contributions by Temple (2005) and Corden and Findlay (1975) and our overview of the literature highlighted a number of 'sources of growth' hypothesised in the dual economy literature. Our empirical analysis emphasised the contribution of the recent panel time-series econometric literature, which suggests to adopt unobserved common factors to deal with the cross-sectional dependence commonly found in macro panel data. In addition we took the nonstationarity of observable and unobservable factor inputs into account and emphasised the importance of parameter heterogeneity — across countries as well as sectors.

We draw the following conclusions from our first, crude attempt at highlighting the importance of structural makeup and change in the empirical analysis of cross-country growth and development:

(i) Empirical analysis of growth and development at the cross-country level — most commonly conducted using the Penn World Tables (Heston et al., 2009) — gains considerably from the separate consideration of modern and traditional sectors that make up the economy. Our analysis of agriculture and manufacturing versus a stylised aggregated economy suggests that the latter yields severely distorted empirical results. Across multiple empirical specifications and estimators we could show that the capital coefficient for aggregated data far exceeds that obtained from separate sector regressions. Analysis of PWT data in parallel with the aggregated data suggested that this finding is not an artefact of our stylised empirical setup. To the best of our knowledge this is the first time that these matters are investigated empirically at this level of aggregation. Our analysis was enabled by the unique data on agricultural and manufacturing investment and capital stock developed by Crego et al. (1998) — a dataset which deserves far greater attention than it presently receives.[16]

(ii) Monte Carlo study results.

(iii) Implications for future work.

[16]Don Larson and collaborators have more recently developed an updated version of this dataset (1967-2003), albeit limited to 30 developing and developed countries. In future work we plan to use this new version and a matched manufacturing dataset to investigate the robustness of our results.

REFERENCES

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

Aghion, P., & Durlauf, S. (Eds.). (2005). *Handbook of economic growth* (Vol. 1) (No. 1). Elsevier.

Aghion, P., & Howitt, P. (1992). A model of growth through creative destruction. *Econometrica*, *60*(2), 323-51.

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

Andrews, D. W. K. (2005). Cross-Section Regression with Common Shocks. *Econometrica*, *73*(5), 1551-1585.

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

Arnold, J., Bassanini, A., & Scarpetta, S. (2007). *Solow or Lucas?: Testing Growth Models Using Panel Data from OECD Countries* (OECD Economics Department Working Papers No. 592).

Azariadis, C., & Drazen, A. (1990). Threshold externalities in economic development. *Quarterly Journal of Economics*, *105*(2), 501-26.

Bai, J. (2009). Panel Data Models with Interactive Fixed Effects. *Econometrica*, *77*(4), 1229-1279.

Bai, J., & Kao, C. (2006). On the estimation and inference of a panel cointegration model with cross-sectional dependence. In B. H. Baltagi (Ed.), *Panel data econometrics: Theoretical contributions and empirical applications.* Amsterdam: Elsevier Science.

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

Bai, J., & Ng, S. (2002). Determining the Number of Factors in Approximate Factor Models. *Econometrica*, *70*(1), 191-221.

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

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.

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)

Barbier, E. B., & Rauscher, M. (2007). Trade and development in a labor surplus economy. *The B.E. Journal of Economic Analysis & Policy*, *7*(1).

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

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. (2001). International data on educational attainment: Updates and implications. *Oxford Economic Papers*, *53*(3), 541-63.

Barro, R. J., & Sala-i-Martin, J. (1995). *Economic growth.* New York: MacGraw-Hill.

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

Binder, M., & Offermanns, C. J. (2007). *International investment positions and exchange rate dynamics: a dynamic panel analysis* (Discussion Paper Series 1: Economic Studies Nos. 2007,23). Deutsche Bundesbank, Research Centre.

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

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

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

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

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.

Cass, D. (1965). Optimum growth in an aggregative model of capital accumulation. *Review of Economic Studies*, *32*(3), 233-240.

Cavalcanti, T., Mohaddes, K., & Raissi, M. (2009). *Growth, development and natural resources: New evidence using a heterogeneous panel analysis* (Cambridge Working Papers in Economics (CWPE) No. 0946). (November 2009)

Chen, E. K. (1997). The Total Factor Productivity debate: Determinants of economic growth in East Asia. *Asian-Pacific Economic Literature*, *11*(1), 18-38.

Chudik, A., Pesaran, M. H., & Tosetti, E. (2009). *Weak and Strong Cross Section Dependence and Estimation of Large Panels* (Cambridge Working Papers in Economics (CWPE) No. 0924). (June 2009)

Clemens, M., & Bazzi, S. (2009). *Blunt Instruments: On Establishing the Causes of Economic Growth* (Working Papers No. 171).

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

Coe, D. T., Helpman, E., & Hoffmaister, A. W. (1997). North-South R&D spillovers. *Economic Journal*, *107*(440), 134-49.

Conley, T., & Ligon, E. (2002). Economic Distance and Long-run Growth. *Journal of Economic Growth*, *7*(2), 157-187.

Corden, W. M., & Findlay, R. (1975). Urban unemployment, intersectoral capital mobility, and development policy. *Economica*, *42*, 59-78.

Córdoba, J. C., & Ripoll, M. (2009). Agriculture and aggregation. *Economics Letters*, *105*(1), 110-112.

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

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).

Dollar, D., & Kraay, A. (2002). Growth is good for the poor. *Journal of Economic Growth*, *7*(3), 195-225.

Dollar, D., & Kraay, A. (2004). Trade, growth, and poverty. *Economic Journal*, *114*(493), F22-F49.

Durlauf, S. N. (1993). Nonergodic economic growth. *Review of Economic Studies*, *60*(2), 349-66.

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.

Eberhardt, M., & Teal, F. (2008). *Modeling Technology and Technological Change in Manufacturing: How do Countries Differ?* [CSAE Working Paper, WPS/2008-12].

Eberhardt, M., & Teal, F. (2010). Econometrics for Grumblers: A New Look at the Literature on Cross-Country Growth Empirics. *Journal of Economic Surveys*. (forthcoming)

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)

Frankel, J. A., & Romer, D. (1999). Does trade cause growth? *American Economic Review*, *89*(3), 379-399.

Gollin, D., Parente, S., & Rogerson, R. (2002). The Role of Agriculture in Development. *American Economic Review*, *92*(2), 160-164.

Greenaway, D., Morgan, W., & Wright, P. (2002). Trade liberalisation and growth in developing countries. *Journal of Development Economics*, *67*(1), 229-244.

Grossman, G. M., & Helpman, E. (1991). *Innovation and growth in the global economy.* Cambridge, Mass: MIT Press.

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, L. C. (1992). How robust is robust regression? *Stata Technical Bulletin*, *1*(2).

Harris, J. R., & Todaro, M. P. (1970, March). Migration, unemployment & development: A two-sector analysis. *American Economic Review*, *60*(1), 126-42.

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)

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

Johnson, S., Larson, W., Papageorgiou, C., & Subramanian, A. (2009). *Is Newer Better? Penn World Table Revisions and Their Impact on Growth Estimates* (NBER Working Papers No. 15455). National Bureau of Economic Research, Inc.

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.

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)

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.

Kourtellos, A., Stengos, T., & Tan, C. M. (2008). *THRET: Threshold Regression with Endogenous Threshold Variables* (University of Cyprus Working Papers in Economics No. 3-2008).

Kuznets, S. (1961). Economic growth and the contribution of agriculture: notes on measurement. *International Journal of Agrarian Affairs*, *3*(1), 56-75.
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.
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)
Lucas, R. J. (1988). On the mechanics of economic development. *Journal of Monetary Economics*, *22*(1), 3-42.
Mankiw, G., Romer, D., & Weil, D. N. (1992). A Contribution to the Empirics of Economic Growth. *Quarterly Journal of Economics*, *107*(2), 407-437.
Martin, W., & Mitra, D. (2002). Productivity Growth and Convergence in Agriculture versus Manufacturing. *Economic Development and Cultural Change*, *49*(2), 403-422.
Matsuyama, K. (1992). Agricultural productivity, comparative advantage, and economic growth. *Journal of Economic Theory*, *58*(2), 317-334.
Moscone, F., & Tosetti, E. (2009). A Review And Comparison Of Tests Of Cross-Section Independence In Panels. *Journal of Economic Surveys*, *23*(3), 528-561.
Murphy, K. M., Shleifer, A., & Vishny, R. W. (1989). Industrialization and the Big Push. *Journal of Political Economy*, *97*(5), 1003-26.
Paauw, D. S., & Fei, J. C. (1973). *The Transition in Open Dualistic Economies — Theory and Southeast Asian Experience.* New Haven, CN: Yale University Press.
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. (2001). Purchasing power parity tests in cointegrated panels. *The Review of Economics and Statistics*, *83*(4), 727-731.
Pedroni, P. (2004). Panel Cointegration: Asymptotic and Finite Sample Properties of Pooled Time Series Tests with an Application to the PPP Hypothesis. *Econometric Theory*, *20*(3), 597-625.
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., Shin, Y., & Smith, R. (1999). Pooled mean group estimation of dynamic heterogeneous panels. *Journal of the American Statistical Association*, *94*, 289-326.
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. (2009). *Panel unit root tests in the presence of a multifactor error structure.* (unpublished working paper, September 2009)
Pesaran, M. H., & Tosetti, E. (2007). *Large panels with common factors and spatial correlations.* (IZA Discussion Papers No. 3032)
Ranis, G., & Fei, J. (1961). A theory of economic development. *American Economic Review*, *51*(4), 533-556.
Restuccia, D., Yang, D. T., & Zhu, X. (2008). Agriculture and Aggregate Productivity: A Quantitative Cross-Country Analysis. *Journal of Monetary Economics*, *55*(2), 234-250.
Rivera-Batiz, L. A., & Romer, P. M. (1991). Economic integration and endogenous growth. *The Quarterly Journal of Economics*, *106*(2), 531-55.
Robertson, P. E. (1999). Economic growth and the return to capital in developing economies. *Oxford Economic Papers*, *51*(4), 577-94.
Robinson, S. (1971). Sources of growth in less developed countries: A cross-section study. *Quarterly Journal of Economics*, *85*(3), 391-408.
Rodriguez, F., & Rodrik, D. (2001, March). Trade policy and economic growth: A skeptic's guide to the cross-national evidence. In *Nber macroeconomics annual 2000, volume 15* (p. 261-338). National Bureau of Economic Research, Inc.
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.
Swan, T. W. (1956). Economic growth and capital accumulation. *Economic Record*, *32*(2), 334-61.
Temin, P. (2002). The golden age of european growth reconsidered. *European Review of Economic History*, *6*(01), 3-22.
Temple, J. (1999). The new growth evidence. *Journal of Economic Literature*, *37*(1), 112-156.
Temple, J. (2001). *Structural change and europe's golden age* (CEPR Discussion Papers No. 2861). C.E.P.R. Discussion Papers.
Temple, J. (2005). Dual economy models: A primer for growth economists. *The Manchester School*, *73*(4), 435-478.
Temple, J., & Wößmann, L. (2006). Dualism and cross-country growth regressions. *Journal of Economic Growth*, *11*(3), 187-228.
UNIDO. (2004). *UNIDO Industrial Statistics 2004* [Online database, Vienna: UNIDO].
Vollrath, D. (2009a). *The Agricultural Basis of Comparative Development.* (Paper presented at the North Eastern Universities Development Conference, Boston, November 2009)
Vollrath, D. (2009b). The dual economy in long-run development. *Journal of Economic Growth*, *14*(4), 287-312.
Vollrath, D. (2009c). How important are dual economy effects for aggregate productivity? *Journal of Development Economics*, *88*(2), 325-334.
World Bank. (2008). *World Development Indicators.* Database.
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.

