Laboratoire d'Economie d'Orléans

Document de Recherche

n° 2013-14

« Dépendance individuelle forte et faible : une analyse en données de panel de la diffusion internationale de la technologie »

Cem ERTUR
Antonio MUSOLESI

Laboratoire d'Economie d'Orléans – UMR CNRS 7322 Faculté de Droit, d'Economie et de Gestion, Rue de Blois, B.P. 26739 – 45067 Orléans Cedex 2 - France
Tél : 33 (0)2 38 41 70 37 – 33 (0)2 38 49 45 04 – Fax : 33 (0)2 38 41 73 80
E-mail : leo@univ-orleans.fr - http://www.univ-orleans.fr/leo/

Dépendance individuelle forte et faible : une analyse en données de panel de la diffusion internationale de la technologie

Cem ERTUR† et Antonio MUSOLESI‡

26 novembre 2013

Résumé

Cet article propose un nouvel examen de la diffusion internationale de la technologie, sous l'angle des effets de débordement géographiques de R&D entre pays, en traitant explicitement le problème de la dépendance en coupe transversale dans des modèles de type Coe, Helpman et Hoffmaister (2009).

En appliquant des tests de racine unitaire en panel de différentes générations, nous montrons d'abord que quand le nombre de retards de la composante autorégressive des tests de type Dickey et Fuller augmenté ou quand le nombre de facteurs communs est estimés dans le cadre d'un processus de sélection de modèle, les variables d'intérêt que sont la productivité totale des facteurs, la R&D étrangère, la R&D domestique et le stock de capital humain apparaissent comme étant stationnaires. Ceci nous permet alors d'estimer des modèles relevant de deux approches complémentaires, l'une fondée sur la prise en compte d'un processus d'erreur spatialement autocorrélés et l'autre fondée sur les facteurs communs non observables. Ces approches intègrent différents types de dépendance en coupe transversale en relation avec les concepts de dépendance faible ou forte récemment développés dans la littérature.

Les résultats obtenus conduisent à des interprétations contrastées : alors que dans le modèle avec erreurs spatialement autocorrélés, la R&D locale et la R&D étrangère affectent significativement la productivité totale des facteurs, dans le modèle à facteurs communs non observables seule la R&D étrangère est significative. Ce dernier résultat souligne l'importance des effets de débordements au-delà même de ceux liés directement à la technologie, en effet d'autres effets de débordement globaux ou l'effet de chocs communs semblent l'emporter sur l'impact de la R&D et du stock de capital humain domestiques. Ce résultat rejoint celui obtenu par Eberhardt et al. (2013) dans l'estimation d'une fonction de production de connaissance à la Griliches (1979) et pourrait indiquer l'ampleur des effets de la globalisation.

Classification JEL : C23 ; C5 ; F0 ; O3.

Mots-Clés : Modèle de données de panel ; corrélation en coupe transversale ; économétrie spatiale ; modèles à facteurs ; racine unitaire ; diffusion internationale de la technologie.

† *Université d'Orléans, CNRS, LEO UMR 7322, F45067 Orléans – France*
(email : cem.ertur@univ-orleans.fr)

‡ *GAEL UMR 1215, INRA - Université de Grenoble (France) - 1241 rue des résidences - Domaine Universitaire 38400 Saint Martin d'Hères – France*
(Antonio.Musolesi@grenoble.inra.fr)

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

CEM ERTUR† and ANTONIO MUSOLESI‡

November 26, 2013

Abstract

This paper provides an econometric examination of geographic R&D spillovers among countries by focusing on the issue of cross-sectional dependence. By applying several unit root tests, we first show that when the number of lags of the autoregressive component of augmented Dickey Fuller test-type specifications or the number of common factors is estimated in a model selection framework, the variables (total factor productivity and R&D capital stocks) appear to be stationary. Then, we estimate the model using two complementary approaches, focusing on spatial autoregressive errors and unobserved common correlated factors. These approaches account for different types of cross-sectional dependence and are related to the concepts of weak and strong cross-sectional dependence recently developed in the literature.

JEL classification: C23; C5; F0; O3.

Keywords: panel data; cross-sectional correlation; spatial models; factor models; unit root; international technology diffusion; geography.

†*LEO, University of Orleans UMR 7322 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)

‡*GAEL UMR 1215, INRA- University of Grenoble (France) - 1241 rue des résidences - Domaine Universitaire 38400 Saint Martin d'Hères –France*
(Antonio.Musolesi@grenoble.inra.fr)

1 Introduction

Since the seminal paper by Coe and Helpman (1995), henceforth CH (recently revisited by Coe et al., 2009), 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 analyse 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 or geography (Keller, 2002).

The present paper aims 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 can plausibly be present in CH-type specifications; however, this correlation complicates standard estimation and inference. Specifically, the main goal of this paper is to contrast the spatial approach (Lee, 2004; Lee and Yu, 2009, 2010,) with multifactor error models (Pesaran, 2006). These two approaches account for different types of cross-sectional dependence and are related to the concepts of weak and strong cross-sectional dependence recently developed in the literature; however, it seems difficult to obtain an a priori precise knowledge of the type of cross-sectional dependence that links the cross-sectional units. As this challenge may exist even for other empirical frameworks, the interest of this paper may go beyond the analysis of international technology diffusion. Furthermore, several factors have led to a focus on geographical proximity as a channel for technology diffusion. First, it is theoretically consistent. Keller (2002) and Eaton and Kortum (2002) show that transport costs or geographical barriers create associations between international technology diffusion and geographical distance. Second, the geographic localisation 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, in spite of the theoretical consistency and empirical relevance of geography. Finally, and perhaps most importantly in the context of this paper, which focuses on the methodological issue of cross-sectional dependence, traditional channels of international technology diffusion might create endogeneity 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, might be endogenous with respect to TFP (see, e.g., Hong and Sun, 2011). In contrast, geographical 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, geographical distance may be considered an exogenous proxy for certain endogenous measures of socioeconomic, institutional, cultural or language-based similarities that might enhance the diffusion of technology.

In a preliminary analysis, we study the order of integration of the variables of interest using several tests, most of which allow for cross-sectional dependence (Choi, 2006; Phillips and Moon, 2004; Bai and Ng, 2004; Pesaran, 2007). Our results clearly indicate that when the number of lags of the autoregressive component of augmented Dickey Fuller (ADF)-type specifications or the number of common factors is estimated in a model selection framework, the variables appear to be stationary. This finding is consistent with a subset of the literature that emphasises that many macroeconomic variables may appear to have unit roots due to the low power of the standard tests (Rudesbusch, 1993; Perron, 1989, 1997; Bierens, 1997, Fleissig and Strauss, 1999) but is new in the literature on international R&D spillovers. Based on this result, we adopt estimation techniques that have not been specifically designed to model non-stationary variables.

First, cross-sectional correlation is posed as a result of spatial spillover effects and is modeled by adopting a spatial panel econometric framework allowing for spatially autoregressive disturbances. The estimation is performed using Lee and Yu's (2010) quasi-maximum likelihood (QML) estimator based on a data transformation that eliminates fixed effects, which avoids the incidental parameter problem discussed in Neyman and Scott (1948) that occurs when adopting a direct maximum likelihood (ML) approach. The latter jointly estimates the common slope's parameters and fixed effects. Lee and Yu (2010) focus on a spatial econometric framework and show that when the time dimension is finite, the direct ML estimate of the variance parameter is inconsistent (even if the estimates of the other common parameters are consistent and asymptotically normal), whereas their QML approach provides a consistent estimate of the variance parameter. The authors also demonstrate that the slope's coefficient estimates from the direct approach are identical to the corresponding estimates from the transformation approach.

Standard panel data models are primarily designed to address individual heterogeneity, which can be inherently spatial; however, these models do not allow for individual interactions or spatial autocorrelation. In other words, in the fixed-effects framework, heterogeneity due to individual characteristics, for instance, "absolute" geographical localisation, is easily taken into account by demeaning. However, heterogeneity due to differentiated feedback effects from cross-sectional interactions cannot be addressed. Such effects may arise due to, for instance, the "relative" geographical localisation of countries with respect to each other. In this case, effective estimation requires explicit modeling of spatial autocorrelation. Debarsy and Ertur (2010) label this particular type of heterogeneity *interactive heterogeneity*, which is genuinely spatial by nature. This distinction aims to avoid confusion with what is traditionally called spatial heterogeneity in the literature. This is actually standard individual heterogeneity arising from spatial structural instability in the coefficients or residual variance. Spatial panel data models are specifically designed to address both types of heterogeneity. Pure individual heterogeneity can be captured by fixed effects, whereas interactive heterogeneity can be captured by impact coefficients computed from the reduced form of the spatial model that accounts for the interaction structure between countries.

Cross-sectional dependence may also 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 aimed at raising 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. The empirical setup adopted in this paper builds on 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). In such a framework, the unobserved factors are viewed as nuisance variables introduced to model the cross-sectional dependence in a parsimonious manner, whereas the main focus is on the estimation and inference of the slope parameters.

