

THE MAGNITUDE OF THE TASK AHEAD: MACRO IMPLICATIONS OF HETEROGENEOUS TECHNOLOGY

BY MARKUS EBERHARDT*

School of Economics, University of Nottingham and Centre for Economic Policy Research
AND

FRANCIS TEAL

*Centre for the Study of African Economies, Department of Economics, University of Oxford and
Institute of Labor Economics (IZA)*

The empirical growth literature is dominated by accounting and regression methods that assume common production technology across countries. Our empirical model relaxes this assumption and further allows unobservable determinants of output (total factor productivity, TFP) to differ across countries and time, while accounting for endogeneity and cross-section correlation arising from global shocks. Using manufacturing sector data for 48 economies, we show that the assumption of common technology is rejected in our regressions. We illustrate that the erroneous choice of homogeneous technology has a substantial impact on the patterns and magnitudes of the resulting TFP estimates.

JEL Codes: O14, O47, C23

Keywords: cross-country analysis, heterogeneous technology, total factor productivity, common-factor model

1. INTRODUCTION

“We compare this [input] index with our output index and call any discrepancy ‘productivity’... It is a measure of our ignorance, of the unknown, and of the magnitude of the task that is still ahead of us.”
Griliches (1961, 446)

“As a careful reading of Solow (1956, 1970) makes clear, the stylized facts for which this model was developed were not interpreted as universal properties for every country in the world. In contrast,

Note: We are grateful to Michael Binder, Steve Bond, Francesco Caselli, Steven Durlauf, David Hendry, John Muellbauer, Hashem Pesaran, Ron Smith, Jon Temple, and participants at conferences and workshops in Oxford, Stockholm, Brighton, Maastricht, and Frankfurt for helpful comments and suggestions. Comments from three anonymous referees and the editor have helped improve the paper substantially. All remaining errors are our own. Eberhardt acknowledges financial support from the U.K. Economic and Social Research Council [PTA-031-2004-00345, PTA-026-27-2048, ES/K008919/1].

*Correspondence to: Markus Eberhardt, School of Economics, University of Nottingham, Sir Clive Granger Building, University Park, Nottingham NG2 2RD, U.K.
(markus.eberhardt@nottingham.ac.uk).

the current literature imposes very strong homogeneity assumptions on the cross-country growth process as each country is assumed to have an identical... production function.” Durlauf *et al.* (2001, 929)

It is an unfortunate misconception that the canonical neoclassical growth model simultaneously developed by Solow (1956) and Swan (1956) necessarily implies that *all* economies in the world, rich or poor, industrialized or agrarian, possess the same production technology. As the above quotes show, there are prominent critics of this assumption, while Solow himself suggested that “whether simple parameterizations do justice to real differences in the way the economic mechanism functions in one place or another” was certainly worth “grumbling” about (Solow, 1986, p. S23). Nevertheless, the notion that cross-country empirical analysis should, in the case of accounting exercises, adopt or, in the case of regression analysis, aim to arrive at a *common* capital coefficient of around 0.3 is deeply ingrained in the minds of growth economists.

Any doubters of this common-technology view (cf. common long-run equilibrium, common convergence process, and common dynamics) are typically referred to a study by Gollin (2002), which provides strong evidence that the observed labor share of aggregate output of around 0.7 varies only little across a diverse set of countries once mismeasurement of labor income in less developed economies is accounted for. Note that Gollin (2002) does not conclude that these income shares are identical across countries, but that his data corrections result in considerable reduction in their variation and that there is no correlation between income and the remaining differences. Nevertheless, Gollin’s findings are typically taken to mean that under the reasonable assumption of constant returns to scale and the perhaps somewhat less reasonable assumption of perfect competition, cross-country growth and levels accounting exercises can assume a common capital coefficient of 0.3 and focus their energies on chipping away at other dimensions of the “measure of our ignorance” (see Caselli, 2005; Hulten, 2010).

In this paper, we revisit the issue of whether technology is common across countries.¹ Using annual data for the manufacturing sector in 48 developing and developed countries for 1970 to 2002 (UNIDO, 2004), we show in panel time-series regressions that technology differences are of crucial importance for understanding cross-country differences in labor productivity and their causes. Our preferred empirical models further emphasize the importance of time-series properties of output, inputs, and TFP (Bond *et al.*, 2010), as well as of accounting for unobserved heterogeneity, which manifests itself as cross-country correlations arising from global shocks and local spillover effects (Chudik *et al.*, 2011). Like the existing cross-country growth literature, our preferred empirical implementations address concerns over endogeneity and reverse causality. We find that once these empirical aspects are accounted for, we obtain average technology estimates (capital coefficients) that are close to 0.3 with favorable residual diagnostics, whereas if we adopt the common-technology assumption the estimates are substantially different from 0.3 and residual testing indicates serious misspecification. Our

¹We refer to “technology heterogeneity” to indicate differential production-function parameters on observable inputs across countries, with unobservables captured as TFP.

conclusion of technology heterogeneity is further supported by formal parameter homogeneity tests.

A second feature of our study is the focus on manufacturing instead of aggregate economy data. The central importance of this industrial sector for successful development has become a widely recognized “stylized fact” in development economics. Yet in contrast to the literature on cross-country growth regressions using aggregate economy (Durlauf *et al.*, 2005) or agriculture data (Mundlak *et al.*, 2012; Eberhardt and Teal, 2013a,b, and references therein), there is comparatively little empirical work dedicated to the analysis of the manufacturing sector in a large cross-section of countries—with the exception of studies on the dual economy model (e.g. Martin and Mitra, 2002; Eberhardt and Teal, 2013a,b), and recent work by Dani Rodrik (Rodrik, 2013; McMillan *et al.*, 2014), cross-country empirical analysis at the sectoral level is typically limited to the investigation of OECD economies (Bernard and Jones, 1996a,b; Eberhardt *et al.*, 2013). If manufacturing matters for development, it is self-evidently important to learn about the production process and its drivers in this industrial sector.

Our findings have two important implications for productivity analysis both at the sectoral and the aggregate economy level: first, like firms in different industries, different countries are characterized by different production technologies. Attempts at estimating cross-country production functions in pooled models, where by construction the same technology is imposed on all countries, are misspecified and yield biased estimates for the technology parameters and thus any TFP estimates derived from them. Second, merely allowing for technology heterogeneity is also insufficient to capture the complex production process at the country level: in a globalizing world, economies interact through trade, cultural, political, and other ties, and at the same time are affected differentially by global phenomena such as the 1970s oil crises or the emergence of China as a major economic player. This creates a web of interdependencies within and across economies, leading to the breakdown of crucial assumptions for standard panel estimators employed in existing cross-country studies. Our empirical strategy accommodates this interplay of endogeneity, heterogeneity, and commonality to provide evidence for the fundamental forces driving manufacturing development across the globe.

In the following two sections, we discuss the challenges and potential solutions to modeling production technology (and TFP) in panel data, focusing on the cross-country growth literature. There are, however, at least two separate literatures that we need to mention briefly to avoid a distorted view of existing empirical approaches to the same problem: first, the spatial econometric literature, with the most relevant contributions being the spatially augmented Solow model (Ertur and Koch, 2007) and a Schumpeterian perspective on growth and development (Ertur and Koch, 2011), allowing for a more flexible specification for TFP evolution; and, second, the empirical literature adopting stochastic frontier and data envelopment methods to analyze efficiency and productivity (going back to Schmidt and Sickles, 1984). This approach can also accommodate flexible modeling of TFP (in the form of $\alpha_i + \lambda'_i f_i$ as laid out in Section 3; see Lovell, 1996; Kneip *et al.*, 2012) and common technology in groups β_g rather than for the entire sample of countries (Battese *et al.*, 2004).² The incorporation of these empirical approaches in our

²We are grateful to a referee for providing us with a whirlwind tour of this literature.

analysis would go beyond the scope of our paper. Interested readers are referred to these papers for an entry point to the respective literatures.

The remainder of the paper is structured as follows. Section 2 motivates technology heterogeneity, non-stationarity, and cross-section dependence. Section 3 lays out the empirical framework and discusses econometric identification. Section 4 introduces our data. Regression results are presented in Section 5, while their implications for productivity analysis are discussed in Section 6. Section 7 concludes.

2. MODELING TECHNOLOGY IN PANEL DATA

In this section, we motivate the concerns with which we approach the estimation of cross-country production functions. We begin by motivating technology heterogeneity and then discuss salient time-series and cross-section properties of the data.

The “new growth” literature provides justification for heterogeneous-technology parameters across countries. This strand of the theoretical growth literature argues that production functions differ across countries and seeks to determine the sources of this heterogeneity (Durlauf *et al.*, 2001). This can intuitively be taken to mean that countries can choose an “appropriate” production technology from a menu of feasible options. Representative examples from this literature include the work by Azariadis and Drazen (1990), Durlauf (1993), and Banerjee and Newman (1993). A simpler justification for heterogeneous production functions is offered by Durlauf *et al.* (2001), who argue that the Solow model was not intended to be valid in a common specification for *all* countries, but may still be a good way to investigate *each* country, by allowing for parameter differences *across* countries. A more formal treatment of technology heterogeneity is provided in Mundlak *et al.* (2012) and linked to the empirical framework that we adopt here in Eberhardt and Teal (2013b).

In the long run, macro variables such as value added or capital stock often appear to represent “nonstationary” processes in at least some countries (Lee *et al.*, 1997; Pedroni, 2007). In empirical practice, many studies establish that real value series typically behave as $I(1)$ processes (Nelson and Plosser, 1982; Lee *et al.*, 1997). Pedroni suggested that variable (non-)stationarity should not be seen as a “global” property, valid for all times, but as a “feature which describes local behaviour of the series within sample” (Pedroni, 2007, p. 432).