Tables and Figures

Table I: Evolution of agricultural VA- and employment-share

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	

Notes: † 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.

Table II: Pooled regression models for agriculture and manufacturing

Panel (A): Unrestricted returns to scale

	Agriculture				Manufacturing			
	[1] **POLS**	[2] **2FE**	[3] **CCEP**	[4] **FD2FE**	[5] **POLS**	[6] **2FE**	[7] **CCEP**	[8] **FD2FE**
log labour	-0.059	-0.205	-0.203	-0.113	0.043	0.069	0.089	0.125
	[7.06]**	[10.03]**	[1.73]	[3.13]**	[3.56]**	[3.68]**	[1.77]	[6.81]**
log capital pw	0.618	0.654	0.484	0.633	0.897	0.855	0.511	0.720
	[74.18]**	[42.29]**	[11.24]**	[21.00]**	[55.53]**	[32.93]**	[8.90]**	[23.95]**
log land pw	0.012	-0.151	-0.092	-0.001				
	[1.07]	[4.89]**	[0.64]	[0.01]				
Implied RS†	DRS	DRS	CRS	DRS	IRS	IRS	CRS	IRS
Implied β_L‡	0.323	0.346	0.516	0.254	0.146	0.214	0.489	0.405
$\hat{e}$ integrated♮	I(1)	I(1)	I(0)	I(0)	I(1)	I(1)	I(0)	I(0)
CD test p-value♯	0.00	0.00	0.57	0.00	0.44	0.55	0.00	0.00
R-squared	0.94	0.86	1.00	-	0.84	0.67	1.00	-

Panel (B): Constant returns to scale imposed

	Agriculture				Manufacturing			
	[1] **POLS**	[2] **2FE**	[3] **CCEP**	[4] **FD2FE**	[5] **POLS**	[6] **2FE**	[7] **CCEP**	[8] **FD2FE**
log capital pw	0.644	0.724	0.493	0.660	0.920	0.865	0.510	0.767
	[85.54]**	[48.86]**	[11.84]**	[22.70]**	[71.30]**	[34.11]**	[11.75]**	[25.60]**
log land pw	0.009	-0.005	0.108	0.002				
	[0.70]	[0.15]	[1.57]	[0.02]				
Implied β_L‡	0.348	0.281	0.399	0.338	0.080	0.135	0.490	0.233
$\hat{e}$ integrated♮	I(1)	I(1)/I(0)	I(0)	I(0)	I(1)	I(1)	I(0)	I(0)
CD test p-value♯	0.00	0.00	0.71	0.00	0.00	0.00	0.00	0.00
R-squared	0.94	0.85	1.00	-	0.84	0.66	1.00	-
Observations	928	928	928	879	928	928	928	879

Notes: Dependent variable: value-added per worker (in logs). All variables are suitably transformed in the 2FE and FD2FE equations. Estimators: POLS — pooled OLS, 2FE — 2-way Fixed Effects, CCEP — Common Correlated Effects Pooled version, FD2FE — 2-way Fixed Effects with variables in first difference. We omit reporting the estimates on the intercept term. t-statistics reported in brackets are constructed using White heteroskedasticity-robust standard errors. *, ** indicate significance at 5% and 1% level respectively. $N = 41$, average $T = 22.6$ (21.4 for FD2FE). Time dummies are included explicitly in [1] and [5] or implicitly in [2],[4],[6] and [8]. Cross-section average augmentation in [3] and [7].

† Returns to scale, based on significance of log labour estimate. ‡ Based on returns to scale result. ♮ Order of integration of regression residuals, determined using Pesaran (2007) CIPS (full results available on request). ♯ Based on Pesaran (2004) CD-test (full results for this and other CSD tests available on request).

Table III: Pooled regression models for manufacturing (HC-augmented)

	Panel (A): Unrestricted returns				Panel (B): CRS imposed			
	[1] **POLS**	[2] **2FE**	[3] **CCEP**	[4] **FD2FE**	[5] **POLS**	[6] **2FE**	[7] **CCEP**	[8] **FD2FE**
log labour	0.005	0.029	0.121	0.162				
	[0.62]	[0.88]	[1.91]	[4.62]**				
log capital pw	0.692	0.851	0.533	0.654	0.695	0.839	0.472	0.558
	[44.38]**	[22.14]**	[8.00]**	[14.56]**	[49.18]**	[24.30]**	[8.87]**	[13.85]**
Education	0.226	-0.006	0.152	0.095	0.226	0.014	0.234	0.220
	[11.91]**	[0.21]	[2.04]*	[1.53]	[11.80]**	[0.71]	[3.67]**	[3.91]**
Education^2	-0.009	0.002	-0.006	-0.005	-0.009	0.001	-0.010	-0.010
	[6.22]**	[1.39]	[1.32]	[1.10]	[6.11]**	[0.98]	[2.55]*	[2.41]*
Implied RS†	CRS	CRS	CRS	IRS				
Implied β_L‡	0.308	0.149	0.467	0.508	0.305	0.162	0.528	0.443
$\hat{e}$ integrated$^{\natural}$	I(1)	I(1)	I(0)	I(0)	I(1)	I(1)	I(0)	I(0)
CD test p-value$^{\sharp}$	0.87	0.18	0.58	0.00	0.88	0.04	0.08	0.00
Mean Education	5.87	5.87	5.87	5.94	5.87	5.87	5.87	5.94
Returns to Edu	12.2%	1.9%	8.4%	4.1%	12.2%	2.7%	11.6%	10.5%
[t-statistic]$^{\flat}$	[19.88]	[1.30]	[3.11]	[1.54]	[20.20]	[2.30]	[5.25]	[4.62]
Observations	860	860	860	817	860	860	860	817
R-squared	0.91	0.57	1.00	-	0.91	0.57	1.00	-

Notes: We include our proxy for education in levels and as a squared term. Returns to Education are computed from the sample mean ($\bar{E}$) as $\beta_E + 2\beta_{E2}\bar{E}$ where β_E and β_{E2} are the coefficients on the levels and squared education terms respectively. $\flat$ computed via the delta-method. For more details on other diagnostics see Notes in Table II.

Table IV: Heterogeneous parameter models (robust means)

PANEL (A): UNRESTRICTED RETURNS TO SCALE	*Agriculture*			*Manufacturing*		
	[1] **MG**	[2] **FDMG**	[3] **CMG**	[4] **MG**	[5] **FDMG**	[6] **CMG**
log labour	-1.936	-0.414	-0.533	-0.125	-0.154	0.094
	[2.50]*	[0.48]	[0.91]	[0.90]	[1.36]	[1.12]
log capital pw	-0.053	0.135	0.526	0.214	0.139	0.545
	[0.28]	[0.61]	[2.76]**	[1.38]	[0.84]	[6.34]**
log land pw	-0.334	-0.245	-0.352			
	[1.09]	[0.85]	[1.12]			
country trend/drift	0.018	0.010		0.014	0.019	
	[1.81]	[1.22]		[2.54]*	[3.35]**	
Implied RS†	DRS	CRS	CRS	CRS	CRS	CRS
Implied β_L‡	n/a	n/a	0.474	n/a	n/a	0.455
reject CRS (10%)	27%	12%	20%	44%	12%	39%
panel-t Labour	-3.17**	-0.93	-1.02	-2.98**	-2.92**	4.68**
panel-t Capital	0.89	0.95	8.10**	4.14**	0.09	16.15**
panel-t Land	-0.32	0.23	-0.02			
panel-t trend/drift	14.95**	5.41**		16.23**	8.35**	
sign. trends (10%)	20	7		19	10	
$\hat{e}$ integrated♮	I(0)	I(0)	I(0)	I(0)	I(0)	I(0)
abs correl.coeff.	0.23	0.22	0.25	0.24	0.22	0.23
CD-test (p)♯	(.00)	(.00)	(.51)	(.00)	(.00)	(.01)

