Spatial autoregressive spillovers vs unobserved common factors models. A panel data analysis of international technology diffusion

CEM ERTUR† and ANTONIO MUSOLESI‡

December 17, 2012

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 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 the R&D capital stocks) appear to be stationary. Then, we estimate the model using two complementary approaches, focusing on generalised spatial autoregression and unobserved common correlated factors. These approaches 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.

JEL classification: C23; C5; F0; O3.

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

†*LEO, University of Orleans UMR 6221 CNRS Faculté de Droit d'Economie et de Gestion, Rue de Blois - B.P. 6739 45067 Orléans Cedex 2 –France*

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

‡*CESAER UMR 1041, French National Institute for Agricultural Research, 26 bd Docteur Petitjean, 21079 Dijon Cedex – France*

(email:amusolesi@enesad.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 the total factor productivity (TFP) is an increasing function of the cumulative research and development (R&D). In particular, CH analyse the role of international trade. By assuming that some intermediate inputs are traded internationally while some 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 trade partners. The influence of this approach is a result of its plausibility with respect to endogenous growth theory (Keller, 2004) and its versatility in allowing the consideration of alternative channels of international technology diffusion, such as foreign direct investments, FDI, (Lichtenberg and Van Pottelsberghe, 2001), bilateral technological proximity and patent citations between countries (Lee, 2006), language skills or geography (Keller, 2002).

This paper aims to contribute to the empirical literature on R&D spillovers among countries by focusing on the issue of cross-sectional dependence. This because cross-country correlation from a variety of sources can plausibly be present in CH-type specifications; however, this complicates standard estimation and inference. To our knowledge, this issue has not been addressed in previous studies. Specifically, the main goal of this paper is to contrast a generalised spatial autoregression approach (Lee and Yu, 2010) with multifactor errors models (Pesaran, 2006; Eberhardt and Bond, 2009). These two approaches 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, whilst, at the same time, it seems to be difficult having an a priori precise knowledge on the kind of cross section dependence that links the cross-sectional units. Since this may be true 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 of technology diffusion. First, it is theoretically consistent. Keller (2002) and Eaton and Kortum (2002) show that transport costs or geographical barriers create links 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-run 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 important in the context of this paper which aims at focusing on a methodological issue, that of cross sectional dependence, is that it can be argued that traditional channels of international technology diffusion could create endogeneity problems when included in econometric specifications. For instance, the international trade, the FDI or the patent activity of a country could depend on its technological level. In contrast, geographical distance is generally conceded to be exogenous. Moreover, geographical distance may be considered an exogenous proxy for some endogenous measures of socioeconomic, institutional, cultural or language-based similarities that could enhance the diffusion of technology.

As 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 root due to the low power of the standard tests (Rudesbusch, 1993; Perron, 1989, 1997; Bierens, 1997, Fleissig and Strauss, 1999). Based on this result, we adopt estimation techniques that have not been specifically designed to model nonstationary variables.

First, cross-sectional correlation is posed as a result of *spatial spillover effects* and is modelled by adopting a general spatial panel econometric framework. The framework simultaneously allows spatial autoregression (SAR) and spatially autoregressive disturbances (this latter kind of spatial model is also known as the spatial error model, SEM). The estimation is performed using Lee and Yu's (2010) quasi-maximum likelihood (QML) estimator, which has been shown to be consistent. Standard panel data models are primarily designed to address individual heterogeneity, which can be inherently spatial; however, they 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 could arise due to, for instance, "relative" geographical localisation of countries with respect to each other. In this case, proper estimation requires explicit modelling of spatial autocorrelation. Debarsy and Ertur (2010) label this particular type of heterogeneity as *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 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, while interactive heterogeneity can be captured by impact coefficients or elasticities. These coefficients and elasticities are computed from the reduced form of the spatial autoregressive model taking into account 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 influence TFP differentially across countries. Such factors could 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 empirical setup adopted in this paper builds on the correlated common effects (CCE) approach by Pesaran (2006), which has also 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, while the main focus is on the estimation and inference of the slope parameters.

It is interesting to note that these two approaches are also related to the notions of *weak* and *strong cross-sectional dependence* and *weak* and *strong factors* recently developed and discussed in Chudik et al. (2011). While the spatial models, under a standard set of regularity conditions, entail weak cross-sectional correlation (Breitung and Pesaran, 2008; Pesaran and Tosetti, 2011), the CCE approach explicitly introduces a finite number of *strong* factors that are possibly correlated with the regressors but does not explicitly introduce *weak* factors. An appealing feature of such an approach is that it provides consistent estimates under a variety of situations. In particular, it has been shown by Pesaran and Tosetti (2011) that the CCE approach is still 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 that a process which 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 such factors is strong. The authors then 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 spillovers effects.

Spatial autoregressive models and unobserved common factors models have been developed separately in the literature. Although several tests to detect the presence of cross-sectional dependence have been proposed (Moscone and Tosetti, 2009, for a review and a simulation study), to the best of our knowledge, there does not yet exist a test to discriminate between spatial spillovers and unobserved common factors. Indeed, some of these tests are built in a spatial econometric framework and may allow one to choose among alternative spatial models (see for example, Debarsy and Ertur, 2010); some others, such as the CD test proposed by Pesaran (2004), are built using the sample pairwise correlation of residuals and are more suitable for detecting strong 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 one to make an a priori choice between weak and strong cross-sectional correlations. Thus, comparing the estimation results obtained using the two approaches mentioned above while keeping in mind the theoretical and simulation results provided by the literature may allow us to shed new light on the relative importance of different mechanisms by which technology spillovers across countries may occur. It also may provide an indication of the bias that could arise when estimating the model without accounting for a particular source of cross-country dependence. Because having an a priori knowledge on the kind of cross section dependence which links the cross-sectional units may be difficult in practice in many other empirical frameworks (e.g. production functions, innovation functions, labour demand equations, etc), the interest of this paper may go beyond the analysis of international technology diffusion.

The results of the estimations suggest that cross sectional correlation play a relevant role when is introduced in the econometric model either using a spatial model or as a results of unobserved common factors. However, the two approaches produce very different predictions in economic terms. The inclusion of *spatial autoregressive spillovers* decreases the parameter associated with the R&D capital stocks (which remain positive and significant) by approximately 20% to 50%, depending on the specification. This result indicates that the standard CH/K specifications are biased upward, while the estimated spatial autoregressive parameter is positive and highly significant. Estimating the model by allowing for unobserved common factors further decreases the parameters associated with domestic and foreign R&D capital stocks and confirm recent works suggesting that unobserved common factors (possibly strong factors) have a key role in explaining the evolution of productivity and innovation (Mastromarco et al., 2012; Eberhardt et al., 2012).

