The dynamics of knowledge production in European regions

Valter Di Giacinto*

Abstract

The paper uses a novel econometric approach, rooted in the identification and estimation of a spatial panel VAR model, to provide new estimates of the dynamic effects of private and public R&D expenditure on innovative activity in European regions. The positive long-run response of regional patent applications to private R&D investment is confirmed, while the localized returns to non-business research activity are found out to be positive but not statistically significant. Some robust evidence of positive spillover effects across regions is also documented, confirming the previous findings of a sharp decline of external effects as the geographical distance between regions increases. The empirical estimates show a heterogeneous pattern of responses between core and peripheral regions, the latter displaying higher returns to localized R&D activity. Business and non-business R&D are found to be complements in non-core regions, while they are estimated to be essentially substitutes in more advanced areas.

keywords: regional innovation systems, knowledge production function, knowledge spillovers, spatial VAR model, spatial econometrics

JEL classification: C21; C33; F14

This version: November 2012

1. Introduction[1]

Europe 2020, the European Union's growth strategy for the present decade, updating the previous Lisbon Agenda, reaffirms the importance of investment in research and development as a crucial factor to promote innovation and long-run growth in the EU. A quantitative target – 3 per cent of the EU's GDP to be invested in R&D – is explicitly set.

The strong emphasis on R&D expenditure, arguably not an ultimate policy goal per se, is implicitly motivated by the assumption that there is a positive and stable relationship between investment in new knowledge accumulation and the innovative performance of a given economy. In the literature, starting from the seminal work of Griliches (1979), the mathematical function that is assumed to relate the inputs of the innovation process to the final innovative output is called the knowledge production function (KPF).

Knowledge production functions have been studied both at the micro (firm) and the macroeconomic level. The motivation for analyzing aggregate knowledge production functions mainly derives from the emphasis that endogenous growth theory has placed on the role of knowledge spillovers as a crucial factor in promoting innovation and economic growth (Romer, 1990, Aghion and Howitt, 1992). At the same time, there is now a large body of evidence that knowledge diffusion is to some extent geographically limited (Acs and Varga, 2002; Bottazzi and Peri, 2003), a circumstance that has been adduced in order to explain clusters of regions

* Bank of Italy, L'Aquila Branch, Regional Economic Analysis and Research Unit
e-mail: valter.digiacinto@bancaditalia.it

[1] I wish to thank Elena Gennari, Luigi Cannari and participants in the Bank of Italy Workshops in Rome and Perugia for useful comments and suggestions. The views expressed in this paper are the author's and do not necessarily reflect those of the Bank of Italy.

with persistently different levels of growth (Glaeser et al. 1992). In fact, while codified knowledge may be easily communicated across space and time, the transmission of *tacit* knowledge – mostly channeled through face-to-face interactions – is greatly hampered by geographical distance.

The presence of spatial frictions in knowledge diffusion has led a number of authors to focus their theoretical and empirical analyses on regions and regional innovation systems (for a survey, see Döring and Schnellenbach, 2006).

This paper contributes to the empirical literature on knowledge production at the regional level in several respects:

1. Unlike the prevailing approach, which is mostly based on the estimation of static cross-sectional models, a dynamic regional KPF specification is used, allowing the short- and long-run impacts of R&D on innovative active to be separately assessed. In this way the timing of the response of the innovative output to an exogenous shock to R&D can be inferred from the data and need not be imposed a priori.
2. The plausible existence of dynamic feedback effects between the inputs and the outputs of the regional KPF is dealt with by adopting a simultaneous equation approach. In line with the methodology set forth in Bottazzi and Peri (2007), a panel VAR (Vector Autoregressive) model in which all the individual KPF variables are jointly endogenously determined is used for this end in order to achieve a high level of generality and flexibility.
3. A more robust measure of the returns (in terms of innovation output) to R&D is obtained, since the panel data approach makes it possible to control for structural differences between regions by means of regional fixed effects.[2]
4. Business and academic R&D are considered separately within the KPF inputs, and the interaction between the two inputs (complementarity vs. substitutability) is assessed.
5. A more refined measure of knowledge spillovers between regions is produced by properly specifying and estimating a structural spatial VAR (SpVAR) model (Beenstock and Felsenstein, 2007, Di Giacinto, 2010), where specific attention is paid to controling for the influence of shocks simultaneously affecting all regions.
6. Lastly, the heterogeneity of the regional KPFs in Europe is evaluated by fitting a model with different parameters for the sets of core and peripheral EU regions.

The empirical analysis is conducted on a sample of 146 European regions, considering yearly time series covering the period 1993-2008. Following a standard approach, innovative activity is proxied by patent counts.

Before briefly describing the empirical findings, it is important to remark that the impulse-responses presented in the paper, being conditional on area-wide trends, can be deemed to capture only strictly *spatially localized* effects of both public and private R&D. In a small number of cases, research projects produce new knowledge that actually move the technological frontier of the entire continent or even the whole world, thus spurring innovation on a global rather than on a local basis. Such global knowledge spillover effects are deliberately excluded in the empirical model here considered, which focuses on improving the measure of localized returns to R&D and geographically bounded spillover effects.

The main findings produced by the SpVAR model estimation confirm that business sector R&D investment yields substantial positive returns in terms of within-region innovative

[2] In a preliminary panel regression analysis, the paper shows that failure to control for such factors can result in severely biased estimates of the elasticities of new knowledge production to regional R&D investment.

activity. At the same time, the localized impact of non-business R&D expenditure on patenting turns out to be statistically non-significant, at least with respect to the broad measure of innovative activity (total patent counts) considered here.

As expected, the effect of R&D on formal innovation is negligible in the very short run and is shown to be highly persistent over time, displaying a non-negligible influence even at a ten-year horizon. While most of the impact of regional R&D on patenting is observed within-region, knowledge spillover effects between regions are positive and sizeable, although they appear to decay rather quickly with distance.

As regards the local interaction of private and public R&D, the two KPF factors are found out to act essentially as complements in non-core EU regions, where a positive shock to either of to the two inputs is followed by a positive response of the other (crowding in), while there is evidence of substitution (crowding out) in core regions.

The rest of the paper is organized as follows. Section 2 introduces the dynamic KPF panel VAR model. Section 3 extends the baseline model by introducing spatial interactions across system variables, yielding the spatial VAR specification that forms the basis for the subsequent empirical study. The European regional data set utilized for this purpose is detailed in Section 4, which also reports some preliminary regression results. The model specification and the main empirical findings yielded by the baseline spatial VAR model are documented in Section 5. Section 6 introduces an extended specification, allowing for different knowledge production dynamics in core and peripheral EU regions. Section 7 summarizes and concludes.

2. Modeling regional knowledge production dynamics

2.1 The baseline dynamic panel VAR model

In this section the proposed dynamic KPF specification is gradually introduced, moving from a standard static log-linear equation (see e.g. Jaffe, 1989; Feldman and Florida, 1994; Anselin et al., 1997; Acs et al., 2002). When specified on regional panel data, the latter may be expressed as

$$y_{it} = \delta_i + \xi_t + \alpha b_{it} + \beta u_{it} + \varepsilon_{it} \tag{1}$$

where y_{it} is the log of the regional innovation output and where b_{it} and u_{it} respectively denote the logs of the regional business and non-business R&D expenditure flows. The δ_i and ξ_t coefficients are region and time specific constants which account for local differences in the efficiency of the innovative activity (the output level obtained in the *i*-th region for given input levels) and for a common time trend.

Finally, ε_{it} is a zero-mean error term that will generally be correlated with the two R&D inputs.

Regional fixed effects in the panel regression model can proxy for the explanatory variables that are usually included in empirical KPF models to control for differences in the structural features of individual regional economies that are particularly relevant for innovative activity like human, public and social capital endowments, agglomeration economies, the sectoral composition of local economy and the average firm size. As far as the latter are approximately constant over the sample period or evolve according to a common time trend (which is also controlled for) it is hence possible to omit them from the empirical fixed-effects KPF specification.

Unobservable common shocks, when not properly taken into account in the estimating equation, are likely to bias estimation results as they may introduce spurious correlation between the inputs and the output of the KPF. At the same time, common shocks cause model residuals to be cross-sectionally correlated, a feature that has to be properly dealt with when the analysis aims at identifying interaction effects across different regions.

The introduction of region and time fixed effects may also be motivated in relation to the unobservability of the innovation output, a problem which is usually overcome by introducing proxy variables, like patent count statistics. When the following log-linear relation is assumed to link the proxy variable p_{it} to the unobservable innovation output

$$p_{it} = y_{it} + c_i + m_t \tag{2}$$

it is immediate to notice how substituting p_{it} for y_{it} in a standard KPF model would yield an estimating equation featuring both region and time specific effects.

The relation postulated in expression (2) states that the innovation proxy is proportional to the underlying innovation outcome, but the proportion may differ systematically across regions, while shifting symmetrically over time across regions. Both features appear to be highly plausible in the case of patent statistics. It is well known that only a share of total inventions are actually patented and the fraction may well differ across regions according to the structural features of the local economies (Scherer, 1983). At the same time, a common tendency of the propensity to patent new knowledge to shift over time can be envisaged, as a response of economic agents to common shocks affecting the cost of obtaining a new patent or to changes in the overall technological and competitive conditions on product markets, which could make firms more or less active with respect to intellectual property protection.

Equation (1), given its static form, may at most represent the long run equilibrium state of the regional innovation system. In empirical applications, however, it is important to allow for transitory deviations from the steady state equilibrium path.

An empirical specification introducing system dynamics can be obtained by augmenting equation (1) with lags of both the dependent and the independent variables, yielding the following expression

$$y_{it} = \delta_i + \xi_t + \alpha_0 b_{it} + \beta_0 u_{it} + \sum_{h=1}^{p} (\phi_h y_{it-h} + \alpha_l b_{it} + \beta_l u_{it}) + e_{it} \tag{3}$$

where a common lag order p for all variables is assumed without loss of generality.

As is well known, the general dynamic model given in expression (3) nests a number of possible alternative specifications, including the ECM (Error Correction Mechanism) model considered in Bottazzi and Peri (2007) and the convergence equation approach adopted in Crescenzi et al. (2007).

While the dynamic KPF specification may be deemed to improve on the simple static set up of equation (1), it still makes no allowance for the existence of feedback effects across the inputs and the output of the innovative process. This clearly represents a rather severe limitation. Strong dynamic feedback effects may indeed be expected to occur, due to the fact that KPF inputs are not randomly assigned to specific areas but are actually selected according to the optimizing choices of private and public sector agents.

A large literature strand has addressed the analysis of the influence of publicly funded research activities on commercial R&D (see David et al., 2000 for a comprehensive critical review). The existence of feedback effects from private R&D to academic research has also received specific attention in the literature (Jaffe, 1989; Feldman and Florida,1994; Anselin et al., 1997).

On the contrary, the issue of possible feedback effects from innovative activity to private and public R&D investment choices has mostly been neglected in empirical applications. Nonetheless, a number of arguments clearly appear to motivate the existence of reverse causation channels. As noted by Feldman (1994), the cumulative nature of innovation activity suggests that areas which achieved innovative success in the past are likely to attract additional private R&D investment on the expectation of positive localized knowledge spillovers.

The spatial allocation of public R&D expenditure may also be influenced by past regional innovative outcomes. If policy makers aim at maximizing aggregate returns, it should be optimal for them to concentrate funding in locations that have proved to be able of achieving the highest innovative performances in the past. On the contrary, R&D investment should be mainly directed to less innovative regions if public policies target catching up and converge across regions.

The existence of such complex network of feedback effects calls for the implementation of a full-fledged simultaneous equations approach that explicitly recognizes the endogeneity of R&D investment. At the present date, Bottazzi and Peri (2007) represents the only study adopting such a system approach to aggregate KPF estimation. The authors fit a cointegrated VAR model to panel data covering a group of 15 OECD countries and including the national and international knowledge stocks and national R&D employment. In line with expectations, a substantial and positive long run response of R&D employment to a positive shock to the knowledge stock (proxied by accumulated patent counts) is uncovered.

A system approach, extending the model considered in Bottazzi and Peri (2007) by separately considering academic and commercial R&D investment, is set forth here by considering the following recursive system of dynamic simultaneous equations

$$\begin{cases} u_{it} = \delta_{Ui} + \xi_{Ut} + \sum_{h=1}^{p} (\phi_{Uh} y_{it-h} + \alpha_{Uh} b_{it-h} + \beta_{Uh} u_{it-h}) + e_{it}^{U} \\ b_{it} = \delta_{Bi} + \xi_{Bt} + \beta_{B0} u_{it} + \sum_{h=1}^{p} (\phi_{Bh} y_{it-h} + \alpha_{Bh} b_{it-h} + \beta_{Bh} u_{it-h}) + e_{it}^{B} \\ y_{it} = \delta_{Yi} + \xi_{Yt} + \alpha_{Y0} b_{it} + \beta_{Y0} u_{it} + \beta_{Y0} b_{it} + \sum_{h=1}^{p} (\phi_{Yh} y_{it-h} + \alpha_{Yh} b_{it-h} + \beta_{Yh} u_{it-h}) + e_{it}^{Y} \end{cases} \qquad (4)$$

where the vector of the random errors $e_{it} = \{e_{it}^{U}, e_{it}^{B}, e_{it}^{Y}\}'$ is assumed to be distributed as a multivariate white-noise process with diagonal covariance matrix $\Sigma_i = E(e_{it} e_{it}') = diag\{[\varphi_i^{U}, \varphi_i^{B}, \varphi_i^{Y}]\}$. Being mutually orthogonal, the three random terms in system (4) can be given a structural interpretation, respectively as exogenous shocks to the KPF inputs or to the efficiency of the regional innovation system (in the case of the shock in the y_{it} equation).

In the above recursive specification, the identification of the structural shocks can be seen to rest on the following two a priori assumptions:

1. Conditional on past system dynamics, the current evolution of public R&D expenditure is not affected by current shocks to private R&D or to the innovation output.
2. Conditional on past system dynamics and on the current level of public R&D expenditure, private R&D investment is unaffected by unforeseen evolutions in the regional innovation output.

The first assumption does not appear to be stringent, as it is well known that the implementation of public expenditure policies entails considerable delays between the timing of policy decision and its actual implementation. The second assumption is slightly more binding, as in the private sector investment decisions are usually more rapidly implemented than in the public sector. However, it can still be deemed to be credible as far as empirical estimates are based on yearly data, given that the bulk of corporate R&D expenditure is usually set at the beginning of each accounting period.