PANEL (B): CONSTANT RETURNS TO SCALE IMPOSED	*Agriculture*			*Manufacturing*		
	[1] **MG**	[2] **FDMG**	[3] **CMG**	[4] **MG**	[5] **FDMG**	[6] **CMG**
log capital pw	-0.012	0.297	0.547	0.320	0.388	0.550
	[0.07]	[2.14]*	[4.66]**	[2.74]**	[4.02]**	[6.33]**
log land pw	0.360	0.138	0.163			
	[1.30]	[0.71]	[0.90]			
country trend/drift	0.016	0.014		0.011	0.011	
	[2.89]**	[3.09]**		[2.63]*	[3.06]**	
Implied β_L‡	1.012	0.703	0.453	0.680	0.612	0.450
panel-t Capital	5.42**	2.65**	13.68**	10.58**	6.36**	20.03**
panel-t Land	6.74**	1.53	1.24			
panel-t trend/drift	14.87**	5.61**		22.65**	8.39**	
sign. trends (10%)	22	6		31	15	
$\hat{e}$ integrated♮	I(0)	I(0)	I(0)	I(0)	I(0)	I(0)
abs correl.coeff.	0.23	0.22	0.26	0.29	0.22	0.26
CD-test (p)♯	(.00)	(.00)	(.90)	(.00)	(.00)	(.00)
Obs (N)	928 (41)	928 (41)	879 (41)	928 (41)	928 (41)	879 (41)

Notes: Dependent variable: value-added per worker (in logs). All variables are suitably transformed in the FD equations. Estimators: MG — Mean Group, FDMG — MG with variables in first difference, CMG — Common Correlated Effects Mean Group version. We report robust means; estimates on intercept terms are not shown. t-statistics in brackets following Pesaran et al. (2009). Panel-t statistic following Pedroni (2004). Estimates on cross-section averages in [3] and [6] not reported. For other details see Table II.

Table V: Heterogeneous Manufacturing models (HC-augmented)

	Panel (A): Unrestricted			Panel (B): CRS imposed		
	[1] **MG**	[2] **FDMG**	[3] **CMG**	[4] **MG**	[5] **FDMG**	[6] **CMG**
log labour	-0.305	-0.293	0.097			
	[1.20]	[1.50]	[0.62]			
log capital pw	0.059	0.144	0.426	0.352	0.347	0.386
	[0.22]	[0.74]	[3.73]**	[3.25]**	[3.66]**	[3.95]**
Education	-0.478	0.237	1.248	-0.228	0.085	0.668
	[1.02]	[0.81]	[2.66]*	[0.62]	[0.29]	[2.43]*
Education squared	0.050	0.011	-0.098	0.005	-0.019	-0.042
	[1.38]	[0.35]	[2.67]*	[0.13]	[0.67]	[1.95]
country trend/drift	0.016	0.020		0.008	0.013	
	[1.55]	[2.44]*		[1.16]	[2.23]*	
reject CRS (10%)	38%	8%	38%			
Implied β_L ‡	n/a	0.857	0.574	0.648	0.653	0.614
Mean Education	5.82	5.91	5.82	5.87	5.94	5.87
Returns to Edu	-6.3%	-1.3%	10.9%	-6.2%	-2.1%	11.9%
[t-statistic]♭	[1.01]	[0.25]	[1.89]	[1.00]	[0.47]	[1.70]
panel-t Labour	4.49**	-2.51*	1.81			
panel-t Capital	0.30	-0.25	8.62**	7.52**	5.48**	10.19**
panel-t Edu	2.08*	0.93	3.58**	3.08**	0.88	3.38**
panel-t Edu^2	1.93	-0.91	3.31**	2.47*	0.97	2.67**
panel-t trend/drift	12.59**	6.41**		13.89	7.05	
sign. trends (10%)	15	9		17	7	
$\hat{e}$ integrated♮	I(0)	I(0)	I(0)	I(0)	I(0)	I(0)
abs correl. coeff.	0.21	0.22	0.22	0.22	0.22	0.22
CD-test (p)♯	(.00)	(.00)	(.71)	(.00)	(.00)	(.27)
Obs (N)	775 (37)	732 (37)	775 (37)	775 (37)	732 (37)	775 (37)

Notes: All averaged coefficients presented are robust means across i. ♭ The returns to education and associated t-statistics are based on a two-step procedure: first the country-specific mean education value ($\bar{E}_i$) is used to compute $\beta_{i,E} + 2\beta_{i,E2}\bar{E}_i$ to yield the country-specific returns to education. The reported value then represents the robust mean of these N country estimates, s.t. the t-statistic should be interpreted in the same fashion as that for the regressors, namely as a test whether the average parameter is statistically different from zero, following Pesaran et al. (2009). For other details see Tables IV and V.

Table VI: Pooled regression models for aggregated and PWT data

PANEL (A): UNRESTRICTED RETURNS TO SCALE

	Aggregated data				Penn World Table data			
	[1] **POLS**	[2] **2FE**	[3] **CCEP**	[4] **FD2FE**	[5] **POLS**	[6] **2FE**	[7] **CCEP**	[8] **FD2FE**
log labour	0.011	-0.096	0.036	-0.013	0.034	-0.138	-0.201	0.019
	[1.50]	[4.49]**	[0.52]	[0.54]	[7.43]**	[4.74]**	[1.75]	[0.94]
log capital pw	0.829	0.792	0.655	0.820	0.742	0.700	0.684	0.729
	[108.41]**	[64.71]**	[21.71]**	[66.28]**	[114.77]**	[49.71]**	[16.90]**	[50.08]**
Implied RS†	CRS	DRS	CRS	CRS	IRS	DRS	CRS	CRS
Implied β_L‡	0.171	0.111	0.345	0.180	0.292	0.162	0.316	0.271
$\hat{e}$ integrated♮	I(1)	I(1)	I(1)/I(0)	I(0)	I(1)	I(1)	I(1)/I(0)	I(0)
CD test p-value♯	0.98	0.01	0.07	0.00	0.02	0.00	0.02	0.00
R-squared	0.96	0.88	1.00	-	0.96	0.82	1.00	
Observations	928	928	928	879	922	922	922	873

PANEL (B): CONSTANT RETURNS TO SCALE IMPOSED

	Aggregated data				Penn World Table data			
	[1] **POLS**	[2] **2FE**	[3] **CCEP**	[4] **FD2FE**	[5] **POLS**	[6] **2FE**	[7] **CCEP**	[8] **FD2FE**
log capital pw	0.825	0.823	0.672	0.821	0.730	0.745	0.656	0.726
	[120.85]**	[72.25]**	[23.14]**	[66.91]**	[130.53]**	[62.33]**	[20.61]**	[50.88]**
Implied β_L‡	0.175	0.177	0.328	0.179	0.270	0.256	0.344	0.274
$\hat{e}$ integrated♮	I(1)	I(1)	I(1)/I(0)	I(0)	I(1)	I(1)	I(0)	I(0)
CD test p-value♯	0.91	0.86	0.05	0.00	0.00	0.00	0.03	0.00
R-squared	0.96	0.88	1.00	-	0.96	0.81	1.00	
Observations	928	928	928	879	922	922	922	873

Notes: See Table II for details.

Table VII: Heterogeneous parameter models (robust means)

Panel (A): Unrestricted returns to scale	*Aggregated data*			*Penn World Table data*		
	[1] **MG**	[2] **FDMG**	[3] **CMG**	[4] **MG**	[5] **FDMG**	[6] **CMG**
log labour	-0.233	-0.169	0.057	-0.442	-1.089	-0.172
	[0.55]	[0.51]	[0.31]	[0.74]	[2.35]*	[0.45]
log capital pw	0.233	0.289	0.651	0.625	0.976	0.715
	[1.28]	[1.71]	[7.00]**	[4.64]**	[6.40]**	[5.49]**
country trend/drift	0.026	0.022		0.011	-0.005	
	[2.93]**	[2.57]*		[1.12]	[0.83]	
Implied RS†	CRS	CRS	CRS	CRS	DRS	CRS
Implied β_L‡	n/a	n/a	0.349	0.375	n/a	0.285
reject CRS (10%)	56%	15%	29%	74%	26%	51%
panel-t Labour	-0.77	-0.16	4.12**	-0.65	-4.42**	-4.36**
panel-t Capital	5.97**	1.83	22.39**	24.66**	18.12**	26.16**
panel-t trend/drift	23.44**	9.31**		16.65**	7.41**	
sign. trends (10%)	27	13		30	12	
$\hat{e}$ integrated$^\natural$	I(0)	I(0)	I(0)	I(0)	I(0)	I(0)
abs correl.coeff.	0.24	0.23	0.23	0.25	0.19	0.24
CD-test (p)$^\sharp$	(.00)	(.00)	(.00)	(.00)	(.00)	(.00)
Observations	928	928	879	922	922	873

Panel (B): Constant returns to scale imposed	*Aggregated data*			*Penn World Table data*		
	[1] **MG**	[2] **FDMG**	[3] **CMG**	[4] **MG**	[5] **FDMG**	[6] **CMG**
log capital pw	0.324	0.222	0.745	0.681	0.892	0.785
	[2.12]*	[2.09]*	[11.78]**	[8.38]**	[7.47]**	[12.59]**
country trend/drift	0.013	0.018		0.001	-0.004	
	[2.69]*	[4.65]**		[0.23]	[1.24]	
Implied β_L‡	0.676	0.778	0.255	0.319	0.108	0.215
panel-t Capital	11.61**	2.68**	40.06**	34.32**	18.49**	51.35**
panel-t trend/drift	21.26**	8.72**		19.33**	8.75**	
sign. trends (10%)	25	11		27	12	
$\hat{e}$ integrated$^\natural$	I(0)	I(0)	I(0)	I(0)	I(0)	I(0)
abs correl.coeff.	0.29	0.23	0.26	0.32	0.23	0.30
CD-test (p)$^\sharp$	(.00)	(.00)	(.07)	(.00)	(.00)	(.00)
Observations	928	928	879	922	922	873

Notes: See Table IV for details.