Spatial error models and the unobserved common factors approach are also related to the recently developed concepts of weak and strong cross-sectional dependence. Although the literature does not provide a unique definition of "weak" and "strong" dependence (see, e.g., Chudik et al., 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 a 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. An appealing feature of this approach is that it provides consistent estimates under a variety of situations. In particular, Pesaran and Tosetti (2011) have shown that the CCE approach is valid even under a generalised data generating process (DGP) with an error term that is the sum of a multifactor structure and a spatial process. Moreover, Chudik et al. (2011) demonstrate by simulation that the CCE approach is robust to the presence of both a limited number of *strong* factors, such as a global policy, and an infinite number of *weak* factors, which can represent spillover effects.

Spatial models and unobserved common factor models have been developed separately in the literature. Although several tests to detect the presence of cross-sectional dependence have been proposed (see Moscone and Tosetti, 2009, for a review and a simulation study), to the best of our knowledge, there is no direct test to discriminate between spatial spillovers and unobserved common factors. Indeed, some of these tests have been developed in a spatial econometric framework and may enable, for example, the discrimination between a standard panel data model under the null hypothesis, i.e., no cross-sectional correlation, and a spatial error model entailing weak cross-sectional correlation under the alternative (see, e.g., Debarsy and Ertur, 2010). Others, such as the CD test proposed by Pesaran (2004), use the pair-wise correlation coefficients on the residuals of a panel regression to test the null of cross-sectional independence against a "generic" alternative of cross-sectional dependence. The main conclusion arising from the thorough simulation study by Moscone and Tosetti (2009) is that *"the choice of the appropriate test should be supported by a priori information (e.g., from the economic theory) about the way that statistical units may be correlated"*. We argue that, when analysing 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 factor models. Comparing the estimation results obtained using the two approaches described above while considering the theoretical and simulation results provided by the literature may allow us to shed new light on the relative importance of the different mechanisms through which technology spillovers across countries may occur. Recently, Pesaran (2012) showed that the implicit null of the CD test depends on the relative expansion rates of N and T and might represent weak cross-sectional correlation rather than cross-sectional independence; in such a case, the CD test makes it possible to detect strong cross-sectional correlation if the null is rejected. We thus also argue that using both types of tests may be of substantial empirical relevance because it may allow researchers to discriminate among cross-sectional independence and weak and strong correlation, whereas the separate use of these tests only allows for the discrimination between two of the three possible situations. Because having a priori knowledge on the type of cross-sectional dependence that links the cross-sectional units may be difficult in practice in many other empirical frameworks (e.g., production functions, innovation functions, labor demand equations, etc.), the interest of this paper may go beyond the analysis of international technology diffusion.

The remainder of the paper is organised as follows: section 2 describes the baseline model and provides some preliminary fixed-effects estimates. Section 3 investigates the level of integration of the variables under examination. Section 4 extends the benchmark specification by allowing for cross-sectional dependence, and section 5 presents the results. Finally, section 6 summarises and concludes.

2 Model specification and preliminary results

2.1 Benchmark econometric model

The baseline econometric model is an extended version of that adopted by CH, as modified by Coe et al. (2009) by including human capital on the right-hand side of the equation:

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

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 take into account 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 linearised by taking logs:

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

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 technological diffusion are trade, FDI and language skills. For instance, Coe et al. (2009) 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 the present paper, to construct foreign R&D capital stock, we follow the theoretical literature on geographical spillovers such as Keller (2002) and Eaton and Kortum (2002) and propose a specification of foreign R&D that incorporates the notion that the impact of foreign R&D is a decreasing function of geographical 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, $\omega_{ij} = \exp(-\varphi d_{ij})$, such that

$$S^f_{it} = \sum_{j \neq i} \exp(-\varphi d_{ij}) S^d_{jt} \tag{4}$$

where d_{ij} represents the geographic distance between country i and country j. 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.[1] Combining equations (4) and (5):

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

If there are positive geographical spillovers (if foreign R&D enhances domestic productivity, $\gamma > 0$), then a positive value of φ indicates that the impact of such spillovers decreases non-linearly with distance, whereas a negative value of φ suggests that the benefits of foreign R&D are increasing with distance. Finally, $\varphi = 0$ indicates that the impact of spillovers does not depend on the distance separating two countries. To allow the impact of foreign R&D capital to differ between the G7 countries and the others, the *benchmark* specification we adopt is a simple variant of eq.(6) that has been widely used in the literature:

$$\begin{aligned} \log f_{it} = {} & \alpha_i + \beta \log S_{it}^d + \gamma_{G7} \mathbf{1}_{G7} \log \sum_{j \neq i} \exp(-\varphi_{G7} d_{ij}) S_{jt}^d + \\ & + \gamma_{NOG7} \mathbf{1}_{NOG7} \log \sum_{j \neq i} \exp(-\varphi_{NOG7} d_{ij}) S_{jt}^d + \delta \log H_{it} + e_{it} \end{aligned} \tag{7}$$

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

Because one of our main objectives is the comparison of the results with previous studies on international R&D spillovers, which do not consider the issue of cross-sectional correlation, our main source is the CH data set, which has been widely used in the literature (see Table 1). This data set is a balanced panel of 21 OECD countries plus Israel observed over the period 1971-90. 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 (2001), as in Coe et al. (2009). Finally, the distance between two countries is calculated as the spherical distance between capitals.

2.2 Preliminary estimation results

Table 2 summarises the results obtained by estimating the benchmark specifications presented above. Columns (ii) to (v) show the estimated parameters from the specification in eq.(6), where

[1] with slope 0.134 for $0 < Edu \leq 4$, 0.101 for $4 < Edu \leq 8$, and 0.068 for $Edu > 8$.

the model is estimated using Nonlinear Least Squares as in Keller (2002). The output elasticity of domestic R&D capital stock, β, is estimated to be 0.067 and is statistically significant. This result is in line with the related empirical literature, such as Coe and Helpman (1995), Coe et al. (2009), 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 found by Coe et al. (2009).[2] Next, we focus on the output elasticity of foreign R&D capital stock incorporated into the geographical technology transfer channel and how the effectiveness of such spillovers decreases with distance (i.e., we focus on the parameters γ and φ). The output elasticity of foreign R&D capital stock incorporated into the geographical technology transfer channel (γ) is estimated to be 0.042 and is significant at the 1% level. In other words, we find evidence of positive (but small in magnitude) geographical spillovers across countries. It is interesting to compare this result with those obtained using alternative technology transfer channels. There is a large body of literature focusing on trade and FDI that generally finds a larger point estimate but presents conflicting results regarding statistical significance (Table 1). Conversely, analyses of technology diffusion via language skills are rare. Musolesi (2007) finds a significant and quite high estimate (approximately 0.2) for the coefficient associated with foreign R&D incorporated into language skills. The positive estimate of φ suggests that the impact of such spillovers decreases with distance. This result is consistent with Bottazzi and Peri (2003), who find that R&D spillovers are small in magnitude and highly localised in European regions. Finally, we turn to the estimates of the benchmark specification in eq.(7), which allows the output elasticity with respect to foreign R&D to differ between large and small countries (Table 2, columns vi to xi), and focus on the parameters $\gamma_{G7}, \gamma_{NOG7}, \varphi_{G7} and \varphi_{NOG7}$. Such a specification will be extended in the following sections to accommodate cross-sectional dependence. Clearly, both elasticities are significant, and the effect of foreign R&D on TFP is much higher for G7 than for non-G7 countries ($\hat{\gamma}_{G7} = 0.170, \hat{\gamma}_{NOG7} = 0.026$). This result is similar to that of Lichtenberg and Van Pottelsberghe (2001), who focus on FDI spillovers. We also find that the effectiveness of such spillovers decreases with distance more quickly for G7 than for non-G7 countries ($\hat{\varphi}_{G7} > \hat{\varphi}_{NOG7}$). In other words, the spillovers are more localised for G7 countries than for smaller countries. These results suggest the existence of substantial differences between richer and poorer countries in terms of how effective they are in adopting foreign technology. Richer countries are, according to our results, better at adopting foreign technology than poorer countries. This pattern can be seen as consistent with the existence of a minimum level of absorptive capacity allowing a country to benefit from foreign technology (see, 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).

[2]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%

3 Panel Unit Root Tests

We first present the results obtained using 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). These tests are comparable because they allow the same degree of heterogeneity,[3] both tests combine the information obtained from the N independent individual tests and (at least when linear trends are included in the deterministic component and the errors are serially correlated) both tests obtain their asymptotic properties by first sending T to infinity and then N to infinity, $(T, N) \rightarrow_{\text{seq}} \infty$.[4]

In performing the tests, we make the following choices: i) because the series are clearly trending, linear time trends have been included in the deterministic component, and ii) the selection of the lag order of the autoregressive components k has to be performed carefully because it is well known that ADF-type tests are highly sensitive to this choice. There is, of course, a delicate balance between choosing a k that is sufficiently large to allow for serially uncorrelated residuals and, simultaneously, sufficiently small such that the model is not overparameterised. Moreover, this choice is crucial because if k is overestimated this may decrease the power of the test while if it is underestimated, this may invalidate the asymptotic distribution of the test. Therefore, the order of the (individual) AR components has been chosen using alternative criteria (AIC, SBC, HQIC) subject to a maximum lag of 3; this maximum seems to be a reasonable point of departure given the annual frequency of the data and the number of observations available for estimation.