While the above identifying restrictions represent the preferred maintained assumptions in the paper, the robustness of empirical findings to alternative orderings of the three endogenous variables is properly assessed at the estimation stage.

When jointly evaluated, the three equations in the system detailed in (4) define a structural vector autoregressive (VAR) model, which can be expressed in matrix notation as follows

$$\mathbf{B}_0\boldsymbol{\zeta}_{it} = \boldsymbol{\delta}_i + \boldsymbol{\xi}_t + \mathbf{B}_1\boldsymbol{\zeta}_{it-1} + ... + \mathbf{B}_p\boldsymbol{\zeta}_{it-p} + \eta_{it} \tag{5}$$

where the simultaneous interaction matrix B_0 has the following lower triangular structure, stemming from the recursiveness hypothesis discussed above

$$\mathbf{B}_0 = \begin{bmatrix} 1 & & \\ -\beta_{B0} & 1 & \\ -\beta_{Y0} & -\alpha_{Y0} & 1 \end{bmatrix} \tag{6}$$

and where the following notation has been introduced

$$\boldsymbol{\zeta}_{it} = \{u_{it}, b_{it}, y_{it}\}', \ \boldsymbol{\delta}_i = \{\delta_{Ui}, \delta_{Bi}, \delta_{Yi}\}', \ \boldsymbol{\xi}_t = \{\xi_{Ut}, \xi_{Bt}, \xi_{Yt}\}'$$

$$\mathbf{B}_h = \begin{bmatrix} \beta_{Uh} & \alpha_{Uh} & \phi_{Uh} \\ \beta_{Bh} & \alpha_{Bh} & \phi_{Bh} \\ \beta_{Yh} & \alpha_{Yh} & \phi_{Yh} \end{bmatrix}, \quad h = 1,\ldots, p.$$

2.2 The spatial VAR model

The main rationale for the choice of studying knowledge production functions at the regional level, instead of the individual firm level, can be related to the expected occurrence of spillover effects promoting the diffusion and the accumulation of knowledge across agents jointly located inside a given area (see Audretsch and Feldman, 2004, for a survey of the topic).

On this respect, as long as knowledge externalities are strictly confined within regional boundaries, the panel VAR specification introduced in Section 2 may be deemed to provide an adequate description of the local knowledge production dynamics.

However, when knowledge spillovers or other types of spatial interactions substantially cross regional boundaries, the within-region panel VAR approach suffers from a misspecification problem, due to the omission of explanatory variables (namely the KPF inputs and output in remaining regions) that are likely to be correlated with the local levels of the system variables.

A number of empirical studies have now documented how the innovative output of individual European regions is significantly and positively associated to the level of the R&D inputs or the level of the innovative output recorded in geographically close regions (Bottazzi and Peri, 2003; Moreno et al., 2005; Marrocu et al., 2011). This evidence can be interpreted as providing some indications that spillovers actually cross regional boundaries in the EU, at least when administrative instead of functionally identified borders are considered in the analysis.

In order to obtain unbiased estimates of the regional KPFs, solving the omitted variables problem outlined above, and to gather some new empirical evidence on the size and geographical reach of spatial spillover effects across regional innovation systems, in this section it is shown how the dynamic regional KPF model of Section 2 can be extended by introducing interaction terms across different regions.

To facilitate model exposition, in each given period all cross-sectional observations relating to the three KPF variables are first stacked in a single 3*N*-dimensional vector $\mathbf{z}_t$, where $\mathbf{z}_t = [u_{1t},...,u_{Nt}, b_{1t},...,b_{Nt}, y_{1t},..., y_{Nt}]'$.

The following multi-regional VAR model composed of 3*N* equations is subsequently assumed to govern the spatio-temporal dynamics of the regional knowledge production system

$$\mathbf{C}_0\mathbf{z}_t = \Delta + \Xi_t + \mathbf{C}_1\mathbf{z}_{t-1} + ... + \mathbf{C}_p\mathbf{z}_{t-p} + \mathbf{e}_t \tag{7}$$

where the following positions have been made

$$\boldsymbol{\delta} = [\delta_{U1},...,\delta_{UN}, \delta_{B1},...,\delta_{BN}, \delta_{Y1},...,\delta_{YN}]'$$

$$\Xi_t = \boldsymbol{\xi}_t \otimes \boldsymbol{\iota}_N, \qquad \text{with } \boldsymbol{\iota}_N \text{ equal to an } N\text{-dimensional vector of ones}$$

$$\mathbf{e}_t = [e_{1t}^U, ..., e_{Nt}^U, e_{1t}^B, ..., e_{Nt}^B, e_{1t}^Y, ..., e_{Nt}^Y]'$$

$$E(\mathbf{e}_t \mathbf{e}'_t) = \Omega = diag\{[\omega_1^U, ..., \omega_N^U, \omega_1^B, ..., \omega_N^B, \omega_1^Y, ..., \omega_N^Y]\}$$

and where the coefficients matrices have the following block structure

$$\mathbf{C}_0 = \begin{bmatrix} \mathbf{A}_{11}^{(0)} & 0 & \\ \mathbf{A}_{21}^{(0)} & \mathbf{A}_{22}^{(0)} & \\ \mathbf{A}_{31}^{(0)} & \mathbf{A}_{31}^{(0)} & \mathbf{A}_{33}^{(0)} \end{bmatrix} \tag{8}$$

$$\mathbf{C}_h = \begin{bmatrix} \mathbf{A}_{11}^{(h)} & \mathbf{A}_{12}^{(h)} & \mathbf{A}_{13}^{(h)} \\ \mathbf{A}_{21}^{(h)} & \mathbf{A}_{22}^{(h)} & \mathbf{A}_{13}^{(h)} \\ \mathbf{A}_{31}^{(h)} & \mathbf{A}_{32}^{(h)} & \mathbf{A}_{33}^{(h)} \end{bmatrix} \tag{9}$$

with

$$\mathbf{A}_{qq}^{(0)} = \mathbf{I}_N - \sum_{m=1}^{K} \lambda_{qq,0m} \mathbf{W}^{(m)}, \; q=1,\dots,3 \tag{10}$$

$$\mathbf{A}_{qs}^{(0)} = \sum_{m=0}^{K} \lambda_{qs,0m} \mathbf{W}^{(m)}, \; q=1,\dots,3, \; s=1,\dots,q-1 \tag{11}$$

$$\mathbf{A}_{qs}^{(h)} = \sum_{m=0}^{K} \lambda_{qs,hm} \mathbf{W}^{(m)}, \quad h=1,\dots,p, \; q,s=1,\dots,3 \tag{12}$$

h=0,1,…,p, m=0,1,…,k, q,s =1,…,3.

The assumptions regarding the block structure of coefficient matrices $\mathbf{C}_h$, h=0,1,…, p, qualify the above multi-regional VAR specification as a *spatial* VAR model (Beenstock and Felsenstein, 2007), in the structural block-recursive representation given in Di Giacinto (2010).

While in a standard time series VAR model the coefficient matrices would usually be unrestricted, in the spatial VAR approach the latter are constrained by imposing a specific spatial structure on the underlying set of regional interactions.

By referring to the standard approach in spatial econometrics, spatial structure is embedded in model specification by means of a set of spatial weights matrices $\mathbf{W}^{(m)}$, m=1,…, k, where the order m measures the degree of spatial displacement (also referred to as spatial *lag*) between the different areas. Regions that are connected at a low spatial lag order are assumed to be more closely located to each other compared to regions that are connected at higher spatial lag orders (see Anselin and Smirnov, 19, and LeSage, 1999, for a comprehensive treatment of the definition and computation of spatial lag operators). By assumption, each region is connected with itself at spatial order m=0.

In the above detailed spatial VAR specification, the weights matrices $\mathbf{W}^{(m)}$ can be deemed to identify the range of *potential* interactions across spatially connected areas, while the individual λ coefficients gauge the *actual* strength and sign of spatial interactions both across time and across different system variables.

The temporal and spatial lag orders of the spatial VAR model do not have to be imposed a priori, but are usually dictated by the available sample evidence. Information criteria like Akaike's AIC, Schwarz's BIC and Hannan and Quinn's criterion provide a standard reference to this purpose in time series VAR modeling (see, e.g., Lütkepohl, 2007, and the references therein) that can be extended in straightforward manner to the SpVAR model specification.

By allowing for higher order spatial lags to be introduced when needed, the SpVAR specification here considered ensures a great degree of flexibility in empirical applications, as it does not impose strong restrictions on the spatial range of regional interactions, letting the data mostly "speak for themselves" on this respect. The SpVAR specification can thus accommodate a wide range of possible spatial interaction patterns, ranging from externalities that rapidly cut-off with distance to spillovers that are possibly highly persistent in space.

At the same time, when all λ coefficients take zero value, spatial interactions are entirely ruled out and the spatial VAR model reduces to the non spatial panel VAR model considered in Section 2.

The SpVAR model can be shown to nest as special cases some dynamic spatial panel model specifications that have received attention in the spatial econometrics literature. When P=S=1, singling out the patents equation from the SpVAR system yields the following expression

$$\begin{aligned}\mathbf{y}_t &= G_{Y,t} + \lambda_{31,00}\mathbf{u}_t + \lambda_{31,01}\mathbf{W}^{(1)}\mathbf{u}_t + \lambda_{32,00}\mathbf{b}_t + \lambda_{32,01}\mathbf{W}^{(1)}\mathbf{b}_t + \lambda_{33,01}\mathbf{W}^{(1)}\mathbf{y}_t + \mathbf{e}_t^Y + \\ &+ \lambda_{31,10}\mathbf{u}_{t-1} + \lambda_{32,10}\mathbf{b}_{t-1} + \lambda_{31,11}\mathbf{W}^{(1)}\mathbf{u}_{t-1} + \lambda_{32,11}\mathbf{W}^{(1)}\mathbf{b}_{t-1} + \lambda_{33,01}\mathbf{y}_{t-1} + \lambda_{33,11}\mathbf{W}^{(1)}\mathbf{y}_{t-1}\end{aligned} \quad (13)$$

where the following straightforward notation has been introduced to allow for a direct comparison with the existing literature.

Apart from the deterministic component, equation (13) is formally equivalent to the most general space-time autoregressive distributed lag model specification considered in Elhorst (2001). In the present context, it may be deemed to convey a dynamic spatial version of the simple panel KPF model given by expression (1).

However, an important feature introducing a difference between single-equation dynamic spatial panel modeling approach and the SpVAR approach lies in the assumptions about the explanatory variables (the two R&D inputs, in this case). While the latter are usually assumed to be strictly exogenous in single-equation modeling, they are only predetermined in the SpVAR specification here considered.

When individually singled out, the business and academic R&D equations from an SpVAR(1,1) model may be shown to have the following expressions

$$\begin{aligned}\mathbf{b}_t &= G_{B,t} + \lambda_{21,00}\mathbf{u}_t + \lambda_{21,01}\mathbf{W}^{(1)}\mathbf{u}_t + \lambda_{22,01}\mathbf{W}^{(1)}\mathbf{b}_t + \mathbf{e}_t^B + \\ &+ \lambda_{21,10}\mathbf{u}_{t-1} + \lambda_{21,11}\mathbf{W}^{(1)}\mathbf{u}_{t-1} + \lambda_{22,10}\mathbf{b}_{t-1} + \lambda_{22,11}\mathbf{W}^{(1)}\mathbf{b}_{t-1} + \lambda_{23,01}\mathbf{y}_{t-1} + \lambda_{23,11}\mathbf{W}^{(1)}\mathbf{y}_{t-1}\end{aligned} \quad (14)$$

$$\mathbf{u}_t = G_{U,t} + \lambda_{11,01}\mathbf{W}^{(1)}\mathbf{u}_t + \mathbf{e}_t^U + \\ + \lambda_{11,10}\mathbf{u}_{t-1} + \lambda_{11,11}\mathbf{W}^{(1)}\mathbf{u}_{t-1} + \lambda_{12,10}\mathbf{b}_{t-1} + \lambda_{12,11}\mathbf{W}^{(1)}\mathbf{b}_{t-1}\lambda_{13,01}\mathbf{y}_{t-1} + \lambda_{13,11}\mathbf{W}^{(1)}\mathbf{y}_{t-1} + \tag{15}$$

demonstrating how both $\mathbf{r}_t^B$ and $\mathbf{r}_t^U$ are affected by past dynamics of the innovative output but not by its current evolutions, conditional on past system dynamics.

When explanatory variables are endogenous, as is the case of the dynamic spatial KPF equation given in expression (13), the full system of dynamic feedbacks must be taken into account when measuring the impact of a change in one explanatory variable on the remaining variables. In the time series VAR literature, impulse-response coefficients, which relate the dynamics of the individual endogenous variables to the impact of unforeseen shocks affecting one of the system variables, are considered to this purpose. The approach can be readily extended to the context of the SpVAR model, as shown in Di Giacinto (2010).

Provided matrix $\boldsymbol{C}_0$ is invertible, a condition that can always be achieved by placing sufficient restrictions on the space of admissible values of autoregressive coefficients, the reduced form expression of the SpVAR model can be defined in the usual way, by setting

$$\mathbf{z}_t = \tilde{\mathbf{G}}_t + \tilde{\mathbf{C}}_1\mathbf{z}_{t-1} + \ldots + \tilde{\mathbf{C}}_p\mathbf{z}_{t-p} + \tilde{\mathbf{e}}_t \tag{16}$$

with $\tilde{\mathbf{G}}_t = \mathbf{C}_0^{-1}\mathbf{G}_t$, $\mathbf{G}_t = \Delta + \Xi_t$, $\tilde{\boldsymbol{C}}_h = \boldsymbol{C}_0^{-1}\boldsymbol{C}_h$, h=1,...,p, and $\tilde{\mathbf{e}}_t = \mathbf{C}_0^{-1}\mathbf{e}_t$.

The reduced form VAR expression subsequently provides the basis for the computation of space-time impulse-responses. When the stability condition holds (see, e.g., Lütkepohl, 2007), the SpVAR model admits the following Moving Average (Wold) representation

$$z_t = \sum_{h=1}^{\infty}\Psi_h\left[\tilde{\mathbf{G}}_t + \tilde{\mathbf{e}}_{t-h}\right] = \sum_{h=1}^{\infty}\tilde{\Psi}_h\left[\mathbf{G}_t + \mathbf{e}_{t-h}\right] \tag{17}$$

with $\tilde{\Psi}_h = \Psi_h\mathbf{C}_0^{-1}$.