The remainder of the paper is organised as follows: section 2 describes the baseline model and provides some preliminary fixed-effects estimates. Section 3 studies the level of integration of the variables under examination. Section 4 extends the benchmark specification by allowing 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 the one adopted by CH, as modified by Coe et al. (2009) by including human capital on the right hand side of the equation:

$$f_{it} = e^{(\alpha_i + \varepsilon_{it})} \left(S_{it}^d\right)^\beta \left(S_{it}^f\right)^\gamma H_{it}^\delta \tag{1}$$

where f_{it} is 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_{it}^d, and foreign, S_{it}^f) and human capital (H_{it}); and ε_{it} is the error term. The foreign capital stock S_{it}^f is defined as the weighted arithmetic mean of S_{jt}^d for $j \neq i$:

$$S_{it}^f = \sum_{j \neq i} \omega_{ij} S_{jt}^d \tag{2}$$

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

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

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 this 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 idea that the impact of foreign R&D is negatively related to the geographical distance from foreign economies. Therefore, 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, $\omega_{ij} = \exp(-\varphi d_{ij})$, so that:

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

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 is converted into a measure of human capital stock through the 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 labour 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), we obtain the baseline specification to be estimated:

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

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, while 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. A simple variant of eq.(6) widely adopted in the literature allows the impact of foreign R&D capital to differ between the largest seven countries and the others:

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

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

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 that has been widely used in the literature (see Table 1). It is a balanced

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

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 2 countries has been 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 baseline specification in eq.(6) , where 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 0.069 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 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 could be 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 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. 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] Finally, we turn to the estimates of the specification that allows the output elasticity with respect to foreign R&D to differ between large and small countries (Table 2, columns vi to xi). 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 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 ones.

[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 upwards. 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 serially uncorrelated residuals and, at the same time, sufficiently small such that the model is not overparametrised. 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 it 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 the variables. Third, the average number of lags obtained with AIC, BIC or HQIC is generally about 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, labelled P, and Z. They differ in whether they use the inverse chi-square or the inverse normal distribution of the p-values.[5] The Fisher-type statistics (in table 3) fully confirm the

[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), while 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] Though 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.

[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 with 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.

IPS tests.

Recent work has shown the importance of taking into account the 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 oversized but the tests explicitly allowing for cross-sectional dependence yield a lower frequency of type I error. 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 where the pairwise cross-sectional errors' covariances are different 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 $u_{it} = \zeta_i^{\prime}\mathbf{f}_t + v_{it}$, where $\mathbf{f}_t$ is a $m \times 1$ vector of unobserved common factors and $\zeta_i^{\prime}$ 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 ($\Delta\bar{y}_t$) and lagged levels ($\bar{y}_{t-1}$) of the individual series, which are $\sqrt{N}$- consistent estimators for $\bar{\zeta} f$ and $\bar{\zeta}\sum_{j=0}^{t-1} f_j$, respectively. This expression is the cross-sectionally augmented Dickey-Fuller (CADF) statistic; 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 criteria, 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 nonstationary, while 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 and $T = 20$, the CIPS tests have low power (but correct size)[8], suggesting that the series could, 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. They 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

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

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)[9]. 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 nonstochastic 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, but Choi's test performs well under the more restrictive assumption that the cross-sectional units are homogeneously influenced by the common factor, and in few cases, it beats 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 performed with the AIC (or with other nonreported 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 amount of literature has emerged since Nelson and Plosser (1982), initially corroborating their finding of a unit root for most of the macroeconomic variables. Yet, subsequent works have shown that this finding is due to *"lack of power even against distant alternatives"* (Rudesbusch, 1993). 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.

[9]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, they have low power when deterministic trends are included.

4 Handling cross-sectional correlation

4.1 Spatial autoregressive spillovers

A subset of the the literature on panel data has addressed the issue of cross-sectional correlation by considering models with spatial autoregression (Baltagi et al., 2009; Lee and Yu, 2009, 2010). In our empirical framework, this type of specification could be useful. Consider the baseline specification (4). A country's TFP (which is an output of the unobserved level of technology) depends on the R&D (an input of technology) of the other countries; however, it could be argued that the TFP of a country also depends on the TFP of other countries. This foreign TFP measure incorporates neighbours' observed and unobserved inputs to technology, including institutions, culture, history, etc.[10] To incorporate this type of spatial effect into our specification, we adopt Lee and Yu's (2010) generalised specification – labeled SARAR(1,1) - which considers a SAR model with spatially autocorrelated disturbances. Now eq. (6) can be rewritten as:

$$\begin{aligned} \log f_{it} &= \alpha_i + \rho \sum_{j \neq i} w_{ij} \log f_{jt} + \beta \log S^d_{it} + \gamma \log \sum_{j \neq i} \exp(-\varphi d_{ij}) S^d_{jt} + \delta \log H_{it} + v_{it} \qquad (8) \\ v_{it} &= \lambda \sum_{j \neq i} w_{ij} v_{jt} + \varepsilon_{it} \end{aligned}$$

where ρ and λ are the spatial autoregressive parameters. For $\lambda = 0$ and $\rho \neq 0$ eq. (8) is a spatial autoregressive model (SAR); conversely, with $\lambda \neq 0$ and $\rho = 0$, it is a spatially autocorrelated error model (SEM) and with both $\lambda = 0$ and $\rho = 0$, it reduces to the baseline *"a-spatial"* specification (6). The results obtained from the spatial econometric specifications have to be interpreted by looking at the so-called *reduced form*. In matrix form, stacking over all individuals for time period t, we have:

$$\begin{aligned} \boldsymbol{y}_t &= \alpha + \rho \mathbf{W}_N \mathbf{y}_t + \beta \boldsymbol{SD}_t + \gamma \boldsymbol{SF}_t + \delta \boldsymbol{H}_t + v_t \qquad (9) \\ \boldsymbol{v}_t &= \lambda \mathbf{W}_N \boldsymbol{v}_t + \boldsymbol{\varepsilon}_t \qquad t = 1, ..., T \end{aligned}$$

where $\boldsymbol{y}_t$ represents the $N \times 1$ vector of log TFP, $\boldsymbol{W}_N$ 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, $\boldsymbol{SD}_t$ is the $N \times 1$ vector of log domestic R&D capital stocks, $\boldsymbol{SF}_t$ is the $N \times 1$ vector of log foreign R&D capital stocks with typical element: $\log S^f_{it} = \log \sum_{j \neq i} \exp(-\varphi d_{ij}) S^d_{jt}$. After some manipulations, eq. (6) can be rewritten in its *reduced form* representation:

