Appendix:

The Magnitude of the Task Ahead: Macro Implications of Heterogeneous Technology

Markus Eberhardt* and Francis Teal

Final version: July 4, 2018

Conditionally Accepted: *The Review of Income and Wealth*

Contents

*Correspondence: Markus Eberhardt, School of Economics, University of Nottingham, Sir Clive Granger Building, University Park, Nottingham NG2 2RD, England; email: markus.eberhardt@nottingham.ac.uk

A Data

Data for output, value-added, material inputs and investment in manufacturing, all in current local currency units (LCU), are taken from the UNIDO Industrial Statistics 2004 (UNIDO, 2004), where material inputs were derived as the difference between output and value-added. The labour data series is taken from the same source, which covers 1963-2002. The capital stocks are calculated from investment data which has been transformed into constant US$ following the 'perpetual inventory' method (Klenow and Rodriguez-Clare, 1997). In order to make data in monetary values internationally comparable, it is necessary to transform all values into a common unit of analysis. We follow the transformations suggested by Martin and Mitra (2002) and derive all values in 1990 US$, using current LCU and exchange rate data from UNIDO, and GDP deflators from the UN Common Statistics database (UN, 2005), for which data are available from 1970-2003. Since our model is for a small open economy, we prefer using a single market exchange rates (LCU-US$ exchange rate for 1990) to purchasing-power-parity (PPP) adjusted exchange rates, since the latter are more appropriate when non-traded services need to be accounted. The resulting panel is unbalanced and has gaps within individual country time-series — see Figure TA-1. We have a total of $n = 1,194$ observations from $N = 48$ countries (min $T = 11$, max $T = 33$, average $T = 24$). Descriptive statistics are presented in Table TA-1 and the sample country make-up is detailed in Table TA-2. We do not carry out any interpolation to fill these gaps and do not account for missing observations in any way. The preferred empirical specifications presented in the main section of our paper are based on heterogeneous parameter models, where arguably the unbalancedness (around 25% 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. As a robustness check we also produced a 'cleaned' dataset where we applied mechanical 'cleaning rules' in order to address the most serious issues of measurement error,[1] which created a sample of $n = 872$ observations for $N = 38$ countries. The empirical results for this sample are virtually the same to those from the larger sample (available on request).

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

Table TA-1: Descriptive statistics

VARIABLES IN LEVEL TERMS

Variable	obs	mean	median	std. dev.	min.	max.
levels						
value-added	1,194	5.47E+10	9.04E+09	1.78E+11	1.76E+07	1.50E+12
labour	1,194	1,469,186	502,214	2,924,524	5,552	1.97E+07
capital	1,194	1.32E+11	2.61E+10	3.12E+11	5.78E+07	2.27E+12
logs						
value-added	1,194	22.70	22.93	2.15	16.68	28.04
labour	1,194	12.92	13.13	1.79	8.62	16.79
capital	1,194	23.72	23.98	2.22	17.87	28.45
annual growth rate						
value-added	1,128	3.9%	3.5%	12.3%	-78.3%	92.7%
labour	1,128	1.7%	0.7%	8.1%	-38.8%	78.1%
capital	1,128	4.1%	3.1%	4.4%	-2.4%	47.8%

VARIABLES IN PER WORKER TERMS

Variable	obs	mean	median	std. dev.	min.	max.
levels						
value-added	1,194	76,932	45,865	72,843	2,007	346,064
capital	1,194	25,305	17,867	19,385	1,660	91,011
logs						
value-added	1,194	9.78	9.79	0.92	7.41	11.42
capital	1,194	10.80	10.73	1.00	7.60	12.75
annual growth rate						
value-added	1,128	2.2%	2.5%	10.8%	-90.3%	74.4%
capital	1,128	2.5%	2.5%	7.9%	-68.0%	45.4%

