Supplementary Appendices to: Weak and Strong Cross-Sectional Dependence: a Panel Data Analysis of International Technology Diffusion

CEM ERTUR

Univ. Orléans, LEO, UMR 7322 CNRS, Faculté de Droit d'Economie et de Gestion, 11 Rue de Blois - F-45067 Orléans-France

(email: cem.ertur@univ-orleans.fr)

ANTONIO MUSOLESI

Department of Economics and Management, University of Ferrara –Italy and SEEDS

(email: antonio.musolesi@unife.it)

Appendix A: Panel Unit Root Tests

We focus on second-generation tests allowing for cross-sectional dependence. Recent work has demonstrated the importance of accounting for cross-sectional dependence when testing the unit root hypothesis. Pesaran's (2007) simulations show that tests assuming cross-sectional independence tend to over-reject the null hypothesis if cross-sectional dependence is present. Baltagi et al. (2007) find that when spatial autoregression is present, first-generation tests become oversized, but the tests explicitly allowing for cross-sectional dependence yield a lower frequency of type I errors. As Pesaran (2007) notes, subtracting the cross-sectional averages from the series before applying the panel unit root test can mitigate the impact of cross-sectional dependence even if cross-sectional demeaning would not generally be effective in conditions under which the pairwise cross-sectional error covariances differ across individuals. Moreover, while weak cross-sectional dependence can be addressed with a simple correction of the tests, the presence of strong cross-sectional dependence is more problematic, causing the test statistics to be divergent (Westerlund and Breitung, 2009). As we clearly documented the presence of strong cross-sectional dependence in our data, it is crucial to apply the so-called second generation unit root tests (Bai and Ng, 2002, 2004; Moon and Perron, 2004; Pesaran, 2007, Pesaran et al. 2013).

Bai and Ng (2004) propose decomposing the panel into deterministic, common and idiosyncratic components as follows:

$$y_{it} = D_{it} + \zeta_i^{'}\mathbf{f}_t + v_{it},$$

where D_{it} is the deterministic component with individual effects and eventually individual trends, $\zeta_i^{'}\mathbf{f}_t$ the common component, with r unobserved factors, and v_{it} the idiosyncratic component. Such a decomposition allows us to consider factors as objects of interest and to determine not only whether the data are stationary but also whether the eventual nonstationarity derives from a nonstationary common component, a nonstationary idiosyncratic component or the nonstationarity of both components. More precisely, Bai and Ng (2004) also assume the following:

$$(\mathbf{I} - L)\,\mathbf{f}_t = \mathbf{C}\,(L)\,u_t,$$
$$(1 - \rho_i)\,v_{it} = \mathbf{B}_i\,(L)\,\epsilon_{it}$$

where $\mathbf{C}\,(L) = \sum_{j=0}^{\infty} C_j L^j$ and $\mathbf{B}_i\,(L) = \sum_{j=0}^{\infty} B_{ij} L^j$. The idiosyncratic component is $I(1)$ if $\rho_i = 1$ and is stationary if $\rho_i < 1$. There are r_0 stationary common factors and r_1 common stochastic trends, such that $r_0 + r_1 = r$, the total number of factors; the rank of $\mathbf{C}\,(1)$ is r_1. The goal of Bai and Ng (2004) is to determine r_1 and test whether $\rho_i = 1$ when neither $\mathbf{f}_t$ nor v_{it} is observed. This approach is known as the PANIC (panel analysis of nonstationarity in idiosyncratic and common components) approach. A preliminary issue that arises involves determining how many common factors, r, are necessary to capture the existing cross-sectional dependence. To this end, we employ the information criteria suggested by Bai and Ng (2002). They are constructed in a similar spirit as the AIC and BIC criteria for time series, involving a trade-off between some measure of fit and a penalty for complexity. As the number of factors increases, the fit must improve, but the penalty also increases. We compute all criteria, but following the relevant literature (Bai and Ng, 2002; Moon and Perron, 2007; Hurlin, 2010), we pay particular attention to the IC2 and BIC3 criteria that are expected to minimize the risk of overestimating the number of factors.[1] These criteria are applied to factors estimated by principal components on first differences (Bai and Ng, 2004). Recent literature suggests that a small number of unobserved common factors is sufficient to explain most of the variations in many macroeconomic variables (see, e.g., Stock and Watson, 2002; Pesaran et al., 2013; Moon and Perron, 2007 and Hurlin, 2010). We begin our analysis by following this literature and apply the above-discussed criteria by imposing the condition that the maximum number of factors is 6, as in Pesaran (2007). Note that Stock and Watson (2002) also find that 6 factors account for much of the variance in their time series. Another practical reason for such a choice is that Bai and Ng (2004) report the critical values of the tests used determine how many of these factors are nonstationary up to 6 factors. In nearly all cases, the criteria suggest that the number of unobserved factors, r, equals the maximum number we allowed. This is the same result as in Pesaran (2007) and Pesaran et al. (2013), and it is not completely surprising given our sample sizes (see also Gutierrez, 2006). This suggests that the number of factors could be even higher than 6. However, given the possibility that the criteria overestimate the number of factors and the number of observations available, we provide our main results without allowing the maximum number of factors to be greater than 6 (**Table A1**). As a robustness check (available upon request), we also perform the procedure of Bai and Ng (2004) to test the stationarity in the common component, to identify the number of nonstationary common factors (if they exist) and to test the stationarity in the idiosyncratic component for different values of r in the range $1 - 20$.