In our general empirical model, we emphasize a view of TFP as a “measure of our ignorance” (Abramowitz, 1956), incorporating a wider set of factors that can shift the production possibility frontier (for instance, “resource endowments, climate, institutions, and so on”; Mankiw *et al.*, 1992, pp. 410–11). This is in contrast to the notion of TFP as a definitive efficiency index, as commonly adopted in the microeconomic literature of productivity analysis. Furthermore, it is important to allow for the possibility that TFP is *in part* common to all countries; for example, representing the global dissemination of non-rival scientific knowledge or global shocks, such as the 1970s oil crises. Alternatively, we can think of multiple economic, social, political, and cultural ties between countries from which commonality (cross-section correlation) may arise. The individual evolution paths of the unobservables making up TFP should not be restrained to follow simple

linear trends, but instead be allowed to evolve in a non-linear and even non-stationary fashion. For instance, a number of empirical papers report that their measures of TFP display non-stationarity, whether analyzed at the economy level (Bond *et al.*, 2010) or at the sectoral level (Bernard and Jones, 1996b). At the same time, a highly flexible approach to empirical modeling using annual data raises the question of how business cycles influence or distort the empirical estimates (Eberhardt and Teal, 2011). All of these concerns point to the adoption of a multi-factor TFP structure that allows for common as well as country-specific elements and is uniquely suited for the analysis of productivity (Bai, 2009).

Existing empirical work has primarily concerned itself with the (potential) endogeneity of regressors in the empirical framework (e.g. Caselli *et al.*, 1996; Bond *et al.*, 2001), an issue that is given considerably more attention in the literature than the data properties or the potential misspecification of the empirical regression model. While the empirical methods adopted here can address the simultaneity between TFP shocks and input accumulation, we resort to an alternative estimation approach following Pedroni (2000) to rule out the potential of reverse causality and assure ourselves that these regressions represent production-function models and not investment or labor-demand equations in disguise. Thus, in addition to incorporating much desirable technology heterogeneity, our empirical analysis also addresses the major concerns that have occupied the existing literature.

3. THE EMPIRICAL MODEL AND IDENTIFICATION

Our regression analysis adopts a common-factor representation for a standard log-linearized Cobb–Douglas production-function model. We discuss all its features in detail below. Formally, for time periods $t = 1, \dots, T$, countries $i = 1, \dots, N$, and inputs $m = 1, \dots, k$, let

$$(1) \quad y_{it} = \sum_{m=1}^k \beta_{mi} x_{mit} + u_{it}, \quad u_{it} = \alpha_i + \lambda'_i \mathbf{f}_t + \varepsilon_{it},$$

$$(2) \quad x_{mit} = \pi_{mi} + \delta'_{mi} \mathbf{g}_{mt} + \rho_{1mi} f_{1mt} + \dots + \rho_{nmi} f_{nmt} + v_{mit},$$

$$(3) \quad \mathbf{f}_t = \rho' \mathbf{f}_{t-1} + \epsilon_t, \quad \text{and} \quad \mathbf{g}_t = \kappa' \mathbf{g}_{t-1} + \epsilon_t,$$

where $\mathbf{f}_{\cdot mi}$ is a subset of \mathbf{f}_t , y_{it} represents value added and \mathbf{x}_{it} represents the observable inputs including labor and capital stock (all in logarithms). The technology parameters β_i can differ across countries but are assumed constant over time.³

³The latter assumption is clearly restrictive, but given the focus on cross-country technology heterogeneity against the background of data restrictions in the time-series dimension, we cannot relax this assumption for the *heterogeneous* regression models. For the *pooled* models, we ran separate regressions using pre- and post-1985 subsamples. The estimates for pooled ordinary least squares (POLS), common correlated effects pooled (CCEP), and first-difference ordinary least squares (FD-OLS) are virtually identical for the two subperiods. Period estimates for the fixed-effects (FE) estimator differ somewhat but 95 percent confidence bounds still show *considerable* overlap.

For unobserved TFP, we employ a country-specific TFP level α_i in combination with a set of common factors \mathbf{f}_t , with country-specific factor loadings λ_t . In equation (2), we provide an empirical representation of the observable inputs (here, capital and labor), which are modeled as linear functions of the unobserved common factors \mathbf{f}_t and \mathbf{g}_t , with respective country-specific factor loadings. These factors introduce cross-section correlation in the observables and unobservables. Some unobserved common factors driving the variation in the observable inputs, \mathbf{g}_t , do not affect value added directly, but some others, \mathbf{f}_t , do.⁴ For illustration, a large global shock such as the recent Global Financial Crisis does not just affect economic performance, y , but also the evolution of inputs, such as investment in physical capital. The presence of the same unobserved factors in the input and output equations induces endogeneity in that the regressors are correlated with the unobservables in the production-function equation (u_{it}), making it difficult to identify β_i separately from λ_t and ρ_i (Kapetanios *et al.*, 2011). Equation (3) specifies the evolution of the common factors, which includes the potential for non-stationary factors ($\varrho = 1, \kappa = 1$) and thus non-stationary inputs and output.

The most important features of this setup are (i) the potential heterogeneity in the impact of observables and unobservables on output across countries ($\alpha_i, \beta_i, \lambda_i$), (ii) the potential non-stationarity of observables and unobservables ($y_{it}, \mathbf{x}_{it}, \mathbf{f}_t, \mathbf{g}_t$), and (iii) the endogeneity of observable inputs created by the common-factor structure. These properties have an important bearing on estimation and inference in macro panel data, which are at the heart of this paper. In the following, we illustrate how assumptions over these aspects give rise to different empirical estimators, relying on the classification in Table 1.⁵

If the data are demonstrably non-stationary, any specification choice carries implicit assumptions about the long-run equilibrium relationship in the data: any pooled regression model assumes that the cointegrating relationship is *identical across all countries* in the sample (common technology), whereas a heterogeneous model assumes that the cointegrating relationship *differs across countries*. Note that if the econometrician makes the wrong decision here and estimates a pooled model for what is a heterogeneous cointegrating relationship, then the empirical results are likely spurious *by construction*.⁶ Spurious results indicating serious empirical misspecification can, however, be detected by investigating residuals for non-stationarity or by implementing formal cointegration tests—we apply both strategies below.

⁴We introduce \mathbf{g}_t to avoid the impression that all macro variables are by assumption driven by the identical set of unobserved common factors \mathbf{f}_t , which would be wildly unrealistic.

⁵We use the following abbreviations: POLS, pooled OLS; 2FE, two-way fixed effects; FD, first-difference estimator; FE, country fixed effects; CCEP, common correlated effects pooled estimator (Pesaran, 2006); IFE, interactive fixed-effects estimator (Bai, 2009) CD-MG, cross-sectionally demeaned mean-group estimator; MG, mean group (Pesaran and Smith, 1995); GM-FMOLS, group-mean fully modified OLS Pedroni (2000); CMG, common correlated effects mean group (Pesaran, 2006); AMG, augmented MG, described in detail in the Appendix (in the Online Supporting Information).

⁶This is very easy to show: since our specification choice of homogeneity—imposing a common parameter, say β —is wrong, we enter linear combinations of the non-stationary observables $(\beta_i - \beta)x_{it}$ in the error terms, which are thus non-stationary by construction.

TABLE 1
TECHNOLOGY HETEROGENEITY AND UNOBSERVED COMMON FACTORS

Factor loadings, λ		Homogeneous	Heterogeneous	
Factors, f		Unrestricted	Linear	Unrestricted
Technology, β	Homogeneous	POLS, 2FE, FD $\alpha_{[i]} + \lambda f_i$	FE with trends $\alpha_i + \lambda_i t$	CCEP, IFE $\alpha_i + \lambda'_i f_i$
	Heterogeneous	CD-MG $\alpha_i + \lambda f_i$	MM $\alpha_i + \lambda_i t$ DKM, Pedroni	(CD) AMG, CMG ET, EHS

Notes: In addition to the various estimators, we provide examples of empirical applications in the cross-country growth literature, which adopted these implementations: MRW, Mankiw *et al.* (1992); Islam, Islam (1995); CEL, Caselli *et al.* (1996); MM, Martin and Mitra (2002); CD, Costantini and Destefanis (2009); DKM, Durlauf *et al.* (2001); Pedroni, Pedroni (2007); ET, Eberhardt and Teal (2013a); EHS, Eberhardt *et al.* (2013). A number of these references are in parentheses: Caselli *et al.* (1996) use the Arellano and Bond (1991) estimator, while Costantini and Destefanis (2009) adopt the Bai and Kao (2006) estimator; however, their empirical specifications nevertheless fit into the respective cells in our schematic presentation. For each case, we report the algebraic representation of how TFP is modeled when using this estimator. α refers to TFP levels, and the combination of λ and f_i (potentially non-linear) or t (linear) to TFP evolution over time—for the encompassing model, see equation (1). $\alpha_{[i]}$ is intended to highlight that in the POLS model we have α , but in the 2FE and FD models we have α_i —to further separate these cases would have constituted table overkill.

Assumptions about unobservable TFP also have direct implications for specification and thus identification: if TFP is non-stationary, then we face the difficulty that the estimation of the cointegrating relationship would somehow need to account for an *unobservable* process. Again, if the econometrician makes the wrong decision here in terms of specification—common versus idiosyncratic TFP evolution or a mix of the two—then the regression results may be spurious. If, on the other hand, TFP is assumed stationary, then deterministic components (year dummies, linear trends) should go a long way toward accounting for its impact and we can still estimate a cointegrating relationship between observable inputs and output (Pedroni, 2007). Our empirical implementation allows us to represent different scenarios for the specification of TFP that are representative of our assumptions about the heterogeneity or homogeneity of TFP evolution.