In the specific case of the three variable SpVAR model considered in the present application, the $\tilde{\Psi}_h$ matrices, h=0,1,…, can be shown to feature the following block structure

$$\tilde{\Psi}_h = \begin{bmatrix} \tilde{\Psi}_{UU}^{(h)} & \tilde{\Psi}_{UB}^{(h)} & \tilde{\Psi}_{UY}^{(h)} \\ \tilde{\Psi}_{BU}^{(h)} & \tilde{\Psi}_{BB}^{(h)} & \tilde{\Psi}_{BY}^{(h)} \\ \tilde{\Psi}_{YU}^{(h)} & \tilde{\Psi}_{YB}^{(h)} & \tilde{\Psi}_{YY}^{(h)} \end{bmatrix} \tag{18}$$

where each NxN block has individual elements

$$\tilde{\psi}_{kr}^{(h)}(i,j) = \frac{\partial z_{ikt+h}}{\partial e_{jrt}}, \qquad k,r=1,...,3 \qquad i,j=1,...,N \tag{19}$$

measuring the response of the k-th endogenous variable on location i at time $t+h$ to a one-off unit shock to the r-th variable on location j and time t.

While partial derivatives are here taken with respect to the shocks driving the system variables rather than the variables themselves, the space-time impulse-response coefficients as defined in (19) bear a close analogy to the spatial multipliers recently considered in LeSage and Pace (2009, p. 74) and extended to a space-time modeling environment by Parent and LeSage (2010). In line with the above contributions, the response coefficients listed on the main diagonal of the $\tilde{\Psi}_h$ matrices may be shown to provide a measure of the effects recorded after h periods *within* the region where the shock initially originated, while the off-diagonal terms gauge the dynamic spillover effects *outside* the region.

A number of statistics useful in summarizing the complex information by the impulse-response matrices can be derived by averaging individual response coefficients as given in expression (19) across regions at varying orders of displacement both across space and time (see Di Giacinto, 2010).

3. The data set and some preliminary regression results

The empirical analysis discussed in this section is conducted on the regional time series data distributed by Eurostat through its web portal. Following a standard practice in the estimation of regional KPF models, innovation is proxied by regional patent counts statistics. While the latter is well known to provide an imperfect measure of the innovative output, it is generally accepted and, in a comparative study considering also a more direct innovation measure, Acs et al. (2002) have confirmed the overall robustness of patent counts as a proxy of innovative activity. The series considered in the analysis refers to the total number of patent applications to the EPO in a given year, individual applications being allocated to the European regions according to the inventor’s residence.

The local amount of resources devoted to R&D activities is measured by the yearly intramural expenditure flows (GERD) expressed at purchasing power parities and constant 2000 prices. R&D expenditure is subdivided between the business sector (BERD) and a residual non business sector, which includes research activities carried out by specific public institutes, universities and private non profit organizations.

All variables are normalized by the regional population level, in order to control for the different size of individual regions, and are subsequently expressed in logs.

As regards the spatial scale of the analysis, in line with previous studies, the data mostly refer to the NUTS2 level of the European geographical classification. However, due to restricted data availability, NUTS1 regions were considered for Belgium, Finland, Portugal and the United Kingdom, and NUTS0 data (national aggregates) were employed in the case of smaller countries (Ireland, Latvia, Lithuania, Slovenia and Slovakia).

Overall, time series for 146 regions, belonging to 15 countries[3], and covering the 1993-2008 period were considered. While patent data series featured almost no missing data, R&D

[3] The countries considered are the following: Austria, Belgium, Finland, France, Germany, Ireland, Italy, Latvia, Lithuania, Netherlands, Portugal, Slovenia, Slovakia, Spain and the United Kingdom. Overseas French and Portuguese regions were not included in the sample due to their remoteness and their peculiarities.

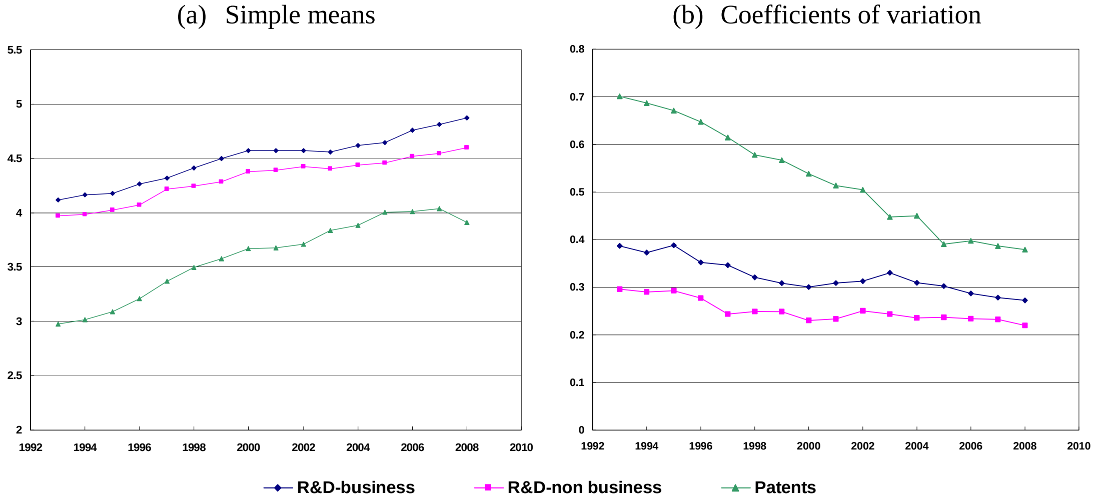
statistics were lacking for about one quarter of the observations. To make the data usable for VAR model estimation purposes, the missing R&D figures had thus to be imputed first. Considering that national aggregate R&D figures were generally available for all countries and years, the missing regional data were estimated by first interpolating regional shares over time (a second order polynomial for each region was utilized for this purpose) and then subdividing the national total across regions according to their respective shares. The interpolated shares were properly rebalanced in each year so that the sum of regional aggregates matches exactly the national total. Some missing R&D data for French and Italian regions could be recovered by referring to national statistical sources (Insee and Istat) and thus no data imputation was required in this case.

The average time dynamics for the indicators of R&D and innovation activity at the regional level, assembled according to the above procedure, are displayed in panel (a) of Figure 1. A clear upward tendency is observed for all the three indicators, the business and non business R&D series showing almost parallel trends, while patents counts show on average a steepest increase up to the year 2004, but subsequently appear to stabilize.

In order to gain some information on the cross-sectional dispersion of the three indicators, the time series of the respective coefficients of variation are displayed in panel (b) of Figure 1.

At the beginning of the sample period, patenting activity showed the highest level of dispersion across regions. However, over the 1993-2008 period, relative dispersion almost halved for the patent count indicator, declining from 70 to less than 40 per cent of the sample mean. A considerable decrease of regional heterogeneity is also observed for the two R&D indicators. Overall, there appears to be evidence of rather strong (sigma) convergence of innovative activity across the set of EU regions considered in the analysis. This feature is not entirely new, as (beta) convergence was also reported in Crescenzi et al. (2007).

FIGURE 1: Summary statistics for the regional R&D and patents series.*



* The graphs plot cross-sectional statistics computed on log-levels data for the 146 EU regions in the panel.

A catching up of less developed regional economies to the innovative performance of more technologically advanced areas might underlie the observed reduction of the cross-regional dispersion of innovative activity indicators.

A large literature strand has addressed regional disparities within Europe. In this context, a core-periphery pattern has been documented with reference to both income and productivity levels (see, e.g., Dignan, 1995; Martin, 1998; Dall'Erba and Le Gallo, 2008), accessibility (Puga, 2002) and foreign direct investment (Kottaridi, 2005).

The synthetic indicator presented in the 2009 Regional Innovation Scoreboard (available at http://www.proinno-europe.eu/metrics) appears to unveil a broadly similar core-periphery pattern also in the case of regional innovative activity, albeit with a few notable exceptions (Finland, Sweden and, to a lesser extent, Ireland).

To what extent the evidence of convergence documented above can be related to a reduction of the core-periphery gap with respect to innovative performance represents thus an issue that appears to be worthwhile investigating.

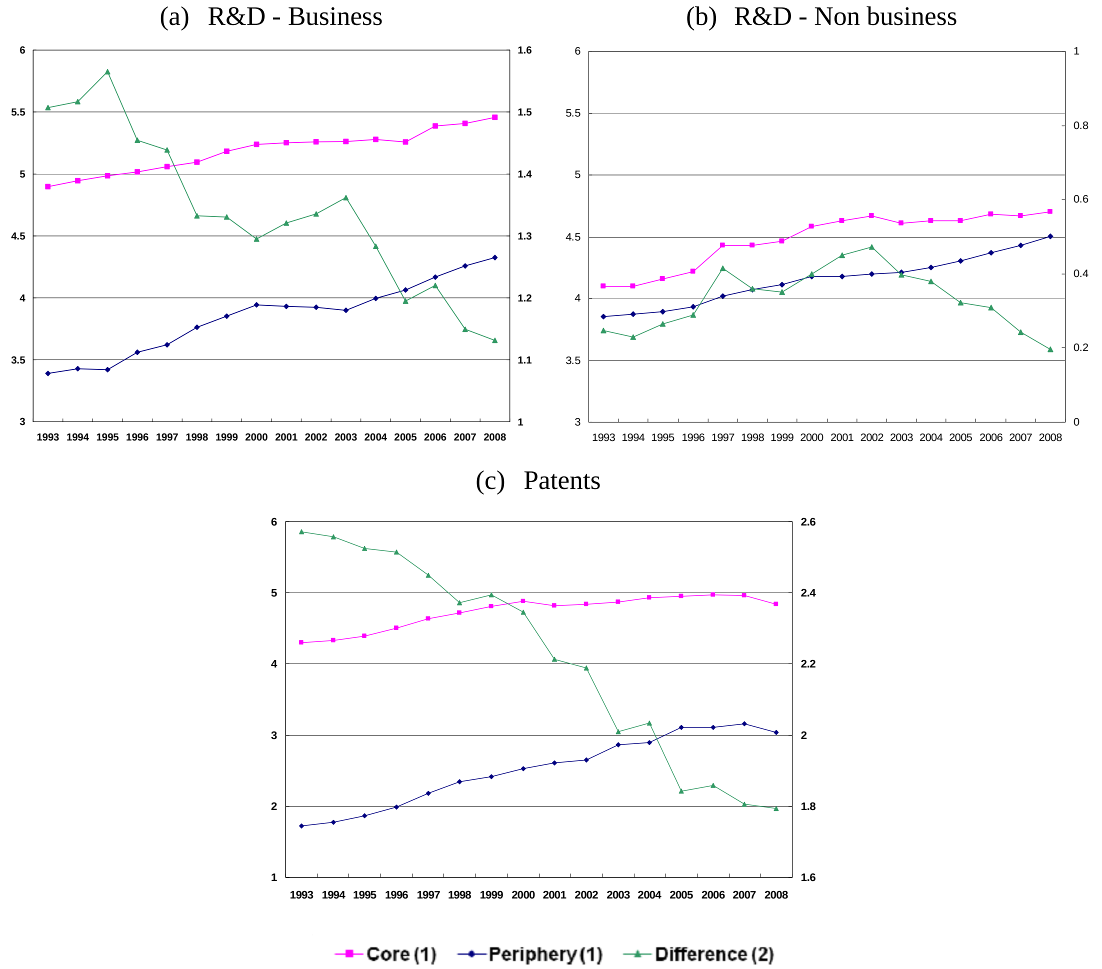
The economic core of the EU is generally identified with the area comprising most of German and French regions, Belgium, Luxembourg and the Netherlands, Southern UK and Northern Italy. A similar approach is followed here in order to derive an operational definition of the set of core regions within the considered sample. More specifically, the core area is identified as a circle centered at the geographical barycentre of the regional sample, as approximately measured by the simple average of the coordinates of the individual regional capitals. According to this procedure, the centre of the regional sample turned out to be located in the vicinity of Strasbourg, which is also the location chosen as the reference EU center in Dignac (1995). A 600 kilometers radium is subsequently considered in order to allow the London region to be included in the circular core area.

While the approach is clearly denoted by a certain degree of arbitrariness, it has the appeal of being grounded mostly on geographical, and hence arguably exogenous, rather than on economic considerations. At the same time, a partition operated along these lines has the advantage of yielding two sets of mostly spatially contiguous core and periphery regions, a feature that is desirable for modeling purposes that will be discussed later in Section X. (i.e. in order to allow for the estimation of differentiated spatial spillovers effects in the core and the peripheral areas).

Figure 2 displays the average time dynamics of regional R&D and patent counts are separately plotted for the 79 core and 67 periphery regions identified according to the above procedure. As expected, periphery regions substantially lag core regions according to all three indicators of innovative activity, although the gap is much smaller for non business R&D, showing how public R&D investment tends to be more evenly distributed across the EU territory.

The graphs provide some visual evidence that a substantial catching up of the non core areas has been continuously in place over the last two decades. The recovery has been particularly strong with respect to the patenting and commercial R&D indicators. The smaller gap suffered by peripheral regions in non commercial research expenditure shows a tendency to rise up the year 2002 but then it also falls systematically, reaching a lower level in 2008 compared to the one recorded at the beginning of the period.

FIGURE 2: R&D and patents dynamics in core and periphery regions. *



* The graphs plot cross-sectional means computed on log-levels data.
(1) Left hand axis. (2) Right hand axis.

Such widely differentiated aggregate dynamics could be driven, at least to some extent, by structural differences in the functioning of local innovation systems in core and periphery areas, regarding, for instance, the degree of complementarity or substitutability of public and private R&D or the intensity of localized knowledge spillovers. Some empirical evidence on this respect can be obtained by fitting separate KPF models to core and periphery regions, a task that will be subsequently undertaken in Section 6.

Some preliminary estimation results are given in Table 1, useful in order to provide some first insights on the data set while allowing a comparison of results with previous studies to be made.

As a first exercise, a pooled cross-section static KPF was initially estimated by OLS. A high, close to unit, elasticity of patent counts to business R&D is obtained in this case, while the elasticity estimate is slightly negative but statistically significant for non commercial R&D (column (a) of Table 1). The underlying estimating equation, however, is well known to suffer from misspecification due to an omitted variable problem. As a first extension, time fixed effects were introduced. Estimation results, displayed in column (b), show no major changes in this case. However, when as a further extension regional fixed effects are also considered, by estimating a panel regression equation as given in expression (1), estimation results display about a sharp drop of business R&D elasticity, to 0.27, while non business R&D elasticity becomes positive and significant, although rather small in size (0.05; see column (c) of Table

1). As regards commercial R&D, elasticity estimates for this model spefication appear to match very closely the values reported by Moreno et al. (2005) and Marrocu et al. (2011) on cross-sectional analyses based on similar EU regional data, although covering a different number of regions.