Notably, the nonstationarity of the idiosyncratic components can be tested without knowing whether the factors are stationary, and vice versa. All that we need to know is the total number of factors, r. For this reason, given the considerable uncertainty that surrounds the number of factors, we perform the tests for a large range of possible values of r. To test the nonstationarity of idiosyncratic components, Bai and Ng (2004) proceed by pooling individual $ADF\ t$ statistics obtained on defactored components. Pooling, however, requires cross-sectional independence of the idiosyncratic components. As the idiosyncratic components in a factor model can be only weakly correlated across units, by construction, while the factors involve strong correlation, the pooled tests based on defactored components appear likely to satisfy the required cross-sectional independence assumption. The two Fisher-type statistics proposed by Bai and Ng (2004), denoted P_e^c and Z_e^c, provide strong evidence for rejection of the null hypothesis of nonstationarity of the idiosyncratic components for all variables. For domestic R&D, foreign R&D and human capital, the null is rejected irrespective of the value of r, the total number of common factors, whereas for TFP, the null is rejected 19 out of 20 times. Note, however, that the pooled test for the idiosyncratic component is valid only for the intercept case. This because only in such a case the limiting distribution of the individual ADF tests obtained on defactored components is a DF-type distribution (see also Gengenbach et al., 2009, p. 121). The two Fisher-type statistics are consequently used only for the intercept case. The rejection of the nonstationarity of the idiosyncratic component does not imply that the series are stationary, as some of the common factors may be nonstationary. We have already attempted to determine the total number of factors

[1]According to Bai and Ng (2002), the IC2 selects the true number of factors and dominates the other criteria. The BIC3 has been shown (Bai and Ng, 2002) to perform better than the others when both T and N are small and are roughly the same size; this result holds even if the BIC3 does not satisfy the conditions for consistency when either N or T dominates the other exponentially.

using information criteria on first differences, and the next task is thus to determine how many of these factors are nonstationary. For this purpose, we follow Bai and Ng (2004) and proceed as follows. For $r = 1$, we use a standard ADF test; its rejection indicates that the unique common factor is stationary. For $r > 1$, we consider the MQ_f and MQ_c statistics. The limiting distributions of these statistics are nonstandard, and critical values are reported in Bai and Ng (2004) for up to 6 factors. The results provide a very clear picture: for all variables, regardless of the test used, the number of nonstationary common factors, r_1, is almost always equal to the total number of common factors, r. This is the same result obtained by Hurlin (2010). The application of the PANIC approach by Bai and Ng (2004) thus suggests that the variables are nonstationary and that this property is the result of multiple nonstationary common factors combined with stationary idiosyncratic components.