One of the central focal points of the cross-country growth empirical literature over the past two decades has been the endogeneity of inputs and, closely related, potential reverse causality in the estimation equation. The former implies that the capital and labor inputs of our production function are correlated with unobservable TFP; conceptually, it seems highly plausible that technical progress does not merely affect output directly, but also affects the choice of factor inputs—and similarly for other aspects of TFP, such as common shocks. Reverse causality implies that although we have written down a production function, we may run the risk of this representing a misspecified investment or labor-demand equation. In the existing literature, identification in the face of these difficulties is typically argued to be achieved through instrumentation, in panel models frequently employing the own-instrumentation strategy of the GMM estimators by Arellano and

Bond (1991) and Blundell and Bond (1998). These estimators, however, assume common technology and stationary variable series, as well as cross-section independence, and their identification strategy is invalid if any of these assumptions are violated (Pesaran and Smith, 1995).⁷

Our own empirical implementation allows us to adopt a flexible approach to dealing with this endogeneity problem, in that we employ unobservable common factors \mathbf{f}_t that induce the correlation with observable inputs in all countries. Note that TFP is, of course, a catch-all, in that shocks such as the 1970s oil crises affect both output and inputs directly, with no means for the existing empirical analysis of production functions to distinguish this type of shock from technological progress through knowledge accumulation and diffusion. Furthermore, shocks may not always be global in nature—for instance, extreme weather episodes leading to productivity shocks in only a small set of countries—so that it is important to emphasize that the common-factor framework also allows us to account for common factors that are more “local” in their impact. In our preferred implementation, the resulting endogeneity problem will be tackled by *accounting for the presence* of the unobservables in the empirical specification. In alternative implementations, (i) these factors are estimated (Bai, 2009), or (ii) a time-series econometric estimation approach (fully modified OLS) corrects for the endogeneity bias arising in this setup, but with more restrictive assumptions about common factors and thus TFP evolution (Phillips and Hansen, 1990; Pedroni, 2000). In order to tackle reverse causality, we will resort to a combination of the implementation dealing with the common factors and the “fully modified OLS” approach.

Conveniently, we can employ residual diagnostic tests to investigate whether our implementation has successfully captured the systematic relationships in unobservable TFP: focusing on the time-varying aspects of TFP, there is much to be said for interdependence across countries, whereby, for instance, knowledge created in one country spills over imperfectly to other countries. These spillovers induce dependence between unobservable TFP across countries and, since TFP is also correlated with the observable variables of the model, between labor and capital inputs across countries. By investigating whether residual series are cross-sectionally correlated, we can highlight the extent to which we have been able to deal with the dependence caused by the unobservable factors and thus, indirectly, whether we have addressed the endogeneity concern: if the residuals are white noise, we know that the empirical results do not suffer from endogeneity bias.

As this discussion highlights, the choice between estimating a pooled and a heterogeneous model as well as the treatment of TFP in this context is not some minor specification choice but a matter of great importance. We expect to see significant differences in estimates when moving between results for pooled and heterogeneous estimators, as well as between models that make different assumptions about the nature of TFP. We expect to see that things go very wrong if we make bad specification choices: parameter estimates may have nonsensical magnitudes or turn out to be insignificant, residuals will be non-stationary, and further diagnostic tests will indicate other serious shortcomings. This line of argument is the reason

⁷This criticism extends to the various control function estimators used in the microeconomic literature on production functions. For details, see the discussion in Eberhardt and Teal (2011).

why, below, we also present results for estimators that we would dismiss on theoretical grounds as unreliable or biased: if the assumptions implicitly made by adopting these estimators are seriously violated, then our diagnostic tests should pick this up.

We use Table 1 to categorize the various estimators adopted in our study and to provide some examples of previous work in the cross-country growth literature. With reference to equation (1), we also highlight the assumptions made about the TFP process in each case. The estimators assuming homogeneous technology in the upper panel of the table differ in their assumptions about the TFP process. The CCEP estimator by Pesaran (2006) and the IFE of Bai (2009) assume that TFP evolution differs across countries but can have common elements. The former represents an augmented version of a standard fixed-effects model, in which cross-section averages of all variables, i.e. $\bar{y}_t = N^{-1} \sum y_{it}$ and $\bar{\mathbf{x}}_t = N^{-1} \sum \mathbf{x}_{it}$, are introduced in the pooled regression to capture the unobserved common factors. In order to account for heterogeneity in the impact of these factors across countries, the coefficients on the cross-section averages are allowed to differ for each country. An alternative is provided by the IFE estimator (Bai, 2009), which is in the tradition of implementations that first estimate the common factors using principal component analysis (PCA) and then include them in the regression equation, which is then estimated iteratively until convergence is achieved (e.g. Bai and Ng, 2009).⁸ In the past, one criticism of this approach focused on the necessity to employ information criteria prior to estimation to establish the number of “relevant” common factors in the data. Recent theoretical work by Moon and Weidner (2015), however, showed that the assumption of too many common factors has a minimal impact on the consistency of the estimator. In contrast, the pooled OLS (POLS), two-way fixed-effects (2FE), and first-difference OLS (FD) estimators all assume common TFP evolution, captured by common year effects, but represent different assumptions about country-specific TFP levels: for 2FE and FD, these are—as in the CCEP—assumed to differ across countries, as for instance in Islam (1995), whereas they are assumed common in the POLS—matching the original Mankiw *et al.* (1992) assumption.

Non-stationarity has different implications for this set of pooled estimators: for POLS and 2FE, we assume homogeneous cointegration. Since both estimators account for time fixed effects,⁹ there is nothing to prevent us from including unobserved TFP in this cointegrating relationship, provided that it is common to all countries. If our specification choice is correct, the estimates from these models under cointegration would be super-consistent, implying that endogeneity would not lead to first-order bias in these models (Engle and Granger, 1987). The FD estimator is unaffected by non-stationarity, since the differencing of the estimation equation renders its observables and unobservables stationary by construction. At the same time, we are prevented from making any statements about a “long-run equilibrium” relationship from the FD estimate. The CCEP estimator *theoretically* yields consistent, but not super-consistent, estimates of β or the mean of β_i , regardless of whether our choice of homogeneous cointegration is correct (Kapetanios

⁸A related approach by Kneip *et al.* (2012) instead combines non-parametric methods with PCA to obtain the common factors.

⁹For POLS in the form of year dummies, in the case of 2FE, the mathematically equivalent data transformation into deviations from the cross-section mean.

et al., 2011). However, in practice it is often found that this estimator yields very different estimates from its mean-group version (see below) and concerns for heterogeneity misspecification remain. The Bai (2009) IFE assumes that the common factors are stationary, although since the PCA estimation is implemented on the differenced data, consistency may well extend to the non-stationary case, although to the best of our knowledge no theoretical results are available.

All models in the lower panel of the table allow for heterogeneous technology and are implemented in two steps: the first step represents some country-specific regression, while the second step consists of the averaging of country-specific estimates across the sample. All of these models thus represent “mean-group”-type estimators, named after the seminal contribution by Pesaran and Smith (1995). Again, they differ in their assumptions about the TFP process, where we have to distinguish both the commonality and the nature of TFP growth over time (all models allow for different TFP levels across countries): the estimators in the first and third columns (CD-MG; AMG, and CMG) allow for TFP to evolve in an unrestricted fashion, which includes the possibility of non-stationary TFP. In the latter case, they can accommodate cointegration between inputs, output, and TFP. These implementations, however, differ in their assumption about the commonality of TFP: in the CD-MG, TFP evolution is assumed to be common to all countries in the sample, whereas in the AMG and CMG it is allowed to differ. These models are implemented by use of data in deviation from the cross-section means (CD-MG) or by augmentation of the country-specific estimation equation with cross-section averages of all variables (CMG: see Pesaran, 2006; Chudik *et al.*, 2011; Kapetanios *et al.*, 2011) or alternative estimated placeholders (AMG: see Bond and Eberhardt, 2013; see also the Appendix, in the Online Supporting Information)—estimation is always by OLS.

In contrast, the heterogeneous estimators in the second column (MG and GM-FMOLS) of the diagram assume constant TFP growth and thus stationary TFP: these estimators adopt linear trends to capture TFP evolution over time and *require* a cointegrating relationship between inputs and output. Although the parameter estimates are in this case super-consistent, it was found that corrections for endogeneity and dynamic misspecification—both leading to second-order bias—as implemented in the “fully modified OLS” (FMOLS) estimator, are necessary in finite samples (Phillips and Hansen, 1990).

As was indicated above, for the AMG and CMG estimates we cannot rule out reverse causality, which represents a major shortcoming. In order to address this, we simply adopt FMOLS versions of these estimators, thus using augmented estimation equations, where the augmentations are cross-section averages or other placeholders. This empirical strategy can address endogeneity, serial correlation, and reverse causality *even in the case of non-stationary TFP*.

Inference for the pooled estimators builds on standard White heteroskedasticity-robust standard errors,¹⁰ with the exception of the CCEP, where we employed the bootstrap. Inference in the heterogeneous parameter models follows Pesaran and Smith (1995), employing a non-parametric variance estimator to construct

¹⁰Standard errors for the capital coefficients increase to 0.05 in the POLS and to 0.11 in the FD models if we cluster by country—those for the 2FE model are unchanged since the Stata® implementation that we adopt clusters standard errors.

standard errors and t -ratios—the exception here is the group mean version of the FMOLS estimator, which obtains “panel t -statistics” as $\bar{t}_{\beta^*} = N^{-1/2} \sum_i t_i$, where t_i is the t -ratio in country i and N is the number of countries.

4. DATA AND DATA PROPERTIES

For our empirical analysis, we employ aggregate sectoral data for manufacturing from developed and developing countries for the period 1970–2002 (UNIDO, 2004)—data from the same source (albeit at a higher level of disaggregation) were recently used by Rodrik (2013) to investigate cross-country convergence in manufacturing value added. Our sample represents an unbalanced panel of 48 countries with an average of 24 time-series observations (minimum 11, maximum 33).¹¹ Basic descriptive statistics and the sample makeup are detailed in the Appendix. The data allow us to estimate production functions with manufacturing-sector value added as output, and labor force and capital stock in manufacturing as inputs—the latter is created from data on gross fixed-capital formation following the standard perpetual inventory methodology. Our focus here is on value-added specifications, although we also considered gross-output specifications, results for which can be found in Eberhardt (2009). Further discussion of the data and their construction is confined to the Appendix.

