GRES

Groupement de Recherches Economiques et Sociales
http://www.gres-so.org

Cahiers du GRES

An Exploration of Local R&D Spillovers in France

Jacques MAIRESSE
CREST-ENSAE, UNU-MERIT & NBER
mairesse@ensae.fr

Benoît MULKAY
LEREPS – GRES & Université de Montpellier 1
benoit.mulkay@orange.fr

Cahier n° 2008 - 15

Septembre 2008

UNE EXPLORATION DES SPILLOVERS GEOGRAPHIQUES DE R&D EN FRANCE

Résumé

Cet article est un essai d'évaluation de l'existence et de l'importance des « spillovers » géographiques de R&D en France. Nous considérons un modèle de fonction de production étendue (Cobb-Douglas et Translog) avec deux stocks de capital de R&D, mesurés respectivement pour chaque bassin d'emploi et pour les bassins d'emploi voisins. Nous estimons ce modèle sur 312 bassins d'emploi pour 1999 ; d'abord pour l'ensemble de l'économie marchande non agricole et ensuite pour cinq grandes branches industrielles. Les estimations des élasticités de la productivité au capital de R&D sont significatives et plausibles à la fois intra-zone et entre zones voisines, ainsi qu'intra-industrie mais non entre industries différentes.

Mots-clés : Productivité, R&D, Spillovers Locaux de R&D, Econométrie Spatiale

AN EXPLORATION OF LOCAL R&D SPILLOVERS IN FRANCE

Abstract

This paper is an attempt to assess the existence and magnitude of local research spillovers in France. We rely on the model of an extended production function (Cobb-Douglas and Translog) with both local and neighborhood R&D capital stocks. We estimate this model on 312 employment areas as of 1999, first for the whole economy, then separately for five large manufacturing industries. We find estimates of R&D capital elasticities with respect to productivity which are significant and plausible both within own-area and across neighboring areas, as well as within own-industry but not across different industries.

Keywords: Productivity, R&D, Local R&D Spillovers, Spatial Econometrics

JEL : O30 ; O32 ; O47 ; C21

1 Introduction

Assessing the local spillover impacts of firms' R&D investments on the various dimensions of economic development: productivity, employment, innovation, …, both in the geographic area where they are located and in neighboring areas, is one the most difficult and important challenge of recent empirical investigations in the economics of research and innovation[1]. Since the seminal book of KRUGMAN [1991] and the renewal of economic geography, these issues and the related ones of understanding the determinants and consequences of the localization and agglomeration of firms' activities have received increasing attention. Firms tend to locate where the factors of production are abundant and less expensive, or where the demand for their products is strongest. They have, however, to balance production costs and costs of transportation. Many authors recognize that various types of externalities play also a major role in the localization of firms, arising from particular historical and geographical contexts, from policies of regional planning, from the agglomeration of natural, human and other economic resources, and in particular from that of specific knowledge assets leading to local increasing returns.

As emphasized by GRILICHES [1992], the search for knowledge spillovers is specially challenging. While other externalities can be assessed more or less directly, even if not easily, knowledge spillovers are not directly observed. Economists can only try to measure the effects of knowledge flows and stocks variables on outcome variables like numbers of innovations or patents, and labor or total factor productivity. A related problem is to evaluate the spatial extent of such spillovers. Other major problems are in trying to understand and to analyze the underlying channels and "mechanisms" by which they operate, and in particular to characterize the conditions allowing firms to benefit from them[2].

In this exploratory econometric analysis, we basically try to identify local knowledge spillovers by estimating the effects of firms' R&D investments on productivity at the aggregate level of some 300 French employment areas for 1999. We do so by relying on the framework of an extended production with both local and neighborhood R&D capital stocks, in addition to the more traditional factors of production of labor and physical capital[3]. We specify this production function both as a simple Cobb-Douglas function and a more general Translog function, and we estimate it for the French economy as a whole as a spatial autoregressive regression, as well as for five large manufacturing industries. On the basis of this framework and our data, we can distinguish between local R&D spillovers within the range of employment areas themselves and within the range of neighboring areas. We thus focus on estimating as our two main parameters of interest the elasticities of productivity with respect to R&D capital, both "within own-area" and "across neighboring areas": first for the whole economy in Section 3, and then separately by manufacturing industry in Section 4. In this last Section, we also try to distinguish between local R&D spillovers "within own-industry" and "across other industries".

[1] See AUDRETSCH – FELDMAN [2004] for a survey, and AUTANT-BERNARD – MAIRESSE – MASSARD [2007] for a summary account of recent empirical studies published in a special issue of Papers in Regional Science on "Spatial Knowledge Diffusion through Collaborative Networks".

