SEEDS

Sustainability Environmental Economics and Dynamics Studies

Working Paper Series

Weak and strong cross-sectional dependence: a panel data analysis of international technology diffusion.

by

Cem Ertur and Antonio Musolesi

19/2015

SEEDS Working Paper 19/2015
December 2015
by Cem Ertur and Antonio Musolesi.

This paper is a substantially revised version of the SEEDS Working Paper 4/2015 published in February 2015.

Weak and strong cross-sectional dependence: a panel data analysis of international technology diffusion

CEM ERTUR† and ANTONIO MUSOLESI‡

December 8, 2015

Abstract

This paper provides an econometric examination of geographic R&D spillovers among countries by focusing on the issue of cross-sectional dependence and in particular on the different ways – weak and strong – that it may affect the model. A preliminary analysis based on the estimation of the exponent of cross-sectional correlation provides a clear-cut result with an estimate of a that is very close to unity, not only indicating the presence of strong cross-sectional correlation but also being fully consistent with the factor literature, which typically assumes that $a = 1$. Moreover, second-generation unit roots tests suggest that while the unobserved idiosyncratic component of the variables under study is stationary, the unobserved common factors appear to be nonstationary. Consequently, a factor structure appears to be preferable to a spatial error model, and in particular, the correlated common effects approach is employed because, among other advantages, it is still valid in the more general case of nonstationary common factors. According to the results, richer countries benefit more from both domestic R&D and geographic spillovers than poorer countries. Finally, comparing the results to those obtained from a spatial model provides some interesting insights on the possible bias occurring when allowing only for weak correlation while strong correlation is present in the data.

JEL classification: C23; C5; F0; O3.
Keywords: panel data; cross-sectional correlation; factor models; spatial models, heterogeneous slopes, unit root; international technology diffusion; geography.

†*LEO, University of Orleans UMR 6221 CNRS Faculté de Droit d'Economie et de Gestion, Rue de Blois - B.P. 6739 45067 Orléans Cedex 2 –France*
(email: cem.ertur@univ-orleans.fr)

‡*Department of Economics and Management, University of Ferrara –Italy*
(antonio.ausolesi@unife.it)

1 Introduction

Since the seminal paper by Coe and Helpman (1995), recently revisited by Coe et al. (2009) – henceforth CH and CHH, respectively – there has been an increasing interest in international technology diffusion. CH test the predictions of models of innovation and growth (Grossman and Helpman, 1991) in which total factor productivity (TFP) is an increasing function of cumulative research and development (R&D). In particular, CH analyze the role of international trade. By assuming that some intermediate inputs are traded internationally, whereas others are not, they relate TFP to both domestic and foreign R&D and construct the foreign R&D capital stock as the import-share-weighted average of the domestic R&D capital stocks of the trading partners. The influence of this approach is based on its plausibility with respect to endogenous growth theory (Keller, 2004) and its versatility in allowing for the consideration of alternative channels of international technology diffusion, such as foreign direct investment (FDI) (Lichtenberg and Van Pottelsberghe, 2001), bilateral technological proximity and patent citations between countries (Lee, 2006), language skills (Musolesi, 2007) or geographic proximity (Keller, 2002).

The goal of the present paper is twofold. First, it is intended to contribute to the empirical literature on R&D spillovers among countries by focusing on the issue of cross-sectional dependence. The rationale is that cross-country correlation from a variety of sources is plausibly present in CH-type specifications; however, this correlation complicates standard estimation and inference. Second, as this issue can be relevant in many other empirical frameworks, we also seek to provide an illustration of new methodologies in this area.

Cross-sectional dependence can be introduced as a result of a finite number of *unobservable (and/or observed) common factors* that may have different effects on TFP across countries. Such factors might include, for instance, aggregate technological shocks, national policies intended to raise the level of technology or oil price shocks that may influence TFP through their effects on production costs. The heterogeneous effect of these factors may be the result, for instance, of country-specific technological constraints. A model with a multifactor error structure can be estimated by adopting the correlated common effects (CCE) approach developed by Pesaran (2006), which has been further developed and proved to be valid in a variety of situations (Chudik et al., 2011; Pesaran and Tosetti, 2011; Kapetanios et al., 2011). Cross-sectional correlation can be alternatively regarded as a result of spatial effects and can be modeled by adopting a spatial panel econometric framework allowing for spatially correlated disturbances. The estimation can be performed using Lee and Yu's (2010) quasi-maximum likelihood (QML) estimator. Some remarks are in order.

First, the unobserved common factors approach and the spatial error model are related to the recently developed concepts of weak and strong cross-sectional dependence. Although the literature does not provide a single definition of 'weak' or 'strong' dependence (see, e.g., Chudik et al., 2011, Robinson, 2011, and Sarafidis, 2009), it is interesting to note that the spatial models, under a standard set of regularity conditions, entail weak cross-sectional correlation regardless of the definition adopted (Breitung and Pesaran, 2008; Pesaran and Tosetti, 2011; Sarafidis and Wansbeek, 2012), whereas the type of dependence arising from the factor model depends on the adopted definition of weak/strong dependence and the limiting properties of averaged factor loadings (Sarafidis and

Wansbeek, 2012). A related concept is that of strong and weak factors, recently proposed by Chudik et al. (2011). The authors demonstrate that there is a direct relationship between the concept of weak/strong factors and their conception of weak/strong dependence. They demonstrate that a process that is the sum of a finite number of common factors and an idiosyncratic error term is cross-sectionally strongly dependent at a given point in time if at least one of those common factors is strong. Specifically, the CCE approach explicitly introduces a finite number of *strong* factors, entailing strong dependence, but does not explicitly introduce *weak* factors.

Second, the most general model would obviously allow for both forms of dependence – weak and strong – as suggested by some recent papers (Pesaran and Tosetti, 2011; Chudick et al., 2011; Bresson and Hsiao, 2011; Bailey et al. 2015b). However, the CCE approach has been shown to be valid even in this case. In particular, Pesaran and Tosetti (2011) prove that the CCE estimator provides consistent estimates of the slope coefficients and their standard errors under a generalized data generating process (DGP) with an error term that is the sum of a multifactor structure and a spatial process, i.e., when both forms of cross-correlation – weak and strong – characterize the DGP. Moreover, Chudik et al. (2011) extended the CCE approach by allowing for the presence of both a limited number of strong factors and a large number of weak or semi-strong factors and then show that, even under this extended framework, the CCE method still provides consistent estimates of the slope coefficients.

Third, while both factor and spatial models allow for cross-sectional correlation, the motivations underlying such models differ meaningfully. In the first approach, the unobserved factors are regarded as nuisance variables introduced to allow for cross-sectional dependence and to capture information in a parsimonious way, whereas the main focus is on the estimation and inference of the slope parameters. Moreover, this set-up introduces endogeneity due to unobservables, whereby the explanatory variables are allowed to be correlated with the factors. Bai (2009) and Sarafidis and Wansbeek (2012) provide many examples of circumstances under which this may occur, such as production and cost function specifications. This can also be relevant in the CH specification, as stressed by Keller (2004, p.763). The spatial models, instead, are intended to model interactions among cross-sections, such as spillover or network effects. Such cross-section interactions will produce differentiated impact coefficients computed from the reduced form of the spatial model (see, e.g., Debarsy and Ertur, 2010). Moreover, spatial models have been shown to be relevant in many contexts related to this paper such as neoclassical and endogenous growth models (Ertur and Koch, 2007, 2011).

We argue that, when analyzing international R&D spillovers, there are neither theoretical reasons nor well-established empirical evidence allowing for an a priori choice between the spatial approach and the factor model. If the data exhibit only weak correlation, this could indicate the necessity of modeling spatial interactions, while if strong correlation is present, this would suggest that relevant variables have not been accounted for in the original formulation or that both forms of correlation may affect the model. In this paper, we first focus on testing and measuring cross-sectional correlation. This appears highly relevant from both a theoretical and a policy perspective. According to the obtained results, we finally estimate the model with the most suitable approach and provide some new results.

The remainder of the paper is organized as follows: section 2 describes the baseline model. Section 3 extends the benchmark specification by allowing for cross-sectional dependence. Section 4 presents the results. Finally, section 5 concludes.

2 Baseline econometric specification

The baseline econometric model is an extended version of that adopted by CH, as modified by CHH by including human capital on the right-hand side of the equation:

$$f_{it} = \exp(\alpha_i + e_{it}) \left(S^d_{it}\right)^\theta \left(S^f_{it}\right)^\gamma H^\delta_{it} \tag{1}$$

where f_{it} is the total factor productivity of country $i = 1, ..., N$ at time $t = 1, ..., T$; α_i are individual fixed effects that account for unobserved time-invariant characteristics, which are allowed to be freely correlated with both R&D capital stocks (domestic, S^d_{it}, and foreign, S^f_{it}) and human capital (H_{it}); and e_{it} is the error term. The foreign capital stock S^f_{it} is defined as the weighted arithmetic mean of S^d_{jt} for $j \neq i$:

$$S^f_{it} = \sum_{j \neq i} \omega_{ij} S^d_{jt} \tag{2}$$

where w_{ij} represents the weighting scheme. The model is then linearized by taking logs:

$$\log f_{it} = \alpha_i + \theta \log S^d_{it} + \gamma \log \sum_{j \neq i} \omega_{ij} S^d_{jt} + \delta \log H_{it} + e_{it} \tag{3}$$

It is interesting to note that this simple empirical specification can be derived from an endogenous growth model (see, e.g., Keller, 2004, p. 762). However, as noted by Lichtenberg and Van Pottelsberghe (2001, p. 490), *International technological spillovers have no widely accepted measures"*. According to Keller (2004), the main channels of technology diffusion are trade, FDI and language skills. For instance, CHH and Lichtenberg and Van Pottelsberghe (1998) use alternative definitions of w_{ij} based on imports, Lichtenberg and Van Pottelsberghe (2001) focus on FDI, and Musolesi (2007) adopts a weighting scheme that takes language skills into account. More recently, Spoalore and Warcziag (2009) suggest genetic distance as a barrier to the diffusion of development.

In this paper, we focus on geographic proximity as a channel for technology diffusion for many reasons. First, it is theoretically consistent. Keller (2002) and Eaton and Kortum (2002) show that international technology diffusion is related to geographic distance due to transport costs or geographic barriers. Second, the geographic localization of international technology diffusion can have economically relevant implications. Specifically, it can affect the process of convergence across countries (Grossman and Helpman, 1991), the agglomeration that takes place in an economy (Krugman and Venables, 1995) and the long-term effectiveness of macroeconomic policies aimed at technological progress (Keller, 2002). Third, there have been far fewer studies on geographic international R&D spillovers than on spillovers via other channels, such as trade or FDI, despite the theoretical consistency and empirical relevance of geography. Finally, and perhaps most important in the context of this paper, which focuses on the methodological issue of cross-sectional dependence, traditional channels of international technology diffusion might create reverse causality problems when included in econometric specifications. For instance, a country's international trade, FDI or patent activity may depend on its technological level and, in turn, may be endogenous with respect to TFP (see, e.g., Hong and Sun, 2011). In contrast, geographic distance is generally considered exogenous, *"Global technology spillovers favor income convergence, and local spillovers tend to lead to divergence, no matter through which channel technology diffuses... An advantage of this is that geography is arguably exogenous in this process"* (Keller, 2004, p.772). Moreover, geographic distance may be considered an exogenous proxy for certain endogenous measures of socioeconomic, institutional, cultural or linguistic similarities that might enhance the diffusion of technology. Following Keller (2002), we propose a specification of foreign R&D that incorporates the notion that the impact of foreign R&D is a decreasing function of geographic distance from foreign economies. Therefore, the foreign R&D capital stock for each country i is obtained by weighting the domestic R&D capital stocks of every other country $j \neq i$ in the sample using an exponential distance decay function, $\omega_{ij} = \exp(-d_{ij})$, such that:[1]

$$S_{it}^{f} = \sum_{j \neq i} \exp(-d_{ij}) S_{jt}^{d} \tag{4}$$

Finally, to construct the stock of human capital, we use the average number of years of schooling in the population over 25 years old. Following Hall and Jones (1999), this parameter is converted into a measure of human capital stock through the following formula:

$$H_{it} = \exp\left[g\left(Edu_{it}\right)\right] \tag{5}$$

where Edu_{it} is the average number of years of schooling and the function $g\left(Edu_{it}\right)$ reflects the efficiency of a unit of labor with Edu years of schooling relative to one with no schooling. Following Psacharopoulos (1994) and Caselli (2005), it is assumed that $g\left(Edu_{it}\right)$ is piecewise linear, which implies a log-(piecewise)linear relationship between H and Edu. Combining equations (4) and (5):

$$\log f_{it} = \alpha_i + \theta \log S_{it}^{d} + \gamma \log \sum_{j \neq i} \exp(-d_{ij}) S_{jt}^{d} + \delta \log H_{it} + e_{it} \tag{6}$$

The homogeneity of the slope parameters implicit in the use of a pooled specification as in equation (6) has, however, been questioned from both an econometric (Pesaran and Smith, 1995, Maddala et al. 1997) and an economic perspective. From an economic standpoint and closely related to the present topic, a theoretical justification for heterogeneous slope parameters across countries can be found in the 'new growth' literature, which argues that technology differs across countries. As Brock and Durlauf (2001, p. 8-9) remark, *"the assumption of parameter homogeneity seems particularly inappropriate when one is studying complex heterogeneous objects such as countries"*. Moreover, Durlauf et al. (2001) also suggest that the explanatory power of the Solow growth model is substantially enhanced by allowing for country-specific production functions. Concerning

[1] This particular choice of the spatial decay pattern, i.e., an exponential distance decay function with a distance decay parameter fixed to one, is very common in the spatial econometric literature and is also consistent with Keller (2002) concerning international technology diffusion. Moreover, the distance decay parameter cannot be estimated neither in the factor nor in the spatial model and has to be fixed in both cases.

international technology diffusion in particular, some studies have challenged the assumption of a common technology by arguing that technologies are specific to particular combinations of inputs. Basu and Weil (1998) propose a learning-by-doing model to describe the process of technological advancement whereby firms improve the productivity not only of the specific capital-labor mix that they are using (as in Atkinson and Stiglitz, 1969) but also of similar techniques. At the aggregate level, this model leads to a situation in which a follower country can use the technology of the leading country only if the former has a sufficiently high level of development and, simultaneously, spillovers are heterogeneous across countries.