$$\boldsymbol{y}_t = (\boldsymbol{I}_N - \rho \boldsymbol{W}_N)^{-1} \alpha + (\boldsymbol{I}_N - \rho \boldsymbol{W}_N)^{-1} (\beta \boldsymbol{SD}_t + \gamma \boldsymbol{SF}_t + \delta \boldsymbol{H}_t) + (I - \rho \boldsymbol{W}_N)^{-1} (\boldsymbol{I}_N - \lambda \boldsymbol{W}_N)^{-1} \boldsymbol{\varepsilon}_t \quad (10)$$

where $(\boldsymbol{I}_N - \rho \boldsymbol{W}_N)^{-1}$ is the so-called *global* spatial multiplier. The multiplicative process arises because if $|\rho| < 1$ it follows that $(\boldsymbol{I}_N - \rho \boldsymbol{W}_N)^{-1} = \boldsymbol{I}_N + \rho \boldsymbol{W}_N + \rho^2 \boldsymbol{W}_N^2 + \rho^3 \boldsymbol{W}_N^3 + ...$ This reduced

[10] Moreover, it has been shown that such specifications are consistent with theoretical growth models (Ertur and Kock, 2007).

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

form has two important implications. *First*, in conditional mean, the total factor productivity in a country i will be affected not only by the domestic R&D capital stock in i but also by those in all the other countries through the inverse spatial transformation $(\boldsymbol{I}_N - \rho \boldsymbol{W}_N)^{-1}$. This is the so-called *global interaction effect*. This is also true for the foreign R&D capital stock and the human capital stock.

More formally, let us write the $N \times N$ matrices of the partial derivatives of $\boldsymbol{y}_t$ with respect to the explanatory variables $\boldsymbol{SD}_t$, $\boldsymbol{SF}_t$ and $\boldsymbol{H}_t$, which we call the impact matrices:

$$\Xi_y^{SD} \equiv \frac{\partial \boldsymbol{y}_t}{\partial \boldsymbol{SD}_t'} = (\boldsymbol{I}_N - \rho \boldsymbol{W}_N)^{-1}\beta = (\boldsymbol{I}_N + \rho \boldsymbol{W}_N + \rho^2 \boldsymbol{W}_N^2 + \rho^3 \boldsymbol{W}_N^3 + ...)\beta \quad (11)$$

$$\Xi_y^{SF} \equiv \frac{\partial \boldsymbol{y}_t}{\partial \boldsymbol{SF}_t'} = (\boldsymbol{I}_N - \rho \boldsymbol{W}_N)^{-1}\gamma = (\boldsymbol{I}_N + \rho \boldsymbol{W}_N + \rho^2 \boldsymbol{W}_N^2 + \rho^3 \boldsymbol{W}_N^3 + ...)\gamma \quad (12)$$

$$\Xi_y^{H} \equiv \frac{\partial \boldsymbol{y}_t}{\partial \boldsymbol{H}_t'} = (\boldsymbol{I}_N - \rho \boldsymbol{W}_N)^{-1}\delta = (\boldsymbol{I}_N + \rho \boldsymbol{W}_N + \rho^2 \boldsymbol{W}_N^2 + \rho^3 \boldsymbol{W}_N^3 + ...)\delta \quad (13)$$

The diagonal elements of these matrices represent the *direct elasticities* including "own spillover" effects, which are inherently heterogenous in presence of spatial autocorrelation due to differentiated interaction terms in the $\boldsymbol{W}_N$ matrix. This type of heterogeneity is called *interactive heterogeneity*, in opposition to standard individual heterogeneity in panel data models (Debarsy and Ertur, 2010). Using obvious notations, letting $\mathbf{X}_{t,i}$ be an explanatory variable, either $\boldsymbol{SD}_{t,i}$, $\boldsymbol{SF}_{t,i}$ or $\boldsymbol{H}_{t,i}$, we have:

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

Note that the own derivative for country i includes the *feedback* effects where country i affects country j and country j also affects country i as well as longer paths which might go from country i to j to k and back to i. The off-diagonal elements of these matrices represent *indirect elasticities*.[12]

The magnitude of those direct and indirect elasticities will depend on: (1) the degree of interaction between countries, which is governed by the $\boldsymbol{W}_N$ matrix, (2) the parameter ρ, measuring the strength of spatial correlation between countries and (3) the parameters β, γ or δ respectively.

Moreover, for each of theses three impact matrices, the sum across the i^{th} row represents the total impact on the TFP of country i of a change of the corresponding explanatory variable across all the n countries of the sample. The sum down the j^{th} column yields the total impact on TFP over all the n countries of the sample of a change of the explanatory variable in country j. The average direct impact is therefore defined as $n^{-1}tr(\widehat{\Xi}_{\mathbf{y}}^{\mathbf{X}})$ whereas the average total impact is defined as $n^{-1}\iota'\widehat{\Xi}_{\mathbf{y}}^{\mathbf{X}}\iota$ where ι is the $n \times 1$ sum vector. Note that in our SAR model with a row-standardized interaction matrix the average total impact further simplifies to $(1-\rho)^{-1}\beta$. Finally the average indirect impact is by definition the difference between the average total impact and the average direct impact.

Second, one can easily see that a random shock due to unobservable factors affecting total factor productivity (i.e., a shock in the disturbances) in a specific country i does not only affect the total

[12] 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 some applications.

factor productivity in i, but it also has an impact on the total factor productivity in all of the other countries through a more complex mechanism than that presented above: this is the so-called *spatial diffusion process of a random shock*. For the SARAR specification, this process can be expressed as:

$$\Xi_y^\varepsilon \equiv \frac{\partial \boldsymbol{y}_t}{\partial \boldsymbol{\varepsilon}_t'} = (I - \rho \boldsymbol{W}_N)^{-1}(\boldsymbol{I}_N - \lambda \boldsymbol{W}_N)^{-1} \tag{15}$$

A possible method for estimating spatial panel econometric models consists of using the direct maximum likelihood approach (Elhorst, 2009). Lee and Yu (2010) show that this approach provides consistent estimates for coefficients on regressors. However, it 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 from the direct approach are identical to the corresponding estimates from the transformation approach.