In preparation for our regression analysis in Section 5, we carried out a range of variable unit root tests—the detailed results are presented in the Appendix. Despite all the problems related to panel unit-root testing, as well as considering the present data dimensions and characteristics, we can conclude that these results strongly suggest that the variable series in levels are non-stationary $I(1)$. We further applied the Pesaran (2015) test for weak cross-section dependence to our model variables. The results presented in the Appendix suggest that all series are subject to strong dependence.

5. REGRESSION RESULTS

The results in Panel A of Table 2 are based on estimating *pooled* models with variables in levels or first differences, including year dummies or in the CCEP country-specific period averages, following Pesaran (2006). Estimates for the capital coefficient in these regressions with constant returns to scale imposed are statistically significant at the 5 percent level or 1 percent level. For all three estimators in levels, the regression diagnostics (not reported) suggest serial correlation in the error terms, while constant returns to scale are rejected at the 1 percent level of significance except for POLS. Further, the OLS and 2FE residuals are found to be non-stationary, suggesting that the empirical results reported are

¹¹We do not carry out any interpolation to fill gaps in the time series and do not account for missing observations in any way. Our preferred empirical specifications are based on heterogeneous parameter models, where arguably the unbalancedness (around 25 percent of observations in the balanced panel are missing) comes less to bear on the estimation results than in the homogeneous models due to the averaging of estimates. For details on missing observations, see Table A.2 in the Appendix.

TABLE 2
THE MAIN REGRESSION RESULTS

PANEL A: POOLED MODELS	[1]	[2]	[3]	[4]		
Estimator	POLS	2FE	CCEP	FD		
Dependent variable	ly	ly	ly	Δ ly		
log capital pw	0.7895 [0.01]**	0.6752 [0.066]**	0.5823 [0.037]**	0.3195 [0.089]**		
Δ log capital pw						
Diagnostics						
CRS: p -value	0.96	0.00	0.00	0.72		
$(\gamma - \beta/k) I(1)$: p -value	0.99	0.99	0.99	0.00		
$\hat{\epsilon} I(1)$: p -value	1.00	0.78	0.00	0.00		
$\hat{\epsilon} CD$: p -value	0.15	0.05	0.03	0.39		
RMSE	0.462	0.135	0.113	0.103		
PANEL B: HETEROGENEOUS MODELS (AVERAGE ESTIMATES)	[1]	[2]	[3]	[4]	[5]	[6]
Estimator	MG	CD-MG	AMG	AMG	CMG	CMG
Dependent variable	ly	ly _{β}	ly- $\hat{\mu}_i$	ly	ly	ly
log capital pw	0.1789 [0.08]**	0.5295 [0.056]**	0.2896 [0.074]**	0.2982 [0.08]**	0.4663 [0.070]**	0.3125 [0.085]**
Common dynamic process				0.8787 [0.202]**		
Country trend	0.0174 [0.003]*		0.0001 [0.003]	0.0023 [0.004]		0.0108 [0.004]**
Diagnostics						
CRS: p -value	0.90	0.20	0.99	0.96	0.05	0.98
$(\gamma - \beta/k) I(1)$: p -value	0.99	0.66	0.99	0.99	0.51	0.99
$\hat{\epsilon} I(1)$: p -value	0.00	0.60	0.00	0.00	0.00	0.00
$\hat{\epsilon} CD$: p -value	0.00	0.25	0.96	0.30	0.02	0.82
RMSE	0.100	0.123	0.097	0.091	0.100	0.088

Notes: The regressions are for $N=48$ countries, $n = 1,194$ ($n = 1,128$) observations in the levels (first-difference) specifications. Values in brackets are White heteroskedasticity-consistent standard errors in Panel A, except for [3], where we present bootstrapped (100-replication) standard errors, and standard errors following Pesaran and Smith (1995) in Panel B. We indicate statistical significance at the 5 percent and 1 percent levels by * and **, respectively. Intercept estimates as well as average estimates on cross-section averages in Model [3] of Panel A and Models [5] and [6] of Panel B are omitted to save space. Dependent variable: ly; log value added per worker; ly _{β} , log value added per worker in deviation from the cross-section mean (ditto for capital-stock pw); Δ ly, growth rate of value added (per worker); $\hat{\mu}_i$ in Panel B is derived from the year-dummy coefficients of a pooled regression (CRS imposed) in first differences (FD), as described in the main text. Models [1], [2], and [4] in Panel A contain $T-1$ year dummies (for [4], in first differences). For all diagnostic tests (except RMSE), we report p -values. (i) The null hypothesis for the “CRS” Wald tests is constant returns. (ii) $(\gamma - \beta/k) I(1)$ reports analysis of regression residuals *incorporating TFP*, using a Pesaran (2007) CIPS test with two lags and a null of non-stationarity (full results are available on request). For this and the following CIPS test, we adopted a pragmatic approach in setting the lags equal to 2; shorter lags may be insufficient to capture serial correlation, while longer lags will demand too much from the data, given the very moderate time-series dimension (and further force us to drop country series). (iii) $\hat{\epsilon} I(1)$ reports results for a Pesaran (2007) CIPS test with two lags and a null of non-stationarity (full results are available on request). (iv) The Pesaran (2015) CD test has the null of cross-sectional weak dependence. Due to data restrictions (unbalanced panel with missing observations), we are forced to drop two (eight) countries from the sample to compute this test for the levels (FD) residuals. (v) RMSE is the root-mean-squared error.

potentially spurious. Cross-section dependence is present in all residual series to a greater or lesser extent, with the 2FE and CCEP models rejecting weak cross-section dependence at the 5 percent level. The POLS results in Model [1] suggest that failure to account for time-invariant (TFP-level) heterogeneity across countries yields biased results: at around 0.8, the capital coefficient is considerably inflated. Accounting for country-specific intercepts in Model [2] reduces these coefficient estimates somewhat. The same parameter in the CCEP results in Model [3] is yet lower still, around 0.6. The OLS regression in first differences in Model [4] yields quite different results: the capital coefficient is now around 0.3, CRS cannot be rejected, and the AR(1) tests (not reported) show only first-order serial correlation for this model, which is to be expected given that the errors are in the first differences. This echoes the favorable performance in simulation exercises to capture the *average* of a heterogeneous-technology coefficient (see Bond and Eberhardt, 2013, and related online Appendix). However, recall that the first-difference specification cannot be interpreted as a long-run equilibrium equation and we may well be capturing short-run (business-cycle) fluctuations in these results. Nevertheless, it appears that the FD estimator obtains sound diagnostics and a theory-consistent technology estimate—this indicates that accounting for non-stationarity (of factor inputs and TFP) plays a crucial role in estimating cross-country production functions.¹²

We can make use of the year dummy coefficients derived from the pooled FD model to obtain an estimate of the common dynamic process $\hat{\mu}_t$: an estimate of the average TFP evolution (for details, see the Appendix). Figure 1 illustrates the evolution path of this common dynamic process for the unrestricted and CRS models. The graphs show severe slumps following the two oil shocks in the 1970s, while the 1980s and 1990s indicate considerable upward movement.¹³ If we follow the “measure of our ignorance” interpretation of TFP, then a decline in global manufacturing TFP as evidenced in the 1970s should not be interpreted as a decline in knowledge, but a worsening global manufacturing *environment*, which seems plausible.

In the following, we relax the assumption implicit in the pooled regressions that all countries possess the same production technology. At the same time, we maintain that common shocks and/or cross-sectional dependence have to be accounted for in some fashion. Unweighted averages of country parameter estimates are presented in Panel B of Table 2.¹⁴ The *t*-statistics for the country-regression averages reported are measures of dispersion for the sample of country-specific estimates, following Pesaran and Smith (1995).

¹²Simulation exercises (Bond and Eberhardt, 2013) generally highlight the favorable performance of the FD estimator in standard non-stationary panel setups. However, while this may yield an unbiased estimate of average technology, country-specific TFP estimates are nevertheless biased if the “true” technology differs across countries.

¹³These graphs are “data specific”: for years where data coverage is good; this can be interpreted as “global,” whereas in later years (ten countries have data for 2001, only two for 2002, omitted from the graph) this interpretation collapses.

¹⁴Robust means weighing down outliers yield very similar results, with kernel estimates of the distribution of capital coefficients showing no influential outliers.

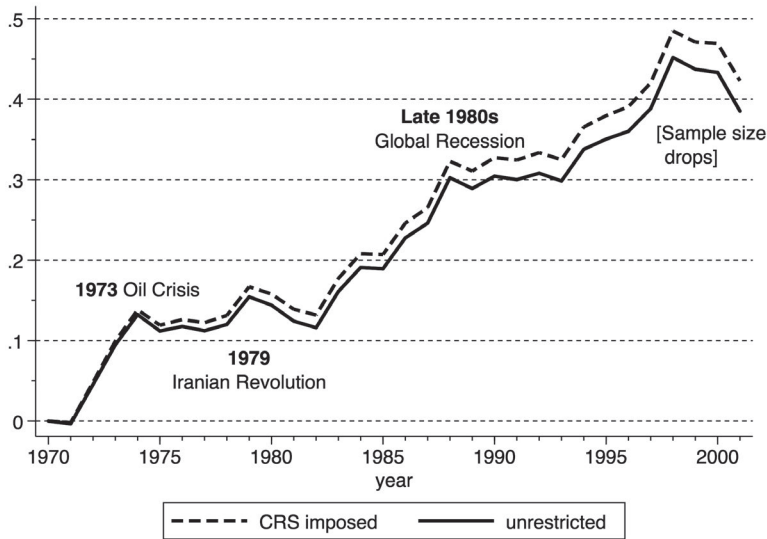


Figure 1. The Evolution of “Average” TFP
 Notes: Derived from results in column [4], Panel A of Table 2.