In this paper, we adopt two alternative specifications. First, to allow the impact of the explanatory variables to differ between the G7 countries and the others, the *benchmark* specification we adopt is a simple variant of equation (6) that has been widely used in the literature (see, e.g., Lichtenberg and Van Pottelsberghe, 2001):

$$\log f_{it} = \alpha_i + \theta_{G7}\mathbf{1}_{G7}\log S^d_{it} + \theta_{NOG7}\mathbf{1}_{NOG7}\log S^d_{it} + \gamma_{G7}\mathbf{1}_{G7}\log\sum_{j\neq i}\exp(-d_{ij})S^d_{jt} + \tag{7}$$

$$+\gamma_{NOG7}\mathbf{1}_{NOG7}\log\sum_{j\neq i}\exp(-d_{ij})S^d_{jt} + \delta_{G7}\mathbf{1}_{G7}\log H_{it} + \delta_{NOG7}\mathbf{1}_{NOG7}\log H_{it} + e_{it}$$

$$\text{with:}\mathbf{1}_{G7} = \begin{cases} 1 & \text{if country } \in \text{ G7 group} \\ 0 & \text{if country } \notin \text{ G7 group} \end{cases}, \text{ and: } \mathbf{1}_{NOG7} = 1 - \mathbf{1}_{G7}$$

Second, we consider variants of equation (7) by adopting methods allowing for heterogeneous slopes such as the mean group estimator (Pesaran and Smith, 1995) and focus on the averages of country-specific estimates within the two groups of countries. Although approaches allowing for country-specific parameters are conceptually appealing, considering the *benchmark* specification in equation (7) would be useful both for comparison purposes with previous studies and for avoiding the problem of parameter estimate instability caused by the estimation of several parameters with relatively short time series (Baltagi et al., 2002, 2004). In the following, for ease of exposition, equation (7) can then be expressed as:

$$y_{it} = \alpha_i + \beta'\mathbf{x}_{it} + e_{it} \tag{8}$$

where:

$$\begin{aligned} y_{it} &= \log f_{it} \\ \mathbf{x}_{it} &= [\mathbf{1}_{G7}\log S^d_{it}, 1_{NOG7}\log S^d_{it}, \mathbf{1}_{G7}\log\sum_{j\neq i}\exp(-d_{ij})S^d_{jt}, \\ &\quad \mathbf{1}_{NOG7}\log\sum_{j\neq i}\exp(-d_{ij})S^d_{jt}, \mathbf{1}_{G7}\log H_{it}, \mathbf{1}_{NOG7}\log H_{it}]' \\ \beta &= [\theta_{G7}, \theta_{NOG7}, \gamma_{G7}, \gamma_{NOG7}, \delta_{G7}, \delta_{NOG7}]' \end{aligned}$$

Our main source is the CHH data set. This data set is a balanced panel of 24 countries observed over the period 1971-2004. Our measures of TFP and domestic R&D capital stock come from this data source. The average number of years of schooling used to construct our measure of human capital is taken from Barro and Lee (2013). Finally, the distance between two countries is calculated as the spherical distance between capitals.

3 Strong and weak cross-sectional correlation

To introduce cross-sectional correlation into the benchmark specification defined by equation (8), following Sarafidis (2009) and Sarafidis and Wansbeek (2012), a *general* error structure can be considered:

$$e_{it} = (\varrho_i \odot \mathbf{w}_i)' \xi_t + \varepsilon_{it} = \sum_{j=1}^{m} \varrho_{ij} w_{ij} \xi_{jt} + \varepsilon_{it} \tag{9}$$

where $\odot$ denotes the Hadamard product, $\xi_t = (\xi_{1t}, \xi_{2t}, ..., \xi_{mt})'$ is an $m \times 1$ vector of unobserved common factors, $\varrho_i = (\varrho_{i1}, \varrho_{i2}, ..., \varrho_{im})'$ is an $m \times 1$ vector of factor loadings, $\mathbf{w}_i = (w_{i1}, w_{i2}, ..., w_{im})'$ is am $m \times 1$ vector of deterministic bounded weights and $\varepsilon_{it} \sim i.i.d.\,(0, \sigma_\varepsilon^2)$. Equation (9) allows us to regard the unobserved components, ξ_t, as shocks, the impact of which are a combination of heterogeneous factor loadings (ϱ_i) with a weight scheme ($\mathbf{w}_i$).

Setting $\mathbf{w}_i = \iota$, ι as a vector of ones, equation (9) reduces to the multifactor error structure:

$$e_{it} = \varrho_i' \xi_t + \varepsilon_{it} = \sum_{j=1}^{m} \varrho_{ij} \xi_{jt} + \varepsilon_{it} \tag{10}$$

With $m = N$, the N common factor model (Chudick et al., 2011) is obtained as follows:

$$e_{it} = \sum_{j=1}^{N} \varrho_{ij} \xi_{jt} + \varepsilon_{it}. \tag{11}$$

In matrix form, stacking over individuals, we obtain the following:

$$\boldsymbol{e}_t = \mathbf{P} \xi_t + \boldsymbol{\varepsilon}_t \tag{12}$$

where $\xi_t = \xi_t = (\xi_{1t}, \xi_{2t}, ..., \xi_{Nt})'$ is an $N \times 1$ vector of unobserved factors, $\mathbf{P}$ is an $N \times N$ matrix of associated factor loadings with typical element $\{\varrho_{ij}\}$ and $\boldsymbol{\varepsilon}_t = i.i.d.\,(0, \sigma_\varepsilon^2 \mathbf{I}_N)$.

Imposing appropriate zero restrictions on $\mathbf{w}_i$, at least $w_{ij} = 0$ for $i = j$, homogeneity restrictions on $\varrho_i = \lambda, \forall i$, and setting $\xi_{jt} = e_{jt}$, for $j = 1, ..., N$, with $m = N$, equation (9) reduces to a spatial autoregressive error specification:

$$e_{it} = \lambda \sum_{j=1}^{N} w_{ij} e_{jt} + \varepsilon_{it} \tag{13}$$

that can be rewritten in matrix form as follows:[2]

$$\boldsymbol{e}_t = \lambda \mathbf{W}_N \boldsymbol{e}_t + \boldsymbol{\varepsilon}_t \tag{14}$$

[2]It is worth noting that equation (9) also contains the spatial moving average process, i.e., $e_{it} = \lambda \sum_{j \neq i} w_{ij} \varepsilon_{jt} + \varepsilon_{it}$, and the spatial error component process, $e_{it} = \lambda \sum_{j \neq i} w_{ij} \varsigma_{jt} + \varepsilon_{it}$, where $E(\varsigma_{jt}) = 0$, $V(\varsigma_{jt}) = \sigma_\varsigma^2$ and $E(\varsigma_{jt} \varepsilon_{it}) = 0$ are special cases.

where $\boldsymbol{W}_N$ is defined in the spatial econometrics literature as an $N \times N$ interaction or spatial weight matrix. It is generally not derived from theory but exogenously given to reflect the interaction pattern connecting individuals, which is considered to be time invariant.[3] Its elements are non stochastic, non negative and finite. Under the invertibility condition of $(\boldsymbol{I}_N - \lambda \boldsymbol{W}_N)$, we obtain:

$$\boldsymbol{e}_t = \mathbf{R}_N \boldsymbol{\varepsilon}_t \tag{15}$$

where $\mathbf{R}_N = (\boldsymbol{I}_N - \lambda \boldsymbol{W}_N)^{-1}$. Note that $\boldsymbol{W}_N$ and $\mathbf{R}_N$ must satisfy some regularity conditions given for instance by Lee (2004) for QML estimation.[4] Those conditions primarily require that $\boldsymbol{W}_N$ and $\mathbf{R}_N$ be uniformly bounded in row and column sums both at the true value of the λ parameter and uniformly in λ in a compact parameter space Λ. The true value of the λ parameter is in the interior of Λ.[5]

It may be useful to understand how these two approaches are related to the concepts of weak and strong cross-sectional dependence recently developed in the literature. Forni and Lippi (2001) introduce the notion of an idiosyncratic process to characterize a weak form of dependence that involves both time series and cross-sectional dimensions. More recently, Chudick et al. (2011) (henceforth CPT) propose a new and more widely applicable definition. They consider the asymptotic behavior of weighted averages at each point in time and define a process, $\{z_{it}\}$, to be cross-sectionally weakly dependent at a given point in time if its weighted average at that time, conditional on the information set available in the previous period, $\Im_{t-1}$, converges to its expectation in quadratic mean, as the cross-sectional dimension is increased without bounds for all weights, w, that satisfy certain granularity conditions ensuring that the weights are not dominated by a few individuals,[6] that is, $\lim_{N \to \infty} Var\left(\sum_{i=1}^{N} w_{ij} z_{it} \mid \Im_{t-1}\right) = 0$. Another definition has been recently proposed by Sarafidis (2009) (henceforth, SARA), who defines a process $\{z_{it}\}$ to be cross-sectionally weakly dependent if $\sum_{j \neq i} |Cov\left(z_{it}, z_{js} | F_{ij}\right)| < \infty$, where F_{ij} denotes the conditioning set of all time-invariant characteristics of individuals i and j (Robinson, 2011, and Robinson and Thawornkaiwong, 2012, provide a very similar definition).

Spatial error models satisfy, under a standard set of regularity conditions, weak dependence under both definitions. For example, the standard uniform boundedness conditions provided by Lee (2004) are sufficient but not necessary to guarantee weak dependence (see Sarafidis and Wansbeek (2012) for a more detailed discussion). Conversely, the factor approach entails strong dependence under both definitions unless further restrictions are imposed on the factor loadings. To see this relationship, consider the single factor error process $e_{it} = \varrho_i \xi_t + \varepsilon_{it}$, with N and T both being large and noticing that $\sigma_{ij,t} = cov\left(e_{it}, e_{jt} | F_{ij}\right) = \varrho_i \varrho_j \sigma_\xi^2 \neq 0$; therefore, $\sum_{j \neq i} |\sigma_{ij,t}|$ is unbounded, thus entailing strong correlation under SARA. A related and important concept is that of strong and

[3] Exceptions are Lee and Yu (2012) and Behrens et al. (2010). Lee and Yu (2012) consider a time-varying interaction matrix, while Behrens et al. (2010) adopt a theoretically defined matrix.

[4] Those assumptions all originate from Kelejian and Prucha (1998, 1999, 2001).

[5] A slightly different set is given, for instance, by Kelejian and Prucha (2010) for GMM estimation along with a detailed discussion of the parameter space.

[6] The definition of an idiosyncratic process advanced by Forni and Lippi (2001) and the definition by CPT of weak dependence differ in how the weights used to construct weighted averages are defined.

weak factors (Chudik et al. 2011). Let b be a constant in the range $0 \leq b \leq 1$, and consider the condition $\lim_{N \to \infty} N^{-b} \sum_{i=1}^{N} |\varrho_i| = K < \infty$. According to Chudik et al. (2011), the strong and weak factors correspond to $b = 1$ and $b = 0$, respectively. For $b \in (0, 1)$, the factor ξ_t is said to be semi-strong ($1/2 \leq b < 1$) or semi-weak ($0 < b < 1/2$). Thus, $b = 0$ implies that the factor affects only a fixed number of cross-sectional units, whereas $b < 1$ means that the subset of cross-sectional units affected by the factor grows more slowly than N at a rate depending on b. Under CPT, if there exists at least one strong factor, the underlying process is strongly cross-sectionally dependent; otherwise, it is cross-sectionally weakly dependent. As Chudik et al. (2011) also note, the CCE approach by Pesaran (2006) explicitly introduces a finite number of *strong* factors according to their definition of *strong* and *weak* factors. Thus, it entails strong dependence under both SARA and CPT.