We finally focus on the specification search of spatial econometric models. We adopt the testing procedure recently proposed by Debarsy and Ertur (2010), which consists of sequentially performing five LM tests: First, the joint test A) considers $H_0^a : \rho = \lambda = 0$ against $H_1^a : \rho$ or $\lambda \neq 0$. If the null hypothesis is not rejected, neither of these two types of spatial correlation -SAR or SEM- is relevant, and there is no need to conduct other tests. Otherwise, the following two simple tests have to be performed: The first, marginal test B), considers $H_0^b : \rho = 0$ $(\lambda = 0)$ against $H_1^b : \rho \neq 0$. Under the alternative, the specification is a SAR. The second, marginal test C), considers $H_0^c : \lambda = 0$ $(\rho = 0)$ against $H_1^c : \lambda \neq 0$. Here, under the alternative, the specification is an SEM. If the null hypothesis is rejected in both cases, then one must discriminate between the SAR and the SEM. This can be performed by using two conditional tests: the first, conditional test D), considers $H_0^d : \lambda = 0$ given $\rho \neq 0$ (SAR) against $H_1^d : \lambda \neq 0$. Under the alternative, the specification is a SARAR. The second, conditional test E), considers $H_0^e : \rho = 0$ given $\lambda \neq 0$ (SEM) against $H_1^e : \rho \neq 0$. Under the alternative, the specification is a SARAR. If the null is rejected in both cases, the appropriate specification is a SARAR. If the null hypothesis is rejected only in (D), then the preferred specification is a SEM. Otherwise if the null hypothesis is rejected only in (E), the preferred specification is a SAR.

4.2 Unobserved common factors

In the macro panel data literature, the standard approach to address cross-sectional correlation (and eventually also allowing to deal with slope heterogeneity) 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 since it involves nuisance parameters that increase at a quadratic rate as the cross-sectional dimension of the panel is allowed to rise or if the errors are correlated with the regressors. This has led to consider unobserved factor models (see e.g. Forni et al., 2000; Bai, 2003).

The empirical setup adopted in this paper builds on the framework originally proposed by Pesaran (2006) and further developed and studied very recently (Chudik et al. 2011; Pesaran and

Tosetti, 2011; Kapetanios et al. 2011) where the unobserved factors are viewed as nuisance variables introduced to model the cross-sectional dependence in a parsimonious manner; the main focus is on the estimation and inference of the slope parameters. To save notation, let us define: $y_{it} = \log f_{it}$, $\mathbf{x}_{it}= \left[\log(S^d_{it}), \log(S^f_{it}), \log(H_{it})\right]'$ and assume the following DGP:

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

where $\mathbf{d}_t$ is an $l \times 1$ vector of observed common effects, α_i' is the associated vector of parameters, $\mathbf{x}_{it}$ is a 3×1 vector of explanatory variables. The slope coefficients $\beta_i' = [\beta_i, \gamma_i, \delta_i]$ 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:

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

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 :

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

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

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

where $\mathbf{A}_i$ and $\mathbf{\Gamma}_i$ are $l \times 3$ and $m \times 3$ factor loading matrices and $\mathbf{v}_{it} = (v_{i1t}, v_{i2t}, , v_{i3t})'$. Combining (18) and (19) we finally obtain a system of equations explaining simultaneously TFP, R&D (domestic and foreign) and Human capital:

$$\underset{4\times 1}{\mathbf{z}_{it}} = \begin{pmatrix} y_{it} \\ \mathbf{x}_{it} \end{pmatrix} = \begin{pmatrix} \log\left(f_{it}\right) \\ \log(S^d_{it}) \\ \log(S^f_{it}) \\ \log(H_{it}) \end{pmatrix} = \underset{4\times l}{\mathbf{B}_i'} \underset{l\times 1}{\mathbf{d}_t} + \underset{4\times m}{\mathbf{C}_i'} \underset{m\times 1}{\xi_t} + \underset{4\times 1}{\mathbf{u}_{it}}, \tag{20}$$

where

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

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

$\mathbf{I}_k$ is an identity matrix of order k. In our specific case, $k = 3$.

Some remarks are in order. First, this set-up is sufficiently general that renders a variety of panel data models as special cases. Secondly, as noted by Chudik et al. (2011), the term $\varrho_i^{'}\xi_t$ represents a strong factor structure. Third, specifying a factor loading matrix $\mathbf{C}_i$ of the kind presented above permits a variety of situations since each variable is allowed to be affected in a specific way by each factor since the typical element of such a matrix, say c_{imj}, measures the country specific effect (eventually being zero) of the mth common factor on the jth variable. For exemple, it may allow that some of the unobserved common factors driving the evolution of TFP 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 that some other factors could specifically affect only one variable of the system. [13] Finally, and very importantly, this set-up introduces endogeneity whereby the $\mathbf{x}_{it}$ are correlated with the unobservable e_{it} and as noted by Kapetanios et al. (2011), standard approaches that neglect common factors fail to identify $\beta_i^{'}$; instead, they yield an estimate of $\kappa_i^{'} = \beta_i^{'} + \varrho_i^{'}\mathbf{\Gamma}_i^{'-1}$.[14]

Under the restrictive assumption of homogenous factor loadings, it is possible to consistently estimate the (homogenous) slope parameters by ordinary least squares (OLS); one may adopt either a two-way fixed effects model or a first difference specification added 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. He suggests using $\begin{pmatrix} \mathbf{d}_t^{'} & \overline{\mathbf{z}}_{wt}^{'} \end{pmatrix}$ as observable proxies for the unobserved factors where $\mathbf{z}_{it} = \begin{pmatrix} y_{it} \\ \mathbf{x}_{it} \end{pmatrix}$ and $\overline{\mathbf{z}}_{wt}$ indicates the cross-sectional average, i.e., $\overline{\mathbf{z}}_{wt} = \sum_{j=1}^{N} w_j \mathbf{z}_{jt}$, w_j are weights equal to $1/N$. The individual slopes $\beta_i^{'}$, or their mean can be consistently estimated by running OLS or pooled regressions of y_{it} on $\mathbf{x}_{it}, \mathbf{d}_t$ and $\overline{\mathbf{z}}_{wt}$. The common factor loadings are thus treated as nuisance parameters and are not estimated. 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 CCEP pooled estimator and the Mean Group estimator known as CCEMG, which is obtained by averaging the country-specific estimates following Pesaran and Smith (1995), which thus also permits that the slope parameters are not the same across cross-sections.

Recent papers have shown some other properties of the CCE approach which are very appealing for this work. One appealing properties of the CCE estimators is that they still provide consistent estimates of the slope coefficients and their standard errors under the more general case of multifactor

[13] 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 which are specific to the each regressor.