[2] See for example COHEN – LEVINTHAL [1989], COCKBURN – HENDERSON [1998] or AGRAWAL [2002].

[3] For a presentation of the extended production framework, and an in-depth discussion of its relevance and usefulness as well as many of the conceptual, measurement and econometric issues it raises, see the seminal article of GRILICHES [1979].

Although our results remain exploratory, they are surprisingly encouraging, leading to estimates of R-D capital elasticities both within own-area and across neighboring areas which are statistically significant and seem economically plausible. Local spillovers thus extend largely beyond the average range of employment areas, but they also appear to be limited to neighboring employment areas that does not reach farther than an average 100 km. We also find evidence that local spillovers tend to be mostly industry specific, with significant estimates for R&D capital elasticities within own-industry in all five manufacturing industries, and significant ones for elasticities across other industries for two industries out of the five: consumption goods and equipment goods industries.

Before turning in Sections 3 and 4 to the detailed presentation of our results, we have in Section 2 to explain briefly the construction of the data at the level of the French employment areas in 1999, and comment on some the descriptive statistics for our main variables, stressing in particular the extreme geographical concentration of R&D firms' investments.

Although our results remain exploratory, they are surprisingly encouraging, leading to estimates of R-D capital elasticities both within own-area and across neighboring areas which are statistically significant and seem economically plausible. Local spillovers thus extend largely beyond the average range of employment areas, but they also appear to be limited to neighboring employment areas that does not reach farther than an average 100 km. We also find evidence that local spillovers tend to be mostly industry specific, with significant estimates for R&D capital elasticities within own-industry in all five manufacturing industries, and significant ones for elasticities across other industries for two industries out of the five: consumption goods and equipment goods industries.

Before turning in Sections 3 and 4 to the detailed presentation of our results, we have in Section 2 to explain briefly the construction of the data at the level of the French employment areas in 1999, and comment on some the descriptive statistics for our main variables, stressing in particular the extreme geographical concentration of R&D firms' investments.

2 Data and main descriptive statistics

2.1. Construction of the necessary data at the level of French employment areas for 1999

Many previous studies in order to assess the importance of geographical knowledge spillovers have been relying on regional or departmental data[4]. We investigate this issue here for France at more detailed geographical level which is *a priori* preferable, that of the "employment areas ("bassins d'emploi").

The data we use relate to the non-agricultural business sector excluding financial activities and interim employment, for "Metropolitan" France without Corsica[5]. They are

[4] See for example CICCONE [2002], GAMBARDELLA – MARIANI – TORRISI [2002], or AUTANT-BERNARD – LESAGE [2008].

[5] Financial activities and interim employment are excluded for lack of good coverage in the administrative data we use. Corsica is left out because of geographical distance and its insular situation (and very little R&D investments).

constructed at the level of "employment areas" for the year 1999. Employment areas are economic zones where local firms are likely to hire their workers. They have been precisely defined by INSEE and the Ministry of Labor, first in 1983 and then revised in 1994, on the basis of residence-to-work displacements[6]. Employment areas are much smaller than regions and departments (which correspond respectively to the NUTS 2 and NUTS 3 levels of the European Union classification). There are 341 of them in Metropolitan France (without Corsica), of which we retain only 312 in our analysis, after discarding 29 as unsuitable because they had no or very small R&D investments or very low employment levels (with an estimated R&D capital stock of less than 100 K€, or with less than 5 000 workers)[7].

Our R&D data come from the annual surveys conducted by the Ministry of Research, which give detailed information on firms' internal and external R&D expenditures, numbers of R&D employees, financial sources, … These individual data are allocated to local municipalities on the basis of the postal code (ZIP code) of the firms' main laboratories (there are approximately 36 000 municipalities in France), and then aggregated at the level of the employment areas. Finally, using here only the internal R&D expenditures obtained for the six years 1993 – 1998, and applying the so called permanent inventory method with a 15% depreciation rate, we can construct an R&D capital stock K at the beginning of 1999 for all employment areas. (See Appendix A for more details.) In order to investigate the spatial range of local spillovers beyond the employment areas, besides measuring the local R&D capital stocks K, we have also computed so called "neighborhood R&D capital stocks" such as $K100$ or $K200$. For a given employment area, these are simply computed as the sums of the R&D capital stocks K of all their neighboring employment areas in a "circle" of 100 km or 200 km. (See also Appendix A for more details.)