3.1 Strong correlation: errors with multifactor structure

The empirical setup adopted in this paper builds on the framework originally proposed by Pesaran (2006) and further developed and studied more recently (Chudik et al. 2011; Pesaran and Tosetti, 2011; Kapetanios et al. 2011). Such a framework has a number of appealing features. It is sufficiently general to render a variety of panel data models as special cases, it allows correlated common factors, it does not require specifying the number of factors, and it is computationally very simple. Let us consider the following DGP:

$$y_{it} = \alpha_i^{'} \mathbf{d}_t + \beta_i^{'} \mathbf{x}_{it} + e_{it} \tag{16}$$

where $\mathbf{d}_t$ is an $l \times 1$ vector of observed common effects, $\alpha_i^{'}$ is the associated vector of parameters and $\mathbf{x}_{it}$ is a 4×1 vector of explanatory variables. The 'one-way' specification is simply obtained by setting $\mathbf{d}_t = 1$. The slope coefficients $\beta_i' = [\theta, \gamma_{G7}, \gamma_{NOG7}, \delta]$ can be assumed to be fixed and homogeneous across countries, $\beta_i' = \beta' \; \forall \; i$, or assumed to follow a random coefficients specification: $\beta_i = \beta + \mathbf{v}_i$, $\mathbf{v}_i \sim IID(\mathbf{0}, \Theta_v)$. The errors e_{it} are assumed to have a multifactor structure as in equation (10):

$$e_{it} = \varrho_i^{'} \xi_t + \varepsilon_{it} \tag{17}$$

where ξ_t is an $m \times 1$ vector of unobserved common factors with country-specific factor loadings ϱ_i. Combining (16) with (17), we thus obtain the following:

$$y_{it} = \alpha_i^{'} \mathbf{d}_t + \beta_i^{'} \mathbf{x}_{it} + \varrho_i^{'} \xi_t + \varepsilon_{it} \tag{18}$$

where the idiosyncratic errors, ε_{it}, are assumed to be independently distributed over $(\mathbf{d}_t, \mathbf{x}_{it})$, whereas the unobserved factors, ξ_t, can be correlated with $(\mathbf{d}_t, \mathbf{x}_{it})$. This correlation is allowed by modeling the explanatory variables as linear functions of the observed common factors, $\mathbf{d}_t$, and the unobserved common factors, ξ_t:

$$\mathbf{x}_{it} = \mathbf{A}_i' \mathbf{d}_t + \mathbf{\Gamma}_i' \xi_t + \mathbf{v}_{it} \tag{19}$$

where $\mathbf{A}_i$ and $\mathbf{\Gamma}_i$ are $l \times 6$ and $m \times 6$ factor loading matrices and $\mathbf{v}_{it} = (v_{i1t}, v_{i2t}, , v_{i3t}, v_{i4t})'$. Combining (18) and (19), we finally obtain a system of equations simultaneously explaining TFP, R&D (domestic and foreign) and human capital:

$$\underset{7\times1}{\mathbf{z}_{it}} = \begin{pmatrix} y_{it} \\ \mathbf{x}_{it} \end{pmatrix} = \begin{pmatrix} \log\left(f_{it}\right) \\ \mathbf{1}_{G7}\log S_{it}^{d} \\ \mathbf{1}_{NOG7}\log S_{it}^{d} \\ \mathbf{1}_{G7}\log\sum\limits_{j\neq i}\exp(-d_{ij})S_{jt}^{d} \\ \mathbf{1}_{NOG7}\log\sum\limits_{j\neq i}\exp(-d_{ij})S_{jt}^{d} \\ \mathbf{1}_{G7}\log H_{it} \\ \mathbf{1}_{NOG7}\log H_{it} \end{pmatrix} = \underset{7\times l l\times 1}{\mathbf{B}_i'\mathbf{d}_t} + \underset{7\times m m\times 1}{\mathbf{C}_i'\ \xi_t} + \underset{7\times1}{\mathbf{u}_{it}}, \tag{20}$$

where:

$$\mathbf{u}_{it} = \begin{pmatrix} \mathbf{1} & \beta_i' \\ \mathbf{0} & \mathbf{I}_k \end{pmatrix}\begin{pmatrix} \varepsilon_{it} \\ \mathbf{v}_{it} \end{pmatrix} = \begin{pmatrix} \varepsilon_{it} + \beta_i'\mathbf{v}_{it} \\ \mathbf{v}_{it} \end{pmatrix},$$

$$\mathbf{B}_i = \begin{pmatrix} \alpha_i & \mathbf{A}_i \end{pmatrix}\begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \beta_i & \mathbf{I}_k \end{pmatrix},$$

$$\mathbf{C}_i = \begin{pmatrix} \varrho_i & \mathbf{\Gamma}_i \end{pmatrix}\begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \beta_i & \mathbf{I}_k \end{pmatrix},$$

where $\mathbf{I}_k$ is an identity matrix of order k. In our specific case, $k = 6$.

We adopt the Pesaran CCE (2006) approach, which augments the regression with proxies for the unobserved factors. Pesaran suggests using $[\mathbf{d}_t' \quad \overline{\mathbf{z}}_{wt}']$ as observable proxies for the unobserved factors, and $\overline{\mathbf{z}}_{wt}$ indicates the cross-sectional average: $\overline{\mathbf{z}}_{wt} = \sum_{j=1}^{N} w_j\mathbf{z}_{jt}$, and w_j are weights equal to $1/N$. The individual slopes, $\beta_i^{'}$, or their mean can be consistently estimated by regressing y_{it} on $\mathbf{x}_{it}, \mathbf{d}_t$ and $\overline{\mathbf{z}}_{wt}$. This type of estimator is referred to as a common correlated effect estimator. In particular, Pesaran (2006) proposes two estimators of the individual coefficients' mean, β : the CCE pooled estimator (CCEP) and the mean group estimator known as CCEMG, which is obtained by averaging the country-specific estimates following Pesaran and Smith (1995).

Some remarks are in order. First, and very important, this setup introduces endogeneity, whereby the $\mathbf{x}_{it}$ are correlated with the unobservable e_{it} via the correlation between ξ_t and $\mathbf{x}_{it}$. As Kapetanios et al. (2011) note, standard approaches that neglect common factors fail to identify β_i'; instead, they yield an estimate of:

$$\kappa_i' = \beta_i' + \varrho_i'\mathbf{\Gamma}_i'^{-1}. \tag{21}$$

The estimation bias is thus a function only of the factor loadings ϱ_i' and $\mathbf{\Gamma}_i'$.[7] Second, specifying a factor loading matrix, $\mathbf{C}_i$, of the type presented above permits a variety of situations, as each variable is allowed to be affected in a specific way by each factor because the typical element of such a matrix, say c_{imj}, measures the country-specific effect of the m^{th} common factor on the j^{th} variable.

[7] To see how this may occur, let us rewrite the model for y_{it} as in Kapetanios et al. (2011) equation (52): abstracting from $\mathbf{d}_t$, assuming that k (the number of regressors)$=$ m(the number of common unobserved factors) and that $\mathbf{\Gamma}_i$ is invertible, we can write the following: $y_{it} = \beta_i^{'}\mathbf{x}_{it} + \varrho_i^{'}\mathbf{\Gamma}_i^{'-1}\left(\mathbf{x}_{it} - \mathbf{v}_{it}\right) + \varepsilon_{it} = \kappa_i^{'}\mathbf{x}_{it} + \varkappa_{it}$, where $\kappa_i^{'} = \beta_i^{'} + \varrho_i^{'}\mathbf{\Gamma}_i^{'-1}$ and $\varkappa_{it} = \varepsilon_{it} - \varrho_i^{'}\mathbf{\Gamma}_i^{'-1}\mathbf{v}_{it}$. Therefore, applying least squares to such an equation consistently estimates $\kappa_i^{'}$ rather than $\beta_i^{'}$.

For example, it may allow some of the unobserved common factors driving the evolution of TFP to also drive the variation in R&D and human capital stocks. It could also allow other factors to specifically affect only one variable in the system (see, e.g., Eberhardt and Teal, 2010).

It is finally worth recalling some recent results concerning the validity of the CCE approach when the underlying DGP is also characterized by weak factors or spatial error correlation. Chudik et al. (2011) also extend the CCE approach by allowing for the presence of both a limited number of strong factors and a large number of weak or semi-strong factors and then show that, even under this extended framework, the CCE method still provides consistent estimates of the slope coefficients. Pesaran and Tosetti (2011) prove that the CCE estimator provides consistent estimates of the slope coefficients and their standard errors under the more general case of a multifactor error structure and spatial error correlation (see also Bresson and Hsiao, 2011, for further simulation results), i.e., when both forms of cross-correlation – weak and strong – characterize the DGP:

$$e_{it} = \varrho_i' \xi_t + \lambda \sum_{j \neq i} w_{ij} e_{jt} + \varepsilon_{it}. \tag{22}$$

This appears to be a very appealing feature because it could be that both forms of dependence are present in the data as shown, for instance, by Bailey et al. (2015b).

3.2 Weak correlation: errors with spatial autocorrelation

We consider the following panel model with spatially autocorrelated error terms:

$$\begin{aligned} y_{it} &= \alpha_i + \beta' \mathbf{x}_{it} + e_{it} \\ e_{it} &= \lambda \sum_{j \neq i} w_{ij} e_{jt} + \varepsilon_{it} \end{aligned} \tag{23}$$

where λ is the spatial autoregressive parameter. For $\lambda = 0$, equation (21) simply reduces to the baseline *a-spatial* specification (8). To obtain a better understanding of such a spatial process, it is useful to examine the so-called *reduced form*. In matrix form, stacking over all individuals for time period t, we have the following:

$$\begin{aligned} \boldsymbol{y}_t &= \alpha + \boldsymbol{X}_t \beta + \mathbf{e}_t \\ \boldsymbol{e}_t &= \lambda \mathbf{W}_N \boldsymbol{e}_t + \boldsymbol{\varepsilon}_t \qquad t = 1, ..., T \end{aligned} \tag{24}$$

where $\boldsymbol{y}_t$ represents the $N \times 1$ vector of log TFP, $\boldsymbol{X}_t$ is the $N \times 4$ matrix of explanatory variables and $\boldsymbol{W}_N$ is an $N \times N$ row-normalized interaction matrix.[8] Under the invertibility condition of $(\boldsymbol{I}_N - \varrho \boldsymbol{W}_N)$, equation (24) can be rewritten in its *reduced-form* representation:

$$\boldsymbol{y}_t = \alpha + \boldsymbol{X}_t \beta + (\boldsymbol{I}_N - \lambda \boldsymbol{W}_N)^{-1} \boldsymbol{\varepsilon}_t \tag{25}$$

[8] According to Lee and Yu (2010), it allows us to consider the parameter space for λ to be a compact subset of $(-1, 1)$. Row-normalization also facilitates the interpretation of the results but is not theoretically necessary (Kelejian and Prucha, 2010).

where $(\boldsymbol{I}_N - \lambda \boldsymbol{W}_N)^{-1}$ is the so-called *global* spatial multiplier. This reduced form has the following important implications. First, in the conditional mean, the total factor productivity in country i will only be affected by the domestic R&D capital stock or human capital stock in the same country i and not by those in any other country j, exactly as in the standard *a-spatial* panel data model. *Therefore there are no spatial spillover effects in this model.* Spillovers enter the benchmark model in eq. (7) only through foreign R&D. In contrast, spatial spillovers exist in the spatially lagged endogenous variable model, sometimes referred to as the mixed regressive spatial autoregressive model, or the SAR model in the spatial econometric literature. Second, and more specifically, one can easily see that a random shock due to unobservable factors (i.e., a shock in the disturbances) in a specific country i not only affects TFP in country i, but it also has an impact on TFP in all the countries of the sample through the inverse spatial transformation, $(\boldsymbol{I}_N - \lambda \boldsymbol{W}_N)^{-1}$: this is the so-called *global spatial diffusion process* of a random shock, which can be expressed as follows:

$$\Xi_y^{\varepsilon} \equiv \frac{\partial \boldsymbol{y}_t}{\partial \boldsymbol{\varepsilon}_t'} \equiv (\boldsymbol{I}_N - \lambda \boldsymbol{W}_N)^{-1} = \boldsymbol{I}_N + \lambda \boldsymbol{W}_N + \lambda^2 \boldsymbol{W}_N^2 + \lambda^3 \boldsymbol{W}_N^3 + ... \tag{26}$$

where Ξ_y^{ε} is an $N \times N$ matrix of the partial derivatives of $\boldsymbol{y}_t$ with respect to the disturbance term, $\boldsymbol{\varepsilon}_t$. This matrix is in general full and not symmetric whatever the sparsity and structure of the interaction matrix $\mathbf{W}_n$.

Therefore, in this model, *a random shock in a given country will spill over the entire sample* and will have a *global* impact. It should be clear that the impact of a random shock hitting a given country is by no means "local" in such a model. The diagonal elements of this matrix represent the *direct impacts* of random shocks affecting each of the countries of the sample, including *feedback* effects, which are inherently heterogeneous in the presence of spatial autocorrelation due to differentiated interaction terms in the $\boldsymbol{W}_N$ matrix.[9] This type of heterogeneity is called *interactive heterogeneity*, in opposition to standard individual heterogeneity in panel data models (Debarsy and Ertur, 2010). The off-diagonal elements of this matrix represent the *indirect impacts* of random shocks (see, e.g., LeSage and Pace, 2009; Kelejian et al., 2006, 2008). Considering column j of the impact matrix, Ξ_y^{ε}, a random shock in a given country j will differently affect each of the countries in the sample. This represents the emission side of the spatial diffusion process. Considering row i, random shocks in any of the countries in the sample will each differently affect country i even if they are common or identical. This represents the reception side of the spatial diffusion process. Using obvious notations, we have the following:

$$\frac{\partial \mathbf{y}_{t,i}}{\partial \varepsilon_{t,i}} \equiv (\Xi_{\mathbf{y}}^{\varepsilon})_{t,ii} \qquad \text{and} \qquad \frac{\partial \mathbf{y}_{t,i}}{\partial \varepsilon_{t,j}} \equiv (\Xi_{\mathbf{y}}^{\varepsilon})_{t,ij} \tag{27}$$

The magnitude of those direct and indirect impacts of random shocks will depend on (1) the degree of interaction between countries, which is governed by the $\boldsymbol{W}_N$ matrix, and (2) the parameter λ, measuring the strength of interactions or cross-sectional correlation between countries. Note that feedback effects are at best of second-order importance and may die out rather quickly, as can be easily seen in equation (26), whereas indirect impacts, although presumably small in magnitude,

[9] Specifically, the own derivative for country i includes the *feedback* effects, where country i affects country j and country j also affects country i, and longer paths that might go from country i to j to k and back to i.

should not be a priori neglected. The QML approach proposed by Lee and Yu (2010) can be used to estimate equation (23). This method is based on a data transformation that eliminates the individual fixed effects and does not suffer from the incidental parameter problem discussed in Neyman and Scott (1948), affecting the direct ML approach. In particular, Lee and Yu (2010) show that direct ML provides consistent estimates of regressor coefficients. However, it provides inconsistent estimates of the variance parameter when T is finite.