Next, to further investigate the order of integration of the variables of interest, we follow Moon and Perron (2004), who also allow for r unobserved common factors but propose expressing the panel in an autoregressive form of the following type:

$$\begin{aligned} y_{it} &= D_{it} + y_{it}^0, \\ y_{it}^0 &= \rho_i y_{it-1}^0 + u_{it} \ , \\ u_{it} &= \zeta_i^{'} \mathbf{f}_t + v_{it}. \end{aligned}$$

As in Bai and Ng (2004), data are first defactored, and panel unit root test statistics based on defactored data are then proposed. Moon and Perron (2004), however, consider the factors to be nuisance parameters, and the unit root test is only based on the estimated idiosyncratic components. This is a relevant difference with respect to Bai and Ng (2004). The proposed test statistic uses defactored data obtained by projecting the data onto the space orthogonal to the factor loadings. The authors derive two modified t-statistics – denoted t_a and t_b – which have a Gaussian distribution under the null hypothesis, and they propose the implementation of feasible statistics – t_a^* and t_b^* – based on the estimation of long-term variances. To assess the robustness of the results to the choice of the kernel function used to estimate long-term variances, we compute t_a^* and t_b^* with both quadratic spectral and Bartlett kernels. In Moon and Perron (2004), the abovementioned information criteria to detect the number of common factors are applied to the residuals (rather than to the first differences). Such criteria provide the same results we obtained using PANIC and tend to select the maximum number of factors allowed (i.e., 6). To obtain results fully comparable with PANIC, we also consider a model with an intercept but without a trend in the deterministic component. For the model with an intercept, these tests strongly reject the unit root hypothesis, as in PANIC. The model with a trend is generally of the same direction: the null is rejected for TFP, domestic R&D and human capital, but not for foreign R&D. As with PANIC, we then perform the tests for all possible values of r in the range 1-20. For the model with an intercept, these tests strongly reject the unit root hypothesis in all cases. The model with an intercept and a trend provides a similar result: for TFP, domestic R&D and human capital, in almost all cases, these tests strongly reject the unit root hypothesis of the idiosyncratic components. The main difference concerns foreign R&D, for which the tests generally do not reject the null. Detailed results obtained using the approach by Moon and Perron (2004) are available upon request.

Overall, the two approaches proposed by Bai and Ng (2004) and Moon and Perron (2004) provide a robust picture suggesting that the idiosyncratic component of the variables under investigation is stationary. The PANIC approach also provides strong evidence in favor of nonstationary common factors.