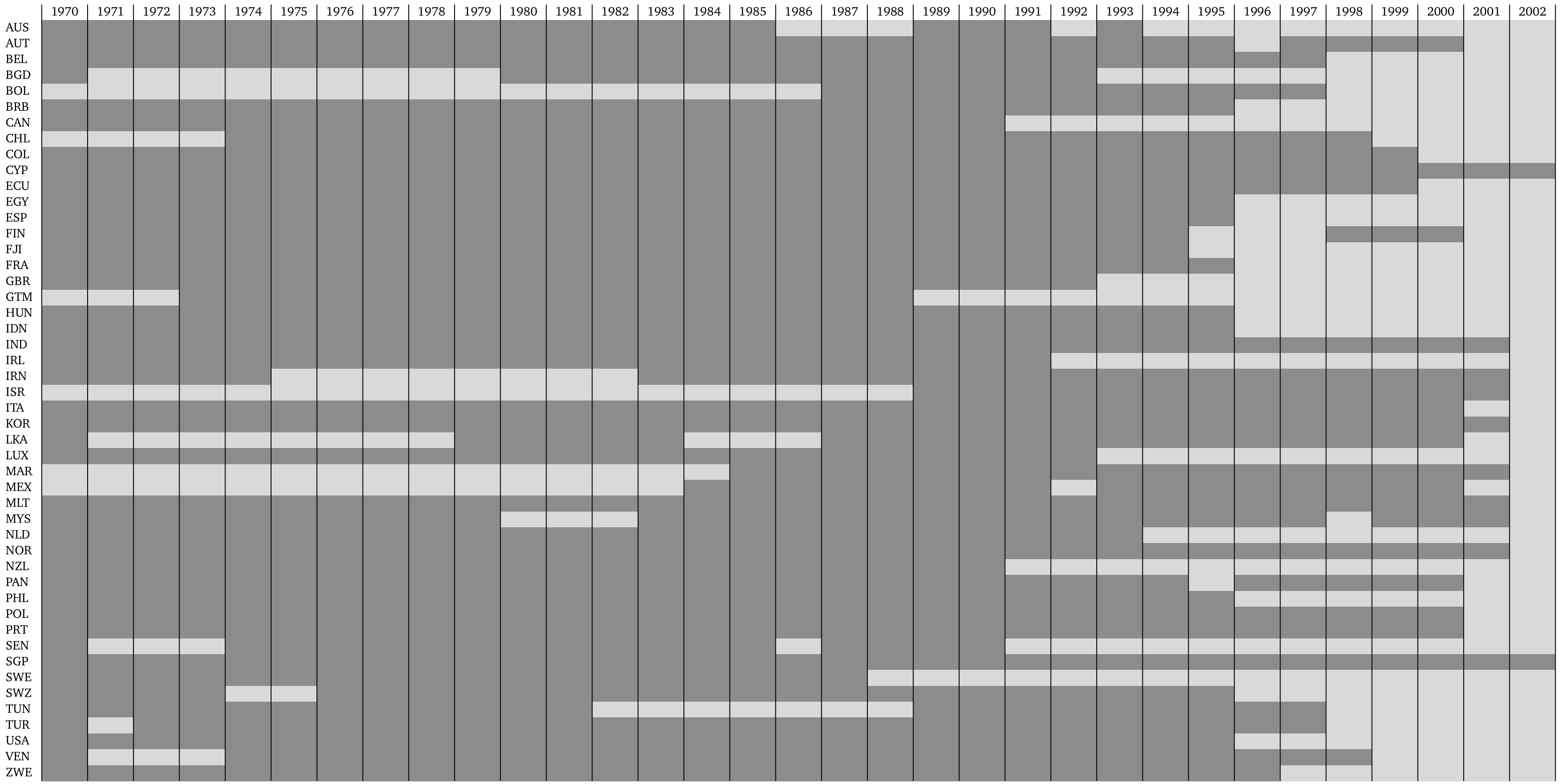
Notes: We report the descriptive statistics for value-added, labour and capital stock for $N = 48$ countries and $n = 1,194$ ($n = 1,128$) observations in the levels (growth) specification. Monetary values are in real US$ (base year 1990). Labour is in number of workers.

Table TA-2: Sample of countries and number of observations

Country	Code	*levels*	*FD*	$t=1$	$t=T$	Gaps
Australia	AUS	20	17	1970	1993	2
Austria	AUT	30	28	1970	2000	1
Belgium	BEL	28	27	1970	1997	-
Bangladesh[♯]	BGD	14	12	1970	1992	1
Bolivia[♮]	BOL	11	10	1987	1997	-
Barbados	BRB	26	25	1970	1995	-
Canada	CAN	21	20	1970	1990	
Chile	CHL	25	24	1974	1998	-
Colombia	COL	30	29	1970	1999	-
Cyprus	CYP	33	32	1970	2002	-
Ecuador	ECU	30	29	1970	1999	-
Egypt	EGY	26	25	1970	1995	-
Spain	ESP	26	25	1970	1995	-
Finland	FIN	28	26	1970	2000	1
Fiji	FJI	25	24	1970	1994	-
France	FRA	26	25	1970	1995	-
United Kingdom	GBR	23	22	1970	1992	-
Guatemala[♯]	GTM	16	15	1973	1988	-
Hungary	HUN	26	25	1970	1995	-
Indonesia	IDN	26	25	1970	1995	-
India	IND	32	31	1970	2001	-
Ireland	IRL	22	21	1970	1991	-
Iran	IRN	24	22	1970	2001	-
Israel[♮]	ISR	13	12	1989	2001	-
Italy	ITA	31	30	1970	2000	-
Korea	KOR	32	31	1970	2001	-
Sri Lanka	LKA	20	17	1970	2000	2
Luxembourg	LUX	23	22	1970	1992	-
Morocco[♯]	MAR	17	16	1985	2001	-
Mexico[♯]	MEX	16	14	1984	2000	1
Malta	MLT	32	31	1970	2001	-
Malaysia	MYS	28	25	1970	2001	2
Netherlands	NLD	24	23	1970	1993	-
Norway	NOR	32	31	1970	2001	-
New Zealand	NZL	21	20	1970	1990	-
Panama	PAN	30	28	1970	2000	1
Philippines	PHL	26	25	1970	1995	-
Poland	POL	31	30	1970	2000	-
Portugal	PRT	31	30	1970	2000	-
Senegal[♯]	SEN	17	14	1970	1990	2
Singapore	SGP	33	32	1970	2002	-
Sweden[♯]	SWE	18	17	1970	1987	-
Swaziland	SWZ	24	22	1970	1995	1
Tunisia	TUN	21	19	1970	1997	1
Turkey	TUR	27	25	1970	1997	1
United States	USA	26	25	1970	1995	-
Venezuela	VEN	26	24	1970	1998	1
Zimbabwe	ZWE	27	26	1970	1996	-
Countries		48	48			
Observations		1,194	1,128			

Notes: Countries highlighted in bold represent the sample used in the second set of GM-FMOLS regressions, Table 5 Panel B ($n = 644$, $N = 26$) of the main text. 'levels' and 'FD' refer to specifications in levels and first differences, respectively; $t = 1$ and $t = T$ report the start and end years of the country series; 'gaps' indicates the number of gaps in the data series.
♮ These countries had to be omitted to compute the Pesaran (2015) CD test for regression models in levels, due to the lack of overlap between the omitted series and the remainder of the sample. ♯ These countries had to be omitted in addition to those already identified to compute the CD test for regression models in first differences.