The IPS test is based on combining individual ADF t statistics. The reported standardised statistic – the W_{t-bar} – has an asymptotically standard normal distribution and has been shown to perform well even in small samples. The results in table 3 have three major implications. First, the test is highly sensitive to the number of lags of the AR component, k. Second, when the number of lags k is chosen with AIC, BIC or HQIC, there is strong evidence against the unit root hypothesis for all variables. Third, the average number of lags obtained with AIC, BIC or HQIC is generally approximately one, which is consistent with the annual frequency of the data.

Next, we use the Fisher-type tests (Fisher, 1932) provided by Maddala and Wu (1999) and Choi (2001) based on combining the p-values of the N independent test statistics. Two statistics are provided here, labeled P and Z. These values differ in whether they use the inverse chi-square

[3] We define τ_i as the coefficient associated with the autoregressive term in the ADF type regression; Levin, Lin and Chu (2002), among others, propose a test assuming that this coefficient is the same for all cross sections. As noted by Maddala and Wu (1999), whereas the null hypothesis (τ=0) is appropriate in some empirical applications, the alternative (τ <0) seems to be too strong to hold in any relevant case. Im, Pesaran and Shin (2003); Maddala and Wu (1999); and Choi (2001) relax the assumption that $\tau_1 = \tau_2 = \ldots = \tau_N$ under the alternative, allowing some of the individual series to have a unit root.

[4] Although the sequential limit results may appear to be more restrictive than the joint limit results obtained by sending T and N to infinity simultaneously, it has been shown that the sequential and joint limit results are identical under additional moment conditions (Phillips and Moon, 1999). As a practical matter, this means that in both cases, a reasonably large number of time periods and cross-sections are required to implement these tests.

or the inverse normal distribution of the p-values.[5] The Fisher-type statistics (in table 3) fully confirm the IPS tests.

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 a cross-sectional correlation is present; Baltagi et al. (2007) find that when spatial autoregression is present, first-generation tests become oversised, but the tests explicitly allowing for cross-sectional dependence yield a lower frequency of type I errors. As noted by Pesaran (2007), subtracting the cross-sectional averages from the series before applying the panel unit root test can mitigate the impact of cross-sectional dependence.[6] Because cross-sectional demeaning could not work in general in conditions under which the pairwise cross-sectional errors' covariances differ across individuals, new panel unit root tests have been proposed. Let us consider a general specification for contemporaneous correlation in the errors by assuming that they can be decomposed as $u_{it} = \zeta_i^{'}\mathbf{f}_t + v_{it}$, where $\mathbf{f}_t$ is a $m \times 1$ vector of unobserved common factors and $\zeta_i^{'}$ is the associated vector of country-specific parameters. v_{it} is an idiosyncratic term. 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 ($\Delta\bar{y}_t = N^{-1}\sum_{i=1}^{N}\Delta y_{it} = \bar{y}_t - \bar{y}_{t-1}$) 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 (or, eventually, the rejection probabilities) are used to develop a modified version of the IPS test (or the Fisher-type tests), named CIPS (CP and CZ). We performed the CADF (CIPS Z_{t-bar}) test, and the results are presented in table 4[7]. Once again, the results are highly sensitive to the choice of the lag order; when this decision is made using a selection criterion, such as the AIC (or SBC or HQIC), the results regarding the order of integration are mixed: the TFP and domestic R&D appear to be non-stationary, whereas foreign R&D and human capital are found to be stationary. It is worth noting that, as shown by Pesaran (2007), in the case of models with linear time trends, high cross-sectional correlation in the data[8] and $T = 20$, the CIPS tests have low power (but correct size)[9], suggesting that the series might,

[5] Choi's (2001) simulation results suggest the use of the Z statistic, which offers the best trade-off between size and power. With the aim of comparing Fisher-type tests to the IPS test, Choi (2001) finds that the Fisher tests are more powerful than the IPS test in finite samples, and Maddala and Wu (1999) confirm this finding even when the errors are cross-correlated.

[6] We have performed both the IPS test and the Fisher-type tests on the demeaned series, and the (non-reported) results are fully consistent with results obtained without demeaning (reported in table 2).

[7] The Z_{t-bar} statistic is a standardised statistic based on the asymptotic moments of the Dickey-Fuller distribution, whereas the W_{t-bar} statistic is based on the means and variances of the individual t statistics. The Z_{t-bar} and W_{t-bar} are asymptotically equivalent.

[8] The CD test has been performed on each series (results available upon request). The results of such a test strongly reject the null hypothesis for all the variables.

[9] Indeed, let $N = 20$, for $T = 20$ and high cross-sectional correlation, the power of CIPS is only 7% (although its size is correct), for $T = 50$, the power increases to 27%, and for $T = 100$ it is 93% (Pesaran, 2007, table VII).

in fact, be stationary.

Next, to further investigate the order of integration of the variables of interest, we follow Moon and Perron (2004), who allow for m unobserved common factors. They consider these factors to be nuisance parameters and propose a test statistic that uses defactored data obtained by projecting the data onto the space orthogonal to the factor loadings. The authors derive two modified t statistics - denoted t_a and t_b - which have a Gaussian distribution under the null hypothesis, and propose the implementation of feasible statistics - t_a^* and t_b^* - based on the estimation of long-run variances. Because the number of common factors is unknown, a common practice is to estimate it in a model selection framework using a penalised criterion. In so doing, we use the information criteria (IC1) suggested by Bai and Ng (2002) and the BIC3 adopted by Bai and Ng (2004) and Moon and Perron (2004)[10]. 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 almost all cases (except for foreign R&D), these tests strongly reject the unit root hypothesis (table 5).

The last test we perform is that proposed by Choi (2006), who uses a two-way error-component model rather than a factor model. To perform this test, the non-stochastic trend component and cross-sectional correlations are eliminated by GLS detrending (Ellitot et al., 1996) and cross-sectional demeaning. Next, three Fisher-type statistics - denoted P_m, Z, L^* - are obtained by combining p-values from the ADF test applied to each (detrended and demeaned) individual time series. From an applied perspective, such an approach can be viewed as complementary to Moon and Perron (2004). Indeed, Gutierrez (2006) has shown through Monte Carlo simulation that Moon and Perron's tests have a better size than Choi's when the common factor influences the cross-sectional units heterogeneously; however, Choi's test performs well under the more restrictive assumption that the cross-sectional units are homogeneously influenced by the common factor, and in a few cases, it outperforms Moon and Perron's test in terms of power. Additionally, for such a test (in table 6), the choice of the lags is crucial, and when such a choice is made with the AIC (or other non-reported criteria), it clearly indicates that the variables are stationary.

In summary, the unit root tests indicate that when the number of lags of the autoregressive component of ADF-type specifications or the number of common factors is estimated in a model selection framework, the variables appear to be stationary. It is worth noting that a large body of literature has emerged since Nelson and Plosser (1982), initially corroborating their finding of a unit root for most the macroeconomic variables. However, subsequent works have shown that this finding is due to a *"lack of power even against distant alternatives"* (Rudesbusch, 1993).

[10]The BIC3 is a modified BIC criterion that has been shown (Bai and Ng, 2002) to perform better than the others when $\min(T, N) \leq 20$ and T and N 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. Moon and Perron's (2004) simulation indicates that with 20 or more cross-sectional units, their tests provide a precise estimation of the number of factors and show good size, especially the test based on the t_b^* statistic. However, the tests have low power when deterministic trends are included.

These papers have also provided evidence against the unit root hypothesis by allowing breaks or nonlinearities in the trend function (Perron, 1989, 1997; Bierens, 1997) or by exploiting the cross-sectional dimension of panel data sets (Fleissig and Strauss, 1999). Our findings are consistent with this strand of the literature.