In order to provide some preliminary insights on the possible underlying heterogeneity of knowledge production functions, the panel regression model was subsequently separately fitted to the set of core and periphery regions as identified above. Higher R&D elasticities are estimated for the panel of peripheral regions with respect to both business and academic R&D, although the latter appears to be rather imprecisely measured (Table 1, columns (d) and (e)).

Overall, the above results appear to provide some support to the hypothesis that, at least as regards the long run equilibrium relation, regional innovation systems may actually differ between the core and periphery of the EU. In both areas, conditional on the structural features of the local economies (as captured by regional fixed effects), strictly decreasing returns to R&D are estimated, but less so for peripheral regions compared to the more advanced core regions.

4. The baseline spatial VAR model

In this section a spatial VAR model as outlined in section 2.2 is fitted to the EU regional panel data. Estimation results are first given for a spatially homogenous specification featuring equal coefficients for all the regions in the panel, although different regional means and variances are allowed for. A specification featuring different coefficients for core and periphery regions is subsequently introduced.

4.1 Model specification and estimation

The empirical specification of a spatial VAR entails a number of subsequent steps (Di Giacinto, 2010). The first step usually involves the assessment of the degree of integration of the individual endogenous variables, by running some panel unit root test procedure.

The Harris and Tzavalis (1999) test, which is based on a pooled ADF statistics whose asymptotic distribution is derived under the assumptions of a large cross-section and fixed time dimension, appears to be well suited to the panel data considered in the present application, where N=146 and T=16. However, to provide a term of comparison, test results are reported also considering the Levin, Lin and Chu (2002) and Im, Pesaran and Shin (2003) panel unit roots test procedures.

The test statistics were computed on data centered on cross-sectional means, to condition on a common time trend, and considering two alternative specifications of the deterministic model component under the null, respectively featuring heterogeneous intercepts and heterogeneous intercepts and linear time trends. In all but one cases, the null hypothesis that the series are integrated is strongly rejected by the individual panel unit root tests (see Table 2), thus providing some robust evidence that the regional R&D and patent series considered in the present analysis are stationary, possibly around a common stochastic trend. Moreover, allowing for region specific deterministic trends does not appear to be strictly required in order to achieve stationarity, as is shown by test results obtained under model specifications including only regional fixed effects.

Some further evidence on the dynamic properties of the three regional time series may be gained from the inspection of the space-time auto and cross correlation functions. The latter were computed on the basis of the pooled version of the bivariate Pearson correlation coefficient across variables lagged in time and space discussed in Martin and Oeppen (1975) and Pfeifer and Deutsch (1980).

As usual, in order to allow for the computation of spatially lagged values, spatial weights matrices up to a given order S have to be first defined. The latter were obtained according to the following procedure. First order spatially contiguous regions were initially identified by introducing the following distance threshold rule

$$\begin{cases} w_{ij}^{(1)} = 1 & \text{if the distance between region } i \text{ and region } j \text{ is smaller than 250 km} \\ w_{ij}^{(1)} = 0 & \text{otherwise,} \qquad\qquad i, j = 1,2,..., N. \end{cases}$$

based on the great circle distance between each couple of regional capitals (in a small number of cases the threshold was risen to 400 km, as no neighboring region could be identified at lower distances).

As is well known, a first order binary contiguity matrix thus specified may be deemed to represent the adjacency matrix of a graph whose nodes are the spatial units (i.e. the regions, in this case) and whose edges connect the couples of units that are contiguous according to the selected rule. The structure of the graph can be subsequently referred to in order to identify higher order spatial neighbors, by applying the following rule

$$\begin{cases} w_{ij}^{(m)} = 1 & \text{if the minumum length path connecting } i \text{ to } j \text{ on the graph is} \\ & \text{composed of } m \text{ edges} \\ w_{ij}^{(m)} = 0 & \text{otherwise} \qquad\qquad i, j = 1,2,..., N, \quad \text{m} = 2,\ldots,S. \end{cases}$$

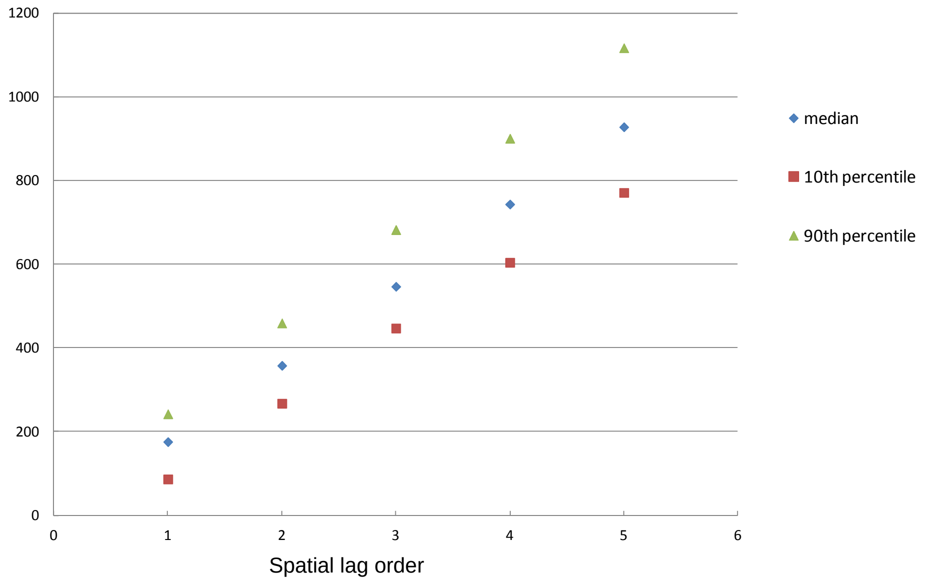
Compared to the alternative approach that is based on the superimposition of concentric rings at increasing distance range around each location, the derivation of higher order spatial weights matrices according to the above procedure has the advantage of avoiding the presence of zero rows in the weights matrices (a situation that can bias estimation results), provided the initial graph is connected and the lag order is not too high.

Each $\mathbf{W}^{(m)}$ identifies the set of m-th order spatial neighbors within the set of all the possible couples of regions in the sample. The distribution of all the bilateral distances between spatially neighboring regions up the 5th order is summarized in Figure 3. It can be noticed how both the median and the 10th and 90th percentiles of the distance distributions monotonically increase as the spatial lag order is incremented. The median distance is equal to about 175 km for 1st order neighbors and then rises almost linearly reaching a value of about 930 km for a lag order equal to 5.

Having ordered spatially neighboring regions according to the above procedure, space-time correlograms were computed up to the 5th lag order both on time and space for all the possible couples of the three model variables. Panel data centered on the respective cross-sectional means and subsequently standardized by region were considered to this purpose, in order to remove both a common time trend and any heterogeneity of regional means and variances. The empirical correlograms are reported in the block matrix displayed in Table 3, where the diagonal blocks refer to the autocorrelation functions of the individual variables and the off diagonal blocks to the bivariate cross-correlations.

For all three indicators the space-time correlograms show a rather quick decay both over time and across space, thus confirming the evidence provided by panel unit root tests that the series are actually stationary once the data are properly transformed.

FIGURE 3: Distribution of bilateral regional distances (in km) at increasing spatial lag orders .



The pattern of the space-time cross-correlation coefficients provide some evidence of a positive and sizeable association between the R&D series and the regional patent statistics, displaying some persistence over time and, especially in the case of non commercial research, over space. A less pronounced, but equally positive and significant association, is observed also between regional business and non business R&D expenditure flows.

Overall, the evidence conveyed but the sample correlograms is compatible with the existence of a complex network of dynamic regional interactions of the type potentially conveyed by an SpVAR model specification as set forth in Section 3.

In order to select the SpVAR specification achieving the best performance on the available sample data, a number of alternative specifications were estimated by maximum likelihood and the preferred specification was selected on the basis of the evidence provided by the AIC and BIC criteria.

The maximum lag order was set equal to 3 in time and 2 in space. Moreover, to provide a term of comparison, a spatial weighting scheme based on the inverse of bilateral distances was also considered as a possible alternative. In this case spatial weights were set according to the expression $w^*_{ij} = d^{-1}_{ij}$, i,j=1,…,N, where d_{ij} is the great circle distance already introduced above. Differently from the approach based on spatial lags of different orders, where distance enters in a discrete fashion through the definition of non overlapping sets of increasingly distant regions, this approach treats distance as a continuous variable and assumes that the degree of potential interaction across regions declines smoothly as distance increases.

Following a standard practice, all spatial weights matrices were row-normalized prior to model estimation, dividing each element by the corresponding row sum.

Some disagreement between the two information criteria is observed, the AIC favoring the less restrictive SpVAR(3,2) specification, while the consistent BIC criterion favors a more parsimonious SpVAR(2,1) model. In all cases, the specifications employing spatial lags define on the basis of nearest neighboring regions outperformed the model specifications based on inverse-distance spatial weights.

The final choice was based on the more restrictive but consistent Bayesian information criterion, and consequently an SpVAR(2,1) model was fitted to the regional data.

Maximum likelihood estimation results for the three SpVAR model equations are separately given in Table 4. As a general remark, it has to be noted that the conditional ML estimators here considered (see Di Giacinto, 2010, for the analytical derivations) are based on standard time series asymptotics and are consistent for fixed *N* and $T \to \infty$, although they are biased for finite *T*. Lee and Yu (2010) derive the analytical expression of the bias of ML estimators for a dynamic spatial panel model specification with time fixed effects that is closely related to the individual equations within the SpVAR model here considered (see Section 2.2). The authors show that the bias is of order $O(T^{-1})$, and is thus rapidly declining as the time series dimension of the panel grows large. They also derive bias-adjusted ML estimators achieving a better performance for small values of *T*. Unfortunately, Lee and Yu bias-adjusted estimators do not apply directly to the case of the SpVAR model, which involves the use of predetermined rather than strictly exogenous explanatory variables and introduces lag orders higher that one both in time and space. Nonetheless, the Monte Carlo simulation results reported in Lee and Yu (2010) prove that the bias is very small, and actually negligible for the parameters measuring the strength of spatial interactions, already for values of *T*=10 and *N*=49, which are smaller than the dimensions of the panel considered in the present study.

ML estimation results show how a large number of the coefficients associated to spatially lagged variables are highly statistically significant, although individual coefficients appear rather small in size. Considering that the empirical specification explicitly controls for regional fixed effects and separate common trends for core and periphery regions, these estimation results may be deemed to provide some new and robust evidence that substantial interaction across regional innovation systems takes place in the context of the production and dissemination of new knowledge in the EU.

As a final diagnostic check, space-time correlograms for model residuals were computed and reported in Table 5. If the SpVAR model is correctly specified residual correlations should be equal to zero at all lag orders. The evidence conveyed by sample estimates show how the model fit is indeed satisfactory, considering that residual correlations are usually very low (albeit sometimes statistically significant) and in no case a value larger than 0.1 is observed.

Once the model is properly specified and estimated, the sign and magnitude of within and between region effects of a localized shock affecting the inputs or the output of the regional knowledge production activity can be assessed by the inspection of the space-time impulse response functions (STIR, Di Giacinto, 2006 and 2010).

4.2 Impulse response analysis

The model-based evidence on the dynamic effects of random shocks to the level of the individual regional KPF system variables can be assessed by inspecting the impulse-response function of the SpVAR system.

The impact of a permanent increase in the system variables is considered to this purpose by referring to accumulated dynamic responses. Accumulated responses appear actually to be the relevant measures from a policy point of view compared to the analysis of transitory, one-off shocks, considering that policy objectives in terms of R&D, like those stated in the Europe 2020 strategy, usually point to a systematic, persistent increases in the level of R&D expenditure.

For clarity of exposition in this section the responses recorded within the own region and in neighboring regions are reported and comment separately. Cross-sectional averages of the within-region effects to a unit shock are displayed first in Figure 4 considering all possible couples of impulse and response variables in the system, yielding a total of nine graphs arranged in a 3x3 array. Confidence bands of ± 1 standard error, yielding approximate 68 percent confidence intervals (an usual choice in time series VAR modeling), complement the point estimates of the response coefficients.

The three plots along the main diagonal of the array portray the responses of patent counts business and non business R&D to their own shock. The pattern of these responses, that are not of great interest per se, show how the impact accumulates gradually over a rather long time span, stabilizing at their respective long run levels about ten years after the initial shock.

The two remaining plots on the first row of Figure 1 display respectively the response of regional patenting activity to a unit shock respectively affecting business and non business R&D expenditure within the region and are thus directly related to the concept of a knowledge production function relating the innovative output to the corresponding inputs.

In line with expectations and previous empirical evidence, the impact on patent applications of a shock affecting business R&D is positive and sizeable in the long run, although it is negligible in the short term (see panel (b) of Figure 4). The accumulated effect of a unit shock is equal to about 0.1 after two years, 0.4 after ten years and eventually reaches a level of 0.6 in the long run. The impact of an exogenous increase of private R&D investment on innovative output within the region where the research activity is carried out hence appears to be substantial. According to these estimates a 10 percent permanent increase of real per capita business R&D expenditure in a region is estimated to induce eventually a 6 percent rise of per capita patent applications.

While a smaller impact of non commercial R&D on patenting activity was largely expected, the evidence conveyed by the SpVAR impulse-response function (plotted in panel (c) of Figure 4) essentially points to the absence of any effect of local non business R&D on regional patent counts at all time horizons. The response estimates are slightly positive in the short run and slightly negative after a few years but they are never statistically different from zero.

This result may seem puzzling when we consider that VAR model response coefficients should capture not only the direct impact of an increase of publicly funded R&D on firm innovative activity but also any indirect impact operating through the additional activation of business R&D.

However, two possible explanations can be advocated to motivate this occurrence. On the one hand, the spillovers of public research activities may be widely diffused in space and hence they could be uncovered only up to a negligible amount by focusing strictly on their impact within the single region. On the other hand, public R&D expenditure may substitute rather than complement commercial R&D, leaving the level of the regional innovation output unchanged.

Some empirical evidence on the possible existence of such displacement effects can be gathered by inspecting the dynamic response of business R&D to an exogenous shock to non commercial R&D, plotted in panel (f) of Figure 4.