Our first observation regarding the averaged country results is that across all specifications the means of the capital coefficients are considerably lower than in the pooled levels models: between 0.2 and 0.5, rather than between 0.6 and 0.9.¹⁵ Closer inspection suggests the following patterns across the heterogeneous parameter regression results: first, the two more restrictive specifications in Models [1] and [2] are misspecified. For the MG, which assumes linear TFP evolution, the residual diagnostics indicate strong cross-section dependence; for the CD-MG, assuming common TFP evolution, the residuals appear non-stationary, so that we cannot rule out that these results are spurious. Second, for the AMG estimators, which account for a flexible TFP process in the estimation equation, the diagnostic test results are favorable and the averaged coefficients around 0.3. Third, the results for the CMG with and without an additional country trend differ considerably, with the former close to the AMG results and the latter slightly larger, around 0.45. Diagnostic tests, however, suggest that the standard CMG suffers from cross-sectionally strongly dependent residuals (see the CD test results of Pesaran, 2015).

Our results imply that: (i) heterogeneous specifications that allow for a combination of commonality and idiosyncrasy in the TFP evolution provide the closest match to the data and the most favorable diagnostics; (ii) the estimated capital coefficients in the preferred empirical specifications are close to 0.3; (iii) TFP

¹⁵The results presented are robust to alternative specifications (all results are available on request): first, we estimated all models in first differences; second, we adopted alternative country-level deterministics (an additional squared trend in the levels models; an additional trend in the models in a first-difference specification); third, we estimated gross-output-based models including material inputs as additional covariate; and fourth, we estimated dynamic ARDL versions of the presented static models.

appears to be non-stationary and thus leads to empirical misspecification in models that ignore this property;¹⁶ and (iv) our preferred results based on favorable residual diagnostics represent a close match between the Pesaran (2006) CMG and the Bond and Eberhardt (2013) AMG estimators.

These conclusions are backed up by the results for an alternative estimator, the Bai (2009) interactive fixed effects, which are presented in Table 3: the capital coefficients are uniformly close to 0.3, except in case of the single-common-factor model in Model [1], for which the estimate is close to 0.6. This model also displays non-stationary residuals. The diagnostics are favorable for all other models, suggesting stationary and only weakly dependent residuals, although it is notable that only the specifications with two and five common factors do not reject constant returns to scale. The relative stability of these results regardless of the number of factors included (other than the case of just a single factor) suggests that the model in column [2], assuming two factors, captures the data quite well. Comparing these results with those for the AMG and CMG, it appears that the standard CMG implementation in Panel B, column [5] of Table 2, which on the basis of containing a single observed covariate (log capital stock per worker) can only capture *a single common factor* f_t , is biased upward.

Finally, the Pedroni (2000) group-mean FMOLS approach, for which results are presented in the Appendix, provides further evidence that failure to account for non-stationary TFP leads to the collapse of the empirical estimates when analyzing cross-country manufacturing production. When we investigate the full sample of 48 countries, we find that the standard GM-FMOLS yields very low and statistically insignificant coefficient estimates for capital, whereas upon inclusion of the common dynamic process, or of cross-section averages, we obtain results that closely match those from the previous table of OLS-based results. Since the FMOLS methodology is robust to reverse causality, this provides assurance that our AMG and CMG estimates represent production-function coefficients and not misspecified investment or labor-demand equations. In a further robustness check, we limit the sample to 26 countries for which individual time-series unit-root and stationarity (DF and KPSS) tests could not reject non-stationarity (the FMOLS approach assumes non-stationarity and cointegration),¹⁷ to show that results do not change in any significant way.

Based on the residual diagnostics, our empirical results thus largely favor models with heterogeneous technology, which account for a combination of heterogeneous and common TFP. The notable exception here is the (pooled) first-difference estimator, which we found relatively unaffected by the failure to explicitly model these features, likely due to the absence of integrated variables and processes once data are differenced. In our minds, the fact that the FD estimator obtains a similar capital coefficient to that in the averaged AMG or CMG results is *in spite*

¹⁶A comparison of results for the unit-root analysis of the regression residuals $\hat{\epsilon}$ and of $y - \hat{\beta}k$ or its heterogeneous-technology variant (which contains $\hat{\epsilon}$ and the common factors) indicates that the POLS, 2FE, and CD-MG models cannot capture non-stationary TFP.

¹⁷We appreciate that the single time-series tests that are typically employed have low power in the present short time series, but this analysis is intended to be indicative of the remarkable robustness of our findings to a reduction in the sample to countries with *plausibly* rather than definitively $I(1)$ variable series.

TABLE 3
HOMOGENEOUS MODELS USING THE IFE OF BAI (2009)

	[1]	[2]	[3]	[4]	[5]
Estimator	IFE	IFE	IFE	IFE	IFE
Dependent variable	ly	ly	ly	ly	ly
Number of factors	1	2	3	4	5
log capital pw	0.5645 [0.0983]***	0.3380 [0.0770]***	0.3251 [0.0637]***	0.3378 [0.0847]***	0.2654 [0.0557]***
Diagnostics					
CRS: p -value	0.10	0.47	0.00	0.00	0.29
$(\gamma - \beta/k) I(1)$: p -value	0.95	0.33	0.38	0.33	0.63
$\hat{\epsilon} I(1)$: p -value	0.37	0.01	0.00	0.00	0.00
$\hat{\epsilon} CD$: p -value	0.23	0.13	0.18	0.22	0.31
RMSE	0.112	0.096	0.086	0.077	0.065

Notes: The results presented are for the Bai (2009) interactive fixed-effects (IFE) estimator. All regression models absorb country and time fixed effects—the results without these fixed effects (available on request) are broadly similar. Values in brackets are absolute standard errors clustered at the country level. The results using a bootstrap procedure are broadly in line with those presented above, albeit less precise—this is not surprising given the unbalanced nature of the panel. All models were estimated in Stata using the `regife` module written by Matthieu Gomez. See the notes to Table 2.

of technology heterogeneity, and not because pooled specifications are favorable. To this end, we also carried out a significant number of formal parameter homogeneity tests (see the Appendix), which confirmed our preference for heterogeneous technology. Since residual testing for stationarity represents a somewhat *ad hoc* cointegration test, we also confirmed this property in our preferred heterogeneous model, adopting the Gengenbach *et al.* (2016) testing procedure (for the results, see the Appendix).

Our general production-function framework provides a number of insights into TFP estimation: first, it seems sensible to allow for maximum flexibility in the structure of the empirical TFP terms; if TFP represents a “measure of our ignorance,” then it makes sense to allow for differential TFP across countries and time, with the latter unconstrained with regard to non-stationarity.

Second, it further makes sense to keep an open mind about the commonality of TFP: while early empirical models (Mankiw *et al.*, 1992; Islam, 1995) assumed common TFP growth for all countries, later studies preferred to specify differential TFP evolution across countries. We believe that the arguments for commonality (the non-rival nature of knowledge, spillovers, and global shocks) and idiosyncrasy (patents, tacit knowledge, and learning by doing) call for an empirical specification that does not rule out either by construction.

Third, following Durlauf *et al.* (2001) and Pedroni (2007), we argue for an empirical specification that allows for parameter heterogeneity across countries and for a shift away from the widespread focus on TFP analysis and toward an integrated treatment of the production technology *in its entirety*, including technology heterogeneity, TFP levels, and growth rates.

We can illustrate the contribution of these three aspects of production technology in Figure 2, where we plot country-specific linear regressions of value added per worker on capital stock per worker for our manufacturing data from 48 countries: in the upper panel, which ignores TFP growth over time, the slopes of these production functions appear to be very similar, reinforcing the notion of a common production technology, whether we assume common or heterogeneous intercept terms (TFP levels) and common or heterogeneous slopes (capital coefficients). The same result obtains if we assume common TFP growth for all countries in the sample. From this, we conclude that common TFP evolution in combination with either common or heterogeneous technology leads to empirical results that run counter to the macro factor share evidence, namely a capital coefficient around 0.7 rather than around 0.3.

In the lower panel, we adjust the value added per worker variable for TFP evolution over time¹⁸ and again plot the country-specific regression lines implied by a production-function model. Thus, allowing for heterogeneous TFP and common shocks, we can see that the fitted regression lines now provide clear evidence of technology heterogeneity, with the average capital coefficient from a heterogeneous parameter model around 0.3, while a pooled model still yields an inflated estimate of 0.79. From this, we can conclude that heterogeneous TFP evolution

¹⁸We compute $ly^{adj} = ly - \hat{c}_i t - \hat{a}_i \hat{\mu}_i^*$ following equation (4) below, using the results for the empirical model in Panel B, column [3] of Table 2.