Figure TA-1: Missing observations



Notes: We indicate the data availability in our sample for each country over the 1970-2002 time horizon, where lighter shading signifies the observation as missing. In total 25% of observations are missing compared with a balanced panel.

B The Common Correlated Effects and Augmented Mean Group Estimators

The CCE estimators, developed by Pesaran (2006) and extended to nonstationary processes in Kapetanios, Pesaran and Yamagata (2011), augment the regression equation with cross-section averages of the dependent ($\bar{y}_t$) and independent variables ($\bar{\boldsymbol{x}}_t$) to account for the presence of unobserved common factors with heterogeneous impact.[2] For the Mean Group version (CMG), the individual country regression is specified as

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

whereupon the parameter estimates $\hat{\boldsymbol{b}}_i$ are averaged across countries akin to the Pesaran and Smith (1995) Mean Group estimator.[3] The pooled version (CCEP) is specified as

$$y_{it} = a_i + \boldsymbol{b}' \boldsymbol{x}_{it} + \sum_{j=1}^{N} c_{0i}(\bar{y}_t D_j) + \sum_{m=1}^{k}\sum_{j=1}^{N} \boldsymbol{c}_{mi}(\bar{x}_{mt} D_j) + e_{it} \tag{TA-2}$$

where the D_j represent country dummies.[4] The former is thus a simple extension to the Pesaran and Smith (1995) MG estimator based on on country-specific OLS regressions, whereas the latter is a standard fixed effects estimator augmented with additional regression terms.

In order to get an insight into the workings of this approach, consider the cross-section average of our model in equation (**??**): as the cross-section dimension N increases, given $\bar{\varepsilon}_t = 0$, we get

$$\bar{y}_t = \bar{\alpha} + \bar{\boldsymbol{\beta}}' \bar{\boldsymbol{x}}_t + \bar{\boldsymbol{\lambda}}' \boldsymbol{f}_t \qquad \Leftrightarrow \qquad \boldsymbol{f}_t = \bar{\lambda}^{-1}(\bar{y}_t - \bar{\alpha} - \bar{\boldsymbol{\beta}}' \bar{\boldsymbol{x}}_t) \tag{TA-3}$$

This simple derivation provides a powerful insight: working with cross-sectional means of y and $\boldsymbol{x}$ can account for the impact of unobserved common factors (TFP) in the production process.[5] Given the assumed heterogeneity in the impact of unobserved factors across countries

[2] Parts of the discussion in this section is taken from Eberhardt and Teal (2013).

[3] Although $\bar{y}_t$ and e_{it} are not independent their correlation goes to zero as N becomes large.

[4] Thus in the MG version we have N individual country regressions with $2k+2$ RHS variables and in the pooled version a single regression equation with $k+(k+2)N$ RHS variables.

[5] Most conservatively the CCE estimators require $\bar{\boldsymbol{\lambda}} \neq 0$, i.e. that the impact of each factor is on average non-zero (Coakley, Fuertes and Smith, 2006). Alternative scenarios (see Pesaran, 2006; Kapetanios et al., 2011) allow for this assumption to be dropped in certain situations but for the sake of generality we maintain it here.

($\boldsymbol{\lambda}_i$) 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 $\bar{\boldsymbol{x}}_t$, and thus implicitly on $\boldsymbol{f}_t$. Simulation studies (Pesaran, 2006; Coakley et al., 2006; Kapetanios et al., 2011; Pesaran and Tosetti, 2011) have shown that this approach performs well even when the cross-section dimension N is small, when variables are nonstationary, cointegrated or not, subject to structural breaks and/or in the presence of local spillovers and global/local business cycles. In the present study we implement two versions of the CCE estimators in the sector-level regressions: a standard form as described above; and a variant which includes the cross-section averages of the input and output variables in the own *as well as* the other sector. The latter specification allows for cross-section dependence *across* sectors, albeit at the cost of a reduction in degrees of freedom. It is conceivable that the evolution of the agricultural sector of developing countries influences that of the wider economy in general and the manufacturing sector in particular, such that this extension is sensible in the dual economy context.

Thus the Pesaran (2006) CCE estimators account for the presence of unobserved common factors by including cross-section averages of the dependent and independent variables in the regression equation and the estimates are obtained as averages of the individual country estimates, following the Pesaran and Smith (1995) MG approach. A related approach which we term the Augmented Mean Group (AMG) estimator (see Bond and Eberhardt, 2013, for details) accounts for cross-section dependence by inclusion of a 'common dynamic process' in the country regression. This process is extracted from the year dummy coefficients of a pooled regression in first differences (FD-OLS) and represents the levels-equivalent mean evolution of unobserved common factors across all countries. Provided the unobserved common factors form part of the country-specific cointegrating relation (Pedroni, 2007), the augmented country regression model encompasses the cointegrating relationship, which is allowed to differ across i.

$$\text{Stage (i)} \qquad \Delta y_{it} = \boldsymbol{b}'\Delta\boldsymbol{x}_{it} + \sum_{t=2}^{T} c_t \Delta D_t + e_{it} \quad \Rightarrow \hat{\boldsymbol{c}}_t \equiv \hat{\mu}_t^{\bullet} \tag{TA-4}$$

$$\text{Stage (ii)} \qquad y_{it} = a_i + \boldsymbol{b}_i'\boldsymbol{x}_{it} + c_i t + d_i \hat{\mu}_t^{\bullet} + e_{it} \quad \hat{\boldsymbol{b}}_{AMG} = N^{-1}\sum_i \hat{\boldsymbol{b}}_i \tag{TA-5}$$