4 Handling cross-sectional correlation

4.1 Spatial autoregressive errors

To introduce cross-sectional correlation in eq. (7), we first consider a spatial panel autoregressive error model. Let us define the following:

$$y_{it} = \log f_{it} \tag{8}$$

$$\mathbf{x}_{it} = [\log(S^d_{it}), \mathbf{1}_{G7} \log \sum_{j \neq i} \exp(-\varphi_{G7} d_{ij}) S^d_{jt}, \mathbf{1}_{NOG7} \log \sum_{j \neq i} \exp(-\varphi_{NOG7} d_{ij}) S^d_{jt}, \log(H_{it})]' \tag{9}$$

and assume the following DGP:

$$\begin{aligned} y_{it} &= \alpha_i + \beta' \mathbf{x}_{it} + e_{it} \\ e_{it} &= \lambda \sum_{j \neq i} w_{ij} e_{jt} + \varepsilon_{it} \end{aligned} \tag{10}$$

where $\beta' = [\beta, \gamma_{G7}, \gamma_{NOG7}, \delta]$ and λ is the spatial autoregressive parameter. For $\lambda = 0$, eq. (10) simply reduces to the baseline *a-spatial* specification (7). 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:

$$\begin{aligned} \boldsymbol{y}_t &= \alpha + \boldsymbol{X}_t \beta + \mathbf{e}_t \\ \boldsymbol{e}_t &= \lambda \mathbf{W}_N \boldsymbol{e}_t + \boldsymbol{\varepsilon}_t \qquad t = 1, ..., T \end{aligned} \tag{11}$$

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-normalised matrix,[11], where the typical element is $w_{ij} = \exp(-\phi d_{ij}) / \sum_j \exp(-\phi d_{ij})$, which measures the strength of the interaction between countries i and j. After some manipulations, eq. (11) can be rewritten in its *reduced form* representation:

$$\boldsymbol{y}_t = \alpha + \boldsymbol{X}_t \beta + (\boldsymbol{I}_N - \lambda \boldsymbol{W}_N)^{-1} \boldsymbol{\varepsilon}_t \tag{12}$$

where $(\boldsymbol{I}_N - \lambda \boldsymbol{W}_N)^{-1}$ is the so-called *global* spatial multiplier. The multiplicative process arises because if $|\lambda| < 1$, it follows that $(\boldsymbol{I}_N - \lambda \boldsymbol{W}_N)^{-1} = \boldsymbol{I}_N + \lambda \boldsymbol{W}_N + \lambda^2 \boldsymbol{W}_N^2 + \lambda^3 \boldsymbol{W}_N^3 + ...$ This reduced form has the following important implication: one can easily see that a random shock due

[11] According to Lee and Yu (2010), it allows us to consider the parameter space for λ to be a compact subset of $(-1, 1)$. It also facilitates the interpretation of the results.

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 "spatial diffusion process of a random shock", which can be expressed as follows:[12]

$$\Xi_y^{\varepsilon} \equiv \frac{\partial \boldsymbol{y}_t}{\partial \boldsymbol{\varepsilon}_t'} \equiv (\boldsymbol{I}_N - \lambda \boldsymbol{W}_N)^{-1} = \boldsymbol{I}_N + \lambda \boldsymbol{W}_N + \lambda^2 \boldsymbol{W}_N^2 + \lambda^3 \boldsymbol{W}_N^3 + ... \tag{13}$$

where Ξ_y^{ε} is a $N \times N$ matrix of the partial derivatives of $\boldsymbol{y}_t$ with respect to the disturbance $\boldsymbol{\varepsilon}_t$.

Therefore, in this model, random shocks spill over the entire sample. The diagonal elements of this matrix represent the *direct impacts* of a random shock, including *feedback* effects, which are inherently heterogeneous in the presence of spatial autocorrelation due to differentiated interaction terms in the $\boldsymbol{W}_N$ matrix.[13] 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 these matrices represent the *indirect impacts* of random shocks.[14] Using obvious notations, we have the following:

$$\frac{\partial \mathbf{y}_{t,i}}{\partial \varepsilon_{t,i}} \equiv (\Xi_{\mathbf{y}}^{\varepsilon})_{t,ii} \qquad \text{and} \qquad \frac{\partial \mathbf{y}_{t,i}}{\partial \varepsilon_{t,j}} \equiv (\Xi_{\mathbf{y}}^{\varepsilon})_{t,ij} \tag{14}$$

The magnitude of those direct and indirect impacts 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 and may die out rather quickly, as can be easily seen in eq. (13), whereas indirect impacts, although presumably small in magnitude, should not be a priori neglected.

Moreover, for the matrix Ξ_y^{ε}, the sum across the i^{th} row represents the global impact on the TFP of country i of identical random shocks affecting all of the n countries in the sample. The sum down the j^{th} column yields the global impact on TFP over all of the n countries in the sample of a random shock affecting country j. The average direct impact is therefore defined as $N^{-1}tr(\widehat{\Xi}_{\mathbf{y}}^{\varepsilon})$, whereas the average total impact is defined as $N^{-1}\iota'\widehat{\Xi}_{\mathbf{y}}^{\varepsilon}\iota$, where ι is the $N \times 1$ sum vector. Finally, the average indirect impact is, by definition, the difference between the average total impact and the average direct impact, i.e., $N^{-1}\iota'\widehat{\Xi}_{\mathbf{y}}^{\varepsilon}\iota - N^{-1}tr(\widehat{\Xi}_{\mathbf{y}}^{\varepsilon})$.[15]

A possible method for estimating spatial panel econometric models consists of using the direct ML approach (see, e.g., Elhorst, 2009). However, the direct ML approach may suffer from the incidental parameter problem discussed in Neyman and Scott (1948), who illustrate the inconsistency

[12] Note that the use of the term diffusion only refers here to the spatial dimension, not to the space-time dimension, and therefore might be misleading.

[13] More 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.

[14] LeSage and Pace (2009) present a comprehensive analysis of those effects along with some useful summary measures in the cross-section setting. Their extension to our panel data setting is straightforward. See, also, Kelejian et al. (2006, 2008) for other applications.

[15] See Le Gallo et al. (2005) for an early study on the diffusion of random shocks in the spatial error model in cross-section.

of the variance parameter for the "a-spatial" linear panel data model when the time dimension is finite. Focusing on spatial panel models, Lee and Yu (2010) show that the direct ML provides consistent estimates of regressor coefficients. However, the direct ML provides inconsistent estimates of the variance parameter when T is finite. Thus, Lee and Yu (2010) propose a consistent QML approach based on a data transformation that eliminates the individual fixed effects. They also demonstrate that except for the variance parameter, the estimates of the direct approach are identical to the corresponding estimates of the transformation approach. The QML approach is used to estimate eq. (10).

4.2 Errors with multifactor structure

Next, we consider a model in which the error term has a multifactor structure. Specifically, 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.[16] Let us assume the following DGP:

$$y_{it} = \alpha_i^{'}\mathbf{d}_t + \beta_i^{'}\mathbf{x}_{it} + e_{it} \tag{15}$$

where $\mathbf{d}_t$ is a $l \times 1$ 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 $\beta_i' = [\beta, \gamma_{G7}, \gamma_{NOG7}, \delta]$ can be assumed to be fixed and homogeneous across countries, $\beta_i' = \beta' \; \forall \; i$, or assumed to follow a random coefficients specification: $\beta_i = \beta + \mathbf{v}_i$, $\mathbf{v}_i \sim IID\left(\mathbf{0}, \Theta_v\right)$. The errors e_{it} are assumed to have a multifactor structure:

$$e_{it} = \varrho_i^{'}\xi_t + \varepsilon_{it} \tag{16}$$

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

$$y_{it} = \alpha_i^{'}\mathbf{d}_t + \beta_i^{'}\mathbf{x}_{it} + \varrho_i^{'}\xi_t + \varepsilon_{it} \tag{17}$$

[16] In the macro panel data literature, the standard approach to addressing cross-sectional correlation has been to adopt a seemingly unrelated regressions (SURE) framework and estimate that system of equations by generalised least squares. The SURE approach, however, is not applicable if the panel has a large cross-sectional dimension because it involves nuisance parameters that increase at a quadratic rate as the cross-sectional dimension of the panel is allowed to rise. Moreover, an often questionable assumption behind this approach is that the errors are uncorrelated with the regressors. This has led to the consideration of unobserved factor models. Coakley et al. (2002) propose a principal component approach requiring, however, that the unobserved factors be uncorrelated with the explanatory variables to be consistent. Other estimators based on principal component analysis have been proposed by Kapetanios and Pesaran (2007) and Bai (2009). For a fixed time dimension, however, both are inconsistent under serial correlation or heteroskedasticity (see, also, Sarafidis and Wansbeek, 2012).

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:

$$\mathbf{x}_{it} = \mathbf{A}_i' \mathbf{d}_t + \mathbf{\Gamma}_i' \xi_t + \mathbf{v}_{it} \tag{18}$$

where $\mathbf{A}_i$ and $\mathbf{\Gamma}_i$ are $l \times 4$ and $m \times 4$ factor loading matrices and $\mathbf{v}_{it} = (v_{i1t}, v_{i2t}, , v_{i3t}, v_{i4t})'$. Combining (17) and (18), we finally obtain a system of equations simultaneously explaining TFP, R&D (domestic and foreign) and human capital:

$$\underset{5\times1}{\mathbf{z}_{it}} = \begin{pmatrix} y_{it} \\ \mathbf{x}_{it} \end{pmatrix} = \begin{pmatrix} \log(f_{it}) \\ \log(S_{it}^d) \\ \mathbf{1}_{G7} \log \sum_{j \neq i} \exp(-\varphi_{G7} d_{ij}) S_{jt}^d \\ \mathbf{1}_{NOG7} \log \sum_{j \neq i} \exp(-\varphi_{NOG7} d_{ij}) S_{jt}^d \\ \log(H_{it}) \end{pmatrix} = \underset{5\times l}{\mathbf{B}_i'} \underset{l\times1}{\mathbf{d}_t} + \underset{5\times m}{\mathbf{C}_i'} \underset{m\times1}{\xi_t} + \underset{5\times1}{\mathbf{u}_{it}}, \tag{19}$$

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, $k = 4$.