Appendix

A-1 Data construction and descriptives

We use a total of four datasets in our empirical analysis, comprising data for agriculture and manufacturing (Crego et al., 1998; UNIDO, 2004; FAO, 2007), an 'aggregated dataset' where the labour, output and capital stock values for the two sectors are added up, and finally a Penn World Table (PWT 6.2) dataset (Heston et al., 2006) for comparative purposes. It is important to stress that the former three datasets differ significantly in their construction from the latter, primarily in the choice of exchange rates and deflation: the former use international (US$-LCU) exchange rates for the year 1990, whereas the Penn World Table dataset comprises Purchasing Power Parity (PPP) adjusted International Dollars taking the year 2000 as the comparative base. The former thus put an emphasis on traded goods, whereas the latter are generally perceived to account better for non-tradables and service. Provided that all monetary values making up the variables used in each regression are comparable (across countries, times), and given that the comparison of sectoral and aggregated data with the PWT is for illustrative purposes, we do not feel there is an issue in presenting results from these two conceptually different datasets. In all cases the results present are for matched observations across datasets: the four datasets are identical in terms of countries and time-periods — we prefer this arrangement for direct comparison despite the fact that more observations are available for individual data sources (e.g. the PWT are now available in the latest version 6.3, covering up to 188 countries for 1950 to 2004, see Heston et al., 2009), which may improve the robustness of empirical estimates. We provide details on the sample makeup in Table A-I. The next two subsections describe data construction. Descriptive statistics for all variables in the empirical analysis are presented in Table A-II.

Table A-I: Descriptive statistics: Sample makeup

#	WBCODE	COUNTRY	OBS	#	WBCODE	COUNTRY	OBS
1	AUS	Australia	20	22	JPN	Japan	28
2	AUT	Austria	22	23	KEN	Kenya	29
3	BEL	Belgium-Luxembourg	22	24	KOR	South Korea	29
4	CAN	Canada	30	25	LKA	Sri Lanka	17
5	CHL	Chile	20	26	MDG	Madagascar	20
6	COL	Colombia	26	27	MLT	Malta	23
7	CRI	Costa Rica	10	28	MUS	Mauritius	16
8	CYP	Cyprus	18	29	MWI	Malawi	23
9	DNK	Denmark	26	30	NLD	Netherlands	23
10	EGY	Egypt	24	31	NOR	Norway	22
11	FIN	Finland	28	32	NZL	New Zealand	19
12	FRA	France	23	33	PAK	Pakistan	24
13	GBR	United Kingdom	22	34	PHL	Philippines	24
14	GRC	Greece	28	35	PRT	Portugal	20
15	GTM	Guatemala	19	36	SWE	Sweden	23
16	IDN	Indonesia	22	37	TUN	Tunisia	17
17	IND	India	29	38	USA	United States	23
18	IRL	Ireland	23	39	VEN	Venezuela	19
19	IRN	Iran	25	40	ZAF	South Africa	26
20	ISL	Iceland	20	41	ZWE	Zimbabwe	25
21	ITA	Italy	21			**Total**	**928**

Notes: Sample makeup for all 4 datasets.

Table A-II: Descriptive statistics

AGRICULTURE DATA

PANEL (A): VARIABLES IN UNTRANSFORMED LEVEL TERMS

Variable	mean	median	std. dev.	min.	max.
Output	1.74E+10	5.91E+09	2.95E+10	3.54E+07	2.24E+11
Labour	9.51E+06	1.21E+06	3.45E+07	3.00E+03	2.33E+08
Capital	6.42E+10	1.01E+10	1.45E+11	2.90E+07	8.64E+11
Land	1.73E+07	3.50E+06	4.06E+07	6.00E+03	1.91E+08
in logarithms					
Output	22.369	22.500	1.737	17.382	26.134
Labour	13.984	14.006	2.011	8.006	19.267
Capital	22.933	23.037	2.276	17.183	27.485
Land	15.089	15.068	1.986	8.700	19.066
in growth rates					
Output	1.75%	1.94%	10.36%	-41.54%	53.86%
Labour	-0.63%	0.00%	3.00%	-28.77%	13.35%
Capital	1.89%	1.25%	3.61%	-5.13%	31.40%
Land	0.06%	0.00%	2.17%	-23.06%	13.57%

PANEL (B): VARIABLES IN PER WORKER TERMS

Variable	mean	median	std. dev.	min.	max.
Output	12,615.6	6,419.6	13,130.6	44.2	57,891.3
Capital	51,847.1	9,661.9	63,427.8	13.1	222,396.5
Land	9.57	2.94	20.25	0.29	110.00
in logarithms					
Output	8.385	8.767	1.817	3.788	10.966
Capital	8.950	9.176	2.694	2.573	12.312
Land	1.105	1.078	1.404	-1.244	4.701
in growth rates					
Output	2.33%	2.52%	10.49%	-43.67%	55.98%
Capital	2.47%	2.00%	4.17%	-7.83%	31.12%
Land	0.70%	0.50%	3.40%	-18.37%	28.77%

MANUFACTURING DATA

PANEL (A): VARIABLES IN UNTRANSFORMED LEVEL TERMS

Variable	mean	median	std. dev.	min.	max.
Output	7.47E+10	8.31E+09	2.07E+11	7.20E+06	1.43E+12
Labour	1.73E+06	4.75E+05	3.42E+06	9.56E+03	1.97E+07
Capital	1.33E+11	1.91E+10	2.97E+11	1.41E+07	1.81E+12
in logarithms					
Output	22.812	22.840	2.292	15.790	27.991
Labour	13.081	13.072	1.653	9.166	16.794
Capital	23.619	23.675	2.269	16.462	28.222
in growth rates					
Output	4.45%	3.83%	10.09%	-40.91%	84.23%
Labour	1.96%	1.13%	6.83%	-38.84%	78.12%
Capital	4.84%	3.62%	4.97%	-5.10%	53.03%

PANEL (B): VARIABLES IN PER WORKER TERMS

Variable	mean	median	std. dev.	min.	max.
Output	26,898.2	20,212.6	22,071.3	753.0	101,933.8
Capital	63,080.3	42,543.9	64,355.0	1,475.5	449,763.4
in logarithms					
Output	9.731	9.914	1.084	6.624	11.532
Capital	10.538	10.658	1.083	7.297	13.016
in growth rates					
Output	2.51%	2.48%	9.00%	-66.95%	73.01%
Capital	2.90%	2.91%	6.59%	-71.65%	42.44%

AGGREGATED DATA

PANEL (A): VARIABLES IN UNTRANSFORMED LEVEL TERMS

Variable	mean	median	std. dev.	min.	max.
Output	9.22E+10	1.69E+10	2.31E+11	1.14E+08	1.55E+12
Labour	1.12E+07	2.31E+06	3.55E+07	2.23E+04	2.40E+08
Capital	1.97E+11	2.79E+10	4.31E+11	1.02E+08	2.25E+12
in logarithms					
Output	23.470	23.553	2.016	18.552	28.069
Labour	14.640	14.653	1.736	10.011	19.297
Capital	24.078	24.052	2.213	18.438	28.442
in growth rates					
Output	3.17%	3.15%	7.37%	-33.87%	42.14%
Labour	0.19%	0.49%	2.56%	-11.39%	19.30%
Capital	3.57%	2.73%	3.62%	-5.00%	25.14%

PANEL (B): VARIABLES IN PER WORKER TERMS

Variable	mean	median	std. dev.	min.	max.
in levels					
Output	19,327.1	10,736.2	19,174.0	72.5	76,031.1
Capital	49,187.4	22,087.4	55,406.5	52.7	236,312.1
in logarithms					
Output	8.830	9.281	1.845	4.284	11.239
Capital	9.438	10.003	2.191	3.964	12.373
in growth rates					
Output	2.95%	3.30%	7.04%	-31.02%	44.49%
Capital	3.38%	3.14%	3.74%	-18.43%	22.16%

PENN WORLD TABLE DATA

PANEL (A): VARIABLES IN UNTRANSFORMED LEVEL TERMS

Variable	mean	median	std. dev.	min.	max.
Output	4.24E+11	1.27E+11	1.01E+12	1.34E+09	7.98E+12
Labour	5.05E+07	1.30E+07	1.19E+08	2.12E+05	8.54E+08
Capital	1.21E+12	3.25E+11	2.93E+12	3.30E+09	2.27E+13
in logarithms					
Output	25.423	25.564	1.716	21.018	29.708
Labour	16.469	16.380	1.627	12.266	20.565
Capital	26.359	26.506	1.801	21.918	30.753
in growth rates					
Output	4.00%	4.00%	4.96%	-37.12%	26.63%
Labour	1.56%	1.43%	1.14%	-1.87%	4.82%
Capital	4.60%	4.19%	2.84%	-1.30%	16.43%

PANEL (B): VARIABLES IN PER WORKER TERMS

Variable	mean	median	std. dev.	min.	max.
Output	11,396.7	10,308.1	8,162.3	594.3	31,074.1
Capital	36,832.4	32,026.3	31,668.2	660.8	136,891.2
in logarithms					
Output	8.945	9.241	1.016	6.387	10.344
Capital	9.868	10.374	1.365	6.493	11.827
in growth rates					
Output	2.44%	2.57%	4.96%	-41.22%	23.19%
Capital	3.04%	2.77%	2.87%	-4.23%	14.26%

Notes: We report the descriptive statistics for value-added (in US$1990 or PPP I$2000), labour (headcount), capital stock (same monetary values as VA in each respective dataset) and land (in hectare) for the full regression sample (n = 928; N = 41).