3.3 Remarks on econometric modeling

In this section, we have discussed factor and spatial models and how they are related to the notions of weak and strong dependence. It is worth stressing that factor and spatial models not only allow for different forms of cross-sectional dependence but also exhibit differences in the motivations of the econometric modeling underlying such models. In the first approach, the unobserved factors are the latent economic forces driving and strongly connecting the economies. These factors are regarded as nuisance variables introduced to allow for cross-sectional dependence and to capture information in a parsimonious way. The main focus is the estimation and inference of the slope parameters. Moreover, this framework can be effectively used to address endogeneity due to unobservables because the explanatory variables are allowed to be correlated with the factors. By contrast, the spatial models are intended to model interactions and interconnections of cross-sections. Such cross-section interactions will produce differentiated impact coefficients computed from the reduced form of the spatial model. More specifically, the spatial error model produces the so-called global spatial diffusion process of a random shock. In this process, a random shock due to unobservable factors (i.e., a shock in the disturbances) in a specific country i not only affects the dependent variable in country i but also impacts this variable in all countries in the sample. This is the result of the so-called global spatial multiplier. Consequently, this approach focuses not only on the slope parameters but also on the estimation of direct and indirect impacts of random shocks. Finally, it is worth recalling that the CCE approach does not allow for the simultaneous modeling of both forms of dependence. Yet, it provides consistent estimates of the slope coefficients and their standard errors under a generalized DGP with an error term that is the sum of a multifactor structure and a spatial process.

4 Results

4.1 Preliminary estimation results

Table 2 summarizes the results obtained by estimating the benchmark specifications presented above. Column (i) shows the estimated parameters from the specification in equation (6), where the model is estimated using LSDV. The output elasticity of domestic R&D capital stock, θ, is estimated to be 0.084 and is statistically significant. This result is in line with the related empirical literature, such as CH, CHH, Lichtenberg and Van Pottelsberghe (2001) and Keller (2002). The estimated coefficient of human capital (δ) is highly significant and of the same order of magnitude as that obtained by CHH. The inclusion of human capital is relevant not only because it affects productivity and the ability of firms to absorb information, but also because it is potentially correlated with R&D; hence, estimating the model without human capital should bias the coefficient associated with R&D upward. In some previous studies (Barrio-Castro et al., 2002; Frantzen, 2000; Engelbrecht, 1997), this bias has been estimated to be approximately 20% to 30%.

Most previous studies focus on rent spillovers, that is, spillovers that originate solely from economic transactions such as trade and FDI (Table 1). While the body of literature focusing on trade broadly finds similar results to ours, as Fracasso and Vitucci (2013) note, analyzing spillovers generated by FDI presents some conflicting results. Lichtenberg and Van Pottelsberghe (2001) report that outward FDI flows are conducive to international knowledge spillovers while inward FDI flows do not have a significant effect. This result has been questioned by subsequent papers obtaining the opposite finding (Lee, 2006; Bitzer and Kerekes, 2008). Analyses of technology diffusion originating from knowledge spillovers that are not necessarily embodied in specific economic transactions (see, e.g., Griliches, 1979, for a discussion of rent and knowledge spillovers) are more rare. Lee (2006) finds that disembodied channels such as bilateral technological proximity and patent citations between countries have a significant and quite substantial effect in the range of 0.15-0.18. This is broadly the same result obtained by Musolesi (2007) focusing on knowledge spillovers incorporated into language skills.

Next, we turn to the estimates of the benchmark specification in equation (8), which allows the output elasticities to differ between large and small countries (Table 2, column ii). Clearly, the effect of domestic R&D on TFP is much higher for G7 than for non-G7 countries ($\widehat{\theta}_{G7} = 0.105$, $\widehat{\theta}_{NOG7} = 0.046$) while the opposite is true when focusing on foreign R&D ($\hat{\gamma}_{G7} = 0.032, \hat{\gamma}_{NOG7} = 0.200$). These results are fully consistent with CH and CHH focusing on trade and with Bitzer and Kerekes (2008) who find that non-G7 countries benefit more from inward FDI than G7 countries, while Lichtenberg and Van Pottelsberghe (2001) find the opposite result for outward FDI. Finally, we observe that the effect of human capital is nearly identical for the two groups of countries ($\widehat{\delta}_{G7} = 0.551$, $\widehat{\delta}_{NOG7} = 0.556$).

When the DGP is characterized by heterogeneous slopes, however, fixed effects estimators yield consistent estimates of the mean coefficients only when the number of cross-sectional units approaches infinity. Following Pesaran and Smith (1995), we thus consider the mean group (MG) estimator (Table 2, column iii). While the estimated coefficients of R&D and human capital are of the same order of magnitude as before, that of foreign R&D increases substantially ($\hat{\gamma} = 0.207$). We also note that when considering the specification with G7 dummy interaction (column iv), the results indicate that human capital – which is constructed using the average number of years of schooling – has a larger effect on non-G7 countries. The MG estimator has, however, only asymptotic justification for $T \to \infty$ and generally suffers from the problem of parameter estimate instability. This, in turn, can produce non-significant estimates because the variance of the MG estimator is simply equal to the variance of the individual OLS estimates scaled by a factor of $1/N$. This is exactly what happens in our case.

To obtain a more in-depth understanding on the issue of slope heterogeneity in a framework characterized by cross-sectional independence, we also applied the shrinkage estimators described in Maddala et al. (1997), i.e., the (two-step) empirical Bayes and the iterative empirical Bayes estimators. These estimators can be regarded as a compromise between the unrealistic homogeneity assumption and the unstable heterogeneous estimates. We also applied the hierarchical Bayes approach (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. The results are interesting, possibly even concerning the use of another MG-type estimator, i.e., the CCEMG. The estimated average parameters are very close to those obtained with the MG but the precision increases substantially, obtaining in most cases significant estimates. This latter result is also reported by Baltagi et al. (2004). Detailed results are available upon request.

4.2 Testing for cross-sectional correlation

A widely adopted test, likely due to its useful small-sample properties, is the CD test developed by Pesaran (2004). An interesting feature of this test is that, as Pesaran (2015) demonstrates, the implicit null hypothesis of the CD test is, in the most common cases, weak cross-sectional correlation. This assumption makes the test more appealing from an applied perspective because when estimating a model, only strong cross-sectional correlation may pose serious problems, that is, inconsistency of estimation.[10] More precisely, let us define ϵ as a measure of the degree to which T expands relative to N, as defined by $T = O(N^{\epsilon})$ for $0 < \epsilon \leq 1$ and a being the exponent of cross-sectional correlation introduced in Bailey et al. (2015a), which can take any value in the range $[0, 1]$. The values of a in the range $[0, 1/2)$ correspond to different degrees of weak cross-sectional dependence, whereas values of a in the range $(1/2, 1]$ relate to different degrees of strong cross-sectional dependence, as discussed in CPT. Pesaran (2015) shows that the implicit null of the CD test is given by $0 \leq a < (2 - \epsilon)/4$. Thus, for ϵ close to zero (T almost fixed as $N \to \infty$), such a null hypothesis is $0 \leq a < 1/2$, whereas in the case in which $\epsilon = 1$ (N and $T \to \infty$ at the same rate, as is roughly the case for the data used in this paper, the implicit null of the CD test is given by $0 \leq a < 1/4$. The CD test has been performed both on the residuals of the benchmark specification and on the macro variables. The CD statistics were 11.090, 83.427, 91.708, 93.363 and 81.145 for the residuals, TFP, domestic R&D, foreign R&D and human capital, respectively. They are all highly statistically significant and strongly reject the null hypothesis, suggesting that the exponent of cross-sectional correlation, a, is in the range $[1/4, 1]$.

To obtain a measure of the degree of such correlation and to discriminate between weak and strong correlation, we adopt the method recently proposed by Bailey et al. (2015a) and compute the bias-corrected version of a for all macro variables under study. As in Bailey et al. (2015a), Holm's approach has been preferred over the Bonferroni procedure. These estimates, along with the 90% confidence bands, are reported in Table 3. The exponent of cross-sectional correlation, a, is estimated as approximately 1 for all variables, while the 90% confidence bands are approximately ± 0.06 around the point estimates such that they lie largely above 0.5 and include unity. As cross-country interconnections concerning technology might have changed over the period, it would be of interest to investigate the possible time variations of a. Therefore, we also provide the results using 15 rolling samples of 20 years. Both the point estimates of a and the associated confidence bands

[10] Note that robust inference methods (see, e.g., Driscoll and Kraay, 1998) can be considered under weak dependence.

appear to be very stable over time, mimicking the full sample results. This is a very clear-cut result not only indicating the presence of strong cross-sectional correlation but also being consistent with the factor literature, which typically assumes that all factors have the same cross-sectional exponent of $a = 1$ (Stock and Watson, 2002; Bai and Ng, 2002). As also suggested by Bailey et al. (2015b), however, this result does not exclude the possibility that both forms of dependence – weak and strong - are present in the data. While Bailey et al. (2015b) model de-factored residuals with spatial tools in order to study interconnections across units, our main goal is obtaining consistent estimates of the slope parameters and the CCE approach can be used. The second implication is that we do not encounter the problems arising when the assumption that $a = 1$ fails, which are discussed in Chudick et al. (2011), Kapetanios and Marcellino (2010) and Onatski (2012). Finally, because estimation and inference on a in Bailey et al. (2015a) is developed only for the case of stationary variables, while the variables in this paper (as discussed in Section 4.3) are likely to contain a unit root (due to nonstationary unobserved common factors), we also estimate a and the corresponding confidence band for first-differenced variables. The point estimates of a for the first-differenced variables are slightly lower than those associated with the variables expressed in levels. They range between 0.816 and 0.954, but the 90% confidence bands still lie substantially above 0.5 in all cases.

4.3 Testing for unit roots

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). We report here the main conclusions of our analysis. While most previous works find evidence of nonstationary variables by applying first-generation tests (see, among others, CHH), we provide a more nuanced and thorough picture. On the one hand, the use of augmented ADF-type specifications (Pesaran, 2007; Pesaran et al., 2013), which by construction exclude the possibility of the factors having unit roots, indicates the rejection of the null hypothesis. On the other hand, the adoption of tests decomposing the panel into deterministic, common and idiosyncratic components (Bai and Ng, 2004; Moon and Perron, 2004) suggest that while the unobserved idiosyncratic component of the variables under study is stationary, the unobserved common factors seem to be nonstationary. The main results are provided in Tables 4, 5 and 6, while a more detailed discussion and all of the results are available in online appendix 1.

These results are highly relevant from an empirical perspective and are related to the recent work by Kapetanios et al. (2011). They partition the vector of observed common factors as $\mathbf{d}_t = \left(\mathbf{d}_{1t}', \mathbf{d}_{2t}'\right)'$, where $\mathbf{d}_{1t}$ is an $l_1 \times 1$ vector of deterministic components and $\mathbf{d}_{2t}$ is an $l_2 \times 1$ vector of unit root observed common factors, with $l_1 + l_2 = l$, and then suppose that the $(l_2 + m) \times 1$ vector of stochastic common effects $\mathbf{h}_t = \left(\mathbf{d}_{2t}', \xi_t'\right)'$ follows a multivariate unit root process. Both analytical results and a simulation study indicate that the CCE approach is still valid when the unobserved factors are allowed to follow unit root processes.

4.4 Main estimation results

We present the results obtained using both pooled and fully heterogeneous estimators, i.e., the CCEP and CCEMG. In our view, they have relative advantages and disadvantages, as *"the truth probably lies somewhere in between. The parameters are not exactly the same, but there is some similarity between them"* (Maddala et al. 1997, p. 91).

While the main focus of this paper is on geographic R&D spillovers, we also present the results obtained using alternative specifications and different definitions of foreign R&D that have been used in previous empirical studies that neglect the role of cross-sectional correlation. This might allow us to provide a richer contribution to the empirical literature on international R&D spillovers.