Finally, we implement the tests proposed by Pesaran (2007) and Pesaran et al. (2013). Instead of basing the unit roots tests on deviations from the estimated factors, they augment standard ADF regressions with cross-sectional averages. In the case of a single unobserved common factor, Pesaran (2007) suggests augmenting the standard (individual) ADF regression with the cross-sectional average of first differences ($\Delta\bar{y}_t = N^{-1}\sum_{i=1}^{N}\Delta y_{it} = \overline{y}_t - \overline{y}_{t-1}$) and lagged levels ($\bar{y}_{t-1}$) of the individual series, which are $\sqrt{N}$- consistent estimators for the rescaled factors $\bar{\zeta}f$ and $\bar{\zeta}\sum_{j=0}^{t-1} f_j$, respectively, where $\bar{\zeta} = N^{-1}\sum_{i=1}^{N}\zeta_i$. This expression yields the cross-sectionally augmented Dickey-Fuller (CADF) statistics; the individual CADF statistics are used to develop a modified version of the IPS test called CIPS. However, Monte Carlo experiments show that Pesaran's CIPS test has desirable small-sample properties in the presence of a single unobserved common factor but exhibits size distortions if the number of common factors exceeds unity. Recently, Pesaran et al. (2013) extend Pesaran's CIPS test to the case of a multifactor error structure. They propose utilizing the information contained in a number of k additional variables, x_{it}, that are assumed to share the common factors of the series of interest, y_{it}. In particular, they propose two tests. The first test, CIPS, is an extension of the test proposed in Pesaran (2007) and is based on the average of t-ratios from ADF regressions augmented by the cross-sectional averages of the dependent variable as well as k additional regressors. The second test, CSB, exploits cross-sectional augmentation for the Sargan–Bhargava test. It is worth noting that the perspective of these tests is quite different from that of Bai and Ng (2004). Indeed, while Bai and Ng (2004) consider whether the source of nonstationarity is due to the common factors and/or the idiosyncratic components, neither of which are observed directly, Pesaran et al. (2013) aim to test for the presence of a unit root in the y_{it} process, which is observed. In doing so, they adopt an autoregressive specification augmented with common factors, and the unit root test is performed by testing whether the autoregressive component of the specification expressed in first difference, δ_i, is 0 for all i against the alternative, which can be expressed as $\delta_i = 0$ for some countries but $\delta_i < 0$ for others. In such a framework, they rule out the possibility of the factors having unit roots because, otherwise, all series in the panel could be $I(1)$ irrespective of whether $\delta_i = 0$ (see Pesaran et al. (2013), p. 96-97). To address the uncertainty surrounding the value of r, we follow Pesaran et al. (2013) and consider the application of the CIPS and CSB tests by allowing the number of factors, $r = k + 1$, to take any value between 1 and 4, and we present the results of these tests for all possible combinations of regressors. Note that 4 is the maximum number of factors for which Pesaran et al. (2013) provide critical values. When $k = 3$, we implicitly assume that the four observed variables used in the econometric analysis, f_{it}, S^d_{it}, S^f_{it}, H_{it}, share the same common factors. Pesaran et al. (2013) set the lag order, $\widehat{p} = \left[4\,(T/100)^{1/4}\right]$; in our case, this rule yields $\widehat{p} = 3$. For small T, as in our case, they also suggest using the CSB test, having higher power than that of CIPS. Adopting the CSB and setting $p = 3$ is thus our preferred choice, with the results are summarized in **Table A2**. The results for both CIPS and CSB and for $p = 1, 2, 3$ are available upon request. For CSB, The test outcomes are as follows. The null hypothesis of a panel unit root is rejected in nearly all cases. This result occurs irrespective of the lag order, p, for all variables under investigation. Few exceptions occur when $r = 1$, a case that is always rejected by the above-presented criteria when choosing the number of factors with

PANIC and Moon and Perron (2004).[2] Moreover, while the results in Table A2 refer to the intercept and trend case, when we focus on the intercept case, the null hypothesis is always rejected at standard levels irrespective of both the lag order and the variables used to augment the ADF regression (even for $r = 1$).

In summary, in this appendix, we focus on second-generation tests allowing for cross-sectional dependence. We consider both augmented ADF-type specifications (Pesaran, 2007; Pesaran et al., 2013) and tests decomposing the panel into deterministic, common and idiosyncratic components (Bai and Ng, 2004; Moon and Perron, 2004). While most previous works find evidence of nonstationary variables by applying first-generation tests (see, e.g., Coe et al. 2009), we provide a more nuanced and thorough picture, ultimately suggesting that while the unobserved idiosyncratic component of the variables under study is stationary, the unobserved common factors appear to be nonstationary. This finding is highly relevant from an empirical perspective and is related to the recent work by Kapetanios et al. (2011) showing that the CCE approach is still valid when unobserved factors are allowed to follow unit root processes.

[2] When the CIPS test is used, the results are mixed and crucially depend both on the variables that are used to augment the ADF regression and on the lag order of the autoregressive component. Rejection of the null hypothesis is more likely to appear when $r > 1$.

Moreover, a high level of sensitivity to the number of lags of the AR component is also found when we consider some first-generation tests, notably the test proposed by Im, Pesaran and Shin (2003) (IPS) and the Fisher-type tests introduced by Maddala and Wu (1999) and further developed by Choi (2001). In particular, we documented that when the number of lags of the autoregressive component of heterogeneous ADF-type specifications is estimated in a model selection framework, the tests generally indicate the rejection of the null. Detailed results are available upon request.

References

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

[2] Bai, J. and Ng, S. (2004). 'A Panic Attack on Unit Roots and Cointegration', *Econometrica*, 72, 1127–1177.

[3] Baltagi, B.H., Bresson, G. and Pirotte, A. (2007), 'Panel Unit Root Tests and Spatial Dependence', *Journal of Applied Econometrics*, 22, 339–360.