Under the restrictive assumption of homogeneous factor loadings and homogeneous slope parameters, it is possible to consistently estimate the slope coefficients by ordinary least squares (OLS): one may adopt either a two-way fixed-effects model or a first difference specification augmented with time dummies (Eberhardt and Bond, 2009). To consistently generalise the model to heterogeneous factor loadings, we adopt the Pesaran CCE (2006) approach, which solves the identification problem by augmenting 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: $\overline{\mathbf{z}}_{wt} = \sum_{j=1}^{N} w_j \mathbf{z}_{jt}$, w_j are weights equal to $1/N$. The individual slopes β_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), which also allows the slope parameters to differ across cross-sections. We focus on the former (and assume $\beta_i' = \beta' \ \forall \ i$) because of its direct comparability with the spatial error model in eq. (11).[17]

Some remarks are in order. First, and very important, this set-up introduces endogeneity, whereby the $\mathbf{x}_{it}$ are correlated with the unobservable e_{it} via the correlation between ξ_t and $\mathbf{x}_{it}$.[18] As noted by Kapetanios et al. (2011), standard approaches that neglect common factors fail to identify β_i'; instead, they yield an estimate of $\kappa_i' = \beta_i' + \varrho_i'\mathbf{\Gamma}_i'^{-1}$.[19] Second, specifying a factor-loading matrix $\mathbf{C}_i$ of the kind 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 (eventually being zero) of the m^{th} common factor on the j^{th} variable. 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. Such factors may be linked to oil price shocks or global policies aimed at raising the level of technology, for example. It could also allow other factors to specifically affect only one variable in the system.[20,21]

4.3 Weak and strong cross-sectional correlation

Before commenting on the estimation results, it may be useful to understand how the spatial error model and the factor model are related. In particular, following Sarafidis (2009) and Sarafidis and Wansbeek (2012), spatial autoregressive error dependence and the factor model can be viewed as

[17] The results using the CCEMG are available upon request. As an alternative to the CCEMG, Eberhardt and Bond (2009) and Eberhardt and Teal (2010) propose the Augmented Mean Group estimator, where the Mean Group group-specific regressions are augmented with a preliminary OLS estimate of a "common dynamic process." These estimates are also available upon request. Note, however, that whereas in Eberhardt and Teal' s (2010) production function framework, such a common dynamic process represents the estimated cross-sectional average of the unobservable TFP (the "residual"), in our empirical setup, where the dependent variable is TFP itself, the common dynamic process does not seem to have a straightforward interpretation.

[18] This can be relevant in the CH/K speciation, as also stressed by Keller (2004, p.763). Bai (2009) and Sarafidis and Wansbeek (2012) provide many examples of circumstances under which this may occur, such as production and cost function specifications.

[19] To see how this may occur, let us rewrite the model for y_{it} as in Kapetanios et al. (2011) eq. (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: $y_{it} = \beta_i'\mathbf{x}_{it} + \varrho_i'\mathbf{\Gamma}_i'^{-1}(\mathbf{x}_{it} - \mathbf{v}_{it}) + \varepsilon_{it} = \kappa_i'\mathbf{x}_{it} + \varkappa_{it}$, 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 κ_i' rather than β_i'.

[20] To make this feature more apparent, Eberhardt and Teal (2010) adopt a scalar notation and replace eq. (19) with $x_{kit} = \pi_{ki}'\mathbf{d}_{kt} + \delta_{ki}'\mathbf{g}_{kt} + \vartheta_{1ki}\xi_{1kt} + ... + \vartheta_{lki}\xi_{lkt} + \omega_{kit}$, where $k = 1, ..., 3$ and $\mathbf{g}_{kt}$ are common factors that are specific to each regressor.

[21] Other relevant results have been provided by Kapetanios et al. (2011), who consider the case of non-stationary common effects. More precisely, they partition the vector of observed common factors as $\mathbf{d}_t = (\mathbf{d}_{1t}', \mathbf{d}_{2t}')'$, where $\mathbf{d}_{1t}$ is a $l_1 \times 1$ vector of deterministic components and $\mathbf{d}_{2t}$ is a $l_2 \times 1$ 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 $\mathbf{h}_t = (\mathbf{d}_{2t}', \xi_t')'$ follow 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.

special cases of a *general* error structure. Let us consider the following error structure:

$$e_{it} = (\varrho_i \odot \mathbf{w}_i)' \xi_t + \varepsilon_{it} = \sum_{j=1}^{m} \varrho_{ij} w_{ij} \xi_{jt} + \varepsilon_{it} \tag{20}$$

where $\odot$ denotes the Hadamard product, $\xi_t = (\xi_{1t}, \xi_{2t}, ..., \xi_{mt})'$ is a $m \times 1$ vector of unobserved common factors, $\varrho_i = (\varrho_{i1}, \varrho_{i2}, ..., \varrho_{im})'$ is a $m \times 1$ vector of factor loadings, $\mathbf{w}_i = (w_{i1}, w_{i2}, ..., w_{im})'$ is a $m \times 1$ vector of deterministic bounded weights and $\varepsilon_{it} \sim i.i.d.\,(0, \sigma^2_\varepsilon)$. The eq. (20) allows us to regard ξ_t as shocks, the impact of which are a combination of a heterogeneous factor loading (ϱ_i) with a weight scheme $(\mathbf{w}_i)$. By setting $\mathbf{w}_i = \iota$, ι as a vector of ones, eq. (20) is reduced to the factor structure in (16), $e_{it} = \varrho_i' \xi_t + \varepsilon_{it} = \sum_{j=1}^{m} \varrho_{ij} \xi_{jt} + \varepsilon_{it}$. By setting $\mathbf{w}_i = \iota$ and $m = N$, the N common factor model (Chudick et al., 2011) is obtained: $e_{it} = \sum_{j=1}^{N} \varrho_{ij} \xi_{jt} + \varepsilon_{it}$. In matrix form, stacking over individuals, we obtain the following:

$$\boldsymbol{e}_t = \mathbf{P} \xi_t + \boldsymbol{\varepsilon}_t \tag{21}$$

where $\xi_t = \xi_t = (\xi_{1t}, \xi_{2t}, ..., \xi_{Nt})'$ is a $N \times 1$ vector of unobserved factors, $\mathbf{P}$ is a $N \times N$ matrix of associated factor loadings and with typical element $\{\varrho_{ij}\}$ and $\boldsymbol{\varepsilon}_t = i.i.d.\,(0, \sigma^2_\varepsilon \mathbf{I}_N)$. Finally, by imposing $m = N$, appropriate zero restrictions on $\mathbf{w}_i$ and homogeneity restrictions on ϱ_i, eq. (20) is reduced to a spatial error specification. In particular, by setting $\xi_{mt} = e_{it}$ when $m = i$, for $m = 1, ..., N$, eq. (20) is reduced to a spatial autoregressive error specification, $e_{it} = \varrho \sum_{j \neq i}^{N} w_{ij} e_{jt} + \varepsilon_{it}$,[22] that can be rewritten as follows:

$$\boldsymbol{e}_t = \varrho \mathbf{W}_N \boldsymbol{e}_t + \boldsymbol{\varepsilon}_t \tag{22}$$

where $\boldsymbol{W}_N$ is an $N \times N$ matrix. Under the inversibility condition of $(\boldsymbol{I}_N - \varrho \boldsymbol{W}_N)$, we obtain:

$$\boldsymbol{e}_t = \mathbf{R} \boldsymbol{\varepsilon}_t \tag{23}$$

where $\mathbf{R} = (\boldsymbol{I}_N - \varrho \boldsymbol{W}_N)^{-1}$. It may be also 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 characterise 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 the 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

[22] It is worth noting that eq. (20) also contains the spatial moving average process, i.e., $e_{it} = \varrho \sum_{j \neq i} w_{ij} \varepsilon_{jt} + \varepsilon_{it}$, and the spatial ertror component process, $e_{it} = \varrho \sum_{j \neq i} w_{ij} \varsigma_{jt} + \varepsilon_{it}$ where $\mathbb{E}(\varsigma_{jt}) = 0$, $V(\varsigma_{jt}) = \sigma^2_\varsigma, \mathbb{E}(\varsigma_{jt}\varepsilon_{it}) = 0$, as special cases.

dominated by a few individuals,[23] 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 it satisfies an *absolutely summable condition*, i.e., that $\sum_{j\neq i} |Cov(z_{it}, z_{jt} \mid \mathcal{F}_{ij})| < \infty$, where $\mathcal{F}_{ij}$ denotes the conditioning set of all time-invariant characteristics of individuals i and j.

Spatial error models satisfy, under a standard set of regularity conditions, weak dependence under both definitions. For example, the standard uniform boundedness condition of the interaction matrix suffices 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 relation, consider the single factor error process $e_{it} = \varrho_i \xi_t + \varepsilon_{it}$. With N and T both large and noticing that $\sigma_{ij,t} = cov(e_{it}, e_{jt} \mid \mathcal{F}_{ij}) = \varrho_i \varrho_j \sigma_\xi^2 \neq 0$; therefore, $\sum_{j\neq i} |\sigma_{ij,t}|$ is unbounded, thus entailing strong correlation under SARA. A related concept is that of strong and weak factors (Chudik et al. 2011). Let b be a constant in the range $0 \leq b \leq 1$, and consider the condition $\lim_{N\to\infty} N^{-a} \sum_{i=1}^{N} |\varrho_i| = K < \infty$. 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 also noted by Chudik et al. (2011), the CCE approach 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.