The response is estimated to be negligible in the short run but it then turns increasingly negative as the time horizon enlarges after the initial shock, attaining a value of -0.3 in the lung run. This pattern is consistent with the operating of a substitution process slowly crowding out private R&D within the region after an exogenous increase of local public R&D expenditure which could hence help explain why a positive shock to non business R&D leaves the innovative output overall unaffected.

The responses of private and public R&D to an autonomous unit shift of local patenting intensity is positive in both cases and highly statistically significant (see panels (d) and (g) of Figure 4). As expected the response is considerably higher for business R&D (slightly above 0.8 in the long run) but it sizeable also for non commercial R&D (0.3 in the long run).

These evidence confirms the substantial endogeneity issue of the inputs of knowledge production functions, a feature already noted and discussed in the related literature. A positive exogenous shock to the Patent series, that can be interpreted as an unforeseen increase of the efficiency of knowledge production (more output is obtained from given input levels) may be expected to feedback on R&D expenditure choices as it increase the expected returns of R&D investment.

The VAR approach, under the identifying assumptions discussed above, allows for the identification of dynamic feedback effects between business and non business R&D running both ways. While the impact of an exogenous shock to public R&D on private R&D was found out to be eventually negative, the opposite situation holds in the case of an autonomous shock to private R&D. As can be gathered by inspecting panel (h) of Figure 4, after an initial slight decline, public R&D expenditure within the region progressively increases after a positive shock to local private R&D, albeit the effect is not pronounced, the long run accumulated response to a one unit shock attaining a value of about 0.2.

The SpVAR-based evidence on the scope, sign and size of spatial spillover effects in the KPF across neighboring regions is summarized in Figures 5 and 6, which respectively plot the average response recorded on first and second order spatial neighbors of a random shock affecting one of the KPF variables in a given region.

The spatial spillover effects on more closely located regions (1^{st} order neighbors) are generally found out to be positive and sizeable, with the only exception of the response of commercial R&D to an unforeseen increase of public R&D which, as already noted for the within-region effect, is eventually negative in the long run.

The response of patent applications in closely neighboring regions to a unit shock to business R&D gradually accumulates up to a value of about 0.3 in the long run, which is both large in size, being equal to half the impact recorded within the originating region, and highly statistically significant .

In a similar fashion, an autonomous increase of patenting activity within a given region is also found out to induce a positive response of patent applications in neighboring regions that is significant and about the same size as recorded for the spillover of private R&D.

At the same time, in line with the evidence above documented for within region effects, shocks to academic R&D turn out to have also negligible spatial spillover effects outside the region.

Moving to estimated second order spatial spillover effects, it can be noticed how while the overall pattern of impulse-responses is preserved and many response coefficients remain statistically significant, the size of the response displays a 5 to 10-fold drop. Based on model estimation results, spatial spillovers from innovative activities would thus mostly be confined within the region where research and patenting activities take place and in immediately adjacent areas.

Overall, the above findings can be deemed to provide some novel empirical evidence showing how: 1) knowledge spillovers and other forms of spatial interactions actually cross regional boundaries, at least at the spatial scale of the present analysis (which is mostly based on the NUTS2 classification level); 2) geographical distance still hampers knowledge diffusion, as most spillover effects are recorded broadly within a 300 kilometers range.

5. The extended spatial VAR model

The SpVAR-based evidence detailed in the previous section assumes the existence of a unique relation linking KPF variables across the whole regional panel. As such, it may fail to convey all the relevant information if the underlying knowledge production dynamics are highly heterogeneous across regions.

In the preliminary analysis set forth in Section 3 it has been shown how when the EU regional panel is subdivided in two groups of core and peripheral regions some discrepancies appeared to show up, as evidenced by the significantly different elasticities yielded by the panel KPF regression analysis.

In order to gather further insights on the possibly peculiar features of the aggregate knowledge production process in core and non core areas of the EU, an extended SpVAR model was fitted to the data, allowing different VAR coefficients across the two broad areas.

The model thus specified allows for potentially highly differentiated behavior of the KPF system dynamics across the two spatial partitions.

The ML estimation results for the extended model are displayed in Table 7. While no large differences in model fit are observed with respect to the baseline model, some noticeable differences in model parameters across the two groups can be detected. In particular, contemporaneous non business R&D expenditure is positive and significant in the Patents equation on non core regions it is not significant in core regions. At the same time, while the public R&D coefficient is significant and negative in the private R&D equation for core regions, it is essentially negligible (although still negative) in the same equation for peripheral regions.

However, while comparing coefficients may yield some initial hints, as usual in VAR modeling, a comprehensive evaluation of any differences in system dynamics across the two areas has to be based on the comparison of the impulse-response functions.

The response functions are separately plotted for core (black line) and non core regions (red line) in Figure 7 for within-region effects and in Figures 8 and 9 for spatial spillover effects.

Starting from dynamic responses internal to the region, it can be immediately gathered how some relevant differences between the core and peripheral areas appear to stand out.

As regards the response of Patents to exogenous shocks to commercial R&D, in line with the evidence conveyed by the preliminary panel regression, the impact of a shock of the same size (1 unit) appears to be substantially larger in peripheral regions (0.8 in the long run) than in core regions (about 0.3). The impact is positive in non core regions also for academic research (about 0.1 in the long run), although it very imprecisely measured, while absolutely no effect of public R&D on regional patent counts is estimated for core regions.

The stronger evidence of a diverging behavior between the two partitions is uncovered with respect to the relation linking business and non business R&D investment. The two types of KPF inputs appear to substitute for each other in core regions, as both cross responses are eventually negative, sizeable and statistically significant (see panels *f* and *h* of Figure 7). On the contrary there is strong evidence of complementarity in non core regions. In this area a positive unit shock to public R&D eventually triggers about a 0.4 increase of business R&D, and a slightly stronger response (about 0.5 in the long run) is observed for public R&D after an exogenous expansion of the local business R&D effort.

While other mechanisms may also be at play, the evidence of mutual *crowding in* of academic and commercial R&D in less advanced regions, as opposed to the *crowding out* observed in more developed core regions may be an important factor in explaining the prior evidence of a larger impact of local R&D investment on knowledge production in non core areas.

As regards the feedback effects of patenting activity on R&D, they are positive and accumulating over time for both core and peripheral regions, although the response recorded in latter is more than double in size compared to the former. Assuming that a positive shock to the patent indicator, given the current and past level of the inputs, represents an exogenous increase in the efficiency of the knowledge production process, the above evidence shows how more R&D investment is activated after the shock in less developed regions compared to more advanced ones.

Coming to the analysis of spatial spillover effects on more closely located regions (1st order neighbors), some substantial dissimilarities also appear to stand out.

While strong positive spillovers on patenting activity in neighboring regions from both private and public R&D expenditure are uncovered in the EU periphery regions, almost no effect is found out in core regions (Figure 8, panels *b* and *c*). Also the spatial transmission of random shocks affecting innovative performance (i.e. shocks to the Patent variable in the model) to both the knowledge production inputs and output appears to be much stronger in peripheral areas compared to the EU centre (see panels *a*, *d*, and *g* of Figure 8).

Moreover, negative (*crowding out*) spillovers from private to public R&D and vice versa are estimated for core regions, while the opposite holds for the remaining group of regions.

The analysis of average responses on 2nd order spatial neighbors confirms how the spatial spillover effects are greatly hampered by distance, but much less so in the EU periphery compared to the core (Figure 9).

6. Summary and conclusions

In this paper a novel approach to the empirical analysis of regional knowledge production systems was considered in order to better identify and estimate the dynamic linkages underlying the production and spatial dissemination of new knowledge within and across the EU regions.

In contrast with most of the previous studies in this field, a dynamic system approach was adopted, and introducing cross-region interactions in a dynamic simultaneous equation panel KPF approach is shown to yield a structural spatial VAR model specification in keeping with recent advances in spatial econometrics.

The empirical analysis was carried out on Eurostat regional data which, after some missing data issues were initially deal with, provided continuous yearly time series for a wide sample of UE regions covering a period of about two decades.

An initial exploratory analysis of the data set showed that regional time series displayed broadly common trends over the sample period, although with some differences of long-term growth rates between regions located close to the EU economic core and more peripheral ones; the latter succeeded in reducing the large gap in innovative performance recorded at the start of the 1990s.

As a further preliminary analysis, also providing some useful terms of comparison with previous studies, a standard Cobb-Douglas knowledge production function was fitted to the data. Controlling for fixed regional effects turned out to have a dramatic effect on estimated elasticities, which display smaller magnitudes than those traditionally estimated, although they are in line with the empirical findings of more recent studies also based on Eurostat regional data. Static knowledge production function elasticities were also estimated to differ significantly between core and non-core EU regions.

Following the preliminary analysis, a first version of the SpVAR was specified and estimated. A recursive causal ordering of the knowledge production function variables, imposing some mild restrictions on the short-run dynamics of the system, was assumed in order to identify exogenous local shocks to the inputs and output of the regional knowledge production system. Space-time impulse-response analysis was subsequently carried out in order to trace the dynamics of the effects of random shocks disturbing any of the variables in the KPF system.

In line with most previous empirical evidence, an exogenous shock to private R&D expenditure is found to exert a positive effect on knowledge production within the region, but it takes considerable time – more than ten years – for the initial shock to fully produce its effects. These are essentially negligible in the short run

The localized impact of non-business R&D expenditure on patent counts turns out to be not statistically significant, at least with respect to the broad measure of innovative activity used here and insofar as strictly spatially localized effects are considered. Nevertheless, academic research may exert an appreciable influence in other directions, e.g. it may affect knowledge accumulation globally rather locally or it may be relevant for specific fields of the innovative activity (basis research) that are not properly accounted for by regional overall patent counts statistics. Discovering such effects clearly goes beyond the scope of the present study.

The system approach to the analysis of regional knowledge production dynamics that was implemented showed the existence of strong feedback effects from past regional innovative performance on local R&D expenditure, confirming the need to control for the endogeneity of R&D properly when performing an empirical KPF analysis.

In order to get some insights into possible heterogeneity of the KPF across the EU area, the SpVAR model was re-estimated allowing for different dynamics between core and non-core regions.

The estimation results showed that the average findings obtained fitting the same model to the whole sample of EU regions mask some deep dissimilarities. On the one hand, innovative output was estimated to be substantially more elastic to R&D inputs in less advanced regions, a feature that may help to explain the convergence of innovative activity between core and non-core regions discovered in the initial descriptive analysis.

On the other hand, sharp differences are found out as regards the complementarity or substitutability of public and private R&D at the regional level. There is strong evidence that the two inputs are complementary in the non-core regions; while the results appear to favour substitution in core regions, where, in particular, academic research is estimated to crowd out business R&D.

Finally, there is evidence of considerable inter-regional knowledge spillover effects. Such spatial externalities were measured in a modelling environment that makes it possible to control for common trends, for fixed regional effects and for the simultaneity of KPF factors, and may thus represent a considerable improvement on previous empirical estimates derived from single equation static cross-sectional KPF models. The range of the spatial propagation of the shocks, however, turned to be rather limited, although less so in the peripheral regions than in the core regions of the EU.

REFERENCES

Acs, Z.J., L. Anselin and A. Varga (2002), "Patents and Innovation Counts as Measures of Regional Production of New Knowledge", *Research Policy*, Vol. 31, pp. 1069-1085.

Acs, Z.J. and A. Varga (2002), "Geography, Endogenous Growth, and Innovation", *International Regional Science Review,* Vol. 25, pp. 132-148 (special issue on Regional Innovaiton Systems, edited by Z.J. Acs and A. Varga).

Aghion, P. and P. Howitt (1992), "A Model of Growth Through Creative Destruction", *Econometrica*, Vol. 60, pp. 323-351.

Anselin, L. and O. Smirnov (1996), "Efficient Algorithms for Constructing Proper Higher Order Spatial Lag Operators". *Journal of Regional Science* Vol. 36, pp. 67-89.

Anselin, L., A. Varga and Z. Acs (1997), "Local Geographic Spillovers between University and High Technology Innovations," *Journal of Urban Economics*, Vol. 42, pp. 442-48.

Audretsch, D.B. and M.P. Feldman, (2004) "Knowledge Spillovers and the Geography of Innovation", in: J.V. Henderson and J.F. Thisse, Editor(s), *Handbook of Regional and Urban Economics*, Elsevier, Vol. 4, pp. 2713-2739.

Baltagi, B.H and C. Kao (2001), "Nonstationary panels, cointegration in panels and dynamic panels: A survey", in B.H. Baltagi, T.B. Fomby, R. Carter Hill (ed.) *Nonstationary Panels, Panel Cointegration, and Dynamic Panels (Advances in Econometrics, Volume 15),* Emerald Group Publishing Limited, pp.7-51.

Beenstock, M. and D. Felsenstein (2007), "Spatial Vector Autoregressions", *Spatial Economic Analysis*, Vol. 2, pp. 167-196.

Bottazzi, L. and G. Peri (2003), "Innovation and Spillovers in Regions: Evidence from European Patent Data", *European Economic Review*, Vol. 47, pp. 687–710.

Bottazzi L. and G. Peri (2007), "The International Dynamics of R&D and Innovation in the Long Run and in the Short Run", *The Economic Journal*, Vol. 117, pp. 486–511.

Crescenzi, R., A. Rodriguez-Pose and M. Storper (2007), "The territorial Dynamics of Innovation: a Europe-United States Comparative Analysis", *Journal of Economic Geography*, Vol. 7, pp. 673-709.

Dall'Erba, S. and J. Le Gallo, (2008),"Regional Convergence and the Impact of European Structural Funds over 1989-1999: A spatial Econometric Analysis," *Papers in Regional Science*, Vol. 87, pp. 219-244.

Di Giacinto, V. (2006), "A generalized Space-Time ARMA Model, with an Application to Regional Unemployment Analysis in Italy", *International Regional Science Review*, Vol. 29, pp. 159-198.

Di Giacinto, V. (2010), "On Vector Autoregressive Modeling in Space and Time", *Journal of Geographical Systems*, Vol. 12, pp. 125-154.

Dignan, T. (1995), "Regional Disparities and Regional Policy in the European Union", *Oxford Review of Economic Policy*, Vol. 11, pp. 64-95.

Döring, T. and Schnellenbach, J. (2006), "What do we Know about Geographical Knowledge Spillovers and Regional Growth?: A Survey of the Literature", *Regional Studies*, Vol. 40, pp. 375-395.

Elhorst, J.P. (2001), "Dynamic Models in Space and Time", *Geographical Analysis*, Vol. 33, pp. 119-140.