A-1.1 Sectoral and aggregated data

Investment data Data for agricultural and manufacturing investment (`AgSEInv`, `MfgSEInv`) in constant 1990 LCU, the US$-LCU exchange rate (`Ex_Rate`, see comment below) as well as sector-specific deflators (`AgDef`, `TotDef`) were taken from Crego et al. (1998).[17] Note that Crego et al. (1998) also provide capital stock data, which they produced through their own calculations from the investment data. Following Martin and Mitra (2002) we believe the use of a single year exchange rate is preferrable to the use of annual ones in the construction of real output (see next paragraph) and capital stock (see below).

Output data For manufacturing we use data on aggregate GDP in current LCU and the share of GDP in manufacturing from the World Bank World Development Indicators (WDI) (World Bank, 2008). For agriculture we use agricultural value-added in current LCU from the same source. We prefer the latter over the share of GDP in agriculture for data coverage reasons (in theory they should be the same, but they are not). The two sectoral value-added series are then deflated using the Crego et al. (1998) sectoral deflator for agriculture and the total economy deflator for manufacturing, before we use the 1990 US$-LCU exchange rates to make them comparable across countries.

Note that the currencies used in the Crego et al. (1998) data differ from those applied in the WDI data for a number of European countries due to the adoption of the Euro: for the latter we therefore need to use an alternative 1990 US$-LCU exchange rate for these economies.[18]

Labour data For agriculture we adopt the variable 'economically active population in agriculture' from the FAO's PopSTAT (FAO, 2007). Manufacturing labour is taken from UNIDO's INDSTAT UNIDO (2004).

Additional data The land variable is taken from ResourceSTAT and represents arable and permanent crop land (originally in 1000 hectare) (FAO, 2007). 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:

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

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.

Capital stock We construct capital stock in agriculture and manufacturing by applying the perpetual inventory method described in detail in Klenow and Rodriguez-Clare (1997b) using the investment data from Crego et al. (1998), which is transformed into US$ by application of the 1990 US$-LCU exchange rate. For the construction of sectoral base year capital stock we employ average sector value-added growth rates g_j (using the deflated sectoral value-added data described above), the average sectoral investment to value-added ratio $(I/Y)_j$ and an assumed depreciation rate of 5% to

[17] Data is available in excel format on the World Bank website at `http://go.worldbank.org/FS3FXW7461`. All data discussed in this appendix are linked at `http://sites.google.com/site/medevecon/devecondata`.

[18] In detail, we apply exchange rates of 1.210246384 for AUT, 1.207133927 for BEL, 1.55504706 for FIN, 1.204635181 for FRA, 2.149653527 for GRC, 1.302645017 for IRL, 1.616114954 for ITA, 1.210203555 for NLD and 1.406350856 for PRT. See Table A-I for country codes.

construct

$$\left(\frac{K}{Y}\right)_{0j} = \frac{IY_j}{g_j + 0.05}$$

for sector j. This ratio is then multiplied by sectoral value-added in the base year to yield K_{0j}. Note that the method deviates from that discussed in Klenow and Rodriguez-Clare (1997b) as they use *per capita* GDP in their computations and therefore need to account for population growth in the construction of the base year capital stock.

Aggregated data We combine the agriculture and manufacturing data to produce a stylised 'aggregate economy': for labour we simply add up the headcount, for the monetary representations of output and capital stock we can do so as well. We are afforded this ability to simply add up variables for the two sectors by the efforts Crego et al. (1998), who have built the first large panel dataset providing data on investment in agriculture for a long timespan.

A-1.2 Penn World Table data

As a means of comparison we also provide production function estimates using data from PWT version 6.2. We adopt Real per capita GDP in International $ Laspeyeres (`rgdpl`) as measure for output and construct capital stock using investment data (derived from Investment Share in Real GDP, `ki`, and the output variable, `rgdpl`) in the perpetual inventory method described above, adopting again 5% depreciation (this time we need to use the data on population from PWT, `pop`, to compute the average annual population growth rate).

TECHNICAL APPENDIX

TA-1 Time-series properties of the data

Table TA-I: Second generation panel unit root tests

PANEL (A): AGRICULTURE DATA

Variables in levels

	log VA pw		**log Labour**		**log Cap pw**	
lags	Ztbar	(p)	Ztbar	(p)	Ztbar	(p)
0	-0.662	(.25)	7.869	(1.00)	7.182	(1.00)
1	-0.326	(.37)	5.392	(1.00)	3.871	(1.00)
2	2.911	(1.00)	7.550	(1.00)	5.490	(1.00)
3	4.817	(1.00)	9.859	(1.00)	5.417	(1.00)
4	7.301	(1.00)	9.686	(1.00)	6.865	(1.00)

	Land pw	
lags	Ztbar	(p)
0	9.432	(1.00)
1	7.223	(1.00)
2	6.069	(1.00)
3	3.266	(1.00)
4	5.339	(1.00)

Variables in growth rates

	VA pw		**Labour**		**Cap pw**	
lags	Ztbar	(p)	Ztbar	(p)	Ztbar	(p)
0	-16.230	(.00)	-2.829	(.00)	-1.550	(.06)
1	-9.960	(.00)	3.394	(1.00)	-0.359	(.36)
2	-4.970	(.00)	5.639	(1.00)	4.161	(1.00)
3	-1.474	(.07)	6.238	(1.00)	5.171	(1.00)
4	3.869	(1.00)	9.043	(1.00)	9.442	(1.00)

	Land pw	
lags	Ztbar	(p)
0	-9.704	(.00)
1	-3.433	(.00)
2	1.324	(.91)
3	3.132	(1.00)
4	6.584	(1.00)

PANEL (B): MANUFACTURING DATA

Variables in levels

	log VA pw		**log Labour**		**log Cap pw**	
lags	Ztbar	(p)	Ztbar	(p)	Ztbar	(p)
0	0.903	(.82)	2.539	(.99)	1.668	(.95)
1	2.631	(1.00)	1.971	(.98)	0.667	(.75)
2	2.513	(.99)	4.240	(1.00)	2.060	(.98)
3	4.022	(1.00)	4.066	(1.00)	3.240	(1.00)
4	9.332	(1.00)	7.207	(1.00)	6.194	(1.00)

Variables in growth rates

	VA pw		**Labour**		**Cap pw**	
lags	Ztbar	(p)	Ztbar	(p)	Ztbar	(p)
0	-18.029	(.00)	-11.824	(.00)	-9.259	(.00)
1	-8.603	(.00)	-6.586	(.00)	-4.928	(.00)
2	-3.585	(.00)	-3.700	(.00)	-2.263	(.01)
3	-1.059	(.14)	-0.176	(.43)	0.847	(.80)
4	2.134	(.98)	4.982	(1.00)	4.511	(1.00)

PANEL (C): AGGREGATED DATA

Variables in levels

	log VA pw		**log Labour**		**log Cap pw**	
lags	Ztbar	(p)	Ztbar	(p)	Ztbar	(p)
0	2.558	(.99)	6.950	(1.00)	5.920	(1.00)
1	3.112	(1.00)	4.292	(1.00)	3.668	(1.00)
2	5.190	(1.00)	4.906	(1.00)	4.177	(1.00)
3	5.361	(1.00)	5.131	(1.00)	4.307	(1.00)
4	7.108	(1.00)	8.155	(1.00)	8.252	(1.00)

Variables in growth rates

	VA pw		**Labour**		**Cap pw**	
lags	Ztbar	(p)	Ztbar	(p)	Ztbar	(p)
0	-15.283	(.00)	-5.625	(.00)	-4.489	(.00)
1	-8.185	(.00)	-2.324	(.01)	-1.073	(.14)
2	-3.429	(.00)	0.035	(.51)	1.154	(.88)
3	-0.640	(.26)	2.637	(1.00)	3.472	(1.00)
4	2.569	(.99)	5.652	(1.00)	6.452	(1.00)

PANEL (D): PENN WORLD TABLE DATA

Variables in levels

	log VA pw		**log Labour**		**log Cap pw**	
lags	Ztbar	(p)	Ztbar	(p)	Ztbar	(p)
0	4.544	(1.00)	-1.069	(.14)	2.802	(1.00)
1	6.126	(1.00)	7.647	(1.00)	6.097	(1.00)
2	6.581	(1.00)	7.215	(1.00)	7.215	(1.00)
3	7.772	(1.00)	6.475	(1.00)	7.576	(1.00)
4	7.578	(1.00)	7.484	(1.00)	8.950	(1.00)

Variables in growth rates

	VA pw		**Labour**		**Cap pw**	
lags	Ztbar	(p)	Ztbar	(p)	Ztbar	(p)
0	-14.287	(.00)	0.711	(.76)	-4.690	(.00)
1	-6.603	(.00)	-1.977	(.02)	-2.437	(.01)
2	-4.112	(.00)	1.784	(.96)	-1.801	(.04)
3	-1.050	(.15)	2.205	(.99)	-0.468	(.32)
4	4.229	(1.00)	3.884	(1.00)	3.656	(1.00)

Notes: We report test statistics and p-values for the Pesaran (2007) CIPS panel unit root test of the variables in our four datasets. In all cases we use $N = 41$, $n = 928$ for the levels data.