Stage (i) represents a standard FD-OLS regression with $T-1$ year dummies in first differences, from which we collect the year dummy coefficients (labelled as $\hat{\mu}^{\bullet}_t$). This process is extracted from the pooled regression *in first differences* since nonstationary variables and unobservables are believed to bias the estimates in the pooled *levels* regressions. In stage (ii) $\hat{\mu}^{\bullet}_t$ is included in each of the N standard country regressions which also include a linear trend term to capture omitted idiosyncratic processes evolving in a linear fashion over time. Alternatively we can subtract $\hat{\mu}^{\bullet}_t$ from the dependent variable, which implies the common process is imposed on each country with unit coefficient. In either case country-specific estimates are averaged across countries following the MG approach. Based on the results of Monte Carlo simulations (Bond and Eberhardt, 2013) we posit that the inclusion of $\hat{\mu}^{\bullet}_t$ allows for the *separate* identification of $\boldsymbol{\beta}_i$ or $\mathbb{E}[\boldsymbol{\beta}_i]$ and the unobserved common factors driving output and inputs, like in the CCE approach. In analogy, we can use $\Delta\hat{\mu}^{\bullet}_t$ in the country equations in first differences and can augment the Swamy (1970) RCM estimator in a similar fashion to yield the Augmented Random Coefficient Model (ARCM) estimators in levels and first differences — results for the ARCM were very similar to those in the AMG and in the interest of space are therefore omitted in the empirical section. We also applied an alternative version of the estimator where the first stage allows for heterogeneous slopes across countries: results for the AMG second stage are next to identical to those presented in Table **??**.

The focus of the CCE estimators is the estimation of consistent $\hat{\boldsymbol{b}}$ and not the nature of the unobserved common factors or their factor loadings: we cannot obtain an explicit estimate for the unobserved factors $\boldsymbol{f}_t$ or the factor loadings $\boldsymbol{\lambda}_i$, since the *average* impact of the factors ($\bar{\boldsymbol{\lambda}}$) is unknown. Our augmented estimators use an *explicit* rather than implicit estimate for $\boldsymbol{f}_t$ from the pooled first stage regression. Compared with the CCE approach we can obtain a simple but economically meaningful construct from the AMG setup: the common dynamic process $\hat{\mu}^{\bullet}_t = h(\bar{\boldsymbol{\lambda}}\boldsymbol{f}_t)$ represents common TFP evolution over time, whereby *common* is defined either in the literal sense, or as the sample mean of country-specific TFP evolution. The country-specific coefficient on the common dynamic process, $\hat{d}_i$ from equation (TA-5), represents the implicit factor loading on common TFP.

Immediate concerns about this augmented estimator relate to the issue of second stage 'regressions with generated regressors' (Pagan, 1984). However, simulation results (Bond and Eberhardt, 2013) suggest that the average standard error of the AMG estimates is of similar magnitude to the empirical standard deviation. A theoretical explanation is provided in Bai and Ng (2008), who show that second stage standard errors need not be adjusted for first stage estimation uncertainty if $\sqrt{T}/N \to 0$, as is arguably the case here.

C Investigating Cross-Section Correlation Properties

Table TA-3: Pesaran (2015) CD test

Variable	**CD-test**	***p*-value**	**corr**	**abs(corr)**
logs				
Value-added pw	68.15	0.000	0.447	0.627
Capital pw	81.12	0.000	0.530	0.639
Labour	16.72	0.000	0.115	0.629
annual growth rates				
Value-added pw	6.86	0.000	0.050	0.218
Capital pw	12.26	0.000	0.085	0.214
Labour	17.69	0.000	0.123	0.214

Notes: We implement the Pesaran (2015) test for null hypothesis of weak cross-section dependence for our main regression variables — in the upper panel the variables are in logs, in the lower panel in first differences of logs (growth rates). Due to the unbalanced nature of the panel it is not possible to obtain correlation coefficients for all 48 countries — in the tests for log variables we are forced to drop two countries (BOL, $n = 11$; ISR, $n = 13$) and in the growth rate versions three countries (BOL, $n = 10$; ISR, $n = 12$; MAR, $n = 16$) in order to make the test feasible. Under the null the test statistic is normally distributed. 'corr' and 'abs(corr)' report the average and average absolute correlation coefficients for each variable.

D Investigating Time Series Properties

Table TA-4: Second generation panel unit root tests

Pesaran (2007) panel unit root tests — CIPS[♯]														
output			**value-added**			**labour**			**capital**			**materials**		
lags	Z[t-bar]	(p)	lags	Z[t-bar]	(p)	lags	Z[t-bar]	(p)	lags	Z[t-bar]	(p)	lags	Z[t-bar]	(p)
0	-1.22	(.11)	0	-1.85	(.03)	0	2.39	(.99)	0	5.11	(1.00)	0	0.29	(.62)
1	0.01	(.51)	1	0.06	(.52)	1	1.26	(.90)	1	3.79	(1.00)	1	0.89	(.81)
1.42	1.13	(.87)	1.96	3.54	(1.00)	1.48	3.74	(1.00)	1.50	4.55	(1.00)	1.65	3.68	(1.00)
2	2.65	(1.00)	2	2.30	(.99)	2	4.21	(1.00)	2	3.96	(1.00)	2	1.05	(.85)
3	7.04	(1.00)	3	3.59	(1.00)	3	4.76	(1.00)	3	7.64	(1.00)	3	4.21	(1.00)
output/worker			**VA/worker**						**capital/worker**			**materials/worker**		
lags	Z[t-bar]	(p)	lags	Z[t-bar]	(p)				lags	Z[t-bar]	(p)	lags	Z[t-bar]	(p)
0	-1.08	(.14)	0	-2.55	(.01)				0	1.92	(.97)	0	0.57	(.72)
1	2.91	(1.00)	1	-0.73	(.23)				1	1.33	(.91)	1	3.74	(1.00)
1.44	5.98	(1.00)	1.65	3.77	(1.00)				1.71	5.92	(1.00)	1.83	9.62	(1.00)
2	5.02	(1.00)	2	2.37	(.99)				2	4.60	(1.00)	2	5.96	(1.00)
3	8.73	(1.00)	3	5.48	(1.00)				3	7.34	(1.00)	3	8.08	(1.00)