Feldman, M.P. (1994), *The Geography of Innovation*, Kluwer Academic, Dordrecht.

Feldman, M.P. and R. Florida (1994), “The Geographic Sources of Innovation: Technological Infrastructure and Product Innovation in the United States”, *Annals of the Association of American Geographers*, Vol. 84, pp. 210-29.

Glaeser, E.L., Kallal H.D., Scheinkman J.A. and Shleifer A. (1992) “Growth in Cities”", *Journal of Political Economy*, Vol. 100, pp. 1126-52.

Griliches, Z. (1979), “Issues in Assessing the Contribution of R&D to Productivity Growth”, *Bell Journal of Economics*, Vol. 10, pp. 92-116.

Harris, R.D.F. and E. Tzavalis (1999), “Inference for unit roots in dynamic panels where the time dimension is fixed”, *Journal of Econometrics*, Vol. 91, pp. 201–226.

Im, K. S., Pesaran M. H. and Shin Y. (2003), “Testing for Unit Roots in Heterogeneous Panels”. *Journal of Econometrics*, Vol. 115, pp. 53-74.

Jaffe, A.B. (1989), “Real Effects of Academic Research”, *American Economic Review*, Vol. 79, pp. 957-970.

LeSage, J.P. (1999), “The Theory and Practice of Spatial Econometrics”, *mimeo*, available at http://www.spatial-econometrics.com.

Levin, A., C.F. Lin and C.S.J. Chu (2002), “Unit root tests in panel data: Asymptotic and finite-sample properties”, *Journal of Econometrics*, Vol. 108, pp. 1–24.

Kottaridi, C. (2005), “The ‘Core–Periphery’ Pattern of FDI-led Growth and Production Structure in the EU”, *Applied Economics*, Vol. 37, pp. 99-113.

Lee, L. and J. Yu (2010), “A Spatial Dynamic Panel Data Model with Both Time and Individual Fixed Effects”, *Econometric Theory*, Vol. 26, pp. 564-597.

Lütkepohl, H. (2007), *New Introduction to Multiple Time Series Analysis*. Berlin: Springer-Verlag.

Marrocu, E., R. Paci and S. Usai (2011), “The Complementary Effects of Proximity Dimensions on Knowledge Spillovers”, CRENoS, Working Papers, No. 2011/21.

Martin, P. (1998), “Can Regional Policies Affect Growth and Geography in Europe?”, *World Economy*, Vol. 21, pp.757–774.

Martin, R.L. and J.E. Oeppen (1975), “The Identification of Regional Forecasting Models Using Space Time Correlation Functions”, *Transactions of the Institute of British Geographers*, Vol. 66, pp. 95-118.

Moreno, R., R. Paci and S. Usai (2005), “Spatial Spillovers and Innovation Activity in European Regions”, *Environment and Planning* A, Vol. 37, pp. 1793–1812.

Parent, O. and J.P. LeSage (2010), “A spatial dynamic panel model with random effects applied to commuting times”, *Transportation Research Part B: Methodological*, Vol. 44, pp. 633-645.

Pfeifer, P. E. and S. J. Deutsch (1980), “A Three-Stage Iterative Procedure for Space-Time Modeling”, *Technometrics*, Vol. 22, pp. 35-47.

Puga, D. (2002), “European Regional Policies in Light of Recent Location Theories”, *Journal of Economic Geography*, Vol. 2, pp. 373-406.

Romer, P.M. (1990) “Endogenous Technological Change”, *Journal of Political Economy*, Vol. 98, pp. S71-102.

Scherer, F.M. (1983), “The Propensity to Patent”, *International Journal of Industrial Organization*, Vol. 1, pp. 107-128

Table 1. Panel regression results: dependent variable = log (Patents) *

Explanatory variables	(a) All regions	(b) All regions	(c) All regions	(d) Core †	(e) Periphery †
log (R&D-business)	0.977 (0.020)	0.972 (0.020)	0.268 (0.045)	0.134 (0.035)	0.242 (0.065)
log (R&D-non business)	-0.024 (0.019)	-0.037 (0.018)	0.054 (0.023)	0.034 (0.016)	0.097 (0.079)
Fixed time effects	No	Yes	Yes	Yes	Yes
Fixed region effects	No	No	Yes	Yes	Yes
R-squared	0.704	0.711	0.956	0.955	0.938
Root-MSE	0.786	0.781	0.315	0.183	0.406
Observations	2,336	2,336	2,336	1,264	2,336

* All variables are normalized by regional population. Robust standard errors are reported in parentheses.
† Estimated elasticities are not directly comparable to those reported under column (c) because separate temporal effects for core and periphery regions are considered in this case.

Table 2. Panel unit root test results (*p*-values in parentheses)

Variables	Harris and Tzavalis		Levin, Lin and Chu*†		Im, Pesaran and Shin†	
	(a)	(b)	(a)	(b)	(a)	(b)
Log (Patents)	-27.41 (0.000)	-27.41 (0.000)	-7.53 (0.000)	-17.55 (0.000)	-5.09 (0.000)	-13.76 (0.000)
log (R&D-business)	-8.08 (0.000)	-8.08 (0.000)	-7.34 (0.000)	-13.80 (0.000)	-0.68 (0.248)	-8.25 (0.000)
log (R&D-non business)	-4.96 (0.000)	-4.96 (0.000)	-12.63 (0.000)	-6.60 (0.000)	-6.18 (0.000)	-2.54 (0.000)

* The Newey-West long run variance estimator is employed, assuming a Bartlett window. † Individual ADF lags selected according to BIC (max lag=2).
(a) The test assumes individual intercepts under the null. Cross sectional means are removed from the data. (b) Equal to (a) but individual time trends are also allowed for.

TABLE 3. Space-Time Auto and Cross-correlations for the three model variables.

	Lag: Time	Patents						Business R&D						Non business R&D					
		Lag: Space						Lag: Space						Lag: Space					
		0	1	2	3	4	5	0	1	2	3	4	5	0	1	2	3	4	5
Patents	0	1.00	0.48	0.39	0.31	0.27	0.06	0.30	0.38	0.39	0.37	0.24	0.09	0.13	0.13	0.15	0.16	0.33	0.24
	1	0.44	0.37	0.33	0.26	0.18	0.06	0.24	0.31	0.32	0.31	0.19	0.09	0.09	0.10	0.12	0.13	0.30	0.23
	2	0.29	0.27	0.24	0.19	0.14	0.05	0.17	0.25	0.25	0.21	0.12	0.05	0.04	0.05	0.08	0.09	0.23	0.20
	3	0.17	0.19	0.19	0.15	0.09	0.03	0.11	0.17	0.18	0.13	0.08	0.04	0.03	0.05	0.07	0.09	0.20	0.15
	4	0.04	0.07	0.07	0.06	0.04	0.00	0.06	0.10	0.11	0.07	0.05	0.04	0.00	0.04	0.04	0.06	0.13	0.09
	5	-0.04	0.03	0.02	0.00	0.01	-0.02	0.02	0.06	0.07	0.03	0.04	0.04	-0.01	0.04	0.03	0.05	0.11	0.09
Bus. R&D	0	0.30	0.38	0.33	0.34	0.19	0.13	1.00	0.45	0.38	0.36	0.32	0.17	0.12	0.12	0.16	0.24	0.28	0.30
	1	0.27	0.34	0.30	0.30	0.19	0.11	0.59	0.35	0.30	0.29	0.25	0.13	0.07	0.09	0.13	0.22	0.26	0.30
	2	0.21	0.26	0.23	0.25	0.16	0.08	0.39	0.28	0.23	0.22	0.18	0.08	0.05	0.05	0.10	0.18	0.23	0.27
	3	0.14	0.17	0.16	0.18	0.10	0.04	0.23	0.18	0.14	0.13	0.08	0.03	0.02	0.01	0.05	0.13	0.15	0.22
	4	0.07	0.07	0.07	0.09	0.04	-0.01	0.09	0.11	0.09	0.08	0.05	0.02	0.01	0.00	0.03	0.08	0.08	0.14
	5	0.02	0.01	0.01	0.03	-0.01	-0.03	-0.02	0.04	0.03	0.02	0.00	-0.02	-0.01	-0.02	0.00	0.03	0.03	0.09
Non bus. R&D	0	0.13	0.14	0.15	0.16	0.19	0.17	0.12	0.12	0.20	0.22	0.19	0.22	1.00	0.46	0.35	0.16	0.05	-0.12
	1	0.11	0.13	0.14	0.14	0.13	0.11	0.12	0.12	0.17	0.17	0.14	0.16	0.59	0.33	0.26	0.11	0.05	-0.07
	2	0.11	0.12	0.12	0.11	0.07	0.05	0.11	0.09	0.12	0.11	0.10	0.09	0.36	0.19	0.17	0.06	0.05	-0.04
	3	0.10	0.08	0.08	0.07	0.04	0.02	0.09	0.05	0.07	0.05	0.06	0.03	0.16	0.08	0.09	0.03	0.05	-0.02
	4	0.08	0.06	0.03	0.02	-0.02	-0.03	0.08	0.05	0.05	0.03	0.04	-0.02	0.03	-0.02	0.01	0.00	0.04	0.03
	5	0.06	0.03	0.00	0.00	-0.03	-0.04	0.06	0.03	0.03	0.00	0.02	-0.04	-0.03	-0.04	0.00	0.01	0.02	0.03

TABLE 4. Information criteria for alternative SpVAR(*P*,*S*) specifications (1)

Model order (*P*,*S*)	*LOG-LIK.*	*AIC*	*BIC*
Contiguity-based weights			
(1,1)	9,646.1	-18,218.2	-15,238.6
(2,1)	9,750.5	-18,390.9	-15,311.5
(3,1)	9,789.2	-18,432.3	-15,253.0
(1,2)	9,680.5	-18,257.0	-15,194.2
(2,2)	9,790.4	-18,422.8	-15,210.2
(3,2)	9,840.8	-18,469.6	-15,107.2
Inverse distance-based weights			
(1,1)	9,628.3	-18,182.6	-15,203.0
(2,1)	9,726.1	-18,342.1	-15,262.7
(3,1)	9,770.1	-18,394.2	-15,214.9

(1) All specifications include a full set of regional dummies and separate time dummies for core and non core regions. Different error variances are allowed for each region.

TABLE 5. Baseline SpVAR(2,1) model: estimation results (*p-values in brackets*) (1)

Explanatory variables	Lag orders: Time	Lag orders: Space	Patents equation	Business R&D equation	Non business R&D equation
Non business R&D	0	0	0.013 (0.316)	-0.028 (0.014)**	–
	0	1	-0.008 (0.734)	0.020 (0.235)	0.142 (0.000)***
	1	0	-0.012 (0.409)	-0.003 (0.853)	0.609 (0.000)***
	1	1	-0.002 (0.927)	0.008 (0.698)	-0.050 (0.011)**
	2	0	0.000 (0.983)	0.017 (0.142)	0.127 (0.000)***
	2	1	0.027 (0.226)	-0.045 (0.008)***	-0.054 (0.002)***
Business R&D	0	0	0.018 (0.388)	–	–
	0	1	0.082 (0.017)**	0.067 (0.003)***	–
	1	0	0.026 (0.282)	0.618 (0.000)***	-0.019 (0.020)**
	1	1	-0.053 (0.134)	-0.038 (0.132)	0.047 (0.001)***
	2	0	-0.013 (0.532)	0.191 (0.000)***	0.023 (0.004)***
	2	1	0.029 (0.363)	0.022 (0.318)	-0.022 (0.099)*
Patents	0	1	0.101 (0.000)***	–	–
	1	0	0.386 (0.000)***	0.034 (0.007)***	0.009 (0.165)
	1	1	0.075 (0.010)***	0.035 (0.045)**	0.012 (0.244)
	2	0	0.185 (0.000)***	0.020 (0.113)	0.016 (0.013)**
	2	1	-0.002 (0.843)	0.002 (0.906)	-0.002 (0.843)
Observations			2.044	2.044	2.044
Pseudo R-squared			0.965	0.970	0.948

(1) All specifications include a full set of regional dummies and separate time dummies for core and non core regions. Different error variances are allowed for each region. ***, ** and * denote significance at the 1, 5 and 10 percent level, respectively.

TABLE 6. Space-Time Auto and Cross-correlations for the SpVAR(2,1) model residuals.