[14] To see how this may happen, let 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: $y_{it} = \beta_i^{'}\mathbf{x}_{it} + \varrho_i^{'}\mathbf{\Gamma}_i^{'-1}(\mathbf{x}_{it} - \mathbf{v}_{it}) + \varepsilon_{it} = \kappa_i^{'}\mathbf{x}_{it} + \varkappa_{it}$, where $\kappa_i^{'} = \beta_i^{'} + \varrho_i^{'}\mathbf{\Gamma}_i^{'-1}$ and $\varkappa_{it} = \varepsilon_{it} - \varrho_i^{'}\mathbf{\Gamma}_i^{'-1}\mathbf{v}_{it}$. So, applying least squares to such an equation consistently estimates $\kappa_i^{'}$ rather than $\beta_i^{'}$.

error structure and spatial error correlation (Pesaran and Tosetti, 2011; see also Bresson and Hsiao, 2011, for further simulation results), i.e., when the error term in eq. (17) can be rewritten as:

$$e_{it} = \varrho_i^{'}\xi_t + \lambda \sum_{j \neq i} w_{ij} e_{jt} + \varepsilon_{it}. \tag{21}$$

Such an approach is also robust to the presence of both a limited number of *strong factors* , such as a global policy or global technological modification, and an infinite number of *weak factors*, which can represent spillover effects (Chudik et al., 2011).

Other very relevant results have been provided by Kapetanios et al. (2011) who consider the case of non stationary common effects. More precisely, they partitioned the vector of observed common factors as $\mathbf{d}_t = \left(\mathbf{d}_{1t}^{'}, \mathbf{d}_{2t}^{'}\right)^{'}$ where $\mathbf{d}_{1t}$ is an $l_1 \times 1$ vector of determinstic components and $\mathbf{d}_{2t}$ is an $l_2 \times 1$ vector of unit root observed common factors, with $l_1 + l_2 = l$ and then suppose that the $(l_2 + m) \times 1$ vector of stocastic common effects $\mathbf{h}_t = \left(\mathbf{d}_{2t}^{'}, \xi_t^{'}\right)^{'}$ follows a multivariate unit root process. Both analytical results and a simulation study indicate that the CCE approach is still valid when the unobserved factors are allowed to follow unit root processes.

Finally, as an alternative to the CCEMG, we also adopt the Augmented Mean Group estimator (Eberhardt and Bond, 2009; Eberhardt and Teal 2010) where the Mean Group group-specific regressions are augmented with a preliminary OLS estimate of a "common dynamic process". 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"); however, in our empirical setup where the dependent variable is TFP itself, the common dynamic process broadly represents the cross-country average of all unobservable factors that may drive TFP.

5 Results

We now turn to the presentation of the results obtained adopting the generalised SARAR spatial panel framework (eq. 8), which has been estimated with the QML approach proposed by Lee and Yu (2010).[15] We first focus on the specification search of spatial econometric models presented in the previous section. The results are presented in table 7. First, the joint LM test highly rejects the null of no spatial correlation. Then, looking at the simple marginal tests, it is not possible to discriminate between SAR and SEM because both tests reject the null hypothesis. Finally, the results of the conditional tests are as follows: for the basic specification, which does not discriminate between G7 and non-G7 countries, the SAR model is the preferred specification (except for $\phi = 1$). For the specification allowing the coefficient with respect to foreign R&D to differ between G7 and non-G7 countries, the SAR model is also the preferred specification whatever the value of ϕ.[16]

[15] 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.

[16] When an endogenous spatial lag is included, there is never evidence of residual spatial correlation; in contrast, when residual spatial correlation is included, there is still evidence of an endogenous spatial autoregressive component.

Next, the model is estimated, and the estimates of the parameters β, γ, δ, ρ and λ as well as the associated average direct impacts are presented in Table 8. Columns (ii) to (vii) present the results for the basic specification, while columns (viii) to (xiii) give the outcomes for the specification in which we include a G7 dummy variable for foreign R&D. We provide the results from estimating both the SAR and SARAR specifications, keeping in mind that the SAR model is generally the preferred specification. A general result is that the estimated coefficients and standard errors are quite robust to the choice of ϕ. Let us now focus on the basic specification. First, for the SAR specification (Columns (ii) to (iv)), compared with the results obtained from the benchmark *a-spatial* specification presented in Table 2, the estimated coefficients of domestic R&D, foreign R&D and human capital decrease by approximately 20% to 65%. Note however that the estimated coefficients in SAR and SARAR models cannot be interpreted as the corresponding elasticities in contrast to the estimated coefficients in the basic a-spatial model. The actual direct elasticities are given by the main diagonal of the impact matrices in equations (11), (12) and (13), which collects the direct partial derivatives computed using the reduced form.[17] Comparing the average direct impacts, which include the feedback effects, to the elasticities estimated in the basic specification, the decrease appears to be less severe for domestic R&D (12% to 16%), and approximately of the same order for foreign R&D and human capital (45% to 62%). This indicates the upward bias of the standard CH-type specification due to the omission of relevant spatial interactions. The estimation of the spatial autoregressive parameter ρ is positive, ranging from 0.24 and 0.44, and it is highly statistically significant. Looking at the SARAR specification (Columns (v) to (vii)) adding spatial autoregression in the disturbances further decreases the average direct impacts (up to 55%). The estimated spatial lag parameter ρ rises to approximately 0.50, while the spatial error parameter λ appears to be negative and significant.

We now focus on the specification with the G7 dummy, in Columns (viii) to (xiii), and in particular on the SAR model (Columns (viii) to (xi)), which is clearly the preferred model according to our specification search. The parameter associated with domestic R&D is found to be lower than that obtained using the *a-spatial* specification and of the same order of magnitude as those obtained in columns (ii) to (iv) for the SAR model without the G7 dummy variable (roughly 0.055-0.06). The magnitude of the average direct impacts are of the same order (0.061-0.064). The coefficient associated to foreign R&D decreases for the G7 countries from 0.17 in the benchmark specification to 0.14 (average direct impact 0.15); for the non-G7 countries the coefficient and the average direct impact associated to foreign R&D are still close to 0.02 but the coefficient is no longer significant. Finally, the spatial parameter ρ is of the same order of magnitude as that obtained from the specification without the G7 dummy variable: it ranges from 0.20 to 0.41.

In summary, the main results from the spatial econometric analysis are as follows: (1) Our specification search clearly favors a SAR model especially when G7 dummy variables are included. (2) Allowing spatial autoregression decreases the parameter associated with the R&D capital stocks and indicates the upward bias of the standard CH/K specifications, while the estimated spatial autoregressive parameter ρ is positive and highly significant. (3) Finally, the estimation of the

[17] The estimated impact matrices are available from the authors upon request.

specification including G7 dummy variables seems to indicate that there is evidence of geographical R&D spillovers only for G7 countries (estimated at 0.14) because the parameter associated with non-G7 countries, γ_G7, is not significant.