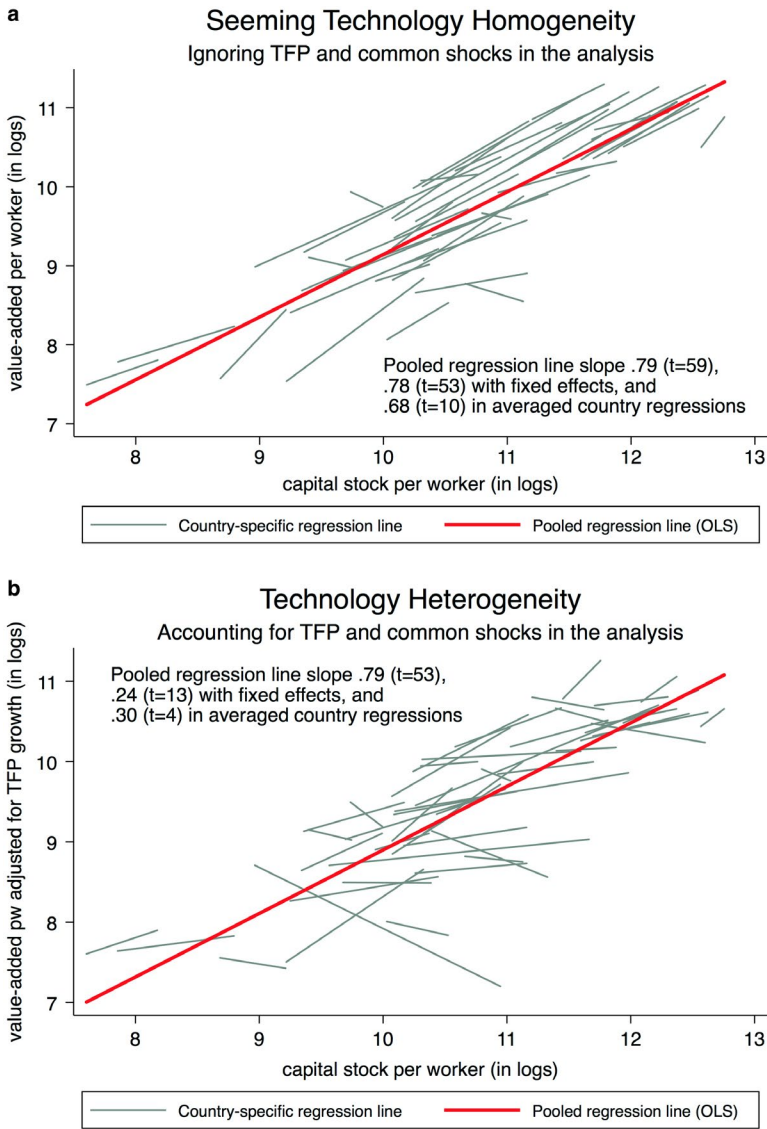


Figure 2. Technology Heterogeneity in the Analysis of Development and TFP

Notes: The graph in panel (a) simply fits a linear regression line from country-specific data on manufacturing value added per worker on manufacturing capital stock per worker (in logs) separately in 48 countries (the teal-colored lines), thus ignoring TFP evolution, spillovers, and common shocks. The red line represents the pooled OLS regression slope. We further report the estimated slopes for the pooled model with and without fixed effects and the mean slope for a naive mean-group model (with country intercepts only). Values in parentheses are *t*-ratios. The graph in panel (b) uses log manufacturing value added per worker adjusted for annual country-specific TFP and plots this variable against log manufacturing capital stock per worker separately for each of the 48 sample countries. Although virtually identical, the pooled regression line, in red here, is for adjusted value added per worker. Again, we report slope coefficients for pooled OLS with and without country fixed effects, and the mean-group result (which is a graphical representation of the AMG result in Panel B, column [3] of Table 2). [Colour figure can be viewed at wileyonlinelibrary.com]

alone yields results that are in conflict with the macro data, whereas the combination of heterogeneous technology and heterogeneous TFP evolution yields a global average of 0.3.

6. TFP IN A HETEROGENEOUS-TECHNOLOGY WORLD

What are the implications of homogeneity misspecification for estimated TFP levels and growth rates? In the following, we provide some insights into the resulting patterns of TFP growth and introduce a new approach to estimate TFP levels, which is necessitated by the adoption of a heterogeneous-technology model. In both cases, we try to establish whether the choice between homogeneous and heterogeneous technology makes a substantial difference to TFP measurement.

In the top-left plot of Figure 3, we compare the distribution of the annual TFP growth estimates from growth accounting (the dashed transparent histogram) and our preferred panel time-series regression (the gray histogram). While both distributions look Gaussian, it is obvious that the accounted TFP growth rates are substantially greater in range. The top-right plot in the same figure fits a linear regression line (with 90 percent confidence bands) for the annual TFP growth rates against value added per worker (in logs). While the *estimated* TFP growth rates from the preferred heterogeneous estimator seem to display a negative relationship with output, a tendency that disappears if we omit the top and bottom 5 percent of the distribution in the bottom-right plot or if we employ total-period averages of TFP growth and value added per worker in the bottom-left plot, the *accounted* TFP growth rates consistently display a positive relationship regardless of censoring or averaging.

We can draw two conclusions from this analysis: first, the range and variance of the common-technology TFP growth estimates are artificially inflated, thus providing an increased likelihood of statistically significant results in further “TFP regressions.” Second, under the assumption of common technology, these TFP growth series are clearly linked to the *level* of development, with richer countries enjoying higher TFP growth.

A further implication of a shift from common to heterogeneous technology is that we require a new methodology to arrive at TFP-level estimates from our preferred country-level regression models: from these regressions, we can obtain estimates for the intercept, technology parameters, idiosyncratic, and common trend coefficients or the parameters on the cross-section averages for AMG and CMG specifications, respectively. One may be tempted to view the coefficients on the intercepts as TFP-level estimates, just as in the pooled fixed-effects case. However, once we allow for heterogeneity in the slope coefficients, the interpretation of the intercept as an estimate for base-year TFP level is no longer valid, as was already recognized by Bernard and Jones (1996a). In order to illustrate our case, we employ a simple linear relationship between value added and capital where the contribution of TFP *growth* has already been accounted for (see equation (4) below).

In Figure 4, we provide scatter plots for the “adjusted” log value added per worker (on the *y*-axis) against log capital per worker (on the *x*-axis) as well as a

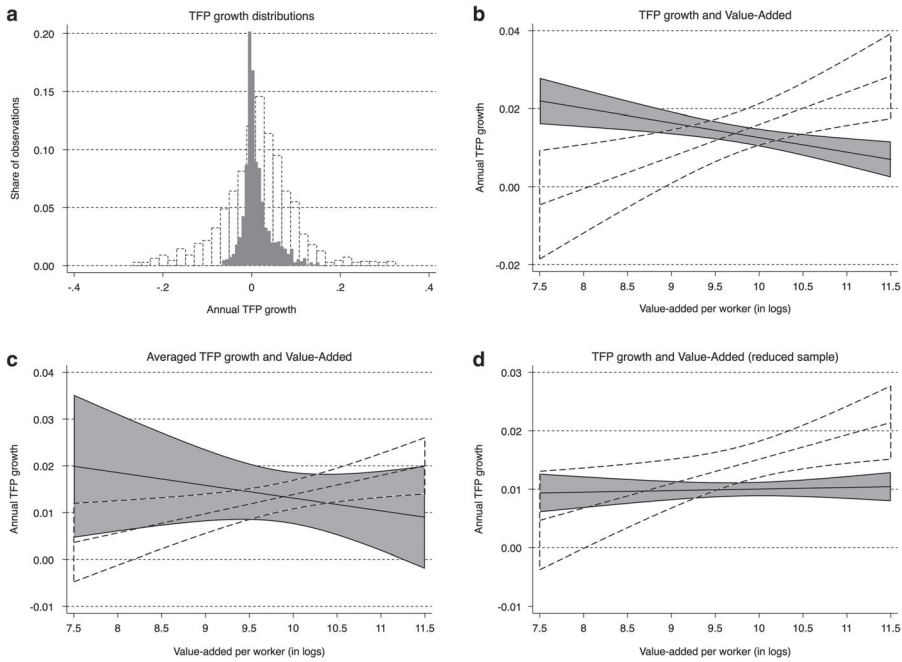


Figure 3. TFP Growth from Regression and Growth Accounting ($\beta^k = .33$)

Notes: We compare the TFP growth estimates derived from our preferred regression model, the AMG estimator (Table 2, Panel B, column [3]—the gray histogram and 90 percent confidence interval; capital coefficients differ across countries), with those obtained from simple TFP growth accounting (transparent histogram and dashed 90 percent confidence intervals; common capital coefficient = 0.33). Clockwise from the top left, the graphs provide (i) histograms for these two sets of estimates, (ii) linear regression lines (and 90 percent confidence intervals) of TFP growth against log value added per worker, (iii) as in (ii) but removing the top and bottom 5 percent of TFP growth estimates as computed in either exercise, and (iv) as in (ii) but using 48 country TFP growth averages.

fitted regression line for these observations in each of the following four countries: in the upper panel, France (circles) and Belgium (triangles); and in the lower panel, South Korea (circles) and Malaysia (triangles). The “adjustment” is based on the country-specific estimates from the AMG regression in Table 2, Panel B: we compute

$$(4) \quad y_{it}^{adj} = y_{it} - \hat{c}_i t - \hat{d}_i \hat{\mu}_i,$$

where \hat{c}_i and \hat{d}_i are the country-specific estimates for the linear trend term and the common dynamic process, respectively. We then plot this variable against log capital per worker for each country separately. This provides a visual equivalent of the estimates for the capital coefficient (the slope) and a candidate TFP-level estimate (the intercept) in the country regression.

The upper panel of Figure 4 shows two countries (France and Belgium) with virtually identical capital-coefficient estimates (slopes). The in-sample fitted regression line is plotted as a solid line, while the out-of-sample extrapolation