Notes: The CIPS test maintains the H_0 of a unit root process; augmentation with lags as indicated. ♯ All variables are in logs. In the third row for each variable in the lower panel we present the CIPS test statistic for 'ideal' lag augmentation of the underlying ADF regression (based on Akaike information criteria); the value for lags reported here is the *average* across countries.

E Robustness Check: Pedroni (2000) Group-Mean FMOLS approach

Table TA-5: Heterogeneous Models using FMOLS (average estimates)

PANEL A: FULL SAMPLE (N=48)

	[1]	[2]	[3]	[4]	[5]
estimator: ‘ ’-**FMOLS**	**GM**	**AMG**	**AMG**	**CMG**	**CMG**
dependent variable	ly	ly-$\hat{\mu}^{\bullet}_{t}$	ly	ly	ly
log capital pw	0.1663	0.2659	0.2937	0.5544	0.3042
	[0.084]	[0.080]**	[0.092]**	[0.069]**	[0.091]**
common process			0.8977		
			[0.257]**		
country trends	0.0171	0.0004	0.0014		0.0108
	[0.003]**	[0.003]	[0.005]		[0.004]**
Panel-t statistics, diagnostics					
capital pw	18.29	14.73	15.36	40.59	15.88
trends	24.94	18.93	12.71		20.70
RMSE	.099	.096	.090	.103	.088

PANEL B: I(1) SAMPLE (N=26)

	[1]	[2]	[3]	[4]	[5]
estimator: ‘ ’-**FMOLS**	**GM**	**AMG**	**AMG**	**CMG**	**CMG**
dependent variable	ly	ly-$\hat{\mu}^{\bullet}_{t}$	ly	ly	ly
log capital pw	0.0816	0.2675	0.2784	0.5528	0.2485
	[0.064]	[0.065]**	[0.090]**	[0.075]**	[0.079]**
common process			0.8034		
			[0.174]**		
country trends	0.0179	-0.0012	0.0019		0.0108
	[0.003]**	[0.003]	[0.005]		[0.004]*
Panel-t statistics, diagnostics					
capital pw	11.45	10.37	9.97	34.96	10.16
trends	23.28	14.63	10.56		17.10
RMSE	.071	.068	.065	.080	.062

Notes: The results in [1] are for the Pedroni (2000) Group-Mean FMOLS estimator; the results in the remaining columns allow for cross-section dependence using either $\hat{\mu}^{\bullet}_{t}$ or cross-section averages in the FMOLS country regressions. In all cases the estimates presented are the unweighted means of the FMOLS country estimates. Intercept estimates as well as average estimates on cross-section averages in [4] and [5] of both panels are omitted to save space. Values in brackets are absolute standard errors following Pesaran and Smith (1995). Panel-t statistics are computed as $N^{-1/2}\sum_i t_i$ where t_i is the country-specific t-ratio for the estimate from the FM-OLS model. Panel B uses observations from only those countries for which variables were determined to be nonstationary (via country-specific ADF and KPSS testing). All models estimated in RATS.

F Diagnostic testing and robustness checks

We first investigated the density estimates for country-specific technology parameters estimated in the levels regressions using standard kernel methods with automatic bandwidth selection. The plots indicate that the distribution of these parameter estimates is symmetric around their respective means and roughly Gaussian, such that no significant outliers drive our results. We further carried out a number of residual diagnostic tests other than the analysis of stationarity and cross-section dependence. A cautious conclusion from these procedures would be that we are more confident about the country regression residuals possessing desirable properties (normality, homoskedasticity) than we are for their pooled counterparts (all results available on request).

Table TA-6: Gengenbach et al. (2016) cointegration tests

ECM-BASED COINTEGRATION TEST

no intercept	**AIC**		**BIC**		**10%**	**5%**	**1%**
$\bar{T}^*_{\alpha_y}$ (truncated)	-2.58	*	-2.75	**	-2.48	-2.55	-2.67
$\bar{T}^*_{\theta}$ (truncated)	25.66	**	25.61	**	12.10	12.43	13.07
avg. lag length	2.0		1.7				
intercept	**AIC**		**BIC**		**10%**	**5%**	**1%**
$\bar{T}^*_{\alpha_y}$ (truncated)	-2.63		-2.78		-2.86	-2.92	-3.03
$\bar{T}^*_{\theta}$ (truncated)	17.04	**	17.12	**	14.08	14.42	15.04
avg. lag length	2.3		1.7				
intercept, trend	**AIC**		**BIC**		**10%**	**5%**	**1%**
$\bar{T}^*_{\alpha_y}$ (truncated)	-2.44		-2.61		-3.23	-3.28	-3.39
$\bar{T}^*_{\theta}$ (truncated)	12.54		13.13		16.23	16.59	17.31
avg. lag length	2.1		1.8				

Notes: The $\bar{T}^*_{\alpha_y}$ and $\bar{T}^*_{\theta}$ statistics are averages of the N t-ratios and F-statistics from the country ECM regressions, where extreme t-ratios/F-statistics have been replaced by bounds (truncated; we used $\varepsilon = .000001$) following the strategy devised in Gengenbach et al. (2016). This paper also provides simulated critical values we present here ($N = 50$) in the supplementary material. Both test statistics are one-sided: for the $\bar{T}^*_{\alpha_y}$ large negative values lead to rejection of the null, whereas for the $\bar{T}^*_{\theta}$ it is large positive values which lead to rejection. H_0 in all cases: no error correction, i.e. no cointegration; lag-length p_i determined using AIC or BIC as indicated; $p_x = 2$ (capital per worker and $\hat{\mu}^\bullet_t$).