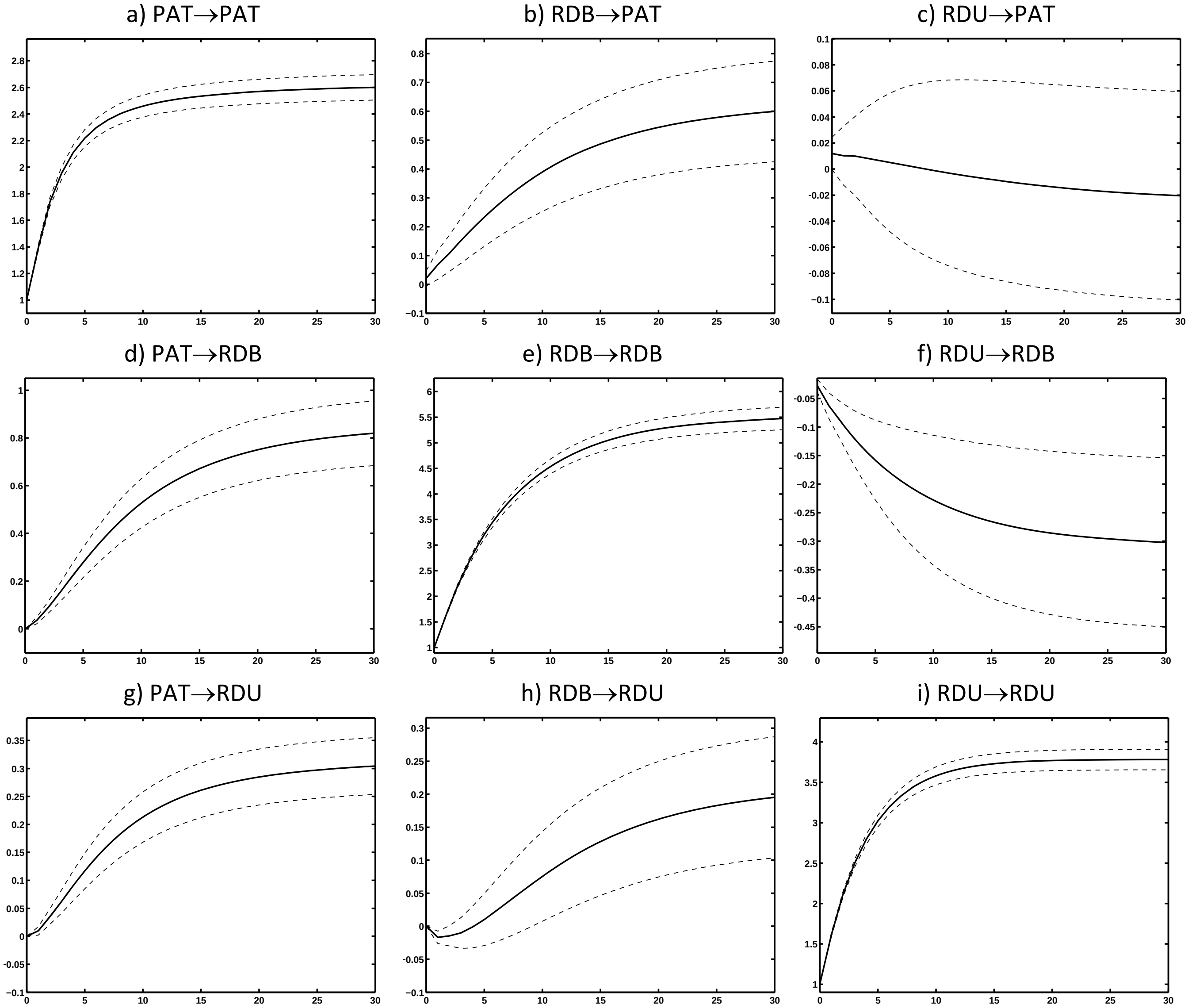
		Patents						Business R&D						Non business R&D					
	Lag: Time	Lag: Space						Lag: Space						Lag: Space					
		0	1	2	3	4	5	0	1	2	3	4	5	0	1	2	3	4	5
Patents	0	1.00	0.08	0.10	0.05	-0.03	-0.06	0.02	0.03	-0.01	0.04	0.00	0.02	0.03	0.02	-0.01	-0.02	0.00	0.04
	1	-0.03	0.03	0.00	-0.01	-0.01	-0.02	0.01	-0.02	0.02	0.02	-0.05	-0.02	0.01	-0.02	0.00	0.00	0.00	-0.03
	2	-0.02	0.02	-0.06	-0.01	0.02	0.00	-0.04	0.02	0.01	0.01	-0.01	-0.04	-0.02	0.04	0.03	-0.01	-0.03	-0.02
	3	-0.05	0.00	0.03	-0.03	0.03	0.01	0.00	-0.01	-0.02	0.01	0.02	0.01	0.04	-0.01	0.00	0.00	0.01	0.01
	4	-0.04	0.00	-0.01	0.00	0.00	0.00	0.00	-0.01	0.02	-0.01	-0.02	-0.01	0.00	0.01	0.01	0.01	-0.02	0.00
	5	0.00	0.00	0.02	0.03	0.00	0.00	-0.01	0.01	0.00	-0.02	-0.01	-0.01	0.02	-0.01	0.00	0.02	0.01	-0.02
Bus. R&D	0	0.02	0.05	-0.02	0.03	-0.03	0.02	1.00	0.04	0.00	0.11	-0.04	0.04	0.02	0.00	0.02	0.05	0.01	-0.02
	1	-0.02	-0.01	-0.01	-0.01	0.00	0.01	-0.04	0.01	-0.03	0.01	-0.01	0.02	0.01	0.02	0.01	-0.01	0.03	0.01
	2	0.01	0.01	-0.01	-0.03	-0.02	0.02	-0.05	0.00	0.00	0.01	-0.01	-0.01	0.00	0.03	-0.02	-0.03	0.05	0.02
	3	0.03	-0.03	-0.03	0.01	0.00	0.01	0.02	0.00	0.00	0.01	-0.04	0.03	-0.03	0.01	0.00	0.01	0.01	-0.05
	4	0.00	0.02	-0.01	0.01	0.00	0.01	-0.02	0.00	0.02	0.00	0.01	0.03	0.03	-0.05	-0.02	0.02	0.02	0.00
	5	0.00	0.00	0.01	0.00	-0.01	0.02	-0.02	0.01	-0.03	0.00	0.01	-0.03	-0.02	0.00	0.01	-0.01	0.00	0.00
Non bus. R&D	0	0.03	-0.02	-0.03	-0.03	-0.02	-0.01	0.02	0.01	0.02	0.02	0.00	-0.01	1.00	0.02	0.02	0.00	0.05	-0.05
	1	0.01	0.02	0.00	-0.06	0.03	0.01	0.00	0.03	0.02	0.03	0.01	0.03	-0.04	0.03	0.07	0.00	-0.02	0.02
	2	-0.01	-0.04	0.00	-0.02	0.01	0.02	-0.03	0.03	-0.01	0.00	0.00	0.00	-0.06	0.01	-0.01	-0.03	0.01	0.01
	3	0.01	0.00	-0.02	-0.02	0.02	-0.01	0.01	0.02	-0.01	-0.01	-0.01	-0.01	0.03	0.03	0.03	0.00	-0.01	0.01
	4	-0.02	0.01	0.00	-0.01	0.04	0.00	0.04	0.00	0.01	-0.02	0.00	0.00	-0.03	0.01	-0.01	0.03	0.03	-0.01
	5	-0.05	0.03	0.01	0.02	0.02	-0.02	0.01	0.02	0.00	0.01	0.02	0.00	-0.05	0.01	-0.01	-0.03	0.00	-0.02

TABLE 7. Extended SpVAR(2,1) model: estimation results (*p-values in brackets*) (1)

Explanatory variables	Lag orders: Time	Lag orders: Space	Patents equation	Business R&D equation	Non business R&D equation
				Core regions	
Non business R&D	0	0	0.005 (0.729)	-0.030 (0.018)**	–
	0	1	-0.015 (0.557)	0.018 (0.387)	0.229 (0.000)***
	1	0	-0.004 (0.821)	0.003 (0.829)	0.545 (0.000)***
	1	1	-0.003 (0.923)	0.008 (0.734)	-0.056 (0.039)**
	2	0	0.001 (0.969)	0.011 (0.373)	0.190 (0.000)***
	2	1	0.019 (0.462)	-0.052 (0.014)**	-0.112 (0.000)***
Business R&D	0	0	-0.014 (0.592)	–	–
	0	1	0.089 (0.196)	-0.009 (0.860)	–
	1	0	0.032 (0.290)	0.600 (0.000)***	-0.036 (0.001)***
	1	1	-0.145 (0.031)**	-0.019 (0.722)	-0.034 (0.543)
	2	0	0.009 (0.716)	0.207 (0.000)***	0.037 (0.000)***
	2	1	0.025 (0.707)	0.118 (0.028)**	-0.050 (0.346)
Patents	0	1	0.109 (0.016)**	–	–
	1	0	0.389 (0.000)***	0.023 (0.247)	0.003 (0.765)
	1	1	0.090 (0.121)	0.085 (0.048)**	0.090 (0.009)***
	2	0	0.145 (0.000)***	0.009 (0.649)	0.025 (0.004)***
	2	1	0.082 (0.121)	-0.032 (0.446)	-0.056 (0.112)
				Non core regions	
Non business R&D	0	0	0.071 (0.048)**	-0.004 (0.876)	–
	0	1	0.077 (0.120)	0.045 (0.169)	0.081 (0.000)***
	1	0	-0.082 (0.053)*	-0.039 (0.247)	0.669 (0.000)***
	1	1	-0.064 (0.299)	-0.01 (0.811)	-0.026 (0.323)
	2	0	0.001 (0.974)	0.056 (0.040)**	0.051 (0.083)*
	2	1	0.059 (0.224)	-0.012 (0.692)	-0.036 (0.121)
Business R&D	0	0	0.072 (0.043)**	–	–
	0	1	0.086 (0.030)**	0.078 (0.003)***	–
	1	0	0.008 (0.840)	0.642 (0.000)***	-0.003 (0.797)
	1	1	-0.029 (0.469)	-0.043 (0.128)	0.044 (0.002)***
	2	0	-0.056 (0.110)	0.175 (0.000)***	0.016 (0.127)
	2	1	0.025 (0.491)	-0.001 (0.965)	-0.010 (0.451)
Patents	0	1	0.105 (0.000)***	–	–
	1	0	0.385 (0.000)***	0.039 (0.017)**	0.016 (0.060)*
	1	1	0.066 (0.048)**	0.020 (0.314)	0.010 (0.348)
	2	0	0.228 (0.000)***	0.022 (0.175)	0.015 (0.073)*
	2	1	-0.002 (0.955)	0.001 (0.974)	-0.004 (0.681)
Observations			2.044	2.044	2.044
Pseudo R-squared			0.964	0.970	0.949

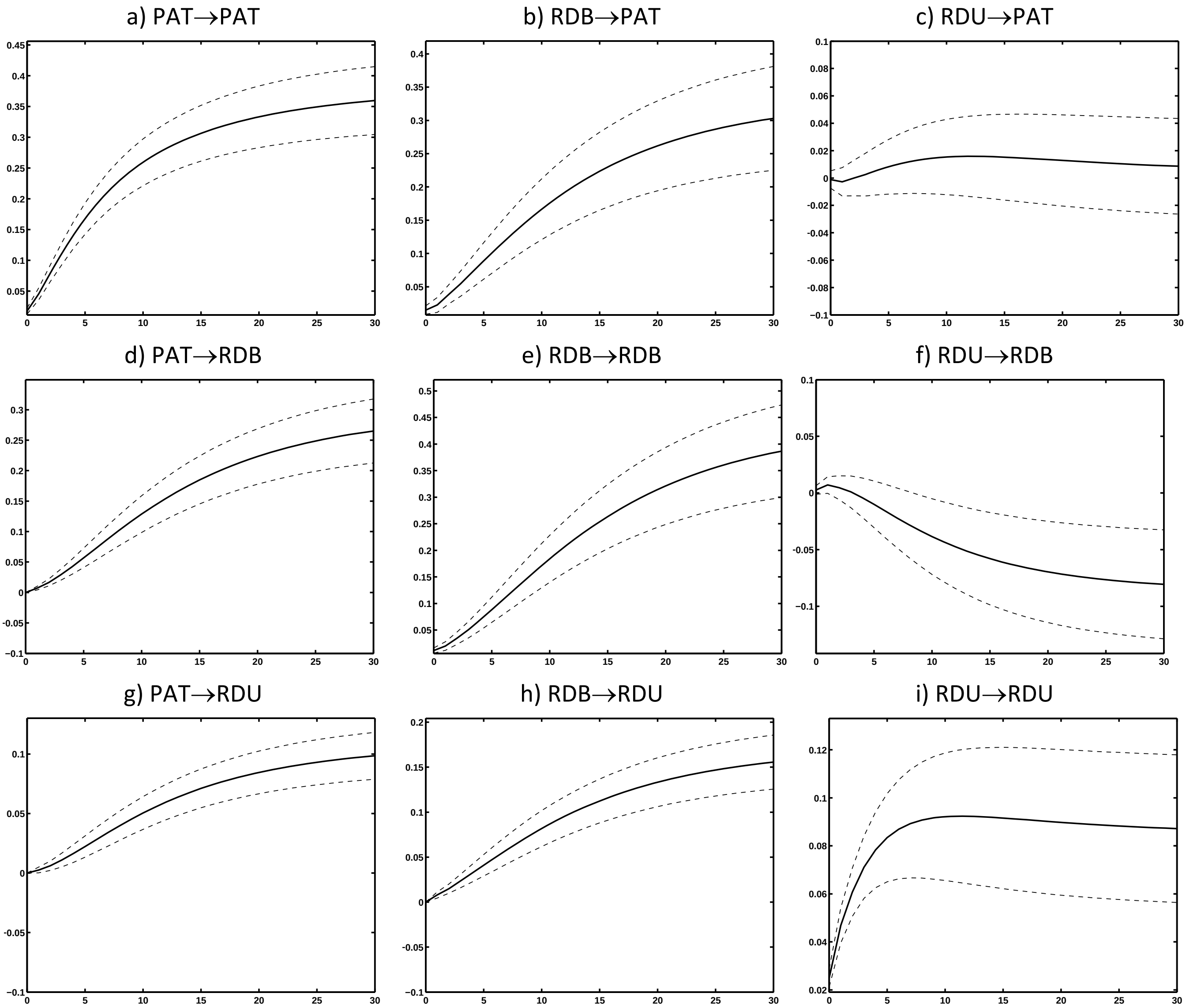
(1) All specifications include a full set of regional dummies and separate time dummies for core and non core regions. Different error variances are allowed for each region. ***, ** and * denote significance at the 1, 5 and 10 percent level, respectively.