Next, we estimate the model by allowing the presence of common factors. Our estimates (in table 9) are structured as follows. We first consider the most constrained case by allowing the common factors to have only homogeneous effects (homogeneous factor loadings) and homogenous slopes. Thus, we estimate the model using the two-way fixed effects estimator (2FE) and the first difference estimator with year dummies (2FD). Next, we adopt the CCE approach to permit heterogeneous effects of the common factors; empirically, we use both the CCEP estimator assuming common slopes and the CCEMG estimator allowing for country-specific slopes. Finally, as an alternative to the CCEMG estimator, the model is estimated using the Augmented Mean Group estimator (AMG). Before running the estimations, we also performed the CD test by Pesaran (2004) on each series and on the residuals of the benchmark specification (from section two). The results of such a test, which is suitable for detecting the presence of strong cross-sectional correlation, strongly reject the null hypothesis of no cross-sectional correlation in all cases (detailed results are available upon request).

The two methods allowing for unobserved homogeneous common factors (2FE, 2FD) provide similar results. All the estimated parameters are lower compared to those obtained with a one-way fixed effect approach and the generalised spatial autoregressive model presented in the previous sections. Only the coefficients associated with the domestic R&D and the interaction between the G7 dummy and the foreign R&D are (or are close to be) statistically significant at standard levels. The coefficients on domestic R&D range from 0.02 to 0.03, while the coefficients on the interaction term G7*Foreign R&D range from 0.08 to 0.09. The estimated effect of foreign R&D for non-G7 countries is very close to zero, whereas the parameter associated with human capital ranges from 0.06 to 0.015. In summary, such estimates indicate that accounting for (homogeneous) common factors make further decrease the estimated coefficients of the R&D capital stocks, which rest positive, even if with rather low significance levels.

Next, we focus on the results obtained estimating models allowing for heterogeneous common factors, that is the CCEP, CCEMG and the AMG estimators. We also complement the results presented in Table 9 by providing alternative CCE estimates using different definitions for the vector $\bar{\mathbf{z}}_{wt}$ of cross-sectional averages or introducing the oil price as an observed common factor (Table 10), similarly to Mastromarco et al. (2012).[18] Column (i) reports the same estimates as in table (9), which include the dependent variables and all the regressors in $\bar{\mathbf{z}}_{wt}$; column (ii) adds the oil price to the previous specification in $\mathbf{d}_t$ as an observed common factor; columns (iii) to (v) differ from (i) because they include a more limited number of regressors in $\bar{\mathbf{z}}_{wt}$; finally, in columns (vi) and (vii), we simplify the empirical specification by eliminating human capital as a regressor, as in the original CH paper (column vi), and by estimating the model including only domestic R&D as a regressor.

[18] Mastromarco et al. (2012) selected a final specification on the basis of statistical significance and parsimony, we are mainly interested in assessing the sensitivity of the CCE estimator to the specific choice of augmented factors that enter the auxiliary regression and thus provide alternative CCE estimations.

The latter modification introduces the assumption, similarly to Eberhardt et al. (2012), that all the spillover effects arise from the common factor representation (column vii).

A first result is that, in many cases, the estimated parameters are not significant at standard levels. A possible explanation is this result may be obtained by looking at the Monte Carlo simulations by Pesaran (2006) made under the assumption that the DGP is characterised by unobserved common factors. For $N = T = 20$, they indicate that, while the naïve estimators (i.e., the estimators that do not account for cross-sectional correlation, such as the LSDV or the MG estimators) are oversized but have high power, the CCE estimators have the correct size but low power. Moreover, Pesaran and Tosetti (2011) (see also Bresson and Hsiao, 2011, for further results) also 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-SEM estimator) but low power compared with the alternative estimators.[19] In summary, these simulations give a possible explanation of the lack of significance. Moreover, since the basic idea behind the CCE approach is to filter the individual specific regressors by means of cross sectional averages such that asymptotically (as $N \to \infty$) the differential effects of the unobserved common factors are eliminated, further investigations using samples with a larger cross sectional dimension could provide further insights.

A relevant general conclusion which can be derived from these estimations is that unobserved common factors may play a key role in explaining the evolution of productivity and innovation. They are consistent with recent empirical literature adopting the CCE approach. Eberhardt et al. (2012) estimated a production function augmented with R&D and find that when unobserved common factors are introduced, the effect of R&D is no more significant whereas Mastromarco et al. (2012) focused on a stochastic frontier model and find that, at best, the CCE estimator handled the effect of FDI compared with a fixed effects specification.[20] Beyond the lack of significance of R&D capital stocks, it is also worth looking at the (high) estimated parameters associated both with the vector $\bar{\mathbf{z}}_{wt}$ of cross-sectional averages in the CCE and with the "common dynamic process" in the AMG.

Looking more in depth at the estimated parameters, it can be noticed that the CCEP provides estimates of the coefficient associated with domestic R&D which are always very close to zero (positive or negative). Not surprisingly, the highest estimate is found when it is assumed that all the spillover effects arise from the common factor representation (column vii). In this case the effect of domestic R&D is estimated to be 0.03. Generally it is also found that the estimated coefficient of domestic R&D increases by decreasing the size of the vector $\bar{\mathbf{z}}_{wt}$ (columns (iii) to (v)). For the foreign R&D, it is found a quite unsatisfactory and not very credible result, that is a negative estimated coefficient ranging from -0.21 to -0.05. When the common slopes assumption is relaxed (CCEMG and AMG),

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

[20] Indeed, similarly to this paper, Mastromarco et al. (2012) estimated the model with various combinations of augmented factors but in the end selected a final empirical specification on the basis of overall statistical significance and parsimony.

more credible results are obtained. Indeed both the domestic and the foreign R&D present positive estimated coefficients, which in some cases have the same magnitude than the results obtained using the spatial models.

Remark 1 *It is interesting to note that even if at first sight a direct comparison between the CCEMG and the SAR(AR) model is not apparent since only the former allows the slope coefficients to vary across cross-sections, a more in deep look at these estimators may allow for some kind of comparison. Precisely, both the estimated coefficients from the CCEMG and the direct average effect from the spatial models are closey related to the average of the country specific partial derivatives of* y_{it} *with respect to* $\mathbf{x}_{it}$*; the CCEMG estimator, say* $\widehat{\beta}_{MG}$ *is the simple average of the individual CCE estimators,* $\widehat{\beta}_i$ *of* β_i *with* $\beta_i' \equiv \frac{\partial y_{it}}{\partial \mathbf{x}_{it}}$, $\widehat{\beta}_{MG} = n^{-1}\sum_{i=1}^{n}\widehat{\beta}_i$*. In the spatial model, as detailed in section 4.1 and using the same notation , we obtain an estimation of* $(\boldsymbol{\Xi}_{\mathbf{y}}^{\mathbf{X}})_{t,ii} \equiv \frac{\partial \mathbf{y}_{t,i}}{\partial \mathbf{X}_{t,i}}$ *noted as* $(\widehat{\boldsymbol{\Xi}}_{\mathbf{y}}^{\mathbf{X}})$ *and then the estimated average direct effect is obtained by averaging the diagonal elements of such a matrix, i.e.* $n^{-1}tr(\widehat{\boldsymbol{\Xi}}_{\mathbf{y}}^{\mathbf{X}})$*.*

It can be also noticed that the CCEMG is rather sensitive to the choice of $\overline{\mathbf{z}}_{wt}$ and to the inclusion of the oil price in $\mathbf{d}_t$ and that the CCEMG and the AMG provide somewhat contradictory results with respect to the relative contribution of domestic/foreign RD capitals to the evolution of TFP. For the CCEMG, the domestic R&D coefficient ranges in most cases from 0.04 to 0.06 whereas the foreign R&D has an estimated coefficient which is often greater than 0.1. The AMG estimator provides a different picture: a higher impact of domestic R&D (0.13) and a lower coefficient of foreign R&D (0.02).