Cointegration tests are commonly carried out as a *pre*-estimation testing procedure, however we have delayed these until *after* estimation since we hypothesise that unobservable TFP forms part of the cointegrating vector. Employing our first stage estimate $\hat{\mu}^\bullet_t$ we carried out a cointegration testing procedure based on the error correction model representation, first introduced by Westerlund (2007) and refined by Gengenbach et al. (2016). Results in Table TA-6 imply that there are good grounds to suggest that value-added per worker, capital per worker and our estimate for TFP are heterogeneously cointegrated.

G Parameter heterogeneity tests

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

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

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

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

$$H_{het} = \pi_b\,k_{it} + \pi_c\,m_{it} + \pi_d\,t\,(+\sum_i \pi_{e,i}) \tag{TA-6}$$

The rationale behind this test is as follows: if factor input parameters were truly heterogeneous across countries, we would expect the pooled regression to produce statistically significant coefficients ($\pi_j \neq 0$). Results are presented in Tables TA-7 and TA-8.

As can be seen the levels regressions imply that capital parameter homogeneity is rejected, while the materials coefficients are more likely to be homogeneous (in the gross output specification). In the VA-specification capital parameter homogeneity is rejected in all models. In contrast the tests for the first difference specifications on the whole do not provide much evidence for heterogeneity, with all covariates insignificant with the exception of the case of CMG in first differences. Note that the kernel densities for the technology parameters underlying the above heterogeneity tests do not differ considerably between levels and FD specifications (FD densities not reported). This stark difference is therefore likely to be driven by the impact of nonstationarity on the test.

Table TA-7: Parameter Heterogeneity — Pooled Tests (levels)

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

Notes: All variables are in logs. Values in brackets are absolute t-statistics. The models underlying the construction of H_{het} are presented in Table **??** in the main text. We indicate statistical significance at the 5% and 1% level by * and ** respectively.

Table TA-8: Parameter Heterogeneity — Pooled Tests (FD)

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
estimator	**ΔMG**	**ΔRCM**	**ΔAMG(i)**	**ΔAMG(ii)**	**ΔCMG(i)**	**ΔCMG(ii)**	**ΔARCM(i)**	**ΔARCM(ii)**
dependent variable$^\sharp$	H_{het}	H_{het}	H_{het}	H_{het}	H_{het}	H_{het}	H_{het}	H_{het}
regressors								
capital pw	0.0989	0.0546	0.0157	0.0069	-0.0791	-0.0202	-0.0162	-0.0106
	[1.11]	[0.61]	[0.18]	[0.08]	[0.88]	[0.22]	[0.18]	[0.12]
drift terms	only 2 sign.	only 2 sign.	only 2 sign.	only 2 sign.	only 2 sign.	only 2 sign.	only 2 sign.	only 2 sign.
obs	1,128	1,128	1,128	1,128	1,128	1,128	1,128	1,128

Notes: See Table TA-7 for details. The results for the country regressions in first difference tested here for parameter heterogeneity are presented in Table **??** in the main text.

Secondly, we report the Swamy (1970) $\hat{S}$ statistic from the gross output and VA regressions in levels and first differences in Table TA-9.[6] For a detailed discussion of this test see Pesaran and Yamagata (2008). Note that the test for the equation in levels is testing heterogeneity of *all* parameters, including the intercepts; since the assumption of heterogeneous TFP levels is rather uncontroversial, this test does not adquately address our interest in the homogeneity of

[6]The levels and FD tests are taken from the regressions in Table **??** of the paper.

technology parameters. We therefore also provide a test for the levels specification where the intercept terms have been dispensed with via transformation of the data into mean-deviations. Estimates for this specification are of course identical to those of the untransformed levels equation.

Table TA-9: Parameter Heterogeneity — Swamy (1970) Tests

	RCM			**ARCM**		
Specification	(a)	(a)'	(b)	(b)'	(c)	(c)'
	Levels	MD	Levels	MD	Levels	MD
	51,123.0 (.00)	1,531.6 (.00)	62,499.9 (.00)	1,440.4 (.00)	69,157.0 (.00)	1726.7 (.00)
	FD		FD		FD	
	191.1 (.00)		153.5 (.00)		258.7 (.00)	

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

The Swamy $\hat{S}$ test rejects parameter heterogeneity for all specifications tested. In general, this test was developed for panels where N is large relative to T. Using Monte Carlo experiments, Pesaran and Yamagata (2008) show that in case of a panel of $T = 30, N = 50$ the test has power but tends to over-reject — a tendency which becomes worse with the number of parameters included in the model.[7] Further, as Pedroni (2007) points out, the Swamy-based tests are not designed for nonstationary panel data.