TA-2 Cross-section dependence in the data

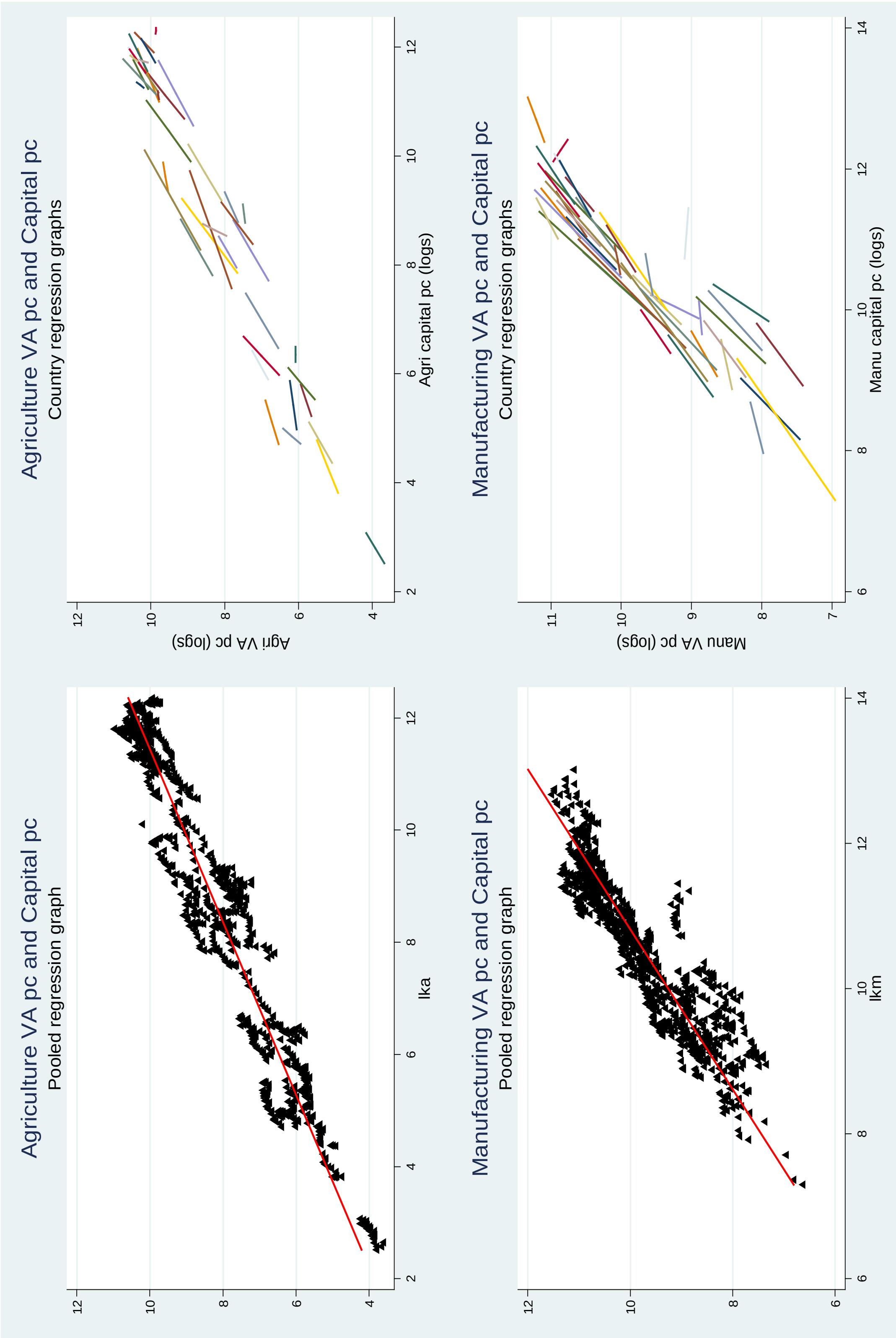
Table TA-II: Cross-section correlation analysis

	Variables in levels				*Variables in first diff.*			
AGRICULTURE DATA	$\bar{\rho}$	$\lvert\bar{\rho}\rvert$	CD	CDZ	$\bar{\rho}$	$\lvert\bar{\rho}\rvert$	CD	CDZ
log VA pw	0.41	0.57	57.65	74.45	0.05	0.23	6.57	6.59
(p)			(.00)	(.00)			(.00)	(.00)
log Labour	-0.01	0.76	-1.10	0.45	0.12	0.52	14.50	22.60
(p)			(.27)	(.65)			(.00)	(.00)
log Cap pw	0.41	0.72	56.06	97.01	0.08	0.40	9.09	11.26
(p)			(.00)	(.00)			(.00)	(.00)
log Land pw	0.02	0.72	2.90	3.49	0.04	0.28	4.96	5.67
(p)			(.00)	(.00)			(.00)	(.00)
MANUFACTURING DATA	$\bar{\rho}$	$\lvert\bar{\rho}\rvert$	CD	CDZ	$\bar{\rho}$	$\lvert\bar{\rho}\rvert$	CD	CDZ
log VA pw	0.43	0.63	66.34	84.24	0.05	0.21	6.27	6.49
(p)			(.00)	(.00)			(.00)	(.00)
log Labour	0.26	0.60	38.19	54.53	0.14	0.25	17.82	18.98
(p)			(.00)	(.00)			(.00)	(.00)
log Cap pw	0.61	0.77	86.11	136.03	0.07	0.22	8.22	9.04
(p)			(.00)	(.00)			(.00)	(.00)
AGGREGATED DATA	$\bar{\rho}$	$\lvert\bar{\rho}\rvert$	CD	CDZ	$\bar{\rho}$	$\lvert\bar{\rho}\rvert$	CD	CDZ
log VA pw	0.61	0.69	83.57	118.17	0.08	0.23	10.65	11.23
(p)			(.00)	(.00)			(.00)	(.00)
log Labour	0.01	0.72	1.36	6.42	0.06	0.31	8.24	9.47
(p)			(.18)	(.00)			(.00)	(.00)
log Cap pw	0.76	0.85	97.16	188.46	0.07	0.29	7.99	9.81
(p)			(.00)	(.00)			(.00)	(.00)
PENN WORLD TABLE DATA	$\bar{\rho}$	$\lvert\bar{\rho}\rvert$	CD	CDZ	$\bar{\rho}$	$\lvert\bar{\rho}\rvert$	CD	CDZ
log VA pw	0.72	0.74	111.55	170.81	0.14	0.20	21.89	19.07
(p)			(.00)	(.00)			(.00)	(.00)
log Labour	0.95	0.95	149.58	298.19	0.11	0.38	16.80	17.57
(p)			(.00)	(.00)			(.00)	(.00)
log Cap pw	0.76	0.86	116.84	219.82	0.26	0.38	39.69	38.66
(p)			(.00)	(.00)			(.00)	(.00)

Notes: In all cases we use $N = 41$, $n = 928$ for the levels data. We report the average correlation coefficient across the $N(N-1)$ variable series $\bar{\rho}$, as well as the average absoulte correlation coefficient $|\bar{\rho}|$. CD and CDZ are formal cross-section correlation tests introduced by Pesaran (2004) and Moscone and Tosetti (2009). Under the H_0 of cross-section independence both statistics are asymptotically standard normal. We investigated two further tests introduced by Moscone and Tosetti (2009), namely CD_{LM} and CD_{ABS}, which yield the same conclusions as the tests presented (detailed results available on request).