6 Summary and concluding remarks

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 a generalized spatial autoregression approach with multifactor errors models. These two approaches have been developed quite separately in the literature but can be viewed as complementary. They indeed 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, whilst, at the same time, it seems to be difficult having an a priori precise knowledge on the kind of cross section dependence that links the cross-sectional units. Since this may be true even for other empirical frameworks, the interest of this paper may go beyond the analysis of international technology diffusion.

Many results are provided. They can be listed as follows: (i) as a preliminary but important result, panel units 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. (ii) We find evidence of positive (but small in magnitude) and localized geographic R&D spillovers. They are higher for G7 than for non-G7 countries, and they decrease with distance more quickly for G7 than for non-G7 countries. (iii) The spatial econometric estimates indicate that introducing G7 dummy variables associated to foreign R&D capital stocks decreases spatial error correlation and the SAR specification is clearly the preferred model. (iv) Allowing spatial autoregression decreases by approximately 20% to 50% the parameter associated with the R&D capital stocks and indicates the upward bias of the standard CH/K specifications, while the estimated spatial autoregressive parameter is positive and highly significant. This indicates that spatial autoregressive spillovers are important. (v) Estimating the model by allowing for unobserved common factors further decreases the parameters associated with domestic and foreign R&D capital stocks. In particular, when adopting the CCE approach and imposing, as in the spatial models, a common slope across countries (the so-called CCEP estimator) the estimated coefficient of domestic R&D is found to be very close to zero and not significant. (vi) We also find that unobserved common factors (possibly strong factors) have a key role in explaining the evolution of TFP. (vii) When the common slopes assumption is relaxed (CCEMG and AMG), higher estimated coefficient are obtained since both the domestic and the foreign R&D present positive estimated coefficients, which in some cases have the same magnitude than the estimates obtained using the spatial models.

In summary, both the preliminary integration analysis and the estimations provide new and interesting insights. A thorough investigation about the order of integration, contrarily to some previous studies, suggests that the variables under study are stationary and, more generally, confirm that detecting the order of integration is an intricate issue which 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 find evidence of stationary variables. The estimations demonstrate that accounting for cross sectional correlation affects greatly the results either using a spatial model or when we introduce in the econometric models unobserved common factors that affect all the cross sectional units. Both approaches suggest that standard *a-spatial* estimates are biased upward but, in the end, produce rather different predictions in economic terms looking both at the point estimates and at their significance levels. To our point of view, a main challenge for the econometric literature could be building panel data models allowing simultaneously spatial autoregression and common unobserved factors, thus allowing the applied econometrician to estimate models incorporating both forms of spatial correlation and to test the most appropriate specification.

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

[5] Baltagi, B. Egger, P. and Pfaffermayr, M. (2009). 'A Generalized Spatial Panel Data Model with Random Effects', Center for Policy Research Working Paper No. 113.

[6] 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.

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

[8] 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.

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

[10] 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.

[11] 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.

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

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

[14] 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.

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

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

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

[18] 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

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

[20] 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.).

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

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

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

[24] 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.

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

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

[27] 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.

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

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

[30] Forni, M. Hallin, M. Lippi, M. and Reichlin, L., (2000). The generalised factor model: identification and estimation. *Review of Economics and Statistics* 82, 540–554.

[31] 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.

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

[33] 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,

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

[35] 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.

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

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

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

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

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

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

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

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

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

[45] 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.

[46] 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.

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

[48] 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.

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

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

[51] 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

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

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

[54] 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.

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

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

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

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

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

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

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

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

Spatial models. LM tests

	Basic specification			Specification with G7 dummy		
	$\phi = 1$	$\phi = 5$	$\phi = 10$	$\phi = 1$	$\phi = 5$	$\phi = 10$
Joint test. $H_0^a : \rho = \lambda = 0$ against $H_1^a : \rho$ or $\lambda \neq 0$.	105.203	58.67	41.10	59.71	30.80	23.33
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Marginal. $H_0^b : \rho = 0$ ($\lambda = 0$) against $H_1^b : \rho \neq 0$.	60.348	35.23	26.34	59.67	29.82	20.04
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Marginal. $H_0^c : \lambda = 0$ ($\rho = 0$) against $H_1^c : \lambda \neq 0$.	18.096	22.75	17.66	44.23	30.10	23.21
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Conditional. $H_0^d : \lambda = 0$ given $\rho \neq 0$ (SAR) against $H_1^d : \lambda \neq 0$.	13.290	2.90	3.81	0.716	0.202	0.237
	(.000)	(0.088)	(0.051)	(0.397)	(0.652)	(0.625)
Conditional. $H_0^e : \rho = 0$ given $\lambda \neq 0$ (SEM) against $H_1^e : \rho \neq 0$.	384.035	362.96	294.81	219.36	112.19	74.03
	(.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Preferred model (5%)	SARAR	SAR	SAR	SAR	SAR	SAR

Notes:

p-value between brackets.