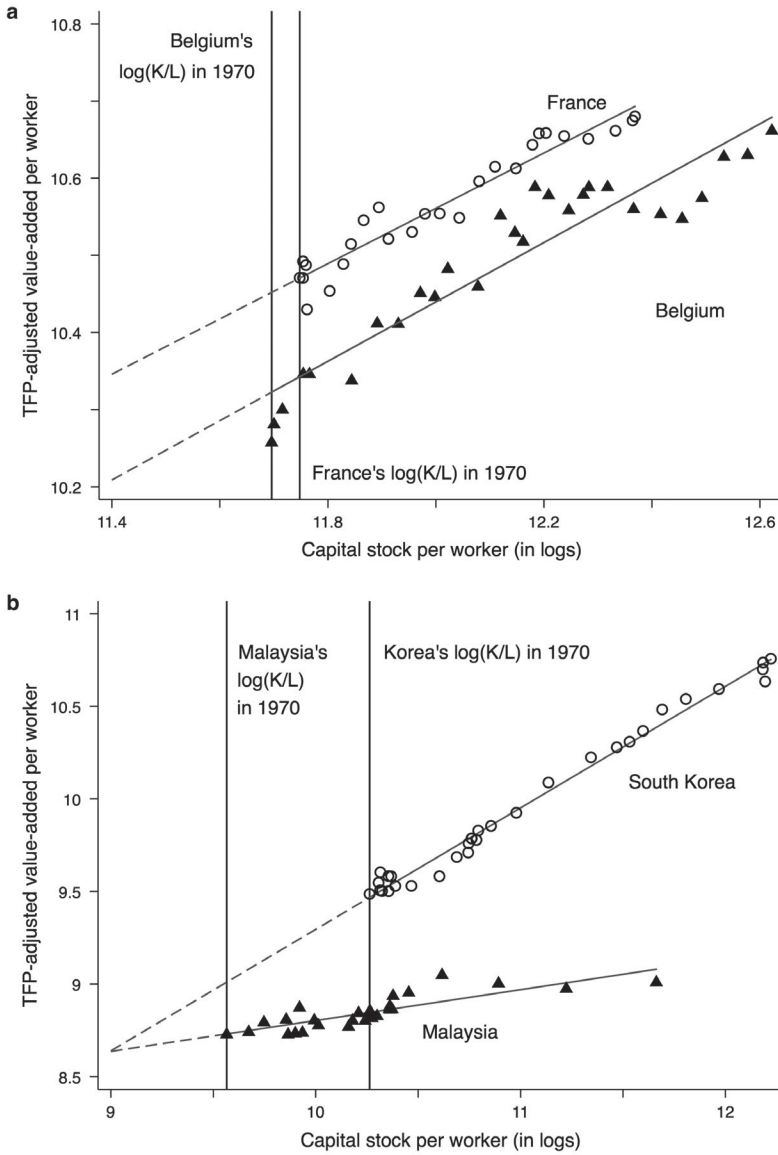


Figure 4. Regression Intercepts and TFP-Level Estimates

Notes: In-sample (solid) and out-of-sample (dashed) linear prediction of the relationship between TFP-adjusted value added per worker (on the y-axis) and capital stock per worker (on the x-axis), all variables being in logarithms—for details on the TFP adjustment, see the main text.

toward the y-axis is plotted in dashes. The country estimates for the intercepts can be interpreted as TFP levels, since these countries have very similar capital-coefficient estimates ($\hat{b}_{FRA} \approx \hat{b}_{BEL} \approx \hat{b}$). In this case, the graph represents the linear model $y_{it}^{adj} = \hat{a}_i + \hat{b} \log(K/L)_{it}$, where \hat{a}_i possesses the *ceteris paribus* property. In contrast, the lower panel shows two countries (Malaysia and South Korea) that

exhibit very different capital-coefficient estimates. In this case, $\hat{\alpha}_i$ cannot be interpreted as possessing the *ceteris paribus* quality, since $\hat{b}_{MYS} \neq \hat{b}_{KOR}$ (i.e. *ceteris non paribus*), or as Bernard and Jones (1996a) put it: “comparing apples to oranges.” In the graph, we can see that Malaysia has a considerably higher intercept term than South Korea, even though the latter’s observations lie above those of the former at any given point in time. This illustrates that once technology parameters in the production function differ across countries, the regression intercept can no longer be interpreted as a TFP-level estimate.

We can suggest an alternative measure for TFP-levels that is robust to parameter heterogeneity. Referring back to the scatter plots in Figure 4, we marked the base-year level of log capital per worker by vertical lines for each of the four countries. We suggest the use of the locus where the solid (in-sample) regression line hits the vertical base-year capital stock level as an indicator of the TFP level in the base year. These *adjusted* base-year and final-year TFP levels are thus

$$(5) \quad \hat{\alpha}_i + \hat{b}_i \log(K/L)_{0,i} \quad \text{and} \quad \hat{\alpha}_i + \hat{b}_i \log(K/L)_{0,i} + \hat{c}_i \tau + \hat{d}_i \hat{\mu}_\tau^*$$

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

Table 4 provides details on absolute rank differences implied by TFP-level rankings for accounting (“Levels”) and regression (“2FE,” “AMG,” and “CMG”) exercises. These descriptives indicate the very substantial differences arising from TFP levels obtained from common versus heterogeneous-technology models.

7. CONCLUDING REMARKS

In this paper, we have investigated how differences in manufacturing-sector technology across countries can be modeled empirically. We have adopted an encompassing framework that allows for the possibility that the impact of observable and unobservable inputs on output differs across countries, as well as for non-stationary evolution of these processes. Our regression framework has enabled us to model a number of characteristics that are likely to be prevalent in manufacturing data from a diverse set of countries: first, we have allowed for technology heterogeneity across countries. The empirical results are confirmed by formal testing procedures to suggest that technology parameters in manufacturing production indeed differ across countries. This finding supports earlier work using aggregate economy data (Durlauf, 2001; Pedroni, 2007): if the production technology differs in cross-country manufacturing, the aggregate economy technology is unlikely to be homogeneous.

Second, we have allowed for unobserved common factors to drive output, but with differential impact across countries, thus inducing cross-section dependence. These common factors are visualized by our common dynamic process, which follows patterns over the 1970–2002 sample period that match historical events. The interpretation of this common dynamic process $\hat{\mu}_t^*$ would be that for

TABLE 4
COUNTRY RANKINGS BY TFP LEVEL

Absolute rank difference between implementations	[1]	[2]	[3]	[4]	[5]	[6]
	AMG—FE	CMG—FE	Levels—FE	Levels—AMG	Levels—CMG	CMG—AMG
Minimum	0.0	0.0	0.0	0.0	0.0	0.0
Mean	10.5	10.2	7.0	5.8	5.4	1.1
Median	10.0	10.0	5.0	5.0	4.5	1.0
IQR	7.5	9.0	7.0	7.0	6.8	1.0
Maximum	33.0	34.0	24.0	17.0	19.0	7.0

Notes: The table provides distributional statistics on the relative TFP-level ranking (by magnitude) derived from the three regression models as well as the levels accounting for 1990: “AMG—FE” is based on the absolute difference between the TFP-level rankings implied by the AMG and FE estimators, and similarly for the other comparisons. FE, two-way fixed-effects estimator (Table 2, Panel A, column [2]); AMG, augmented mean-group estimator (Table 2, Panel B, column [3]); CMG, mean-group version of the Pesaran (2006) CCE estimator (Table 2, Panel B, column [5]); IQR, interquartile range of rank differences.

the manufacturing sector, *similar factors* drive production in all countries, *albeit to different extents*. This is equivalent to suggesting that the “global tide” of innovation can “lift all boats,” but that technology transfer from developed to developing countries is dependent on the recipient’s production technology and absorptive capacity, among other things.

Third, our empirical setup allows for a type of endogeneity whereby the unobservables driving output are also driving the evolution of inputs. This leads to an identification problem, in that standard panel estimators cannot identify the parameters on the observable inputs as distinct from the impact of the unobservables. Monte Carlo simulations (Bond and Eberhardt, 2013) have highlighted the ability of the CMG and AMG estimates to deal with this problem successfully, and our empirical results indicate parity between these two heterogeneous panel estimators. Furthermore, additional analysis has confirmed that the empirical results are robust to the use of an alternative panel time-series econometric approach that further addresses reverse causality. Standard practices to deal with endogeneity (Arellano and Bond, 1991; Blundell and Bond, 1998) are only appropriate in a stationary framework with homogeneous technology (Pesaran and Smith, 1995). The adoption of a non-stationary panel econometric approach that accounts for cross-section dependence is, in our view, a sound empirical strategy to address both of these concerns and should be applied more widely to cross-country productivity analysis.

Our analysis represents a step toward making cross-country empirics relevant to individual countries by moving away from empirical results that characterize the *average country* and toward a deeper understanding of the *differences across countries*, a notion that is clearly echoed elsewhere in the literature (Temple, 1999; Durlauf, 2001; Durlauf *et al.*, 2005, 2001). Cross-country regressions of time averages, in the empirical tradition of Barro (1991) and Mankiw *et al.* (1992), emphasize the variation in the data across countries (“between variation”) and implicitly assume that the processes driving capital accumulation in, say, the United States and Malawi are the same, and that at a distant point in time the latter can feasibly reach the capital-to-labor ratio of the former to achieve the same level of development. However, development is an evolution over time, which requires that apart from recognizing the potential for differences across countries, we analyze the individual evolutionary paths of countries over time (emphasizing the “within variation” in the data). The empirical methods used in this paper enable us to incorporate all of these concerns within one unifying empirical framework. A second conclusion from this study is that the key to understanding cross-country differences in income is not exclusively linked to understanding TFP differences, but requires careful concern for differences in production technology. Since modeling production technology as heterogeneous across countries requires an entirely different set of empirical methods, we have focused on developing this aspect in the present paper and have left empirical testing of rival hypotheses about the patterns and sources of technological differences for future research.