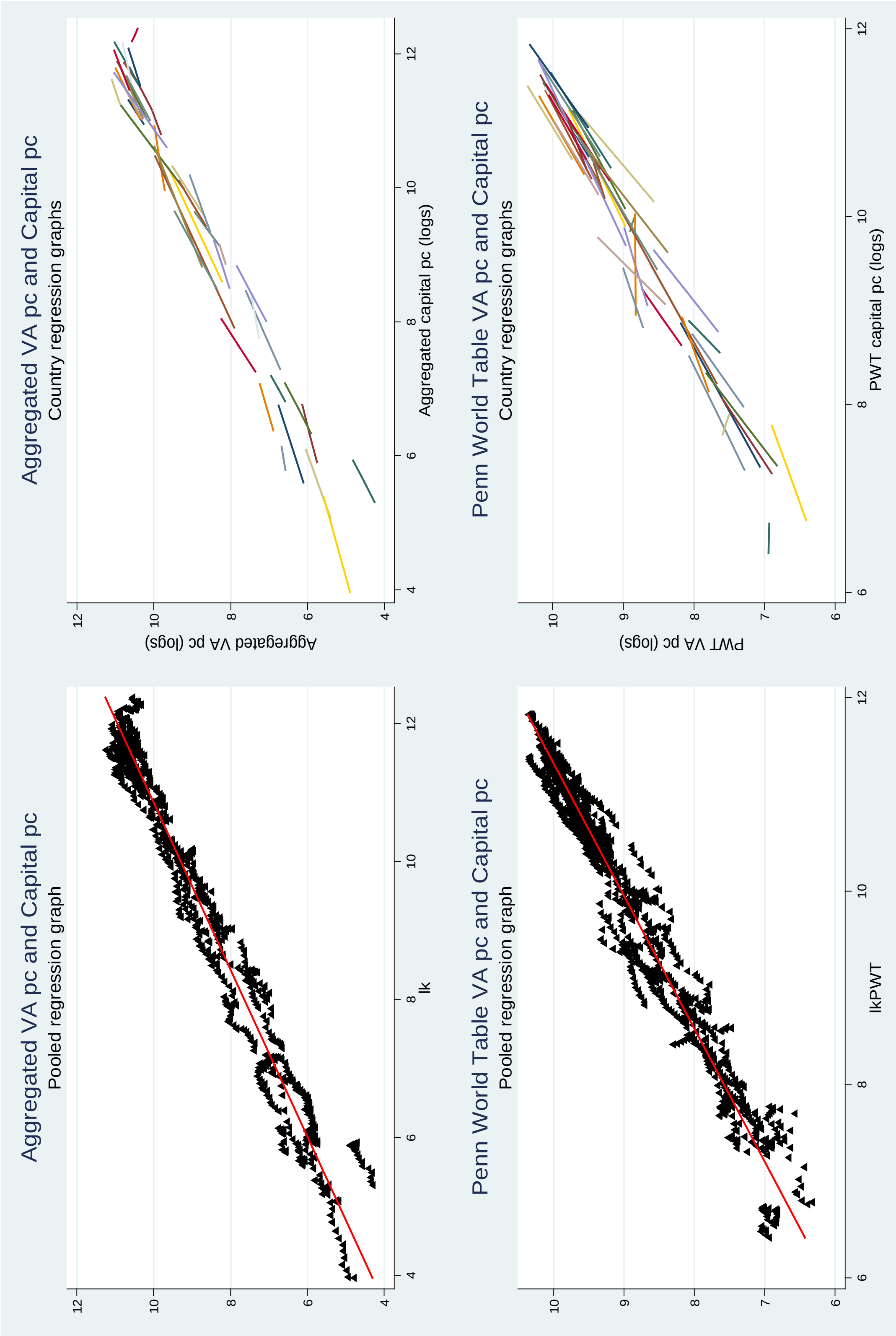
TA-3 Additional tables and figures

Figure TA-I: Scatter plots — Agriculture and Manufacturing data



Notes: Scatter graphs for agricultural (manufacturing) VA per worker plotted against capital per worker. The red line represents a least squares regression line, mimicking a pooled OLS regression model (without TFP growth). The multi-coloured lines represent N regression lines, mimicking a heterogeneous parameter model (without TFP growth).

Figure TA-II: Scatter plots — Aggregated/PWT data



Notes: Scatter graphs for aggregated (Penn World Table) VA per worker plotted against capital per worker. The red line represents a least squares regression line, mimicking a pooled OLS regression model (without TFP growth). The multi-coloured lines represent N regression lines, mimicking a heterogeneous parameter model (without TFP growth).

Table TA-III: Alternative dynamic panel estimators

Panel (A): Agriculture

	Dynamic FE			PMG				CPMG⋆			DGMM	SGMM
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]
EC [y_{t-1}]	-0.293	-0.312	-0.300	-0.460	-0.459	-0.624	-0.466	-0.482	-0.503	-0.455	-1.087	-0.432
	[11.80]**	[12.43]**	[11.91]**	[10.63]**	[9.34]**	[14.29]**	[10.44]**	[10.06]**	[9.74]**	[9.34]**	[2.60]**	[5.38]**
capital pw	0.672	0.684	0.582	0.652	0.714	0.036	0.132	0.501	0.464	0.530	1.135	0.776
	[12.47]**	[12.69]**	[7.50]**	[20.16]**	[18.52]**	[0.57]	[3.01]**	[10.78]**	[11.05]**	[10.83]**	[2.85]**	[12.59]**
land pw	0.124	0.121	0.135	0.136	0.367	0.867	0.361	0.247	0.494	0.228	0.083	-0.247
	[1.30]	[1.29]	[1.45]	[2.90]**	[6.43]**	[8.27]**	[8.05]**	[5.03]**	[8.95]**	[4.73]**	[0.35]	[1.17]
trend(s)†			0.001			0.008	0.012					
			[1.59]			[3.36]**	[12.26]**					
Constant	0.667	0.679	0.896	1.072	0.644	4.273	3.084	1.545	1.402	1.298		0.714
	[5.03]**	[4.75]**	[4.58]**	[10.48]**	[7.53]**	[13.11]**	[10.27]**	[10.38]**	[9.69]**	[9.94]**		[4.21]**
lags [trends]‡	1	2	1 [l-r]	1	2	1 [s-r]	1 [l-r]	1	2	1	i: 2-3	i: 2-3
impl. labour	0.328	0.316	0.418	0.212	-0.081	0.098	0.507	0.253	0.042	0.242	-0.135	0.224
obs	894	857	894	894	857	894	894	894	857	872	857	894

Panel (B): Manufacturing

	Dynamic FE			PMG				CPMG⋆			DGMM	SGMM
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]
EC [y_{t-1}]	-0.196	-0.195	-0.195	-0.219	-0.181	-0.543	-0.214	-0.245	-0.194	-0.272	-2.1959	-0.0414
	[9.40]**	[9.16]**	[9.31]**	[6.59]**	[5.97]**	[4.04]**	[4.13]**	[7.16]**	[6.45]**	[7.33]**	[0.72]	[0.65]
capital pw	0.711	0.708	0.637	1.016	1.044	0.298	1.379	0.598	1.264	0.505	1.866	-1.515
	[12.96]**	[12.34]**	[6.85]**	[29.64]**	[33.09]**	[5.34]**	[26.80]**	[11.58]**	[22.28]**	[9.47]**	[3.25]**	[0.40]
trend(s)†			0.001			0.001	-0.010					
			[1.00]			[0.24]	[6.77]**					
Constant	0.452	0.456	0.588	-0.212	-0.228	3.493	-0.977	0.225	-0.434	0.372		1.042
	[3.87]**	[3.73]**	[3.29]**	[5.43]**	[4.95]**	[3.87]**	[4.18]**	[5.68]**	[5.77]**	[6.48]**		[1.80]
lags [trends]‡	1	2	1 [l-r]	1	2	1 [s-r]	1 [l-r]	1	2	1	i: 2-3	i: 2-3
impl. labour	0.289	0.292	0.363	-0.016	-0.044	0.702	-0.379	0.402	-0.264	0.495	-0.866	2.515
obs	902	880	902	902	880	902	902	902	880	879	880	902

Panel (C): Aggregated data

	Dynamic FE			PMG				CPMG⋆			DGMM	SGMM
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]
EC [y_{t-1}]	-0.172	-0.176	-0.173	-0.279	-0.277	-0.429	-0.284	-0.292	-0.294	-0.317	-0.3803	-0.2426
	[8.59]**	[8.39]**	[8.59]**	[6.89]**	[7.25]**	[9.55]**	[6.72]**	[6.98]**	[7.38]**	[7.48]**	[0.71]	[4.21]**
capital pw	0.705	0.709	0.668	0.974	1.015	0.128	0.899	0.891	0.949	0.905	0.271	0.896
	[15.25]**	[14.65]**	[8.17]**	[36.86]**	[37.38]**	[1.90]	[21.11]**	[24.84]**	[24.92]**	[27.54]**	[0.27]	[22.80]**
trend(s)†			0.000			0.011	0.004					
			[0.54]			[6.07]**	[2.42]*					
Constant	0.390	0.393	0.446	-0.100	-0.200	3.061	0.082	-0.062	-0.169	-0.145		0.120
	[4.96]**	[4.62]**	[3.42]**	[3.73]**	[5.18]**	[9.30]**	[4.20]**	[2.53]*	[4.97]**	[4.58]**		[1.44]
lags [trends]‡	1	2	1 [l-r]	1	2	1 [s-r]	1 [l-r]	1	2	1	i: 2-3	i: 2-3
impl. labour	0.295	0.292	0.332	0.026	-0.015	0.872	0.102	0.109	0.051	0.095	0.729	0.104
obs	879	836	879	879	836	879	879	879	836	879	836	879

Panel (D): Penn World Table data

	Dynamic FE			PMG				CPMG⋆			DGMM	SGMM
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]
EC [y_{t-1}]	-0.098	-0.101	-0.107	-0.333	-0.138	-0.567	-0.392	-0.338	-0.081	-0.347	0.8351	0.0309
	[5.82]**	[6.01]**	[6.22]**	[6.70]**	[4.37]**	[12.63]**	[7.88]**	[6.63]**	[2.56]*	[8.24]**	[1.07]	[0.49]
capital pw	0.538	0.553	0.356	0.923	0.916	0.698	0.652	0.903	-0.125	0.731	0.604	0.863
	[8.14]**	[8.66]**	[3.44]**	[130.34]**	[71.72]**	[65.10]**	[67.96]**	[52.90]**	[1.81]	[86.83]**	[0.60]	[1.88]
trend(s)†			0.001			0.002	0.006					
			[2.44]*			[2.57]*	[19.84]**					
Constant	0.363	0.360	0.567	-0.122	-0.020	1.085	0.935	-0.071	0.456	0.504		0.010
	[5.38]**	[5.29]**	[5.28]**	[4.44]**	[1.63]	[13.05]**	[7.79]**	[3.47]**	[2.99]**	[8.29]**		[0.07]
lags [trends]‡	1	2	1 [l-r]	1	2	1 [s-r]	1 [l-r]	1	2	1	i: 2-3	i: 2-3
impl. labour	0.462	0.447	0.645	0.077	0.084	0.302	0.349	0.097	1.125	0.270	0.396	0.137
obs	914	904	914	914	904	914		914	904	873	904	914

Notes: We report the long-run coefficients on capital per worker (and in the agriculture equations also land per worker). EC [y_{t-1}] refers to the Error-Correction term (speed of adjustment parameter) with the exception of Models [11] and [12], where we report the coefficient on y_{t-1} — conceptually, these are the same, however in the latter we do not impose common factor restrictions like in all of the former models. Note that in the PMG and CPMG models the ECM term is heterogeneous across countries, while in the Dynamic FE and GMM models these are common across i. † In model [6] we include *heterogeneous* trend terms, whereas in [7] a *common* trend is assumed (i.e. linear TFP is part of cointegrating vector). ‡ 'lags' indicates the lag-length of first differenced RHS variables included, with the exception of Models [11] and [12]: here 'i:' refers to the lags (levels in [11], levels and differences in [12] used as instruments. ⋆ In the models in [8] and [9] the cross-section averages are only included for the long-run variables, whereas in the model in [10] cross-section averages for the first-differenced dependent and independent variables (short-run) are also included. Note that in the agriculture equation for Model [10] we drop CRI ($n = 7$) as otherwise no convergence would occur.