It is finally worth recalling some recent results concerning the validity of the CCE approach when the underlying DGP is also characterised by weak factors or spatial error correlation. Chudik et al. (2011) also 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. 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).

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

5 Results

We first focus on the presentation of the results obtained from the spatial panel autoregressive error model (eq. 10). The model has been estimated by the QML approach proposed by Lee and Yu (2010), and the estimates of the parameters $\beta' = [\beta, \gamma_{G7}, \gamma_{NOG7}, \delta]$ and λ are presented in Table 7.[24] A first general result is that all of the estimated coefficients and standard errors but those concerning human capital are very robust to the choice of ϕ. Indeed, the parameter associated with domestic R&D, β, ranges from 0.07 to 0.08; the coefficient associated with foreign R&D for the G7 countries, γ_{G7}, is approximately 0.18, irrespective of value of ϕ, whereas for the non-G7 countries, the estimated coefficient, γ_{NOG7}, ranges from 0.02 to 0.03. In all cases, they are significant at 5%. Only the effect of human capital varies greatly with ϕ, ranging from 0.01 (and not being significant) for $\phi = 1$ to 0.24 (significant at 5%) for $\phi = 10$. These results are very close to those obtained by applying least squares to the benchmark "a-spatial" specification in eq. (7) and presented in Table 2. This similarity is not surprising because spatial error dependence does not affect the consistency of standard "a-spatial" panel data estimators. There are two gains from modeling spatial error dependence. The first arises with respect to estimation efficiency and the validity of inference. Second, this type of estimation provides information about the strength and the significance of spatial error dependence through the parameter λ and its significance level. In particular, the estimated spatial error parameter, λ, is always positive and significant at 1%, ranging from 0.21 to 0.46. Moreover, the estimation of a spatial autoregressive error model (eq. 10) also allows for the construction of appropriate tests statistics. Adopting the LM marginal test proposed by Debarsy and Ertur (2010) allows us to discriminate between an a-spatial model in eq. (7) under the null hypothesis ($H_0 : \lambda = 0$) against a spatial autoregressive error model (10) ($H_1 : \lambda \neq 0$). The results of this test (Table 8) strongly reject the null hypothesis for all values of ϕ.

Before presenting the results obtained using the CCEP approach, it is worth recalling that in standard panel models, which avoid the use of an interaction matrix, other types of tests have been developed (see, e.g., Moscone and Tosetti, 2009, and Sarafidis and Wansbeek, 2012, for recent surveys). These tests use the pair-wise correlation coefficients on the residuals of a panel regression to test the null of cross-sectional independence. A widely adopted test, likely due to its useful small-sample properties, is the CD test developed by Pesaran (2004). A very interesting feature of this test in the context of this paper is that, as shown by Pesaran (2012), 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. 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

[24]The exponential decay parameter of the spatial autoregressive component, ϕ, will take three different values (1, 5 and 10) to check the sensitivity of the results to its value.

al. (2012), which can take any value in the range $[0, 1]$, with 1 indicating the highest degree of cross-sectional dependence, while $a < 0.5$ and $a > 0.5$ correspond to the cases of weak and strong cross-sectional correlation discussed in CPT, respectively. Pesaran (2012) 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 where $\epsilon = 1$ (N and $T \to \infty$ at the same rate, as is roughly the case of 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 on the residuals of the benchmark specification (eq.(7)) (Table 8). The result of this test is a strong rejection of the null hypothesis, suggesting that the exponent of cross-sectional correlation, a, is in the range $[1/4, 1]$.[25]

Next, we adopt the CCEP approach. Our estimates (in Table 9) are obtained using alternative definitions of the vector d_t of observed common factors. The results are structured as follows. In column (i), as in Mastromarco et al. (2012), a linear time trend is used as an observed common factor, such that $\mathbf{d}_t = t$. In column (ii), more flexibility is allowed by adding a squared time trend, such as $\mathbf{d}_t = (t, t^2)'$. Finally, in columns (iii) and (iv), the oil price, denoted p, is added to the previous two specifications such that $\mathbf{d}_t = (t, p)'$ and $\mathbf{d}_t = (t, t^2, p)'$, respectively.

The results provide a robust picture. The coefficient associated with domestic R&D is always very close to zero, ranging from -.043 to 0.026, and is never significant at standard levels. In other words, both the estimated coefficient of domestic R&D and its significance level have decreased substantially with respect to both the "a-spatial" benchmark specification and the spatial autoregressive error model. Conversely, adopting the CCEP method increases the effect of external R&D. For G7 countries, the effect of foreign R&D on TFP ranges from 0.22 to 0.46, which is statistically significant at least at 5% in all cases, whereas for the other countries, the effect ranges from 0.12 to 0.34 but is only statistically significant for the specification in column (iv). Finally, human capital is found to be insignificant at standard levels, although the estimated coefficient ranges between 0.19 and 0.37.

These results are meaningful with respect to both the magnitude of the estimated coefficients and their significance levels. The very low (and not significant) estimate of the coefficient associated with domestic R&D provides additional support for Eberhardt et al. (2012), who estimate a production function augmented with domestic R&D and find that when unobserved common factors are introduced, the effect of R&D is close to zero and no longer significant. The high values of the estimated coefficients for foreign R&D seem to us a new and relevant result. In other words, our results indicate not only that unobserved common factors matter for enhancing productivity growth, and their inclusion in the model decreases the effect of domestic R&D, as already suggested by Eberhardt et al. (2012), but also suggest that all types of spillovers, both

[25] The full correlation matrix of residuals is available upon request; the average absolute value of the its off-diagonal elements is 0.470. Table 8 also reports the results of other tests based on the pair-wise correlation coefficients and precisely provides the tests by Frees (1995) and Friedman (1937). Both are discussed in Sarafidis and Wansbeek (2012). Both tests strongly reject the null hypothesis of cross-sectional independence.

observed (foreign R&D) and unobserved (common factors), may play a key role in explaining TFP growth.

Concerning the inference provided by these estimates, it may be interesting to recall the Monte Carlo simulations by Pesaran (2006) made under the assumption that the DGP is characterised by unobserved common factors. For $N = T = 20$, these simulations indicate that, whereas the naïve estimators (i.e., the estimators that do not account for cross-sectional correlation, such as the LSDV) are oversised but have high power, the CCE estimators have the correct size but have low power. Moreover, Pesaran and Tosetti (2011) (see Bresson and Hsiao, 2011, for additional results) provide interesting simulations under the assumption that the error is generated by a spatial autoregressive model or is a mixture of a spatial process and a multifactor model. In all cases (for $T = N = 20$), the CCE estimators have better size than any others (including the ML spatial error estimator) but low power compared to the alternative estimators.[26]

These simulation results may have relevant implications in our context, especially for the coefficient associated with foreign R&D for the countries that do not belong to the G7 group. The point estimate of this coefficient is high, taking values of 0.12, 0.16, 0.21 and 0.34 (from column (i) to (iv)), and the corresponding t values, which are also relatively high, take the values of 0.93, 1.11, 1.44 and 2.27, which correspond to p-values taking the values of 0.35, 0.26, 0.15 and 0.02 going from column (i) to column (iv), respectively. This finding might suggest that in some cases, our inference erroneously indicates a non-significant effect of foreign R&D for the non-G7 countries. Finally, it is also interesting to note that both the magnitude of the estimated parameter, γ_{NOG7}, and the corresponding significance level increase with the number of observed common factors introduced in the CCEP framework.

6 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 geographical proximity as a channel of technological diffusion. In particular, the main goal of this paper has been to contrast the spatial approach with the multifactor errors model. These two approaches have been developed quite separately in the literature but can be viewed as complementary. They account for different types of cross-sectional dependence and are related to the notions of weak and strong cross-sectional dependence recently developed in the literature. It is also difficult to obtain a priori precise knowledge on the type of cross-sectional dependence that links the cross-sectional units. As this may be true even for other empirical frameworks, the interest of this paper may extend beyond the analysis of international technology diffusion.

[26] These simulations also show that the CCE estimators are superior to all competitors with respect to bias when the errors are a mixture of a spatial and a multifactor process.