The CCEP results are presented in Table 7. There are eight specifications. In the first, we use the same regressors as in our benchmark specification (equation (7)). Then (column (ii)), we estimate the model using a measure of foreign R&D capital incorporating information on bilateral imports and adopt the same definition proposed by Lichtenberg and van Pottelsberghe de la Potterie (1998), $\log S_{it}^{f\text{-}LP}$. In the third and fourth specifications, as in CHH, the foreign R&D variable – $\log S_{it}^{f}$ and $\log S_{it}^{f\text{-}LP}$, respectively– is interacted with m_{it}, i.e., the share of imports in the GDP of country i at time t. In the next specifications, similar to CHH and Lee (2006), we consider simultaneously alternative measures of foreign R&D. Specifically, in column (v) we consider both geographic proximity, $\log S_{it}^{f}$, and trade, $\log S_{it}^{f\text{-}LP}$, as channels of technology diffusion, while in column (vi), we simultaneously introduce the interactive variables, $m_{it} * \log S_{it}^{f}$ and $m_{it} * \log S_{it}^{f\text{-}LP}$. Then, in column (vii), we estimate a similar model to Fracasso and Marzetti (2013) containing both trade-unrelated geographic spillovers, $\log S_{it}^{f}$, and trade-related geographic spillovers, $m_{it} * \log S_{it}^{f}$. Similarly, the eighth specification contains both $\log S_{it}^{f\text{-}LP}$ and $m_{it} * \log S_{it}^{f\text{-}LP}$.

Regarding the CCEP estimates, we note that the coefficient associated with domestic R&D for G7 countries, $\widehat{\theta}_{G7}$, is always positive, ranging between 0.103 and 0.478, and in almost all cases it is significant. This indicates a much larger effect relative the benchmark fixed effects estimates. However, these estimates provide small negative and not significant values of the coefficient associated with non-G7 countries, $\widehat{\theta}_{NOG7}$. This is a new result in the literature on international R&D spillovers. In our view, it is consistent with the stream of literature suggesting that a critical mass of investments in technology and research is necessary to make such research effective (see, e.g., Röller and Waverman, 2001) and complements Eberhardt et al. (2013), who estimate a knowledge capital production function *à la* Griliches - i.e., a production function augmented with domestic R&D – at the industry level and find that when unobserved common factors are introduced, the effect of R&D is close to zero and no longer significant.

Concerning geographic spillovers (columns (i), (v) and (vii)), we find that for G7 countries, the effect of $\log S_{it}^{f}$ on TFP, $\widehat{\gamma}_{G7}$, is high, ranging from 0.304 to 0.374, and is always statistically significant. For the other countries, the estimated elasticity, $\widehat{\gamma}_{NOG7}$, is much lower, ranging from 0.094 to 0.209, and is never statistically significant. This also differs from the standard fixed effects results. Richer countries are, according to these results, better at adopting foreign technology than are poorer countries. This pattern can be considered consistent with the existence of a minimum level of absorptive capacity that allows a country to benefit from foreign technology (e.g., Xu, 2000) and theories describing how technology that is invented in frontier countries is less appropriate for poorer countries (e.g., Basu and Weil, 1998). Finally, human capital ($\widehat{\delta}_{G7}$ and $\widehat{\delta}_{NOG7}$), never significantly affects TFP in any of the specifications. This appears to complement some previous studies on growth that report that when addressing an omitted-variable bias by including in the regression variables accounting for the quality of education, the quantity of education has no further effect (e.g., Hanushek and Kimko, 2000). In our view, this is a particularly relevant result because the factor approach not only allows us to relax the assumption of cross-sectional independence but is also a parsimonious way to capture information from the data and to address the problem of (possibly many) correlated unobserved variables (e.g., Bai-Ng, 2008, p.114).

When focusing on alternative definitions of foreign R&D, some results are provided. First, comparing spillovers incorporated into geographic proximity and bilateral imports (columns (i) and (ii), respectively) suggests that while the largest and most developed countries benefit more from geographic spillovers that are not necessarily embodied in specific economic transactions, smaller countries benefit more from rent spillovers originating from trade. In column (ii), $\widehat{\gamma}_{G7}^{LP} = 0.017$ and is not significant while $\widehat{\gamma}_{NOG7}^{LP} = 0.058$ and is highly significant. This latter result is also consistent with some previous studies focusing on trade and inward FDI (CH, CHH and Bitzer and Kerekes, 2008). Column (v) suggests that foreign R&D capital incorporating geographic spillovers, S_{it}^{f}, performs somewhat better than the one incorporating trade, $S_{it}^{f\text{-}LP}$. Indeed, when both definitions of foreign R&D are included, only S_{it}^{f} continues to be significant and high in magnitude for G7 countries, while $S_{it}^{f\text{-}LP}$ is no longer significant for either group. This result holds even when using the definition of foreign R&D proposed by CH. We focus on $S_{it}^{f\text{-}LP}$ because it outperforms the foreign R&D constructed using the definition of CH. These results are available upon request.

When the foreign R&D variable is interacted with the import share, the results in columns (iii) and (iv) indicate that both $m_{it} * \log S_{it}^{f}$ and $m_{it} * \log S_{it}^{f\text{-}LP}$ positively and significantly affect the two groups of countries, although the magnitude of the impacts is quite low, with the estimated parameters ranging between 0.014 and 0.038. However, when both variables are simultaneously included in the regression, neither of them is still significant (column (vi)). Column (vii), complementing Fracasso and Marzetti (2013), suggests that both geographic spillovers, $\log S_{it}^{f}$, and trade-related geographic spillovers, $m_{it} * \log S_{it}^{f}$, simultaneously affect TFP, but only for G7 countries. When both $\log S_{it}^{f\text{-}LP}$ and $m_{it} * \log S_{it}^{f\text{-}LP}$ are included in the regression (column viii), $S_{it}^{f\text{-}LP}$ has a significant effect only for non-G7 countries while the interactive effect is low for both groups of countries. It is significant for the G7 and almost significant for the others.

Overall, the use of the CCEMG (Table 8) confirms the main findings and provides some further insights. First, note that while the CCE provides some unexpected negative signs, especially for $\widehat{\theta}_{NOG7}$ and $\widehat{\delta}_{G7}$, this problem tends to disappear when considering the CCEMG, which provides estimated parameters with the expected sign in nearly all cases. Only $\widehat{\theta}_{NOG7}$ continues to be generally negative, but it is still never significant and is very close to zero, while $\widehat{\delta}_{G7}$ is now positive (but always insignificant) seven times out of eight and ranges between -0.021 and 0.302. Second, when using the CCEMG, as documented with the standard MG estimator, the significance level decreases substantially with respect to the pooled estimates. This is again because of the instability of the country-specific estimates (detailed country-specific results are available upon request). In general, the most parsimonious specifications yield estimates that are characterized by the highest significance levels. The effect of domestic R&D for the G7, $\widehat{\theta}_{G7}$, decreases with respect to the CCE, now ranging from 0.034 to 0.128, but it is still significant five times out of eight. It is also interesting to note that among the different variables used to capture spillovers, the most significant is the interaction $m_{it} * \log S_{it}^{f}$.

Two remarks are in order. A first remark concerns inference. As the rank condition, $Rank(E(\varrho_i, \mathbf{\Gamma}_i)) = m$, may not hold in practice, this inference is obtained using the nonparametric variance estimator proposed by Pesaran (2006), which can be used irrespective of whether the rank condition holds (see also Pesaran and Tosetti, 2011).[11] Second, note that foreign R&D is a weighted average of domestic R&D, i.e., $\mathbf{1}_{G7} \log \sum_{j \neq i} \exp(-d_{ij}) S_{jt}^{d}$ and $\mathbf{1}_{NOG7} \log \sum_{j \neq i} \exp(-d_{ij}) S_{jt}^{d}$, for G7 and non-G7 countries, respectively, and that the CCE procedure introduces other averages of domestic R&D: $\mathbf{1}_{G7}\overline{\log S_t^d}$ and $\mathbf{1}_{G7}\overline{\log S_t^d}$. One may wonder whether and how these quantities are correlated and if this correlation may affect the results. A closer examination of the data indicates that the correlation between the observed variables, $\mathbf{z}_{it}$, and the cross-sectional averages, $\overline{\mathbf{z}_t}$, is quite low, ranging from 0.025 to 0.693, and in particular, the correlation coefficients between the crosss-sectional averages of domestic R&D and the foreign R&D variables are very low, ranging between 0.037 and 0.086, and are not much affected by the weighting scheme that is adopted to construct foreign R&D, i.e., geographic distance or trade.[12] Even when considering variables not interacted with the G7 dummy, this correlation is quite low and ranges between 0.405 and 0.573 depending on the adopted definition of foreign R&D. These correlations are of the same order of magnitude as the correlations between foreign R&D stocks and the other cross-sectionally averaged variables. The correlation of cross-sectionally averaged variables among themselves is, instead, very high in all cases, ranging between 0.97 and 0.99.

Next, we try to interpret the latent (possibly non-stationary) strong factors by looking at the coefficients of the cross-sectional averages in the CCE regressions.[13] Such variables approximate the unobserved factors. The CCE allows for country-specific effects of these variables. Table 8 provides the cross-sectional averages of the estimated parameters. The results indicate that the cross-sectional average of TFP has a strong effect, whose mean is generally close to one. Among the averaged regressors, it appears that those referring to the G7 countries have a higher average effect compared to the others, in particular concerning domestic R&D. In our view, this result might suggest the existence of a strong asymmetry between the G7 and the other countries in the way the

[11] Failure of the rank condition can occur if there is an unobserved factor for which the average of the loadings tends to a zero vector or if the number of unobservable factors, m, is larger than $k+1$, where k is the number of regressors. Further note that the consistency and asymptotic normality of the CCE estimator are maintained in the rank deficiency case.

[12] $Corr(\mathbf{1}_{G7}\overline{\log S_t^d}, \mathbf{1}_{G7} \log S_{it}^{f}) = 0.0369$, $corr(\mathbf{1}_{G7}\overline{\log S_t^d}, \mathbf{1}_{G7} \log S_{it}^{f\text{-}LP}) = 0.0327$, $corr(\mathbf{1}_{NOG7}\overline{\log S_t^d}, \mathbf{1}_{NOG7} \log S_{it}^{f}) = 0.0614$, $corr(\mathbf{1}_{NOG7}\overline{\log S_t^d}, \mathbf{1}_{NOG7} \log S_{it}^{f\text{-}LP}) = 0.0861$.

[13] These comments refer to the CCEMG. For the CCEP, the geometry of LS requires that the average of country-specific coefficients associated with $\overline{y}_{it}$ be equal to one, while the average of the coefficients associated with any $\overline{x}_{it}$ have the same value as but the opposite sign of the parameter associated with x_{it}.

factors – such as technological shocks or national policies intended to increase the level of technology – and the country-specific TFP are linked (see e.g., Gregory and Head, 1999).[14] It is worth noticing that the use of the CCE approach has been favoured in this paper for a number of reasons. Firstly, it yields consistent estimates in a variety of situations that have been detailed above. Secondly, it does not require any a priori knowledge of the number of unobserved common factors, when in practice there is often considerable uncertainty about it. At the same time, it introduces the unobserved factors as nuisance variables and does not allow for a direct estimation of them. Nevertheless, approaches based on principal component analysis (e.g., Bai, 2009) allow for such an estimation.[15] Consequently, such approaches may be used to shed light on the behaviour of these latent factors. While an in depth investigation of these latent variables is beyond the scope of the present paper, some preliminary descriptive results of the estimated factors using the approach by Bai (2009) are available upon request.

Finally, comparing the results with those obtained using a spatial panel data model with spatially autocorrelated errors, as in equation (23), may provide some interesting insights into the possible bias occurring when allowing only for weak correlation when strong correlation is present in the data. Therefore, considering an interaction matrix with typical element $w_{ij} = \exp(-d_{ij})/\sum_j \exp(-d_{ij})$ to model interactions between countries i and j, a spatial model has been estimated by the QML approach proposed by Lee and Yu (2010), and the estimation results are presented in Table 9. The results are quite similar to those obtained with the a-spatial baseline specification and suggest i) a high and significant effect of human capital (for both the G7 and the other countries), ii) that non-G7 countries benefit more from geographic spillovers than do G7 countries, and iii) that domestic R&D significantly affects both groups of countries. In summary, when using a spatial model instead of a model with multifactor error structure, we obtain very different results concerning the effect of R&D, geographic spillovers and human capital. This is a relevant result both in terms of econometric modeling and with respect to policy implications. This result may be due to the bias induced by strong cross-sectional correlation/unobservable (correlated) common factors, which is not taken into account in the spatial econometric approach.[16]

[14]As indicated by Pesaran (2006), we have proposed the main results introducing the cross-sectional averages for all the regressors. To complement the discussion about the factors, we have run regressions using a subset of the $\overline{\mathbf{x}_t}$. A first result is that when we introduced only $\overline{\log f_t}$ and $\mathbf{1}_{G7}\overline{\log S_t^d}$ which are the cross-sectional averages characterized by the highest average coefficients, the estimated slope parameters associated with the $\mathbf{x}_{it}$ were of the same order of magnitude than those obtained including all the $\overline{\mathbf{x}_t}$. At the same time, introducing all the cross-sectional averages improves the precision of the estimates. Another interesting result is that despite some variability across specifications, in all cases, these estimates confirmed a high positive effect of domestic R&D for the G7 countries, while the effect of the same variable is very close to zero for non-G7 countries. Finally, in all the specifications that we estimated, human capital continued to have no effect for both groups.

[15]Estimating the factors can be useful for many purposes such as forecasting, instrumental variables estimation, etc. (see e.g., Stock and Watson, 2011).

[16]The results are very robust to the choice of the exponential decay parameter of the spatial autoregressive component.