Table TA-IV: Aggregate & PWT data: Pooled models (HC-augmented)

Panel (A): Unrestricted returns

	Aggregated data				*Penn World Table data*			
	[1] **POLS**	[2] **2FE**	[3] **CCEP**	[4] **FD2FE**	[5] **POLS**	[6] **2FE**	[7] **CCEP**	[8] **FD2FE**
log labour	-0.001	-0.058	0.566	0.083	0.040	-0.064	-0.193	-0.032
	[0.14]	[1.97]*	[4.13]**	[2.50]*	[8.99]**	[3.27]**	[1.49]	[1.11]
log capital pw	0.662	0.782	0.677	0.766	0.725	0.680	0.601	0.676
	[97.95]**	[31.50]**	[7.25]**	[25.24]**	[72.79]**	[24.79]**	[9.12]**	[18.96]**
Education	0.243	-0.004	0.086	0.065	0.041	0.043	0.032	0.103
	[16.97]**	[0.15]	[1.24]	[1.22]	[3.42]**	[2.86]**	[0.80]	[3.41]**
Education squared	-0.010	0.003	-0.007	-0.003	-0.001	-0.002	-0.002	-0.006
	[8.05]**	[1.82]	[1.57]	[0.77]	[1.77]	[2.97]**	[0.83]	[2.94]**
Implied RS†	CRS	DRS	CRS	CRS	CRS	DRS	CRS	CRS
Implied β_L‡	0.337	0.160	0.890	0.318	0.315	0.256	0.206	0.292
Mean Education	5.824	5.824	5.824	5.885	5.822	5.822	5.822	5.883
Returns to Edu	12.9%	2.5%	1.0%	3.4%	2.4%	1.9%	0.9%	3.3%
[t-statistic]♭	[22.35]	[1.68]	[0.37]	[1.40]	[6.82]	[2.02]	[0.56]	[2.26]
$\hat{e}$ integrated♮	I(1)	I(1)	I(0)	I(0)	I(1)	I(1)	I(0)	I(1)/I(0)
CD test p-value♯	0.00	0.02	0.59	0.00	0.34	0.22	0.01	0.00
R-squared	0.98	0.87	1.00	-	0.97	0.78	1.00	-
Observations	775	775	775	732	769	769	769	726

Panel (B): Constant returns to scale imposed

	Aggregated data				*Penn World Table data*			
	[1] **POLS**	[2] **2FE**	[3] **CCEP**	[4] **FD2FE**	[5] **POLS**	[6] **2FE**	[7] **CCEP**	[8] **FD2FE**
log capital pw	0.662	0.798	0.485	0.744	0.694	0.706	0.611	0.691
	[102.10]**	[35.45]**	[7.03]**	[25.48]**	[73.08]**	[27.73]**	[10.05]**	[21.13]**
Education	0.243	-0.016	0.210	0.111	0.043	0.037	0.016	0.092
	[16.98]**	[0.62]	[3.00]**	[2.21]*	[3.30]**	[2.44]*	[0.48]	[3.22]**
Education squared	-0.010	0.004	-0.013	-0.005	-0.001	-0.002	-0.002	-0.006
	[8.17]**	[2.75]**	[2.92]**	[1.37]	[0.97]	[2.12]*	[0.95]	[2.79]**
Constant	1.586				1.843			
	[21.62]**				[20.44]**			
Implied β_L‡	0.338	0.203	0.515	0.256	0.306	0.294	0.390	0.309
Mean Education	5.824	5.824	5.824	5.885	5.822	5.824	5.824	5.883
Returns to Edu	12.9%	2.6%	6.5%	5.8%	3.3%	2.0%	-0.6%	2.7%
[t-statistic]♭	[22.41]	[1.68]	[2.56]	[2.56]	[8.62]	[1.99]	[0.42]	[1.98]
$\hat{e}$ integrated♮	I(1)	I(1)	I(0)	I(0)	I(1)	I(1)	I(0)	I(0)
CD test p-value♯	0.00	0.00	0.65	0.00	0.25	0.57	0.02	0.00
R-squared	0.98	0.86	1.00		0.97	0.78	1.00	
Observations	775	775	775	732	769	769	769	726

Notes: We include our proxy for education in levels and as a squared term. Returns to Education are computed from the sample mean ($\bar{E}$) as $\beta_E + 2\beta_{E2}\bar{E}$ where β_E and β_{E2} are the coefficients on the levels and squared education terms respectively. ♭ computed via the delta-method. For more details on other diagnostics see Notes in Table III.

Table TA-V: Aggregate & PWT data: Heterogeneous models with HC

Panel (A): Unrestricted returns to scale

	Aggregated data			*Penn World Table data*		
	[1] **MG**	[2] **FDMG**	[3] **CMG**	[4] **MG**	[5] **FDMG**	[6] **CMG**
log labour	-0.066	0.269	-0.428	-1.609	-2.478	-1.324
	[0.16]	[0.57]	[1.22]	[1.97]	[3.76]**	[2.79]**
log capital pw	-0.070	-0.021	0.453	0.963	1.245	1.122
	[0.26]	[0.07]	[2.47]*	[4.44]**	[5.99]**	[5.52]**
Education	0.601	0.637	0.489	0.123	0.004	-0.012
	[1.29]	[1.75]	[0.98]	[0.52]	[0.02]	[0.05]
Education squared	-0.089	-0.065	-0.063	-0.002	0.004	-0.001
	[1.76]	[1.70]	[1.48]	[0.11]	[0.25]	[0.03]
country trend/drift	0.005	0.005		0.021	0.008	
	[0.33]	[0.29]		[2.25]*	[0.77]	
Mean Education	5.72	5.84	5.72	5.72	5.84	5.72
Returns to edu	-7.1%	-3.2%	-11.1%	-4.5%	0.5%	1.3%
[t-statistic]$^\flat$	[1.33]	[0.65]	[1.24]	[1.33]	[0.18]	[0.43]
Implied RS†	CRS	CRS	CRS	CRS	DRS	DRS
Implied β_L‡	n/a	n/a	0.547	n/a	n/a	n/a
reject CRS (10%)	38%	3%	19%	38%	18%	33%
panel-t Labour	-1.77	0.16	-1.42	-6.37**	-5.60**	-7.30**
panel-t Capital	0.58	0.94	2.79**	15.62**	13.48**	14.39**
panel-t Edu	0.26	1.21	0.86	0.89	0.23	0.68
panel-t Edu ˆ2	-1.07	-1.87	-1.26	-1.55	-0.35	-0.72
panel-t trends	14.73**	10.93**		11.09**	5.83**	
# sign. trends	18	13		18	4	
$\hat{e}$ integrated$^\natural$	I(0)	I(0)	I(0)	I(0)	I(0)	I(0)
abs correl.coeff.	0.24	0.24	0.22	0.23	0.24	0.22
CD-test (p)$^\sharp$	7.23(.00)	7.88(.00)	-0.50(.61)	7.59.00)	9.29.00)	0.98(.33)

Panel (B): CRS imposed

	Aggregated data			*Penn World Table data*		
	[1] **MG**	[2] **FDMG**	[3] **CMG**	[4] **MG**	[5] **FDMG**	[6] **CMG**
log capital pw	0.093	0.151	0.528	0.779	1.052	0.906
	[0.49]	[0.90]	[4.90]**	[5.75]**	[6.43]**	[5.86]**
Education	0.075	0.260	0.683	-0.215	-0.134	0.089
	[0.18]	[0.99]	[1.73]	[1.25]	[0.84]	[0.42]
Education squared	-0.023	-0.023	-0.075	0.013	0.014	-0.023
	[0.65]	[0.89]	[1.57]	[0.82]	[1.13]	[1.16]
country trend/drift	0.017	0.015		-0.001	-0.010	
	[1.96]	[1.33]		[0.21]	[2.08]*	
Implied β_L‡	n/a	n/a	0.472	0.221	n/a	0.094
Mean Education	5.79	5.84	5.79	5.79	5.84	5.79
Returns to edu	-9.3%	-4.0%	3.2%	-1.4%	0.3%	-0.2%
[t-statistic]$^\flat$	-1.34	-0.88	0.50	0.50	0.16	0.05
panel-t Capital	2.96**	1.84	7.63**	16.24**	11.99**	15.70**
panel-t Edu	-2.05*	1.97*	3.78**	-1.80	-1.23	0.74
panel-t Edu ˆ2	0.79	-2.77**	-3.83**	1.20	0.96	-1.11
panel-t trends	15.65**	12.21**		11.57**	7.84**	
# sign. trends	15	13		15	14	
$\hat{e}$ integrated$^\natural$	I(0)	I(0)	I(0)	I(0)	I(0)	I(0)
abs correl.coeff.	0.24	0.24	0.23	0.26	0.24	0.22
CD-test (p)$^\sharp$	8.05(.00)	8.59(.00)	0.11(.92)	9.75(.00)	10.84(.00)	3.12(.00)

Notes: All averaged coefficients presented are robust means across i. $\flat$ The returns to education and associated t-statistics are based on a two-step procedure: first the country-specific mean education value ($\bar{E}_i$) is used to compute $\beta_{i,E} + 2\beta_{i,E2}\bar{E}_i$ to yield the country-specific returns to education. The reported value then represents the robust mean of these N country estimates, s.t. the t-statistic should be interpreted in the same fashion as that for the regressors, namely as a test whether the average parameter is statistically different from zero, following Pesaran et al. (2009). For other details see Tables IV and V.