[4] Breitung, J. and Pesaran, M.H. (2008). 'Unit roots and cointegration in panels', In: Matyas, L., Sevestre, P. (Eds.), *The Econometrics of Panel Data: Fundamentals and Recent Developments in Theory and Practice*, Springer, Berlin.

[5] Choi, I. (2001). 'Unit Root Tests for Panel Data', *Journal of International Money and Finance*, 20, 249-272.

[6] Choi, I. (2006). 'Combination Unit Root Tests for Cross-Sectionally Correlated Panels'. In: D. Corbae, S. N. Durlauf, and B. E. Hansen (eds.): *Econometric Theory and Practice: Frontiers of Analysis and Applied Research: Essays in Honor of Peter C. B. Phillips*. Cambridge University Press, Chapt. 11, 311–333.

[7] Elliott, G., Rothenberg, T. and Stock, J. (1996). 'Efficient Tests for an Autoregressive Unit Root', *Econometrica*, 64, 813–836.

[8] Fisher, R. A. (1932). *Statistical Methods for Research Workers*, Oliver & Boyd, Edinburgh, 4th Edition.

[9] Gengenbach, C., Palm, F. C. and Urbain, J.P. (2009). 'Panel Unit Root Tests in the Presence of Cross-Sectional Dependencies: Comparison and Implications for Modelling', *Econometric Reviews*, 29, 111-145.

[10] Gutierrez, L. (2006). 'Panel Unit-root Tests for Cross-sectionally Correlated Panels: A Monte Carlo Comparison', *Oxford Bulletin of Economics and Statistics*, 68, 519-540.

[11] Hurlin, C. (2010). 'What would Nelson and Plosser find had they used panel unit root tests?', *Applied Economics*, 42, 1515-1531.

[12] Im, K.S., Pesaran, M.H. and Shin, Y. (2003). 'Testing for Unit Roots in Heterogenous Panels', *Journal of Econometrics,* 115, 53–74.

[13] Levin, A., Lin, C. and Chu, C.J. (2002). 'Unit Root Tests in Panel Data: Asymptotic and Finite-sample Properties', *Journal of Econometrics*, 108, 1–24.

[14] Maddala, G.S. and Wu, S. (1999). A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test, *Oxford Bulletin of Economics and Statistics*, 61, 631–652.

[15] Moon, R. and Perron, B. (2004), 'Testing for Unit Root in Panels with Dynamic Factors', *Journal of Econometrics*, 122, 81-126.

[16] Moon, R. and Perron, B. (2007), 'An empirical analysis of nonstationarity in a panel of interest rates with factors', *Journal of Applied Econometrics*, 22, 383-400.

[17] Pesaran, M.H. (2007). 'A Simple Panel Unit Root Test in the Presence of Cross Section Dependence', *Journal of Applied Econometrics*, 22, 265–312.

[18] Pesaran, M.H. (2012). 'On the interpretation of panel unit root tests', *Economics Letters,* 116, 545-546.

[19] Pesaran, M. H., V. Smith and T. Yamagata (2013). 'Panel unit root tests in the presence of a multifactor error structure', *Journal of Econometrics,* 175, 94-115.

[20] Stock, J. H. and M. W. Watson (2002). 'Macroeconomic forecasting using diffusion indexes', *Journal of Business and Economic Statistics,* 20, 147.162.

[21] Westerlund, J. and J. Breitung (2009). 'Myths and facts about panel unit root tests', Working Papers in Economics No 380, University of Gothenburg.

TABLE A1

Bai and Ng test

		Idiosyncratic component		Common factors	
	Number of factors			Nonstationary factors $\widehat{r}_1$	
	$\widehat{r}_{BIC3}$	P_e^c	Z_e^c	MQ_f	MQ_c
		p-value			
Model with intercept					
$\log f$	3	0.07	0.07	3	3
$\log S^d$	6	0	9.45E-14	6	6
$\log S^f$	6	2.21E-10	1.17E-06	6	6
$\log H$	6	0	9.67E-12	6	6
Model with intercept and trend					
$\log f$	3			3	3
$\log S^d$	6			6	6
$\log S^f$	6			6	6
$\log H$	6			6	6

Notes:

$\widehat{r}_{BIC3}$ is the estimated number of common factors, based on BIC3 criterion. We impose that the maximum number of factors is 6.