5 Conclusion

This paper provides an analysis of international technology diffusion by accounting for the role of cross-country correlation when estimating the econometric specification. Theoretical consistency, empirical relevance and exogeneity arguments have allowed us to focus on geographic proximity as a channel of technology diffusion.

Cross-sectional dependence can accounted for using two alternative approaches: unobserved common factors and spatial models. The motivations underlying these two approaches are completely different. The first approach is related to the notion of strong cross-sectional dependence and is based on a parsimonious means of capturing information in a data-rich environment using a small number of unobserved factors, which are allowed to be freely correlated with the conditioning variables. The second approach is related to weak cross-sectional dependence. It is explicitly oriented toward modeling cross-sectional interactions and capturing spatial spillovers or analyzing the global spatial diffusion process of a random shock that occurs in a given country. There are no theoretical or empirical reasons to a priori favor one of them, and the type of cross-sectional dependence should be tested before estimating the most suitable model.

A preliminary analysis based on the CD test proposed by Pesaran (2004, 2015) and on the estimation of the exponent of cross-sectional correlation proposed by Bailey et al. (2015a), a, provides a very clear-cut result in favor of strong correlation. Indeed, first, the CD test strongly rejects the null. Using the results of Pesaran (2015) and given the size of our sample, this result suggests a situation in which $1/4 \leq a \leq 1$, where a is the exponent of cross-sectional correlation introduced in Bailey et al. (2015a), which can take any value in the range $[0, 1]$. The values of a in the range $[0, 1/2)$ correspond to different degrees of weak cross-sectional dependence, whereas values of a in the range $(1/2, 1]$ relate to different degrees of strong cross-sectional dependence, discussed in Chudick et al. (2011). Therefore, we focus on the estimation of a and obtain an estimate of a very close to unity, not only indicating the presence of strong cross-sectional correlation but also being consistent with the factor literature, which typically assumes that $a = 1$. Moreover, before moving to the estimation, we also study the order of integration of the variables of interest using several tests allowing for cross-sectional dependence (Pesaran, 2007; Pesaran et al., 2013; Bai and Ng, 2004; Moon and Perron, 2004). While most previous works find evidence of nonstationary variables by applying first-generation tests (see, among others, 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 result 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 the unobserved factors are allowed to follow unit root processes.

Then, the model is estimated using both pooled and fully heterogeneous estimators, i.e., the CCEP and CCEMG, and some new results are provided. In particular, these estimates suggest that domestic R&D has a strong positive effect for G7 countries while it does not play a significant role the other countries. This might indicate that a critical mass of R&D is necessary to make it truly effective. The findings also cast considerable doubt on the notion that the stock of human capital–

constructed using the average years of schooling – significantly affects TFP. They also complement some previous studies on growth stressing the key role played by the quality of education (e.g., Hanushek and Kimko, 2000) and suggesting that when addressing omitted variable bias, the quantity of education no longer has a significant effect. Specifically concerning international R&D spillovers, some new results are provided. For instance, comparing spillovers incorporated into geographic proximity and bilateral imports suggests that that largest and most developed countries benefit more from geographic spillovers that are not necessarily embodied in specific economic transactions, while smaller countries benefit more from rent spillovers originating from trade.

Finally, the results are compared with those obtained from a spatial error model, providing some interesting insights into the possible bias occurring when allowing only for weak correlation when strong correlation is present in the data.

References

[1] Atkinson, A.B., and Stiglitz J.E. (1969). 'A New View of Technological Change', *Economic Journal*, 79, 573-578.

[2] Bai, J. (2003). 'Inferential theory for factor models of large dimensions', *Econometrica*, 71, 135–17

[3] Bai, J. (2009). 'Panel data models with interactive fixed effects', *Econometrica*, 77, 1229-1279.

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

[5] Bai, J. and Ng, S. (2008), 'Large Dimensional Factor Analysis', Foundations and Trends in Econometrics, 3(2), 89-163.

[6] 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.

[7] 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.

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

[9] Bailey, N., G. Kapetanios, and Pesaran, M.H. (2015a). 'Exponent of Cross-sectional Dependence: Estimation and Inference', *Journal of Applied Econometrics*, on-line.

[10] Bailey, N., Holly, S. and Pesaran, M. H. (2015b). 'A two stage approach to spatiotemporal analysis with strong and weak cross-sectional dependence', *Journal of Applied Econometrics*, on-line.

[11] Barrio-Castro, T.; Lopez Bazo, E. and Serrano Domingo, G. (2002) 'New Evidence on International R&D Spillovers, Human Capital and Productivity in the OECD', *Economics Letters*, 77, 41-45.

[12] Barro, R.J. and Lee, J. (2013). 'A New Data Set of Educational Attainment in the World, 1950-2013', *Journal of Development Economics,* 104,184-198

[13] Basu, S. and Weil, D. (1998). 'Appropriate Technology and Growth', *Quarterly Journal of Economics*, 113, 1025-54.

[14] Behrens, K., Ertur, C. and Koch, W. (2012). 'Dual gravity: using spatial econometrics to control for multilateral resistance', *Journal of Applied Econometrics*, 2012, 27, 773-794.

[15] Bresson, G. and Hsiao, C. (2011), ‘A functional connectivity approach for modeling cross-sectional dependence with an application to the estimation of hedonic housing prices in Paris’, *Advances in Statistical Analysis*, 95, 501-529.

[16] Brock, W.A., and Durlauf, S.N. (2001). ‘Growth Economics and Reality’, *World Bank Economic Review,* 15, 229-72.

[17] Caselli, F. (2005). ‘Accounting for Cross-Country Income Differences’, *Handbook of Economic Growth* (Aghion and Durlauf, eds.), North-Holland.

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

[19] 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.

[20] Chudik, A. and M. H. Pesaran (2015). ‘Large Panel Data Models with Cross-Sectional Dependence: A Survey’. In B. H. Baltagi (Ed.), The Oxford Handbook of Panel Data, pp. 3ñ45. Oxford University Press, New York.

[21] Chudik, A., Pesaran, H. and & Tosetti, E. (2011). ‘Weak and strong cross-section dependence and estimation of large panels’, *Econometrics Journal*, 14, 45-90.

[22] Coe, D.T. and Helpman, E. (1995). ‘International R&D spillovers’, *European Economic Review*, 39, 859–887.

[23] Coe, D.T., Helpman, E. and Hoffmaister, A. (2009). ‘International R&D spillovers and institutions’, *European Economic Review*, 53, 723-741.

[24] Debarsy, N. and Ertur, C. (2010). ‘Testing for Spatial Autocorrelation in a Fixed Effects Panel Data Model’, *Regional Science and Urban Economics*, 40, 453-470.

[25] Driscoll, J. C., Kraay, A. C. (1998). ‘Consistent covariance matrix estimation with spatially dependent panel data’, *The Review of Economics and Statistics*, 80, 549-560.

[26] Durlauf, S. N., Kourtellos, A. and Minkin, A. (2001). ‘The Local Solow Growth Model’, *European Economic Review,* 45, 928-40.

[27] Eaton, J. and Kortum, S. (2002). ‘Technology, Geography, and Trade’, *Econometrica*, 70, 1741-1779.

[28] Eberhardt, M. and Helmers, C. and Strauss, H. (2013). ‘Do spillovers matter when estimating private returns to R&D?’ *The Review of Economics and Statistics*, 95, 436-448.

[29] Eberhardt, M. and Teal. F. (2010). 'Productivity Analysis in Global Manufacturing Production', Economics Series Working Papers 515, Department of Economics, University of Oxford

[30] Engelbrecht, H.J. (1997). 'International R&D spillovers, human capital and productivity in OECD economies: an empirical investigation', *European Economic Review*, 41, 1479–1488.

[31] Ertur, C. and Koch, W. (2007). 'Growth, Technological Interdependence and Spatial Externalities: Theory and Evidence', *Journal of Applied Econometrics*, 22, 1033-1062.

[32] Ertur, C. and Koch, W. (2011). 'A contribution to the theory and empirics of Schumpeterian growth with worldwide interactions', *Journal of Economic Growth*, 6, 215-255.

[33] Forni, M., and Lippi, M. (2001). 'The Generalized Factor Model: Representation Theory', *Econometric Theory*, 17, 1113-1141.

[34] Fracasso, A., and Vittucci Marzetti, V. (2013). 'An empirical note on international R&D spillovers', *Empirical Economics*, 45, 179–191.

[35] Frantzen, D. (2000). 'R&D, human capital and international technology spillovers: a cross-country analysis', *Scandinavian Journal of Economics*, 102, 57–75.

[36] Gregory, A. and Head, A. (1999). 'Common and country-specific fluctuations in productivity, investment, and the current account', *Journal of Monetary Economics*, 44, 423-451.

[37] Grilliches, Z. (1979). 'Issues in assessing the contribution of research and development to productivity growth', *The Bell Journal of Economics*, 10 , 92–116.

[38] Grossman, G., and Helpman, E. (1991). 'Innovation and growth in the global economy', Cambridge, MA, MIT Press.

[39] Hall, R.E. and Jones, C.I. (1999). 'Why Do Some Countries Produce So Much More Output Per Worker Than Others?', *The Quarterly Journal of Economics*, 114, 83-116.

[40] Hanushek, E.A. and Kimko, D.D. (2000). 'Schooling, labor force quality, and the growth of nations', *American Economic Review*, 90, 1184–208.

[41] Hong, E. and Sun, L. (2011). 'Foreign Direct Investment and Total Factor Productivity in China: A Spatial Dynamic Panel Analysis', *Oxford Bulletin of Economics and Statistics*, 73, 6, 771-791.

[42] 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.

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

[44] Kao, C. Chiang, M.H. and Chen, B. (1999). 'International R&D Spillovers: An Application of Estimation and Inference in Panel Cointegration', *Oxford Bulletin of Economics and Statistics*, 61, 691-709.

[45] Kapetanios, G. and M. Marcellino (2010). Factor-GMM estimation with large sets of possibly weak instruments. Computational Statistics and Data Analysis 54 (11), 2655.2675.

[46] Kapetanios, G., Pesaran, M.H. and Yamagata, T. (2011). 'Panels with Nonstationary Multifactor Error Structures', *Journal of Econometrics*, 160, 326–348.

[47] Kelejian, H. H. and Prucha, I. R. (1998). 'A generalized spatial two stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances', *Journal of Real Estate Finance and Economics*, 1998, 17, 99-121.

[48] Kelejian, H. H. and Prucha, I. R. (1999). 'A generalized moments estimator for the autoregressive parameter in a spatial model', *International Economic Review*, 1999, 40, 509-533.

[49] Kelejian, H. H. and Prucha, I. R. (2001). 'On the Asymptotic Distribution of The Moran I Test Statistic with Applications', *Journal of Econometrics*, 2001, 104, 219-257.

[50] Kelejian, H. H. and Prucha, I. R. (2010). 'Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances', *Journal of Econometrics*, 157, 53-67.

[51] Keller, W. (2002). 'Geographic Localization of International Technology Diffusion', *American Economic Review*, 92, 120-42.

[52] Keller, W. (2004). 'International Technology Diffusion', *Journal of Economic Literature*, 42, 752-782.

[53] Krugman, P. and Venables, A. (1995). 'Globalization et the Inequality of Nations', *Quarterly Journal of Economics*, 110, 857-880.

[54] Lee, G. (2006). 'The effectiveness of international knowledge spillover channels, *European Economic Review*, 50, 2075–2088.

[55] Lee, L., and Yu, J. (2010). 'Estimation of Spatial Autoregressive Panel Data Models With Fixed Effects', *Journal of Econometrics,* 154, 165-185.

[56] Lee, L., and Yu, J. (2012). 'QML estimation of spatial dynamic panel data models with time varying spatial weights matrices', *Spatial Economic Analysis*, 2012, 7, 31-74.

[57] 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.

[58] Lichtenberg, F.R. and Van Pottelsberghe de la Potterie, B.P. (1998). 'International R&D spillovers: A comment', *European Economic Review*, 42, 1483-1491.

[59] Lichtenberg, F.R. and Van Pottelsberghe de la Potterie, B.P. (2001). 'Does foreign direct investment transfer technology across borders?', *The Review of Economics and Statistics*, 83, 490–497.

[60] 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.

[61] 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.

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

[63] Musolesi, A. (2007). 'Basic stocks of knowledge and productivity: further evidence from the hierarchical Bayes estimator', *Economics Letters*, 95, 54-59.

[64] Onatski, A. (2012). 'Asymptotics of the principal components estimator of large factor models with weakly influential factors', *Journal of Econometrics,* 168, 244.258.

[65] Neyman, J., and Scott, E.L. (1948). 'Consistent estimates based on partially consistent observations', *Econometrica*, 16, 1 32.

[66] Pesaran, M.H. (2004), General diagnostic tests for cross section dependence in panels, Cambridge Working Papers in Economics No. 0435, Faculty of Economics, University of Cambridge.

[67] Pesaran, M.H. (2006), 'Estimation and inference in large heterogeneous panels with a multifactor error structure', *Econometrica*, 74, 967-1012.

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

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

[70] Pesaran M.H. (2015), 'Testing weak cross-sectional dependence in large panels', *Econometric Reviews,* 34,1089–1117.

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

[72] Pesaran, M. H. and Tosetti, E. (2011). 'Large panels with common factors and spatial correlation', *Journal of Econometrics*, 161, 182-202.

[73] 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.

[74] Psacharopulos, G. (1994). 'Returns to investment in education: A global update', *World Development*, 22, 1325-1343.