TABLE 8

Spatial models. Estimation results

	Basic specification						Specification with G7 dummy					
	SAR			SARAR			SAR			SARAR		
	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)	(x)	(xi)	(xii)	(xiii)
ϕ	1	5	10	1	5	10	1	5	10	1	5	10
φ	9.649	9.649	9.649	9.649	9.649	9.649						
φ_G7							15.22	15.22	15.22	15.22	15.22	15.22
φ_NOG7							7.04	7.04	7.04	7.04	7.04	7.04
β	0.053***	0.055***	0.056***	0.032***	0.039***	0.040***	0.057***	0.060***	0.062***	0.045***	0.087***	0.090***
	(5.023)	(5.099)	(5.175)	(3.005)	(3.426)	(3.498)	(5.869)	(5.978)	(6.055)	(4.486)	(10.649)	(11.662)
γ	0.020*	0.032***	0.034***	0.0195*	0.020	0.023*						
	(1.782)	(2.659)	(2.804)	(1.686)	(1.520)	(1.696)						
δ	0.191*	0.318***	0.348***	0.208**	0.281***	0.295***	0.074	0.204*	0.235**	0.089	0.064	0.133
	(1.712)	(2.763)	(2.993)	(2.174)	(2.743)	(2.860)	(0.706)	(1.879)	(2.132)	(0.907)	(0.643)	(1.406)
ρ	0.445***	0.286***	0.240***	0.614***	0.504***	0.469***	0.412***	0.247***	0.202***	0.534***	-0.448***	-0.485***
	(9.391)	(6.480)	(5.859)	(12.765)	(7.154)	(6.948)	(8.735)	(5.683)	(4.984)	(8.750)	(-7.001)	(-9.762)
λ				-0.402***	-0.293***	-0.286***				-0.229***	0.629***	0.622***
				(-4.673)	(-3.049)	(-3.331)				(-2.133)	(14.12)	(16.17)
γ_G7							0.134***	0.147***	0.149***	0.115***	0.221***	0.218***
							(7.549)	(8.063)	(8.142)	(6.239)	(14.279)	(15.165)
γ_NOG7							0.006	0.017	0.018	0.004	0.056***	0.052***
							(0.563)	(1.487)	(1.619)	(0.440)	(5.314)	(5.278)
Av. Dir. Impact *SD*	0.0565	0.0578	0.0587	0.0368	0.0453	0.0461	0.0607	0.0625	0.0638	0.0500	0.0962	0.1028
Av. Dir. Impact *SF*	0.0220	0.0339	0.0358	0.0225	0.0237	0.0265						
Av. Dir. Impact *SFG7*							0.1412	0.1510	0.1525	0.1263	0.2425	0.2489
Av. Dir. Impact *SFNOG7*							0.0065	0.0176	0.0194	0.0054	0.0614	0.0592
Av. Dir. Impact *H*	0.2030	0.3308	0.3590	0.2399	0.3215	0.3367	0.0788	0.2068	0.2400	0.0984	0.0711	0.1526

Notes:

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

t statistics between brackets

TABLE 9

Estimates allowing unobserved common factors

	Basic Specification					Specification with G7 dummy				
	2FE	2FD	CCEP	CCEMG	AMG	2FE	2FD	CCEP	CCEMG	AMG
	(iii)	(iv)	(v)	(vi)	(viii)	(x)	(xi)	(xii)	(xiii)	(xv)
φ	9.649	9.649	9.649	9.649	9.649					
φ_G7						15.22	15.22	15.22	15.22	15.22
φ_NOG7						7.04	7.04	7.04	7.04	7.04
β	.017	.033	-0.054	-0.263	.130	.031***	.043	-0.143	.074	.134
	(1.50)	(1.18)	(-0.216)	(-1.619)	(0.93)	(2.81)	(1.54)	(-1.437)	(0.308)	(1.02)
γ	-.009	-.013	-0.213	0.533	.022					
	(-0.75)	(-0.44)	(-0.464)	(1.187)	(0.14)					
δ	.151	.064	-0.625	-3.968**	-.606	.135	.082	-.982	-3.05*	-.631
	(1.20)	(0.31)	(-0.457)	(-2.399)	(-1.08)	(1.13)	(0.41)	(-0.870)	(1.75)	(-1.07)
γ_G7						.091***	.075	.251	.179	.201*
						(4.48)	(1.41)	(0.878)	(1.31)	(1.71)
γ_NOG7						-0.010	-.017	-.040	.625***	-.110
						(-0.90)	(-0.55)	(-0.131)	(2.130)	(-1.27)
$\overline{\log f_{it}}$			1.00	0.945				1.00	0.936	
$\overline{\log S_{it}^d}$			0.054	-1.158				0.143	-0.63	
$\overline{\log S_{it}^f}$			0.213	1.100						
$\overline{\log H_{it}}$			0.625	2.950				0.982	2.476	
$\mathbf{1}_{G7}\overline{\log S_{it}^f}$								-0.251	-1.21	
$\mathbf{1}_{NOG7}\overline{\log S_{it}^f}$								0.041	0.39	
CDP					1.050					1.041

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

The values corresponding to $\overline{\log f_{it}}, ..., \overline{\log H_{it}}$ refer to the cross section averages of their estimated coefficients.

The values corresponding to CDP refer to the cross section averages of the estimated coefficients of the "common dynamic processs".

TABLE 10

Additional CCE estimates

	CCEMG							CCEP						
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)
φ	9.649	9.649	9.649	9.649	9.649	9.649	9.649	9.649	9.649	9.649	9.649	9.649	9.649	9.649
β	-0.263	0.001	0.039	0.056	0.060	0.060	0.087	-0.054	-0.044	-0.016	0.002	0.004	-0.021	0.029
	(-1.619)	(0.006)	(0.224)	(0.291)	(0.331)	(0.295)	(0.508)	(-0.216)	(-0.357)	(-0.161)	(0.015)	(0.051)	(-0.275)	(0.378)
γ	0.533	0.368	0.097	0.238	0.034	0.136		-0.213	-0.170	-0.101	-0.055	-0.052	-0.111	
	(1.187)	(0.716)	(0.306)	0.833	(0.129)	(0.503)		(-0.464)	(-0.605)	(-0.741)	(-0.309)	(-0.641)	(-0.574)	
δ	-3.968**	-2.539*	-0.809	-0.631	-0.667			-0.625	-0.735	-0.457**	-0.473**	-0.471***		
	(-2.399)	(-1.819)	(-0.968)	(-1.111)	(-1.109)			(-0.457)	(-0.926)	(-1.978)	(-2.059)	(-2.72)		
$\overline{\mathbf{z}}_{wt}$														
$\log f_{it}$	0.945	0.932	1.035	1.016	1.015	1.038	1.064	1.00	1.00	0.925	0.941	0.951	1.00	1.00
$\log S_{it}^d$	-1.158	-1.012	-0.586	-0.087		-0.074	-0.026	0.054	0.044	0.302	0.098		0.021	-0.029
$\log S_{it}^f$	1.100	0.666	0.657		0.020	-0.036		0.213	0.170	-0.157		0.100	1.111	
$\log H_{it}$	2.950	3.350						0.625	0.735					
$\mathbf{d}_t$														
p_oil_t		-0.00012							0.0001					

Notes:

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

The values corresponding to $\overline{\log f_{it}}, ..., \overline{\log H_{it}}$ and to p_oil_t refer to the cross section averages of their estimated coefficients.