In summary, both the preliminary integration analysis and the estimations provide new and interesting insights. A thorough investigation of the order of integration, contrary to some previous studies on international R&D spillovers, suggests that the variables under study are stationary and, more generally, confirms that detecting the order of integration is an intricate issue that crucially depends on the hypotheses underlying the tests. In our specific case, choosing the number of lags of the autoregressive component of augmented Dickey Fuller test-type specifications or the number of common factors using a model selection framework has been, ex post, crucial to finding evidence of stationary variables. The results of the tests to detect cross-sectional dependence and the estimations suggest that cross-sectional correlation plays a relevant role. Concerning the tests, we first adopted a marginal LM test that strongly rejected a standard panel data model in favor of a spatial error model; second, we performed the CD test that also strongly rejected the null. Using the results by Pesaran (2012) 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. (2012), which can take any value in the range $[0, 1]$, with 1 indicating the highest degree of cross-sectional dependence, and where $a < 0.5$ and $a > 0.5$ correspond to the cases of weak and strong cross-sectional correlation discussed in Chudick et al. (2011), respectively. Concerning estimation and inference, the two approaches produce very different predictions in economic terms. The results obtained using a spatial error model are very close to those obtained by applying least squares to the benchmark "a-spatial" specification. This finding is not surprising because spatial error dependence does not affect the consistency of standard panel data estimators, whereas a possible gain from modeling spatial error dependence arises with respect to estimation efficiency and the validity of inference. For both the "a-spatial" and the spatial error model, domestic and foreign R&D significantly affect TFP, with estimated coefficients that are of the same order of magnitude as those obtained in the related literature. Conversely, estimating the model by allowing for unobserved common factors provides some new findings. The coefficient associated with domestic R&D decreases substantially, approaching zero and becoming non- significant at standard levels; however, the effect of external R&D increases remarkably. In other words, our results indicate not only that unobserved common factors matter for enhancing productivity growth, and their inclusion in the model decreases the effect of domestic R&D, as already suggested by Eberhardt et al. (2012), but more generally suggest that all types of spillovers, both observed (foreign R&D) and unobserved (strong factors), may play a key role in explaining TFP growth.

References

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

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

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

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

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

[6] Bailey, N., G. Kapetanios, and Pesaran, M.H. (2012) 'Exponent of Cross-sectional Dependence: Estimation and Inference', University of Cambridge Working Papers in Economics 1206, Faculty of Economics, University of Cambridge.

[7] 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.

[8] Barro, R.J. and Lee, J. (2001). 'International Data on Educational Attainment: Updates and Implications', *Oxford Economic Papers*, 53, 541-563.

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

[10] Bierens, H. J. (1997). 'Testing the unit root with drift hypothesis against nonlinear trend stationarity, with an application to the US price level and interest rate', *Journal of Econometrics*, 81, 29-64.

[11] Bottazzi, L., and Peri, G. (2003). 'Innovation, Demand, and Knowledge Spillovers: Evidence from European Patent Data', *European Economic Review*, 47, 687–710.

[12] 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.

[13] 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.

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

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

[16] 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.

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

[18] Coakley, J., Fuertes, A.M. and Smith, R. (2002. 'A principal components approach to cross-section dependence in panels', Birkbeck College Discussion Paper 01/2002.

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

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

[21] 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

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

[23] Eberhardt, M. and Helmers, C. and Strauss, H. (2012). 'Do spillovers matter when estimating private returns to R&D?' *The Review of Economics and Statistics*, (In Press.).

[24] Eberhardt, M. and Bond, S. (2009). 'Cross-section dependence in nonstationary panel models: a novel estimator', MPRA Paper 17692, University Library of Munich, Germany.

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

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

[27] Elhorst, J.P. (2009). 'Spatial Panel Data Models', In Fischer M.M., Getis A. (eds.) *Handbook of Applied Spatial Analysis*, pp. 377-407. Springer, Berlin.

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

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

[30] Fleissig, A. R., and Strauss, J. (1999). 'Is OECD real per capita GDP trend or difference stationary? Evidence from panel unit root models', *Journal of Macroeconomics*, 21, 673-90.

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

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

[33] Frees, E. W. (1995). 'Assessing cross-sectional correlations in panel data', *Journal of Econometrics*, 69, 393-414.

[34] Friedman, M. (1937). 'The use of ranks to avoid the assumption of normality implicit in the analysis of variance', *Journal of the American Statistical Association*, 32, 675-701.

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

[36] 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.

[37] 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.

[38] 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.

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

[40] 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.

[41] Kapetanios, G. and Pesaran, M.H. (2007). 'Alternative approaches to estimation and inference in large multifactor panels: small sample results with an application to modelling of asset returns. In: Phillips, G., Tzavalis, E. (Eds.), The Refinement of Econometric Estimation and Test Procedures: Finite Sample and Asymptotic Analysis. Cambridge University Press, Cambridge.

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

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

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

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

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

[47] Lee, L. (2004). 'Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models', *Econometrica,* 72, 1899–1925.

[48] Lee, L., and Yu, J. (2009). 'Some Recent Developments in Spatial Panel Data Models', *Regional Science and Urban Economics*, 40, 255-271.

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

[50] Le Gallo, J., Baumont, C., Dallerba, S., Ertur C. (2005). "On the Property of Diffusion in the Spatial Error Model", *Applied Economics Letters*, 12, 533-536, 2005.

[51] 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.

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

[53] 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.

[54] 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.

[55] Mastromarco, C., Serlenga L. and Shin, Y. (2012). 'Globalisation and technological convergence in the EU', *Journal of Productivity Analysis*, in press.

[56] Moscone; G. and Tosetti, E. (2009). 'A Review And Comparison Of Tests Of Cross-Section Independence In Panels', *Journal of Economic Surveys*, 23, 528-561.

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

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

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

[60] Nelson, C. and Plosser, C. (1982), 'Trends and Random Walks in Macroeconomics Time Series: Some Evidence and Implications, *Journal of Monetary Economics*, 10, 139-162.

[61] Perron, P. (1989). 'The great crash, the oil price shock, and the unit root hypothesis', *Econometrica*, 57, 1361-40

[62] Perron, P. (1997). 'Further evidence on breaking trend functions in macroeconomic variables', *Journal of Econometrics*, 80, 355-385.

[63] 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.

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

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

[66] Pesaran, M. H. (2012). 'Testing weak cross-sectional dependence in large panels', CESifo Working Paper no. 3800.

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

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

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

[70] Rudesbusch, G. (1993). 'The Uncertain Root in real GNP', *American Economic Review*, 83, 264-72.

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

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

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

[74] 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		LSDV	.185-.206
			DOLS	.206-.213
Kao et al. (1999)	22 countries, 1971-90	trade	BC-OLS	.09-.125
			FM-OLS	.075-.103
			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
		FDI		-.06NS-.072
		trade	FD	.067
		FDI		.006NS-.039
Musolesi (2007)	13 countries, 1981-98	trade	HB	.09
		FDI		-0.01NS-.004NS
		language		.23
Lee (2006)	16 countries, 1981-2000	trade	DOLS	-.02NS-.17
		FDI		-.017NS-.034
		patents		.157-.183
Keller (2002)	14 countries, 1970-95	geographic distance	NLS	.843
		language		.565
Engelbrecht (2002)	21 countries, 1971-85	trade	LSDV	.220-.305
Barrio-Castro et al. (2002)	21 countries, 1971-85	trade	LSDV	.094-.225
	21 countries, 1966-95		LSDV	0.016-0.106
			DOLS	.092-.141

Notes:

LSDV: Least Square Dummy Variable ; DOLS: Dynamic Ordinary Least Square; BC-OLS: Bias Corrected OLS;

FM-OLS: Fully Modified OLS; FD: First Difference; HB: Hierarchical Bayes; NLS: Non Linear Least Square; NS: not significant.

TABLE 2

Benchmark estimates

	Basic specification				Specification with G7 dummy					
	β	γ	φ	δ	β	γ_G7	φ_G7	γ_NOG7	φ_NOG7	δ
Coefficient	.067	.042	9.649	.538	.069	.170	15.217	.0267	7.042	.375
Standard errors	.0115***	.0125***	.0125***	.1224***	.0105***	.0183***	.0272***	.0116**	.1801***	.1135***

Notes:

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

TABLE 3

IPS (W_{-t-bar}) and Fisher-type statistics (P and Z)

lag order (k)	Test	$\log f$	$\log S^d$	$\log S^f$	$\log H$
1	W_{-t-bar}	-2.14861 (0.0158)	-1.75052 (0.0400)	-2.33921 (0.0097)	-1.55144 (0.0604)
1	P	64.6462 (0.0229)	70.2988 (0.0071)	71.8629 (0.0050)	60.6723 (0.0483)
1	Z	-2.31288 (0.0104)	-1.58178 (0.0569)	-2.47066 (0.0067)	-1.74831 (0.0402)
2	W_{-t-bar}	0.55338 (0.7100)	0.16992 (0.5675)	-1.15137 (0.1248)	-2.36996 (0.0089)
2	P	29.0843 (0.9593)	49.1626 (0.2740)	47.2918 (0.3397)	65.1072 (0.0209)
2	Z	1.31609 (0.9059)	1.30367 (0.9038)	-0.63744 (0.2619)	-1.75297 (0.0398)
3	W_{-t-bar}	-0.25439 (0.3996)	0.21641 (0.5857)	2.09292 (0.9818)	-2.17194 (0.0149)
3	P	46.5330 (0.3685)	39.1391 (0.6797)	15.3795 (1.0000)	66.5139 (0.0158)
3	Z	0.81728 (0.7931)	1.40485 (0.9200)	3.56924 (0.9998)	-1.33588 (0.0908)
aic	W_{-t-bar}	-3.32945 (0.0004)	-1.84356 (0.0326)	-3.58654 (0.0002)	-2.97393 (0.0015)
aic	P	79.1326 (0.0009)	71.8741 (0.0050)	79.3389 (0.0009)	79.8767 (0.0008)
aic	Z	-3.14469 (0.0008)	-1.1403 (0.1271)	-3.58923 (0.0002)	-2.60986 (0.0045)
		$\bar{k}$= 1.00	$\bar{k}$ =1.55	$\bar{k}$ =1.45	$\bar{k}$ =1.36
bic	W_{-t-bar}	-2.57982 (0.0049)	-1.81427 (0.0348)	-3.68351 (0.0001)	-2.91527 (0.0018)
bic	P	70.0827 (0.0074)	73.6950 (0.0033)	80.9023 (0.0006)	79.4714 (0.0008)
bic	Z	-2.41172 (0.0079)	-1.18474 (0.1181)	-3.68534 (0.0001)	-2.58006 (0.0049)
		$\bar{k}$= 0.77	$\bar{k}$ =1.27	$\bar{k}$= 1.32	$\bar{k}$= 1.32
hqic	W_{-t-bar}	-3.32945 (0.0004)	-1.51511 (0.0649)	-3.58654 (0.0002)	-2.97393 (0.0015)
hqic	P	79.1326 (0.0009)	70.5376 (0.0067)	79.3389 (0.0009)	79.8767 (0.0008)
hqic	Z	-3.14469 (0.0008)	-0.77733 (0.2185)	-3.58923 (0.0002)	-2.60986 (0.0045)
		$\bar{k}$= 1.00	$\bar{k}$ =1.41	$\bar{k}$ =1.45	$\bar{k}$ =1.36