[75] Robinson, P.M. (2011). 'Asymptotic theory for nonparametric regression with spatial data', *Journal of Econometrics,* 165, 5-19.

[76] Robinson, P.M., and Thawornkaiwong, S. (2012). 'Statistical inference on regression with spatial dependence', *Journal of Econometrics* 167, 521-542.

[77] Röller, L.-H., and Waverman, L. (2001). 'Telecommunications Infrastructure and Economic Development: A Simultaneous Approach', *American Economic Review,* 91, 909-923.

[78] Sarafidis, V. (2009). 'GMM Estimation of Short Dynamic Panel Data Models with Error Cross-sectional Dependence', Mimeo.

[79] Sarafidis, V., and Wansbeek, T. (2012). 'Cross-sectional dependence in panel data analysis', *Econometric Reviews,* 31, 483-531.

[80] Spolaore, E. and Wacziarg, R. (2009). 'The Diffusion of Development', *The Quarterly Journal of Economics*, 124, 469-529.

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

[82] Stock, J. H. and M. W. Watson (2011). 'Dynamic factor models'. In: Clements, M.P., Hendry, D.F. (Eds.), Oxford Handbook of Forecasting. Oxford University Press, Oxford.

[83] Xu, B. (2000). 'Multinational Enterprises, Technology Diffusion, and Host Country Productivity Growth', *Journal of Development Economics*, 62, 477-93.

TABLE 1

Some previous studies on R&D international spillovers

Author	Sample	Technology transfer	Method	Foreign R&D
Coe and Helpman (1995)	22 countries, 1971-90	trade	LSDV	.06-.092
Coe et al. (2009)	22 countries, 1971-90	trade	DOLS	.165-.186
	24 countries, 1971-2004	trade	LSDV	.185-.206
	24 countries, 1971-2004	trade	DOLS	.206-.213
Kao et al. (1999)	22 countries, 1971-90	trade	BC-OLS	.09-.125
	22 countries, 1971-90	trade	FM-OLS	.075-.103
	22 countries, 1971-90	trade	DOLS	.044NS-.056NS
Lichtenberg and Van Pottelsberghe (1997)	22 countries, 1971-90	trade	LSDV	.058-.276
Lichtenberg and Van Pottelsberghe (2001)	23 countries, 1971-90	trade	LSDV	.154
	23 countries, 1971-90	FDI	LSDV	-.06NS-.072
	23 countries, 1971-90	trade	FD	.067
	23 countries, 1971-90	FDI	FD	.006NS-.039
Musolesi (2007)	13 countries, 1981-98	trade	HB	.09
	13 countries, 1981-98	FDI	HB	-0.01NS-.004NS
	13 countries, 1981-98	language	HB	.23
Lee (2006)	16 countries, 1981-2000	trade	DOLS	-.02NS-.17
	16 countries, 1981-2000	FDI	DOLS	-.017NS-.034
	16 countries, 1981-2000	patents	DOLS	.157-.183
Keller (2002)	14 countries, 1970-95	geographic distance	NLS	.843
	14 countries, 1970-95	language	NLS	.565
Engelbrecht (2002)	21 countries, 1971-85	trade	LSDV	.220-.305
Bitzer and Kerekes (2008)	17 countries, 1973-2000	trade	FGLS	.009-.02
	17 countries, 1973-2000	FDI	FGLS	-.016-.012
Barrio-Castro et al. (2002)	21 countries, 1971-85	trade	LSDV	.094-.225
	21 countries, 1966-95	trade	LSDV	0.016-0.106
	21 countries, 1966-95	trade	DOLS	.092-.141
Fracasso and Vittucci (2013)	24 countries, 1971-2004	trade	DOLS	.107-.126
	24 countries, 1971-2004	geographic distance	NLS	0.016-0.106

Notes:

LSDV: Least Squares Dummy Variable ; DOLS: Dynamic Ordinary Least Squares; BC-OLS: Bias Corrected OLS; FM-OLS: Fully Modified OLS; FD: First Difference; HB: Hierarchical Bayes; NLS: Non Linear Least Squares; FGLS: Feasible Generalized Least Squares; NS: not significant.

TABLE 2

Benchmark estimation results

	LSDV	LSDV_G7	MG	MG_G7
	(i)	(ii)	(iii)	(iv)
θ	0.084***		0.071	
	(0.0059)		(0.0628)	
γ	0.064***		0.207	
	(0.0082)		(0.1443)	
δ	0.667***		0.627	
	(0.0675)		(0.4844)	
θ_{G7}		0.105***		0.0476
		(0.0325)		(0.0391)
θ_{NOG7}		0.046***		0.022
		(0.0072)		(0..0500)
γ_{G7}		0.032**		0.019
		(0.0129)		(0.0653)
γ_{NOG7}		0.200***		0.188
		(0.0181)		(0.1299)
δ_{G7}		0.551***		0.203
		(0.1153)		(0.2241)
δ_{NOG7}		0.556***		0.424
		(0.1065)		(0.4381)

Notes:

LSDV: Least Squares Dummy Variable ;

LSDV_G7: Least Squares Dummy Variable with G7 interactions;

MG: Mean Group; MG_G7: Mean Group with G7 interactions;

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

Standard errors in brackets.

TABLE 3

Exponent of cross-country correlation of the macro variables

	$\log f$			$\log S^d$			$\log S^f$			$\log H$		
Sample	$\widehat{\alpha}^*_{0.05}$	$\widehat{\alpha}$	$\widehat{\alpha}^*_{0.95}$	$\widehat{\alpha}^*_{0.05}$	$\widehat{\alpha}$	$\widehat{\alpha}^*_{0.95}$	$\widehat{\alpha}^*_{0.05}$	$\widehat{\alpha}$	$\widehat{\alpha}^*_{0.95}$	$\widehat{\alpha}^*_{0.05}$	$\widehat{\alpha}$	$\widehat{\alpha}^*_{0.95}$
Full sample (1971-2004)	0.9450	1.0039	1.0628	0.9467	1.0045	1.0623	0.9478	1.0046	1.0614	0.9273	0.9931	1.0589
Rolling sample 1 (1971-1990)	0.9079	0.9889	1.0699	0.9315	1.0078	1.0841	0.9326	1.0080	1.0833	0.9163	0.9984	1.0804
Rolling sample 2 (1972-1991)	0.9128	0.9887	1.0645	0.9324	1.0078	1.0832	0.9332	1.0080	1.0828	0.9076	0.9935	1.0794
Rolling sample 3 (1973-1992)	0.9158	0.9889	1.0620	0.9335	1.0079	1.0822	0.9339	1.0080	1.0821	0.8931	0.9773	1.0615
Rolling sample 4 (1974-1993)	0.9235	0.9976	1.0716	0.9344	1.0079	1.0813	0.9347	1.0080	1.0813	0.9021	0.9933	1.0845
Rolling sample 5 (1975-1994)	0.9190	0.9951	1.0713	0.9352	1.0079	1.0806	0.9354	1.0080	1.0806	0.8984	0.9881	1.0779
Rolling sample 6 (1976-1995)	0.9207	0.9930	1.0653	0.9357	1.0079	1.0802	0.9359	1.0080	1.0801	0.9002	0.9907	1.0812
Rolling sample 7 (1977-1996)	0.9214	0.9932	1.0650	0.9359	1.0079	1.0800	0.9362	1.0080	1.0798	0.9058	0.9974	1.0891
Rolling sample 8 (1978-1997)	0.9199	0.9951	1.0702	0.9359	1.0080	1.0800	0.9362	1.0080	1.0799	0.9170	1.0069	1.0967
Rolling sample 9 (1979-1998)	0.9236	0.9985	1.0734	0.9356	1.0080	1.0804	0.9361	1.0080	1.0799	0.9087	0.9970	1.0854
Rolling sample 10 (1980-1999)	0.9313	1.0071	1.0828	0.9351	1.0080	1.0809	0.9361	1.0081	1.0799	0.9080	0.9947	1.0814
Rolling sample 11 (1981-2000)	0.9288	1.0072	1.0855	0.9351	1.0080	1.0809	0.9351	1.0081	1.0809	0.9079	0.992	1.0768
Rolling sample 12 (1982-2001)	0.9272	1.0071	1.0870	0.9334	1.0080	1.0825	0.9341	1.0081	1.0819	0.9072	0.9909	1.0746
Rolling sample 13 (1983-2002)	0.9182	0.9981	1.0780	0.9325	1.0080	1.0834	0.9329	1.0081	1.0831	0.9076	0.9909	1.0741
Rolling sample 14 (1984-2003)	0.9165	0.9944	1.0724	0.9318	1.0080	1.0841	0.9319	1.0081	1.0842	0.9093	0.9923	1.0754
Rolling sample 15 (1985-2004)	0.9149	0.9910	1.0670	0.9317	1.0080	1.0842	0.9319	1.0081	1.0842	0.9114	0.9953	1.0793
Full sample (1971-2004), FD	0.8111	0.9012	0.9913	0.7364	0.8169	0.8973	0.8958	0.9548	1.0138	0.8527	0.9399	1.0270

* 90% level confidence bands. FD indicates first-differenced variables.

TABLE 4

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 de-factored 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 pooled test 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 5

Moon and Perron test

	Kernel	$\log f$	$\log S^d$	$\log S^f$	$\log H$
Model with intercept					
$\widehat{r}_{BIC3}$		3	6	6	6
t_a^*	QS	-15.936 (1.7797e-057)	-17.309 (2.0002e-067)	-10.651 (8.6216e-027)	-20.012 (2.1434e-089)
	B	-7.703 (6.6452e-015)	-5.462(2.3436e-008)	-4.646(1.6887e-006)	-5.904 (1.7685e-009)
t_b^*	QS	-16.340 (2.5533e-060)	-18.1155 1.2021e-073)	-11.1456 (2.7606e-029)	-18.899 (58006e-080)
	B	-7.781 (3.5859e-015)	-6.059 (6.8333e-010)	-4.289 (8.9456e-006)	-5.805 (3.3059e-009)
Model with intercept and trend					
$\widehat{r}_{BIC3}$		2	6	6	6
t_a^*	QS	-4.264 (1.0017e-005)	-1.531 (0.0629)	-0.760 (0.2236)	-1.284 (0.0994)
	B	-4.427 (4.7626e-006)	-1.124 (0.1305)	-1.015 (0.1550)	-2.964 (0.0015)
t_b^*	QS	-4.312 (8.0777e-006)	-1.607 (0.0504)	-0.635 (0.2627)	-1.556 (0.0598)
	B	-4.452 (4.2440e-006)	-1.178 (0.1194)	-0.697 (0.2429)	-2.713 (0.0033)

Notes:

p-value between brackets. QS: Quadratic spectral kernel. B: Bartlett kernel. r is the number of factors.

Here, we impose that the maximum number of factors is 6. In almost all cases, the criteria suggest that the number of unobserved factors, r, equals the maximum number we allowed, 6. This is the same result as in Pesaran (2007) and Pesaran et al. (2013) (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 and do not allow the maximum number of factors to be greater than 6.

As a robustness check, available upon request, we also perform the test for all possible values of r in the range $1-20$.

Model with intercept: the null hypotesis is always rejected.

Model with intercept and trend: the null is almost always rejected for $\log f$, $\log S^d$ and $\log H$. It is often not rejected for $\log S^f$.

TABLE 6

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

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

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.

TABLE 7

CCEP

	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
θ_{G7}	0.175	0.423**	0.442**	0.427**	0.251	0.478***	0.103	0.451**
	(0.2080)	(0.1890)	(0.2053)	(0.1999)	(0.2524)	(0.1816)	(0.527)	(0.1986)
θ_{NOG7}	-0.055	-0.060	-0.043	-0.041	-0.101	-0.068	-0.077	-0.057
	(0.0715)	(0.0622)	(0.0386)	(0.0296)	(0.0673)	(0.0451)	(0.0845)	(0.0521)
γ_{G7}	0.329**				0.374*		0.304**	
	(0.1391)				(0.2002)		(0.1516)	
γ_{NOG7}	0.094				0.209		-0.008	
	(0.1571)				(0.1768)		(0.2400)	
γ_{G7}^{LP}		0.017			-0.022			0.0075
		(0.0308)			(0.0200)			(0.0296)
γ_{NOG7}^{LP}		0.058***			0.023			0.042*
		(0.0229)			(0.0229)			(0.0256)
γ_{G7}^{m}			0.031**			-0.0277	0.045***	
			(0.0126)			(0.0639)	(0.0173)	
γ_{NOG7}^{m}			0.014*			-0.060	0.013	
			(0.0085)			(0.0565)	(0.0097)	
$\gamma_{G7}^{m\_LP}$				0.038**		0.057		0.039**
				(0.0169)		(0.0769)		(0.0201)
$\gamma_{NOG7}^{m\_LP}$				0.023**		0.099		0.017AL
				(0.0104)		(0.0716)		(0.0112)
δ_{G7}	-0.450	-0.221	-0.107	-0.254	-0.380	0.236	-0.039	-0.142
	(0.5388)	(0.4454)	(0.5122)	(0.4508)	(0.5871)	(0.5236)	(0.4272)	(0.4623)
δ_{NOG7}	0.222	0.177	0.168	0.216	0.168	0.050	0.045	0.221
	(0.4403)	(0.5500)	(0.4542)	(0.4159)	(0.6794)	(0.4531)	(0.4619)	(0.405)

Notes.

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

Standard errors within brackets are based on the nonparametric variance estimator of eq. (69) in Pesaran (2006).