BIC3 and IC2 provide very similar results.

For the idiosyncratic components, only pooled unit root tests are reported.

Individual ADF t-statistics based on defactored components are available upon request.

The pooled test for the idiosyncratic component is valid only for the intercept case.

In the linear trend case, the limiting distribution of the individual ADF test obtained on defactored components is not a DF-type distribution.

For the common factors components, the estimated number $\widehat{r}_1$of independent stochastic trends is reported (5% level).

The results of the test for all values of r in the range 1-20 are available upon request.

They strongly confirm the finding of nonstationary common factors and stationary idiosyncratic components.

TABLE A2

CSB tests - Pesaran (2007) and Pesaran et al. (2013)

	$\log f$		$\log S^d$		$\log S^f$		$\log H$	
	$\overline{x}_t$	$CSB(\widehat{p})$	$\overline{x}_t$	$CSB(\widehat{p})$	$\overline{x}_t$	$CSB(\widehat{p})$	$\overline{x}_t$	$CSB(\widehat{p})$
$r=1$		0.150		0.069**		0.123		0.112
$r=2$								
	$(\overline{\log S^d})$	0.082*	$(\overline{\log f})$	0.052**	$(\overline{\log f})$	0.070**	$(\overline{\log f})$	0.095
	$(\overline{\log S^f})$	0.079*	$(\overline{\log S^f})$	0.042**	$(\overline{\log S^d})$	0.040**	$(\overline{\log S^d})$	0.081*
	$(\overline{\log H})$	0.098	$(\overline{\log H})$	0.059**	$(\overline{\log H})$	0.038**	$(\overline{\log S^f})$	0.037**
$r=3$								
	$(\overline{\log S^d}, \overline{\log S^f})$	0.042**	$(\overline{\log f}, \overline{\log S^f})$	0.032**	$(\overline{\log f}, \overline{\log S^d})$	0.026**	$(\overline{\log f}, \overline{\log S^d})$	0.062
	$(\overline{\log S^d}, \overline{\log H})$	0.055*	$(\overline{\log f}, \overline{\log H})$	0.041**	$(\overline{\log f}, \overline{\log H})$	0.032**	$(\overline{\log f}, \overline{\log S^f})$	0.029**
	$(\overline{\log S^f}, \overline{\log H})$	0.033**	$(\overline{\log S^f}, \overline{\log H})$	0.023**	$(\overline{\log S^d}, \overline{\log H})$	0.023**	$(\overline{\log S^d}, \overline{\log S^f})$	0.032**
$r=4$								
	$(\overline{\log S^d}, \overline{\log S^f}, \overline{\log H})$	0.016**	$(\overline{\log f}, \overline{\log S^f}, \overline{\log H})$	0.016**	$(\overline{\log S^d}, \overline{\log f}, \overline{\log H})$	0.016**	$(\overline{\log S^d}, \overline{\log S^f}, \overline{\log f})$	0.020**

Notes:

We set the lag order, $\widehat{p} = \left[4\,(T/100)^{1/4}\right]$, as in Pesaran et al. (2013). This gives $\widehat{p} = 3$.

For small T, as in our case, Pesaran et al. (2013) also suggest using the CSB test, which has higher power than that of CIPS.

Model with intercept and trend.

The variables under the heading $\overline{x}_t$ indicate the regressors used for cross-section augmentation in addition to $\overline{y}_t$.

In the case where $r = 1$, no additional regressors are used as in Pesaran (2007).

For the selected lag order, the 5% critical values for the CSB are 0.102, 0.076, 0.054, 0.035 for r=1, 2, 3, 4, respectively.

For the selected lag order, the, 10% critical values are 0.109, 0.083, 0.059, 0.039 for r=1, 2, 3, 4, respectively.

**, * Denote rejection at the 5% and 10% level, respectively.

Appendix B: Additional Estimation Results: Bayesian Approaches and Weighted Mean Group

This appendix aims to provide further insights on the issue of slope heterogeneity.