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

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

where $\hat{\theta}_i$ is the parameter coefficient from the country regression, $\bar{\theta}$ is the unweighted average parameter estimate and $\mathbb{V}ar(\theta_i)$ its variance across all countries. If parameters are similar across countries, the test statistic will be small, whereas if parameters are heterogeneous W_θ will be larger. The validity of this test depends on T being moderate to large. The null for this test is that *all* countries have the same parameter value. Table TA-10 presents the summed Wald statistics for the entire sample, as well as an indication of the share of country-specific tests rejecting the null of equality between country estimate and full sample mean estimate (for both the levels and FD specifications).

The Wald tests reject homogeneity for the factor parameters derived from the levels models

[7] The adjusted Swamy statistic $\tilde{\Delta}_{adj}$ developed by the same authors, although appropriately sized, suffers from low power in a sample of $T = 30, N = 50$, in particular if errors are non-normal.

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

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

$$F = \left(\frac{RSS_{hom} - RSS_{het}}{RSS_{het}}\right)\left(\frac{df_D}{df_N}\right)$$

$$F \sim F(df_N, df_D) \qquad df_N = k \times (N-1) \qquad df_D = N(\bar{T} - k - 1)$$

where k is the number of parameters in each country-regression and RSS_{hom} and RSS_{het} are the sums of the squared residuals of the homogeneous and heterogeneous models respectively — in the former case the mean coefficient estimates are imposed. This tests the full parameter heterogeneity versus the full homogeneity case. We do not compute F-tests for the CMG models, as the parameters on the period-average are not meant to be identical.

The F tests are valid for fixed N, when the regressors are strictly exogenous and the error variances are homoskedastic (Pesaran and Yamagata, 2008).[8] All of the test results presented in Table TA-11 reject parameter homogeneity for the factor input variables at the 1% level of significance. It is intuitive why the test statistics may emphatically reject the null: if the homogeneity restriction is incorrect, the country regressions do not cointegrate under the null, such that the regression errors will be nonstationary. As a result the F-statistic will quickly diverge and reject the null (Pedroni, 2007).

Like in the Swamy $\hat{S}$ Test we are faced with the problem that the tests evaluate the *full* regression model for the null of parameter homogeneity, which is not sensible in the levels regression case since heterogeneous intercepts are commonly accepted in the literature. In order to by-

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

Table TA-10: Parameter Heterogeneity — Wald Tests (levels and FD)

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

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

Notes: Analysis for 1,194 observations (1,128 in the first difference specifications) in 48 countries. The models underlying the construction of the Wald statistics are presented in Table **??** in the main text. In the full sample tests $W_\theta \sim \chi^2(48)$, with 5% and 1% critical values 65.17 and 73.70 respectively ($W_\theta = \sum_i W_{\theta,i}$); for country-specific tests ($W_{\theta,i}$) we apply the 10% critical value of 2.7. The null hypothesis in all cases is parameter homogeneity. For W_θ we indicate statistical significance at the 5% and 1% level by * and ** respectively.

Table TA-11: Parameter Heterogeneity — F-Tests

	[1]	[2]	[3]	[4]	[5]	[6]
estimator	**MG**	**RCM**	**AMG(i)**	**AMG(ii)**	**ARCM(i)**	**ARCM(ii)**
levels						
F	413.7 (.00)	334.1 (.00)	339.9 (.00)	279.4 (.00)	287.5 (.00)	232.3 (.00)
distr	$F(141,1002)$	$F(141,1002)$	$F(141,1002)$	$F(188,954)$	$F(141,1002)$	$F(188,954)$
first differences						
F	2.0 (.00)	1.5 (.00)	2.6 (.00)	2.4 (.00)	1.6 (.00)	1.8 (.00)
distr	$F(94,950)$	$F(94,950)$	$F(94,950)$	$F(141,902)$	$F(94,950)$	$F(141,902)$

Notes: See above text for construction of the Panel F statistic. The models underlying the construction of the F statistics are presented in Table **??** in the main text. The null hypothesis in all cases is parameter homogeneity.

pass this problem we also computed F-statistic for the levels MG and Augmented MG cases where the intercepts have been dispensed with by taking all variables in deviations from the country-mean — all of these reject parameter homogeneity at the 1% level.

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

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

H The growth accounting literature

Empirical studies using TFP growth accounting have a long tradition since Abramowitz (1956), Kendrick (1956) — who coined the term Total Factor Productivity — and Solow (1957). Under standard assumptions value-added growth is decomposed into contributions of inputs and TFP growth, imposing a common capital coefficient β^K

$$\Delta y_{it} = \beta^K \Delta k_{it} + \Delta \mathrm{TFP}_{it} \qquad \Leftrightarrow \qquad \Delta \mathrm{TFP}_{it} = \Delta y_{it} - \beta^K \Delta k_{it} \tag{TA-7}$$

The simple computation as well as function-free nature of this approach represent considerable strengths and in part explain its popularity. The accounted TFP growth is in theory disembodied, Hicks-neutral exogenous technical progress. In practice however, one needs to keep in mind that TFP is a residual, such that it represents a 'catch-all' for output growth that cannot be explained by factor accumulation. Thus if TFP growth is recovered via growth accounting its coefficient "need not represent only technological change and may not represent technological change at all" (Baier, Dwyer and Tamura, 2006, p.27) since measurement error, violations of assumptions and 'incorrect' variable construction can cause considerable bias. Any measurement error in output, labour or capital enters the residual term and thus TFP growth. This may have considerable impact since factor inputs need to account correctly for embodied technical change, which given the difficulty of distinguishing between embodied and disembodied technical progress seems impossible (Lipsey and Carlaw, 2001; Baier et al., 2006). The method further cannot disentangle the underlying endogeneity problem, such that inputs cannot be argued to *cause* output (Gollin, 2010). Violations of the assumptions of constant returns to scale, and private and social marginal product equality can add to further accounting error (Barro, 1999) — conceptually, the simple accounting framework for instance runs counter to the substantial empirical literature on knowledge spillovers (Eberhardt, Helmers and Strauss, 2013). As a result, it is now widely recognised that TFP growth derived from growth accounting "does not really measure technical change" (Caselli, 2008), nevertheless most empirical work takes findings of substantial TFP growth as a very positive and meaningful insight into the growth process.