REFERENCES

- Abramowitz, M., "Resource and Output Trends in the United States Since 1870," *American Economic Review*, 46, 5–23, 1956.
- Arellano, M. and S. R. Bond, "Some Tests of Specification for Panel Data," *Review of Economic Studies*, 58, 277–97, 1991.
- Azariadis, C. and A. Drazen, "Threshold Externalities in Economic Development," *Quarterly Journal of Economics*, 105, 501–26, 1990.
- Bai, J., "Panel Data Models with Interactive Fixed Effects," *Econometrica*, 77, 1229–79, 2009.
- Bai, J. and C. Kao, "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*, Elsevier Science, Amsterdam, 2006.
- Bai, J. and S. Ng, "Panel Cointegration with Global Stochastic Trends," *Journal of Econometrics*, 149, 82–99, 2009.
- Banerjee, A. and A. Newman, "Occupational Choice and the Process of Development," *Journal of Political Economy*, 101, 274–98, 1993.
- Barro, R. J., "Economic Growth in a Cross-Section of Countries," *Quarterly Journal of Economics*, 106, 407–43, 1991.
- Battese, G. E., Prasada Rao, D. S., O'Donnell, C. J., "A Metafrontier Production Function for Estimation of Technical Efficiencies and Technology Gaps for Firms Operating under Different Technologies," *Journal of Productivity Analysis*, 21, 91–103, 2004.
- Bernard, A. B. and C. I. Jones, "Comparing Apples to Oranges: Productivity Convergence and Measurement Across Industries and Countries," *American Economic Review*, 86, 1216–38, 1996a.
- _____, "Productivity Across Industries and Countries: Time Series Theory and Evidence," *Review of Economics and Statistics*, 78, 135–46, 1996b.
- Blundell, R. and S. R. Bond, "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models," *Journal of Econometrics*, 87, 115–43, 1998.
- Bond, S. R. and M. Eberhardt, "Accounting for Unobserved Heterogeneity in Panel Time Series Models," Unpublished mimeo, November 2013.
- Bond, S. R., A. Hoeffler, and J. Temple, "GMM Estimation of Empirical Growth Models," CEPR Discussion Paper 3048, 2001.
- Bond, S. R., A. Leblebicioglu, and F. Schiantarelli, "Capital Accumulation and Growth: A New Look at the Empirical Evidence," *Journal of Applied Econometrics*, 15, 1073–99, 2010.
- Caselli, F., "Accounting for Cross-Country Income Differences," in Ph. Aghion and S. N. Durlauf (eds), *Handbook of Economic Growth*, vol. 1, Elsevier, Amsterdam, 679–741, 2005.
- Caselli, F., G. Esquivel, and F. Lefort, "Reopening the Convergence Debate: A New Look at Cross-Country Growth Empirics," *Journal of Economic Growth*, 1, 363–89, 1996.
- Chudik, A., M. H. Pesaran, and E. Tosetti, "Weak and Strong Cross Section Dependence and Estimation of Large Panels," *Econometrics Journal*, 14, C45–90, 2011.
- Costantini, M. and S. Destefanis, "Cointegration Analysis for Cross-Sectionally Dependent Panels: The Case of Regional Production Functions," *Economic Modelling*, 26, 320–7, 2009.
- Durlauf, S. N., "Manifesto for a Growth Econometrics," *Journal of Econometrics*, 100, 65–9, 2001.
- _____, "Nonergodic Economic Growth," *Review of Economic Studies*, 60, 349–66, 1993.
- Durlauf, S. N., P. A. Johnson, and J. R. W. Temple, "Growth Econometrics," in Ph. Aghion and S. N. Durlauf (eds), *Handbook of Economic Growth*, vol. 1, Elsevier, Amsterdam, 555–677, 2005.
- Durlauf, S. N., A. Kourtellos, and A. Minkin, "The Local Solow Growth Model," *European Economic Review*, 45, 928–40, 2001.
- Eberhardt, M., "Modelling Technology in Agriculture and Manufacturing Using Cross-Country Panel Data," D.Phil. thesis, St John's College, University of Oxford, Oxford, 2009, available online at <https://ora.ox.ac.uk/objects/uuid:d60f62f5-43e2-4473-b899-f4358d758e1e>
- Eberhardt, M., C. Helmers, and H. Strauss, "Do Spillovers Matter when Estimating Private Returns to R&D?" *The Review of Economics and Statistics*, 95, 436–48, 2013.
- Eberhardt, M. and F. Teal, "Econometrics for Grumblers: A New Look at the Literature on Cross-Country Growth Empirics," *Journal of Economic Surveys*, 25, 109–55, 2011.
- _____, "No Mangoes in the Tundra: Spatial Heterogeneity in Agricultural Productivity Analysis," *Oxford Bulletin of Economics and Statistics*, 75, 914–39, 2013a.
- _____, "Structural Change and Cross-Country Growth Empirics," *World Bank Economic Review*, 27, 229–71, 2013b.
- Engle, R. F. and C. W. J. Granger, "Cointegration and Error Correction: Representations, Estimation and Testing," *Econometrica*, 55, 252–76, 1987.

- Ertur, C. and W. Koch, "Growth, Technological Interdependence and Spatial Externalities: Theory and Evidence," *Journal of Applied Econometrics*, 22, 1033–62, 2007.
- _____, "A Contribution to the Schumpeterian Growth Theory and Empirics," *Journal of Economic Growth*, 16, 215–55, 2011.
- Gengenbach, Ch., J. Westerlund, and J.-P. Urbain, "Error Correction Testing in Panels with Common Stochastic Trends," *Journal of Applied Econometrics*, 31, 982–1004, 2016.
- Gollin, D., "Getting Income Shares Right," *Journal of Political Economy*, 110, 458–74, 2002.
- Griliches, Z., "Comment on An Appraisal of Long-Term Capital Estimates: Some Reference Notes by Daniel Creamer," in *Output, Input, and Productivity Measurement, The Conference on Research in Income and Wealth*, National Bureau of Economic Research (NBER), Cambridge, MA, 446–9, 1961.
- Hulten, C. R., "Growth Accounting," in B. H. Hall and N. Rosenberg (eds), *Handbook of the Economics of Innovation*, vol. 2, Elsevier, Amsterdam, 987–1031, 2010.
- Islam, N., "Growth Empirics: A Panel Data Approach," *Quarterly Journal of Economics*, 110, 1127–70, 1995.
- Kapetanios, G., M. H. Pesaran, and T. Yamagata, "Panels with Nonstationary Multifactor Error Structures," *Journal of Econometrics*, 160, 326–48, 2011.
- Kneip, A., R. C. Sickles, and W. Song, "A New Panel Data Treatment for Heterogeneity in Time Trends," *Econometric Theory*, 28, 590–628, 2012.
- Lovell, C. A. K., "Applying Efficiency Measurement Techniques to the Measurement of Productivity Change," *Journal of Productivity Analysis*, 7, 329–40, 1996.
- Lee, K., and M. H. Pesaran and R. P. Smith, "Growth and Convergence in a Multi-Country Empirical Stochastic Solow Model," *Journal of Applied Econometrics*, 12, 357–92, 1997.
- Mankiw, N. G., D. Romer, and D. N. Weil, "A Contribution to the Empirics of Economic Growth," *Quarterly Journal of Economics*, 107, 407–37, 1992.
- Martin, W. and D. Mitra, "Productivity Growth and Convergence in Agriculture versus Manufacturing," *Economic Development and Cultural Change*, 49, 403–22, 2002.
- McMillan, M., D. Rodrik, and I. Verdusco-Gallo, "Globalization, Structural Change and Productivity Growth, with an Update on Africa," *World Development*, 63, 11–32, 2014.
- Moon, H. R. and M. Weidner, "Linear Regression for Panel with Unknown Number of Factors as Interactive Fixed Effects," *Econometrica*, 83, 1543–79, 2015.
- Mundlak, Y., R. Butzer, and D. F. Larson, "Heterogeneous Technology and Panel Data: The Case of the Agricultural Production Function," *Journal of Development Economics*, 99, 139–49, 2012.
- Nelson, C. R. and C. R. Plosser, "Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications," *Journal of Monetary Economics*, 10, 139–62, 1982.
- Pedroni, P., "Fully Modified OLS for Heterogeneous Cointegrated Panels," in B. H. Baltagi (ed.), *Nonstationary Panels, Panel Cointegration, and Dynamic Panels*, JAI Press, New York, 2000.
- _____, "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, 429–51, 2007.
- Pesaran, M. H., "Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure," *Econometrica*, 74, 967–1012, 2006.
- _____, "A Simple Panel Unit Root Test in the Presence of Cross-Section Dependence," *Journal of Applied Econometrics*, 22, 265–312, 2007.
- _____, "Testing Weak Cross-Sectional Dependence in Large Panels," *Econometric Reviews*, 34, 1089–117, 2015.
- Pesaran, M. H. and R. P. Smith, "Estimating Long-Run Relationships from Dynamic Heterogeneous Panels," *Journal of Econometrics*, 68, 79–113, 1995.
- Phillips, P. C. B. and B. E. Hansen, "Statistical Inference in Instrumental Variables Regression with $I(1)$ Processes," *Review of Economic Studies*, 57, 99–125, 1990.
- Rodrik, D., "Unconditional Convergence in Manufacturing," *Quarterly Journal of Economics*, 128, 165–204, 2013.
- Schmidt, P. and R. C. Sickles, "Production Frontiers and Panel Data," *Journal of Business & Economic Statistics*, 2, 367–74, 1984.
- Solow, R. M., "A Contribution to the Theory of Economic Growth," *Quarterly Journal of Economics*, 70, 65–94, 1956.
- _____, *Growth Theory—An Exposition*, Oxford University Press, Oxford, 1970.
- _____, "Unemployment: Getting the Questions Right," *Economica*, 53, S23–34, 1986.
- Swan, T. W., "Economic Growth and Capital Accumulation," *Economic Record*, 32, 334–61, 1956.
- Temple, J., "The New Growth Evidence," *Journal of Economic Literature*, 37, 112–56, 1999.
- UNIDO (United Nations Industrial Development Organization), "UNIDO Industrial Statistics 2004," Online database, UNIDO, Vienna, 2004.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's web site:

Appendix A.1: Data

Table A.1: Descriptive Statistics

Table A.2: Sample of Countries and Number of Observations

Figure A.1: Missing Observations

A.2: The Common Correlated Effects and Augmented Mean Group Estimators

A.3: Investigating Cross-Section Correlation Properties

Table A.3: Pesaran (2015) CD Test

A.4: Investigating Time Series Properties

Table A.4: Second Generation Panel Unit Root Tests

A.5: Robustness Check: Pedroni (2000) Group-Mean FMOLS Approach

Table A.5: Heterogeneous Models Using FMOLS (Average Estimates)

A.6: Diagnostic Testing and Robustness Checks

Table A.6: Gengenbach *et al.* (2016) Cointegration Tests

A.7: Parameter Heterogeneity Tests

Table A.7: Parameter Heterogeneity—Pooled Tests (Levels)

Table A.8: Parameter Heterogeneity—Pooled Tests (FD)

Table A.9: Parameter Heterogeneity—Swamy (1970) Tests

Table A.10: Parameter Heterogeneity—Wald Tests (Levels and FD)

Table A.11: Parameter Heterogeneity—*F*-Tests

A.8: The Growth Accounting Literature Appendix References