We first focus on a framework characterized by cross-sectional independence and extend the analysis, with the results reported in Table 2. We take this step by applying the shrinkage estimators described in Maddala et al. (1997): the (two-step) Empirical Bayes (EB) and Iterative Empirical Bayes (IEB) estimators. These estimators can be viewed as a compromise between the unrealistic homogeneity assumption and unstable heterogeneous estimates. We also apply the hierarchical Bayes approach (HB), which makes use of Markov Chain Monte Carlo methods via Gibbs sampling, and we use the same priors as Hsiao et al. (1999). According to Hsiao et al. (1999), this estimator is asymptotically equivalent to the MG estimator but performs better in small samples. We also adopt a variant of the MG estimator by resorting to estimating weighted averages of the slope parameters (WMG), and we use the same approach previously adopted by Bond et al. (2010), based first on Huber weights and then on biweights (see Hamilton, 1991).

The results (**Table B1**) are as follows. Overall, the use of Bayesian methods yields estimated average parameters that are very close to those obtained with the MG but increase the precision. The HB exhibits the best performance in terms of precision, followed by the IEB, while the EB gives standard errors very close to the MG. This result also complements some previous papers (Baltagi et al., 2002, 2004; Mazzanti and Musolesi, 2013). When adopting the WMG, we instead find a substantial change in the estimated mean parameters – and, in particular, a decrease in the parameters associated with foreign R&D and human capital – as well as a very slight increase in precision compared with the unweighted MG, with estimates that are not statistically significant at standard levels as for the standard MG estimator.

Second, more relevant for the present paper, we focus on a framework allowing for strong dependence and compare the standard unweighted CCEMG with its weighted variant – where the weights are determined as explained above – which we label CCEWMG (**Table B2**). The use of a weighted average in this framework also produces some relevant changes in the estimated mean parameters and a very modest increase in the precision of estimation. More precisely, the change in the mean parameters when adopting the CCEWMG is very pronounced for human capital; moreover, when we use such an estimation method, the estimated mean coefficient for human capital varies greatly (from -0.329 to 0.444) depending on the adopted specification and is never significant. This could further confirm that the quantity of education does not affect TFP. The other estimated mean parameters are less affected by the use of a weighted average. Adopting other ways to endogenously determine the appropriate weights could be a focus for future studies and a means of providing more precise estimates in a heterogeneous panel data environment.

References

[1] Baltagi, B.H., Bresson, G. and Pirotte, A. (2002). 'Comparison of forecast performance for homogeneous, heterogeneous and shrinkage estimators: Some empirical evidence from US electricity and natural- gas consumption', *Economics Letters*, 76, 375-382.

[2] Baltagi, B.H., Bresson, G. and Pirotte, A. (2004). 'Tobin q: forecast performance for hierarchical Bayes, shrinkage, heterogeneous and homogeneous panel data estimators', *Empirical Economics,* 29, 107-113.

[3] Bond, S., A. Leblebiciog and F. Schiantarelli (2010). 'Capital Accumulation and Growth: A New Look at the Empirical Evidence', *Journal of Applied Econometrics*, 25, 1073-1099.

[4] Hamilton, L.C. (1991). 'How robust is robust regression?', *Stata Technical Bulletin,* 2, 21–26.

[5] Hsiao, C. Pesaran, M.H. Tahmiscioglu, A.K. (1999). 'Bayes estimation of short-run coefficients in dynamic panel data models'. In: Hsiao C, Lahiri K, Lee L-F, Pesaran MH (eds) Analysis of panels and limited dependent variable models. Cambridge University Press, Cambridge, 268-296.

[6] Maddala, G.S. Trost, R.P. Li, H. and Joutz, F. (1997). 'Estimation of short-run and long-run elasticities of energy demand from panel data using shrinkage estimator', *Journal of Business and Economic Statistics*, 15, 90-100.

[7] Mazzanti, M. and Musolesi, A. (2013). 'The heterogeneity of carbon Kuznets curves for advanced countries: comparing homogeneous, heterogeneous and shrinkage/Bayesian estimators', *Applied Economics*, 45, 3827-42.