Appendix References

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

Bai, Jushan and Ng, Serena (2008). "Large Dimensional Factor Analysis." *Foundations and Trends in Econometrics*, Vol. 3(2): 89–163.

Baier, Scott L., Dwyer, Gerald P. and Tamura, Robert (2006). "How Important are Capital and Total Factor Productivity for Economic Growth?" *Economic Inquiry*, Vol. 44(1): 23–49.

Barro, Robert J. (1999). "Notes on Growth Accounting." *Journal of Economic Growth*, Vol. 4(2): 119–137.

Bond, Stephen R. and Eberhardt, Markus (2013). "Accounting for unobserved heterogeneity in panel time series models." Unpublished mimeo, November.

Canning, David and Pedroni, Peter (2008). "Infrastructure, Long-Run Economic Growth And Causality Tests For Cointegrated Panels." *Manchester School*, Vol. 76(5): 504–527.

Caselli, Francesco (2008). "Growth Accounting." In: Steven N. Durlauf and Lawrence E. Blume (Editors), "The New Palgrave Dictionary of Economics Online," (Palgrave Macmillan), 2nd edition.

Coakley, Jerry, Fuertes, Ana-Maria and Smith, Ron P. (2006). "Unobserved heterogeneity in panel time series models." *Computational Statistics & Data Analysis*, Vol. 50(9): 2361–2380.

Eberhardt, Markus, Helmers, Christian and Strauss, Hubert (2013). "Do spillovers matter when estimating private returns to R&D?" *The Review of Economics and Statistics*, Vol. 95(2): 436–448.

Eberhardt, Markus and Teal, Francis (2013). "Structural Change and Cross-Country Growth Empirics." *World Bank Economic Review*, Vol. 27: 229–71.

Gengenbach, Christian, Westerlund, Joakim and Urbain, Jean-Pierre (2016). "Error Correction Testing in Panels with Common Stochastic Trends." *Journal of Applied Econometrics*, Vol. 31: 982–1004.

Gollin, Doug (2010). "Agricultural productivity and economic growth." In: P Pingali and R Evenson (Editors), "Handbook of Agricultural Economics," Vol. 18, chapter 71 (Elsevier North-Holland), pp. 3825–3866.

Kapetanios, George, Pesaran, M. Hashem and Yamagata, Takashi (2011). "Panels with Non-stationary Multifactor Error Structures." *Journal of Econometrics*, Vol. 160(2): 326–348.

Kendrick, John W. (1956). "Productivity Trends: Capital and Labor." *The Review of Economics and Statistics*, Vol. 38(3): 248–257.

Klenow, Peter J. and Rodriguez-Clare, Andres (1997). "Economic growth: A review essay." *Journal of Monetary Economics*, Vol. 40(3): 597–617.

Lipsey, Richard G. and Carlaw, Kenneth I. (2001). "What does Total Factor Productivity measure?" Unpublished manuscript.

Martin, Will and Mitra, Devashish (2002). "Productivity Growth and Convergence in Agriculture versus Manufacturing." *Economic Development and Cultural Change*, Vol. 49(2): 403–422.

Pagan, Adrian (1984). "Econometric Issues in the Analysis of Regressions with Generated Regressors." *International Economic Review*, Vol. 25(1): 221–47.

Pedroni, Peter (2000). "Fully modified OLS for heterogeneous cointegrated panels." In: Badi H. Baltagi (Editor), "Nonstationary panels, cointegration in panels and dynamic panels," (Amsterdam: Elsevier).

Pedroni, Peter (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*, Vol. 22(2): 429–451.

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

Pesaran, M. Hashem (2015). "Testing Weak Cross-Sectional Dependence in Large Panels." *Econometric Reviews*, Vol. 34(6-10): 1089–1117.

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

Pesaran, M. Hashem and Tosetti, Elisa (2011). "Large Panels with Common Factors and Spatial Correlations." *Journal of Econometrics*, Vol. 161(2): 182–202.

Pesaran, M. Hashem and Yamagata, Takashi (2008). "Testing Slope Homogeneity In Large Panels." *Journal of Econometrics*, Vol. 142(1): 50–93.

Phillips, Peter and Sul, Donggyu (2003). "Dynamic panel estimation and homogeneity testing under cross section dependence." *Econometrics Journal*, Vol. 6(1): 217–259.

Solow, Robert. M. (1957). "Technical Change and the Aggregate Production Function." *Review of Economics and Statistics*, Vol. 39(3): 312–20.

Swamy, P. A. V. B. (1970). "Efficient Inference in a Random Coefficient Regression Model." *Econometrica*, Vol. 38(2): 311–23.

UN (2005). "UN Common Statistics 2005." Online database, New York: UN.

UNIDO (2004). "UNIDO Industrial Statistics 2004." Online database, Vienna: UNIDO, united nations Industrial development organisation.

Westerlund, Joakim (2007). “Testing for Error Correction in Panel Data.” *Oxford Bulletin of Economics and Statistics*, Vol. 69(6): 709–748.