FIGURE 4. Space-time impulse response functions from the baseline model: within region effects (1)



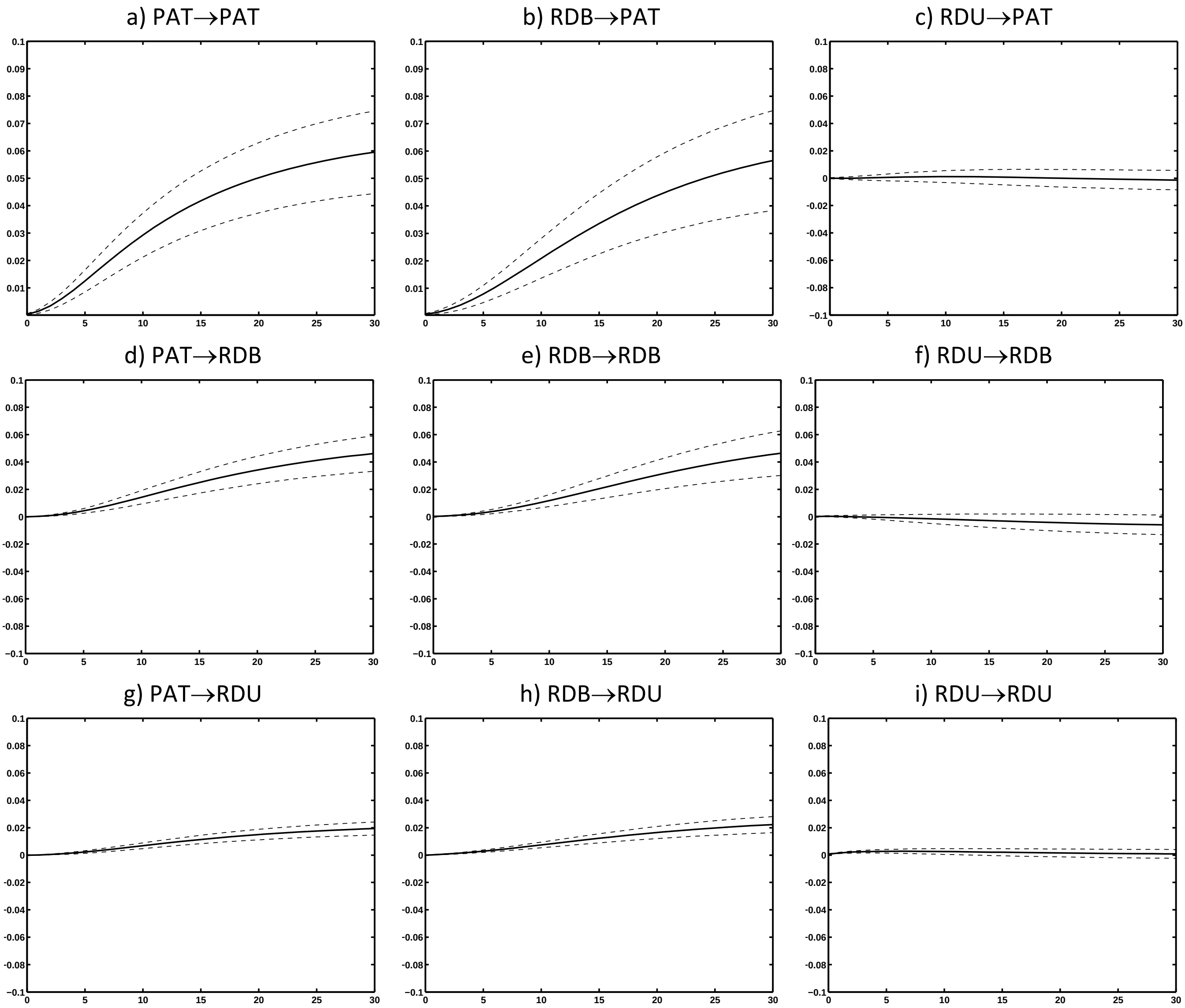
(1) Simple averages of the responses estimated for the individual regions. Dashed lines represent ± 1 standard error bands. Boostrap estimates of the standard errors are considered, based on 100 replications of the sample data.

FIGURE 5. Space-time impulse response functions from the baseline model: spatial spillovers on 1st order neighbors (1)



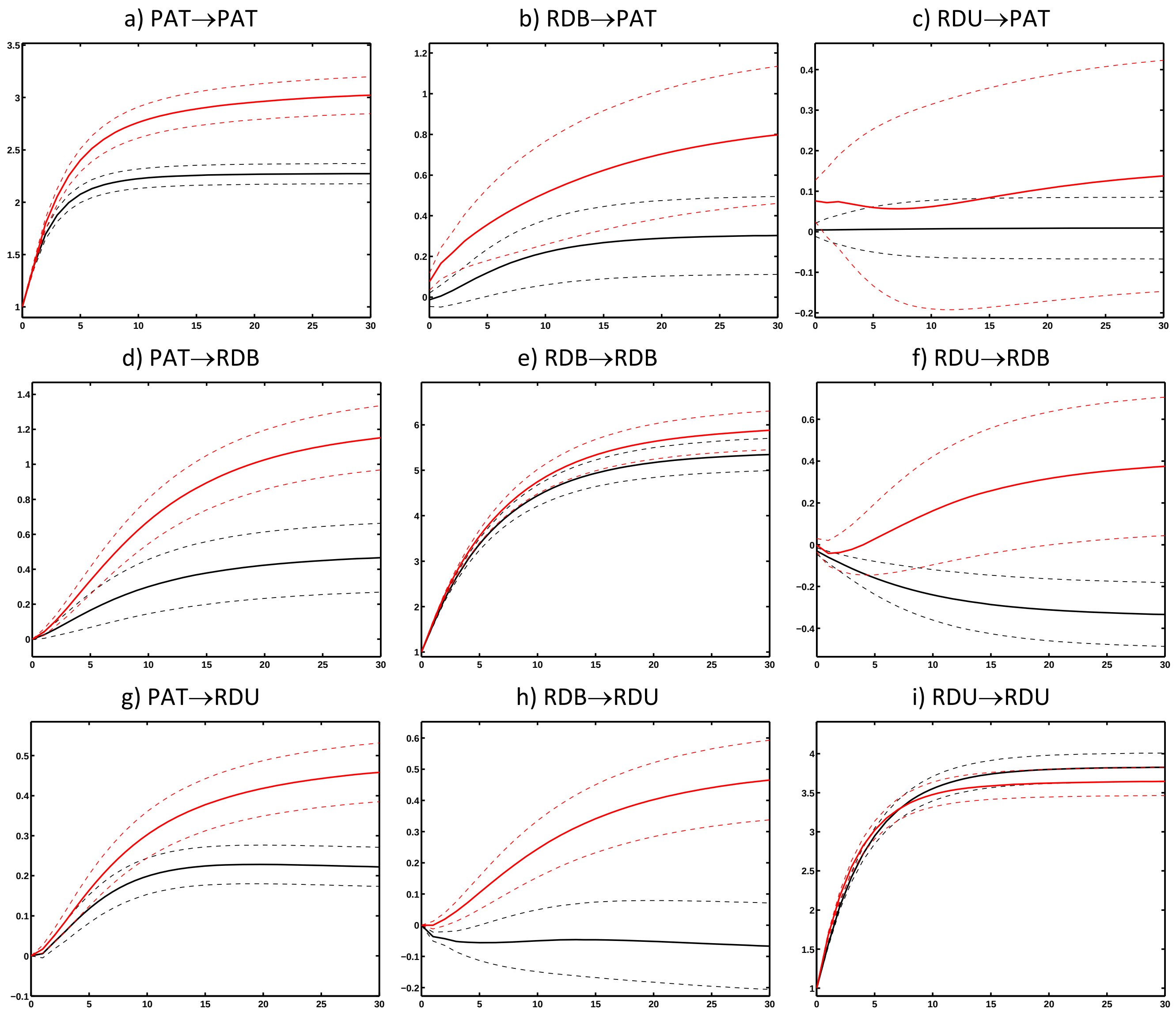
(1) See notes to Figure 4.

FIGURE 6. Space-time impulse response functions from the baseline model: spatial spillovers on 2nd order neighbors (1)



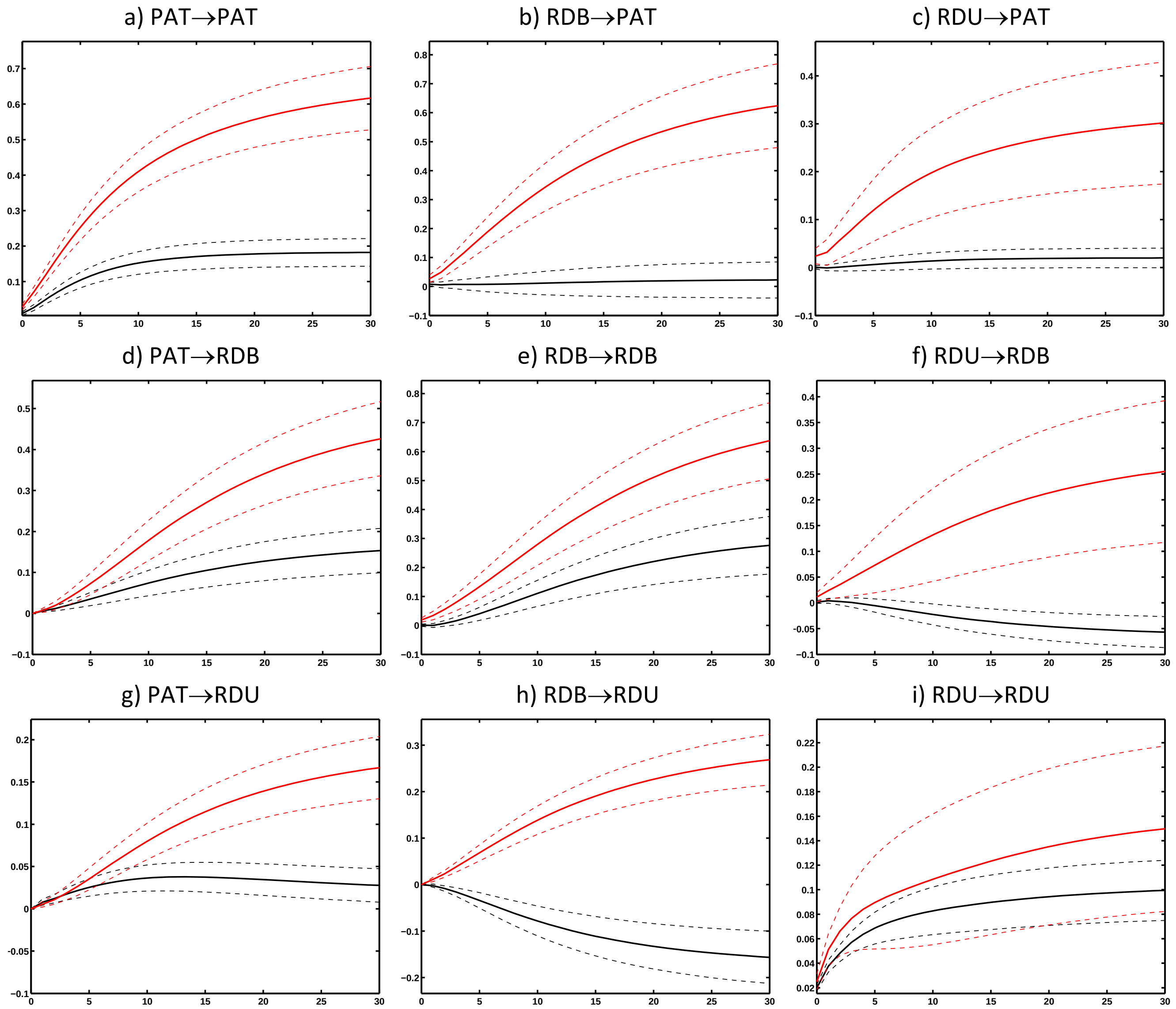
(1) See notes to Figure 4.

FIGURE 7. Space-time impulse response functions from the extended model: within region effects (1)



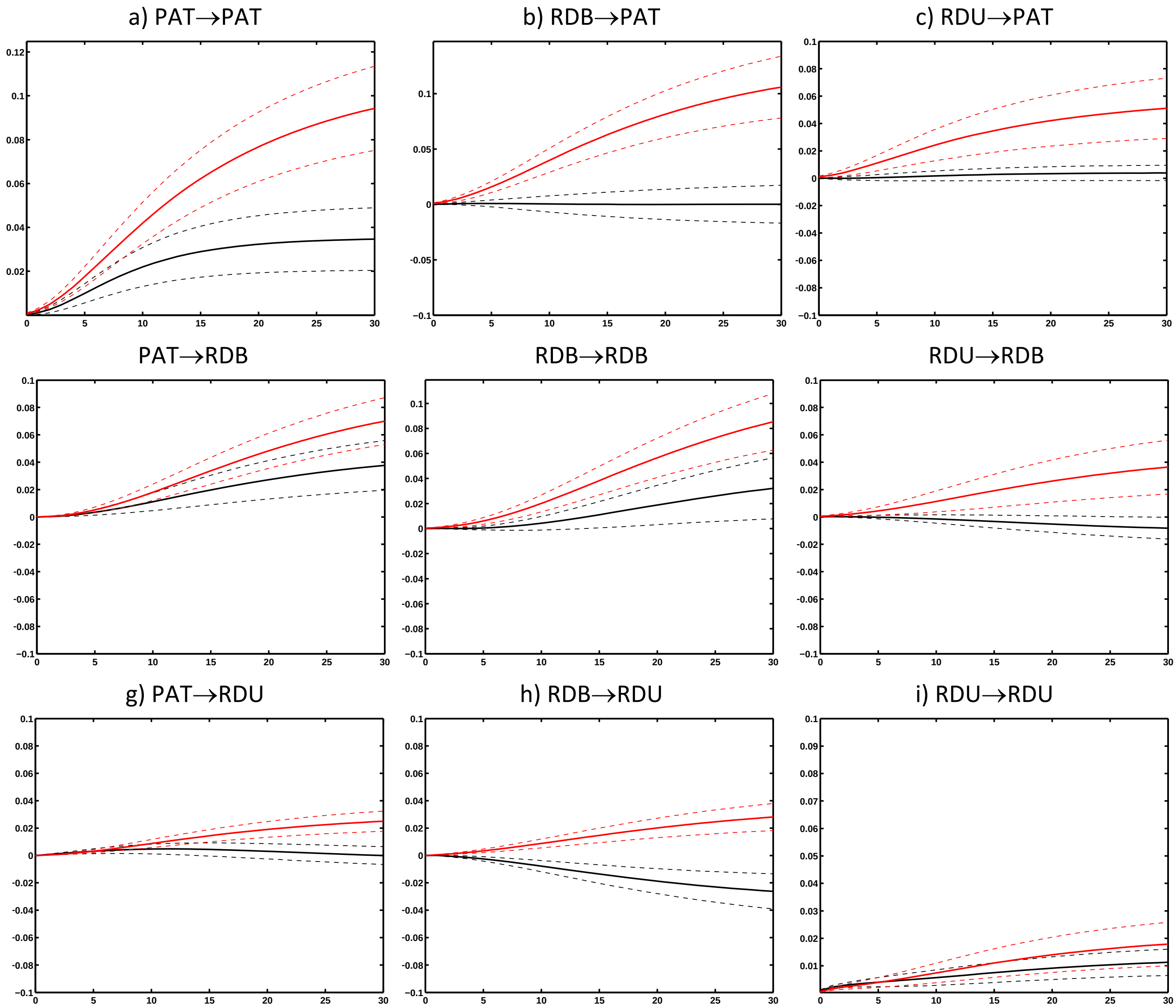
(1) Simple averages of the responses estimated for the regions belonging to the core (black color) and non core (red color) areas. Dashed lines represent ± 1 standard error bands. Boostrap estimates of the standard errors are considered, based on 100 replications of the sample data.

FIGURE 8. Space-time impulse response functions from the baseline model: spatial spillovers on 1st order neighbors (1)



(1) See notes to Figure 7.

FIGURE 9. Space-time impulse response functions from the baseline model: spatial spillovers on 2nd order neighbors (1)



(1) See notes to Figure 7.