[8] Pesaran, M.H. and Smith, R. (1995). 'Estimating long-run relationships from dynamic heterogeneous panels', *Journal of Econometrics,* 68, 79-113.

TABLE B1

Benchmark estimation results: additional results

	MG	EB	IEB	HB	WMG
	(i)	(ii)	(iii)	(iv)	(v)
$\log S^d$	0.071	0.088	0.088*	0.089**	0.067
	(0.062)	(0.060)	(0.046)	(0.044)	(0.065)
$\log S^f$	0.207	0.200	0.203*	0.208**	0.118
	(0.144)	(0.143)	(0.115)	(0.101)	(0.121)
$\log H$	0.627	0.553	0.481	0.465	0.097
	(0.484)	(0.478)	(0.323)	(0.296)	(0.364)

Notes:

MG: Mean Group.

EB: two-step Empirical Bayes.

IEB: Iterative Empirical Bayes.

HB: Hierarchical Bayes.

WMG: Weighted Mean Group.

***, **, *: significant at 1%, 5%, 10%, respectively.

Standard errors in brackets.

TABLE B2

CCEMG and CCEWMG

	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
CCEMG								
$\log S^d$	-0.00734	0.0904	0.0551	0.0558	0.00204	0.0757	0.00973	0.0943
	(0.0931)	(0.0674)	(0.0648)	(0.0656)	(0.0961)	(0.0701)	(0.106)	(0.0623)
$\log S^f$	-0.0755				0.00405		-0.116	
	(0.350)				(0.361)		(0.384)	
$\log S^{f\_LP}$		0.0351**			-0.00332			0.0154
		(0.0178)			(0.0170)			(0.0165)
$m{*}\log S^f$			0.0170*			0.0171	0.0183**	
			(0.0101)			(0.0509)	(0.00714)	
$m*\log S^{f\_LP}$				0.0160		0.00501		0.0150
				(0.0132)		(0.0591)		(0.0138)
$\log H$	0.523	0.453	0.425	0.467	0.398	0.593	0.471	0.716
	(0.675)	(0.586)	(0.638)	(0.620)	(0.699)	(0.640)	(0.746)	(0.652)
CCEWMG								
$\log S^d$	0.0194	0.0717	0.0304	0.0323	0.0475	0.0695	0.0580	0.0726
	(0.0906)	(0.0677)	(0.0627)	(0.0635)	(0.0753)	(0.0695)	(0.0837)	(0.0657)
$\log S^f$	-0.243				-0.0364		-0.159	
	(0.345)				(0.394)		(0.421)	
$\log S^{f\_LP}$		0.0444**			0.0110			0.0350***
		(0.0201)			(0.0156)			(0.0133)
$m{*}\log S^f$			0.0125			-0.0542*	0.0185**	
			(0.00957)			(0.0277)	(0.0077)	
$m*\log S^{f\_LP}$				0.0116		0.0717*		0.0166
				(0.0141)		(0.0427)		(0.0139)
$\log H$	0.444	0.0765	-0.329	-0.214	0.295	-0.101	0.299	-0.110
	(0.710)	(0.528)	(0.434)	(0.460)	(0.737)	(0.487)	(0.825)	(0.417)
CD	-3.179	-3.199	-2.465	-2.663	-3.228	-2.694	-3.132	-2.995
$\widehat{a}$	-0.077	-0.180	0.151	0.045	-0.340	0.217	-0.232	0.092
$\widetilde{a}$	0.406	0.305	0.416	0.349	0.146	0.522	0.216	0.500

Notes.

CCEMG: Common Correlated Effects Mean Group

CCEWMG: Common Correlated Effects Weighted Mean Group

***, **, *: significant at 1%, 5%, 10%, respectively

Standard errors in brackets.

$\widehat{a}$: bias-corrected version of a given by equation (13) in Bailey et al. (2015).

$\widetilde{a}$: robust estimator of a correcting for both serial correlation in the factors and weak cross-sectional dependence in the error term.

We use four principal components when estimating equation (30) in Bailey et al. (2015).