TABLE 8

CCEMG

	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
θ_{G7}	0.035	0.128**	0.074**	0.074*	0.022	0.139***	0.034	0.112**
	(0.0560)	(0.0549)	(0.0323)	(0.0439)	(0.0990)	(0.0518)	(0.0498)	(0.0449)
θ_{NOG7}	-0.112	0.035	-0.011	-0.018	-0.040	-0.011	-0.051	0.015
	(0.0708)	(0.035)	(0.0629)	(0.0647)	(0.0973)	(0.0658)	(0.1010)	(0.0514)
γ_{G7}	0.510				0.389		0.3728	
	(0.4348)				(0.3011)		(0.3454)	
γ_{NOG7}	0.240				0.075		0.273	
	(0.6709)				(0.7861)		(0.670)	
γ_{G7}^{LP}		-0.004			-0.004			-0.0027
		(0.0098)			(0.0065)			(0.0067)
γ_{NOG7}^{LP}		0.007			-0.002			0.008
		(0.0118)			(0.0119)			(0.0127)
γ_{G7}^{m}			0.007**			0.013	0.012**	
			(0.0032)			(0.023)	(0.0049)	
γ_{NOG7}^{m}			0.009			0.030	0.011**	
			(0.0063)			(0.0319)	(0.0060)	
$\gamma_{G7}^{m\_LP}$				0.006*		-0.010		0.004
				(0.0036)		(0.0237)		(0.0036)
$\gamma_{NOG7}^{m\_LP}$				0.0112		-0.017		0.008
				(0.0095)		(0.0374)		(0.0110)
δ_{G7}	0.302	0.057	0.0614	0.039	0.265	0.206	-0.021	0.132
	(1.0812)	(0.4460)	(0.3987)	(0.4172)	(0.7727)	(0.3791)	(0.8400)	(0.3783)
δ_{NOG7}	0.328	0.7969	0.111	0.033	1.036	-0.432	-0.408	0.163
	(0.328)	(0.5193)	(0.6144)	(0.6999)	(0.9833)	(0.657)	(0.9138)	(0.705)
$\overline{\log f_{it}}$	0.970	0.953	0.978	0.959	0.896	1.118	0.962	0.996
$\mathbf{1}_{G7}\overline{\log S_{it}^{d}}$	-1.496	-0.634	-0.426	-0.463	-1.021	-0.521	-1.977	-0.312
$\mathbf{1}_{NOG7}\overline{\log S_{it}^{d}}$	0.056	0.018	0.021	0.035	0.1724	0.090	0.125	0.004
$\mathbf{1}_{G7}\overline{\log S_{it}^{f}}$	0.362				-0.165		0.456	
$\mathbf{1}_{NOG7}\overline{\log S_{it}^{f}}$	-0.545				-0.141		0.013	
$\mathbf{1}_{G7}\overline{\log S_{it}^{f\_LP}}$		0.054			0.042			0.057
$\mathbf{1}_{NOG7}\overline{\log S_{it}^{f\_LP}}$		-0.039			-0.030			-0.009
$\mathbf{1}_{G7}\overline{m_{it}\log S_{it}^{f}}$			-0.061			-0.474	-0.040	
$\mathbf{1}_{NOG7}\overline{m_{it}\log S_{it}^{f}}$			-0.003			0.011	-0.017	
$\mathbf{1}_{G7}\overline{m_{it}\log S_{it}^{f\_LP}}$				-0.023		0.479		-0.002
$\mathbf{1}_{NOG7}\overline{m_{it}\log S_{it}^{f\_LP}}$				-0.018		-0.024		-0.023
$\mathbf{1}_{G7}\overline{\log H_{it}}$	-0.565	0.069	0.065	0.485	-3.737	-0.790	-3.550	-0.946
$\mathbf{1}_{NOG7}\overline{\log H_{it}}$	0.471	-0.489	0.498	0.445	-0.186	0.465	0.668	0.026

Notes.

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

Standard errors in brackets are based on the variance estimator of eq. (58) in Pesaran (2006).

The values corresponding to $\overline{\log f_{it}}, ..., 1_{NOG7}\overline{\log H_{it}}$ refer to the averages of their estimated heterogeneous coefficients.

TABLE 9

Spatial autoregressive Error Model (QML)

	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
θ_{G7}	0.117***	0.062**	0.144***	0.131***	0.082***	0.208***	0.071*	0.052**
	(3.668)	(2.451)	(5.913)	(5.833)	(2.637)	(6.546)	(1.773)	(2.085)
θ_{NOG7}	0.044***	0.050***	0.093***	0.090***	0.037***	0.083***	0.057***	0.063***
	(6.148)	(7.456)	(14.778)	(14.447)	(5.248)	13.283	(8.177)	(9.155)
γ_{G7}	0.011				-0.016		0.028**	
	(0.932)				(-1.304)		(2.082)	
γ_{NOG7}	0.222***				0.142***		0.179***	
	(11.110)				(5.408)		(9.467)	
γ_{G7}^{LP}		0.126***			0.134***			0.124***
		(5.672)			(5.864)			(5.654)
γ_{NOG7}^{LP}		0.214***			0.125***			0.155***
		(11.247)			(4.969)			(7.498)
γ_{G7}^{m}			0.014			-0.386***	0.033*	
			(0.783)			(-4.588)	(1.706)	
γ_{NOG7}^{m}			0.065***			-0.086**	0.052***	
			(10.047)			(-2.163)	(8.237)	
$\gamma_{G7}^{m\_LP}$				0.039**		-0.002		0.023
				(2.002)		(-0.015)		(1.214)
$\gamma_{NOG7}^{m\_LP}$				0.084***		0.470***		0.056***
				(10.393)		(5.108)		(6.554)
δ_{G7}	0.516***	0.289***	0.441***	0.413***	0.251**	0.190***	0.583***	0.292***
	(0.397)	(2.657)	(4.075)	(3.791)	(2.096)	(3.826)	(5.146)	(2.739)
δ_{NOG7}	0.397***	0.123	0.653***	0.588***	0.105	0.527***	0.234**	0.089
	(3.627)	(1.009)	(6.312)	(5.633)	(0.880)	(5.045)	(2.188)	(0.749)
λ	0.225***	0.233***	0.144***	0.199***	0.271***	0.316***	0.152***	0.232***
	(5.336)	(5.556)	(3.264)	(4.643)	(6.648)	(8.059)	(3.458)	(5.529)

Notes.

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

Asymptotic t-statistics in brackets

Appendix: 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 correlation 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 correlation 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 be effective in general in conditions under which the pairwise cross-sectional errors' covariances differ across individuals. Moreover, while weak cross-sectional correlation can be addressed with a simple correction of the tests, the presence of strong cross-sectional correlation is more problematic, causing the test statistics to be divergent (Westerlund and Breitung, 2009). As we clearly documented the presence of strong correlation 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, i.e.:

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

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 determine not only whether the data are stationary but also if 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:

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

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 is to determine how many common factors, r, are necessary to capture the existing cross-sectional correlation. To this end, we will employ the information criteria suggested by Bai and Ng (2002). They are constructed in a similar spirit to 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 of the criteria, but following a 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.[17] These criteria are applied to factors estimated by principal components on first differences (Bai and Ng, 2004). Recent

[17] According to Bai and Ng (2002), the IC2 selects the true number of factors and dominate 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 of 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.

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 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 of 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 4**). As a robustness check that is 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 $1-20$.

It is interesting to note that the nonstationarity of the idiosyncratic components can be tested without knowing whether the factors are stationary, and vice versa. All that one need know is the total number of factors, r. This is why, 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 de-factored components. Pooling, however, requires cross-sectional independence of the idiosyncratic components. As the idiosyncratic components in a factor model can only be weakly correlated across units, by construction, while the factors involve strong correlation, it appears that the pooled tests based on de-factored components are 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 the rejection of the null hypothesis of nonstationarity of the idiosyncratic components for all of the 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, while for TFP, the null is rejected 19 times out of 20. 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 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, while 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-cut picture: for all of the variables, whatever 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 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 de-factored, and then panel unit root test statistics based on de-factored data are 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 de-factored 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 propose the implementation of feasible statistics - t_a^* and t_b^* - based on the estimation of long-run variances. To assess the robustness of the results to the choice of the kernel function used to estimate the long-run 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, that is, 6 (**table 5**). 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 it is not for foreign R&D. As with PANIC, we then perform the tests for all the 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. 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 gives the cross-sectionally augmented Dickey-Fuller (CADF) statistics; the individual CADF statistics are used to develop a modified version of the IPS test named 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 some 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 $\delta_i = 0$. 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, allowing the number of factors, $r = k + 1$, to take any value between 1 and 4 and 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 the critical values. When $k = 3$, we implicitly assume that the four observed variables used in the econometric analysis, f_{it}, S_{it}^{d}, S_{it}^{f}, H_{it}, share the same common factors. Pesaran et al. (2013) set the lag order, $\widehat{p} = \left[4\left(T/100\right)^{1/4}\right]$; in our case, this rule gives $\widehat{p} = 3$. For small T, as in our case, they also suggest using the CSB test, having higher power than that of CIPS. We thus provide our main findings in the paper using the CSB test and stetting $p = 3$ but provide in this appendix the results of both CIPS and CISB and for $p = 1, 2, 3$. These results are summarized in **Table A1**. The test outcomes are as follows. When the CSB test is used, the null hypothesis of a panel unit root is rejected in nearly all cases irrespective of the lag order for all the variables under investigation. Few exceptions occur when $r = 1$, which is a case that is always rejected by the above-presented criteria when choosing the number of factors with PANIC and Moon and Perron (2004). 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.[18] Rejection of the null hypothesis is more likely to appear when $r > 1$.

[18]A high level of sensitivity to the number of lags of the AR component is also found when considering 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] 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.

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

[11] 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.

[12] 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.

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

[14] 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.

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

[16] 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.

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

[18] 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

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

	Lag	$\log f$		$\log S^d$		$\log S^f$		$\log H$	
		$CIPS(\widehat{p})$	$CSB(\widehat{p})$	$CIPS(\widehat{p})$	$CSB(\widehat{p})$	$CIPS(\widehat{p})$	$CSB(\widehat{p})$	$CIPS(\widehat{p})$	$CSB(\widehat{p})$
$r = 1$	1	-2.177	0.181	-2.984**	0.118**	-3.031**	0.117**	-1.928	0.151
	2	-2.157	0.155	-2.409	0.083**	-1.983	0.134	-1.838	0.131
	3	-2.319	0.150	-2.055	0.069**	-2.220	0.123	-1.808	0.112
$r = 2$									
$\overline{x}_t = (\overline{\log S^d})$	1	-2.795**	0.109*	-2.923**	0.112*	-1.786	0.116*	-1.773	0.135
	2	-2.784**	0.092**	-2.515*	0.069**	-0.268	0.135	-1.699	0.113
	3	-2.681**	0.082*	-2.174	0.052**	-0.521	0.070**	-1.880	0.095
$\overline{x}_t = (\overline{\log S^f})$	1	-2.886 **	0.094*	-3.300**	0.085**	-4.138 **	0.067**	-1.949	0.120
	2	-2.585*	0.080**	-2.679**	0.052**	-3.158**	0.045**	-1.896	0.100*
	3	-2.525**	0.079*	-2.606**	0.042**	-3.604 **	0.040**	-1.960	0.081*
$\overline{x}_t = (\overline{\log H})$	1	-2.195	0.122	-2.636	0.098**	-4.254**	0.084**	-2.482	0.075**
	2	-2.007	0.114	-2.085	0.067**	-3.260**	0.059**	-2.664**	0.054**
	3	-1.875	0.098	-1.907	0.059**	-3.849**	0.038**	-3.102**	0.037**
$r = 3$									
$\overline{x}_t = (\overline{\log S^d}, \overline{\log S^f})$	1	-2.877**	0.077**	-2.917**	0.080**	-3.250**	0.070**	-1.817	0.113
	2	-2.736**	0.052**	-2.261	0.050**	-1.001	0.057**	-1.777	0.087
	3	-2.327*	0.042**	-2.072	0.032**	-1.008	0.026**	-1.846	0.062
$\overline{x}_t = (\overline{\log S^d}, \overline{\log H})$	1	-2.973**	0.095**	-2.469	0.089**	-2.766*	0.087**	-2.235	0.071**
	2	-2.789**	0.078*	-1.900	0.052**	-0.877	0.059**	-1.961	0.046**
	3	-2.358*	0.055*	-1.336	0.041**	-1.101	0.032**	-2.129	0.029**
$\overline{x}_t = (\overline{\log S^f}, \overline{\log H})$	1	-3.101**	0.061**	-3.162**	0.065**	-3.723**	0.058**	-2.431	0.067**
	2	-2.561*	0.052**	-2.839**	0.034**	-3.166**	0.036**	-2.672**	0.048**
	3	-2.535 **	0.033**	-2.818**	0.023**	-3.565**	0.023**	-2.750**	0.032**
$r = 4$									
$\overline{x}_t = (\overline{\log S^d}, \overline{\log S^f}, \overline{\log H})$	1	-3.419**	0.058**	-2.567	0.061**	-2.576	0.057	-2.318	0.064**
	2	-2.929**	0.035**	-1.739	0.028**	-0.507	0.039**	-2.760**	0.036**
	3	-2.883**	0.016**	-1.050	0.016**	-0.234	0.016**	-2.024*	0.020**

Notes:

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).

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