Notes:

p-value between brackets.

TABLE 4

CIPS (Z_{-t-bar})

lag order (k)	Test	$\log f$	$\log S^d$	$\log S^f$	$\log H$
1	Z_{-t-bar}	1.697 (0.955)	-1.381 (0.084)	-2.457 (0.007)	1.414 (0.921)
2	Z_{-t-bar}	1.047 (0.852)	2.841 (0.998)	1.582 (0.943)	0.802 (0.789)
3	Z_{-t-bar}	2.044 (0.980)	3.375 (1.000)	-1.661 (0.048)	-3.486 (0.000)
aic	Z_{-t-bar}	0.758 (0.776)	1.319 (0.906)	-1.702 (0.044)	-2.072 (0.019)

Notes:

p-value between brackets.

TABLE 5

Moon and Perron test

	m	kernel	$\log f$	$\log S^d$	$\log S^f$	$\log H$
m*(IC1)			4	4	4	4
m*(BIC3)			2	4	4	4
t_a^*	1	QS	-2.6604(0.0039)	-1.7862 (0.0370)	-0.7206 (0.2356)	-0.2279(0.4099)
	1	B	-2.9104(0.0018)	-1.7852 (0.0371)	-0.7889(0.2151)	-0.2382 (0.4059)
	2	QS	-4.2229(1.2057e-005)	-1.6096 (0.0537)	-0.7984 (0.2123)	-4.5750 (2.3815e-006)
	2	B	-4.4621 (4.0585e-006)	-1.6856 (0.0459)	-0.6783 (0.2488)	-4.5750 (2.3815e-006)
	3	QS	-4.3942 (5.5589e-006)	-0.7444 (0.2283)	-4.2315 (1.1605e-005)	-3.5113 (2.2296e-004)
	3	B	-4.5508 (2.6726e-006)	-0.7703 (0.2206)	-3.7149 (1.0164e-004)	-3.6976 (1.0883e-004)
	4	QS	-4.9478 (3.7529e-007)	-3.6826 (1.1544e-004)	-0.4179 (0.3380)	-3.6500 (1.3112e-004)
	4	B	-5.0309 (2.4409e-007)	-4.6880 (1.3792e-006)	-0.5859 (0.2790)	-4.0252 (2.8465e-005)
t_b^*	1	QS	-2.6563 (0.0040)	-1.1354 (0.1281)	-0.6198 (0.2677)	-0.0747 (0.4702)
	1	B	-2.7187(0.0033)	-1.0977 (0.1362)	-0.7003 (0.2419)	-0.0834 (0.4668)
	2	QS	-4.6476 (1.6787e-006)	-0.9969 (0.1594)	-0.8512 (0.1973)	-5.5644 (1.3150e-008)
	2	B	-4.4694 (3.9228e-006)	-1.1119 (0.1331)	-0.7572 (0.2245)	-4.6502 (1.6583e-006)
	3	QS	-4.6777 (1.4502e-006)	-0.3774 (0.3530)	-5.9500 (1.3410e-009)	-6.6819 (1.1796e-011)
	3	B	-4.6046 (2.0666e-006)	-0.4235 (0.3360)	-3.5446 (1.9659e-004)	-6.3602 (1.0077e-010)
	4	QS	-6.3634 (9.8699e-011)	-4.6224 (1.8964e-006)	-0.3629 (0.3583)	-6.2396 (2.1941e-010)
	4	B	-5.1844 (1.0838e-007)	-4.7808 (8.7295e-007)	-0.3591 (0.3598)	-7.2367 (2.2992e-013)

Notes:

p-value between brackets. QS: Quadratic spectral kernel. B: Bartlett kernel.

TABLE 6

Choi (2006) Fisher-type statistics (P_m, Z, L^*)

lag order (k)	Test	$\log f$	$\log S^d$	$\log S^f$	$\log H$
1	P_m	1.9495 (0.0256)	7.1271 (5.1237e-013)	10.8056 (0)	0.5401(0.2946)
1	Z	-2.4621 (0.0069)	-3.8111 (6.9186e-005)	-6.7776 (6.1086e-012)	-1.3161 (0.0941)
1	L^*	-2.3866 (0.0085)	-4.6426 (1.7204e-006)	-7.5693 (1.8768e-014)	-1.2233 (0.1106)
2	P_m	-1.3238 (0.9072)	1.5406 (0.0617)	7.3856 (7.5939e-014)	0.3634 (0.3581)
2	Z	0.3686 (0.6438)	0.5268 (0.7008)	-5.4012 (3.3094e-008)	-1.0541 (0.1459)
2	L^*	0.3375 (0.6321)	1.1224 (0.8691)	-5.5847 (1.1708e-008)	-0.9689 (0.1663)
3	P_m	-0.1229 (0.5489)	-0.7096 (0.7610)	-0.6469 (0.7411)	-0.1871 (0.5742)
3	Z	0.0251 (0.5100)	1.1386 (0.8726)	-0.5273 (0.2990)	-0.3876 (0.3492)
3	L^*	-0.0452 (0.4820)	1.4084 (0.9205)	-0.4800 (0.3156)	-0.2621 (0.3966)
aic	P_m	2.6412 (0.0041)	5.9129 (1.6806e-009)	12.2388 (0)	1.9548 (0.0253)
aic	Z	-2.5990 (0.0047)	-2.6092 (0.0045)	-7.7401 (4.9682e-015)	-1.2459 (0.1064)
aic	L^*	-2.5389 (0.0056)	-2.7818 (0.0027)	-8.5914 (4.2945e-018)	-1.1490 (0.1253)

Notes:

p-value between brackets.

TABLE 7

Estimation results. Spatial autoregressive error model (QML)

	(i)	(ii)	(iii)
ϕ	1	5	10
φ_G7	15.22	15.22	15.22
φ_NOG7	7.04	7.04	7.04
β	0.078***	0.074***	0.074***
	(8.481)	(8.067)	(7.871)
γ_G7	0.183***	0.180***	0.175***
	(10.527)	(10.358)	(10.032)
γ_NOG7	0.020**	0.032***	0.029***
	(1.979)	(3.123)	(2.792)
δ	0.010	0.182	0.235**
	(0.086)	(1.599)	(2.061)
λ	0.4569***	0.254***	0.207***
	(9.106)	(5.535)	(4.875)

Notes:

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

t statistics between brackets

TABLE 8

Tests of cross sectional independence

Spatial models. LM-test			Pair-wise correlation coefficients		
$\phi = 1$	$\phi = 5$	$\phi = 10$	CD	Friedman	Frees
44.23	30.10	23.21	4.710	40.696	5.235
(0.000)	(0.000)	(0.000)	(0.000)	(0.0061)	
				[0.470]	

Notes:

p-value between brackets. Average absolute value
of the off-diagonal elements between square brackets.
Critical value from Frees' Q distribution: alpha = 0.01 : 0.2468.

TABLE 9

Estimation results. Unobserved common correlated factors (CCEP)

	(i)	(ii)	(iii)	(iv)
φ_G7	15.22	15.22	15.22	15.22
φ_NOG7	7.04	7.04	7.04	7.04
β	0.026	0.020	0.010	-0.004
	(0.408)	(0.203)	(0.146)	(-0.411)
γ_G7	0.317**	0.624***	0.216**	0.464***
	(1.986)	(3.992)	(2.276)	(2.806)
γ_NOG7	0.122	0.161	0.213	0.337**
	(0.925)	(1.110)	(1.439)	(2.274)
δ	0.305	0.372	0.194	0.222
	(0.818)	(0.882)	(0.478)	(0.410)
$\mathbf{d}_t$	$(t)'$	$(t,t^2)'$	$(t,p)'$	$(t,t^2,p)'$

Notes:

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

t statistics between brackets

Standard error based on Newey-West type variance estimator of eq. (74) in Pesaran (2006).