The employment data come from the firms' declarations to the Social Security (i.e. the *Déclarations Annuelles de Données Sociales* or DADS). Being separately available for the different establishments of firms, they can be merged into an INSEE database which is constructed at the establishment level and provides other economic key variables for 1999: total sales, value added, gross earning before interests and taxes, and the book value of fixed assets[8]. Establishments being localized at the municipality level, these variables are aggregated at the level of the employment areas as in the case of R&D. The General Census of Population of 1999 is also a source of complementary information at the level of municipalities and employment areas.[9]

2.2. Main descriptive statistics

Table 1 gives the mean, standard deviation, minimum, median and maximum, as computed over the 312 employment areas, for the main variables in our investigation. It shows the very large dispersion and skewness (asymmetry) of most of these variables in absolute levels (that is before being normalized by size and being taken in logs). While the surface (S) of the largest employment area (Toulouse) is already 140 times that of the smallest one (Vitry-sur-Seine) and the mean surface (1 601 km^2) is about 10% higher than the median

[6] See INSEE [1994].
[7] 10 employment areas are excluded on the basis of these two criteria, 13 only on that of very small R&D investment, and 6 of very small employment.
[8] In fact this establishment database is constructed on the basis of firm level statistics. For mono-establishment firm, this evidently raises no difficulties, but for multi-establishments firms, various imputations methods have been used on the basis of very detailed industry ratios by establishment size and localization.
[9] See JULIA [2003] for more detailed explanations on these aspects of the construction of our database.

surface (1 432 km^2), the employment (L) of the largest employment area (Paris) is of about 190 times that of the smallest one (Gannat), and the mean employment (40 158) is about twice the median employment (20 512). These two max-to-min and mean-to-median ratios are even larger for value-added (Y) and physical capital (C) than for employment, and even much more so for local R&D capital (K) and our prefered measure of neighborhood R&D capital ($K100$). As could be expected, however, when we normalize by employment size and consider labor productivity(Y/L), physical capital (C/L) intensity, and local and neighborhood R&D capital intensities (K/L and $K100/L$), we see that their distribution across employment areas appear much less dispersed and skewed. Going one step further and taking logarithms which is what do when estimating the Cobb-Douglas and Translog production functions regressions, we can also see that their log-distributions become roughly symmetrical.

			Std. Dev.	**Min**	**Median**	**Max**
Surface	*in km²*	1 601	1 015	45	1 432	6 264
Employment (L)	*workers*	40 158	76 202	5 034	20 512	992 637
Value Added (Y)	*in K€*	2 045 815	4 861 684	187 652	915 745	61 077 052
Fixed Capital (C)	*in K€*	3 092 345	8 300 117	224 538	1 314 853	116 760 038
R&D Capital (K)	*in K€*	42 151	216 005	115	2 956	2 598 767
Neighborhood R&D Capital (K100)	*in K€*	1 208 750	145 507	115	2 444	8 815 790
R&D Workers (LRD)	*workers*	220	1 096	0	20	14 086
Y / L	*in K€*	44.842	9.101	33.015	42.391	115.461
C / L	*in K€*	69.546	31.040	35.974	63.558	269.548
K / L	*in K€*	0.467	1.394	0.006	0.148	14.819
K100 / L	*in K€*	56.699	8.717	0.038	10.027	1 051.815
log(Y / L)		3.7874	0.0095	3.4970	3.7469	4.7489
log(C / L)		4.1737	0.0196	3.5828	4.1520	5.5967
log(K / L)		-1.8693	0.0747	-5.1442	-1.9074	2.6959
log(K100 / L)		2.3685	0.0992	-3.2749	2.3053	6.9583
K / Y		0.85%	2.15%	0.02%	0.35%	29.34%
K / C		0.7%	2.1%	0.0%	0.2%	27.5%
LRD / L		0.2%	0.6%	0.0%	0.1%	5.7%

312 Employment Areas.

Table 1 : Main Descriptive Statistics

The distributions of the local and neighborhood log R&D capital stocks per employee : Log(K/L) and Log($K100/L$) remain nonetheless very dispersed across the employment areas, as compared to log labor productivity : Log(Y/L) and to log physical capital stock per employee: Log(K/L). This corresponds to an extremely high concentration of firms R&D activities in a few zones. This geographical concentration of R&D activities is particularly striking since it much more pronounced than that of productive activities. This appears most clearly by looking at the Lorenz curves shown in Figure 1 respectively for the surface (S), total employment (L), value added (Y) and local and neighborhood R&D capital

(*K*) and (*K*100), and by comparing the corresponding Gini coefficients[10]. We can see that the 10 % (i.e. the 31) largest employment areas in terms of surface, employment, value-added and physical capital respectively correspond to 23 %, 47 %, 53 % and again 53 % of the total surface, total employment, total value-added and physical capital respectively, but that the 10 % largest ones in terms of local R&D capital account for as much as 88 % of the total R&D capital stock, while the 10 % largest ones in terms of neighborhood R&D capital account for as much as 71 % of the "total neighborhood R&D capital stock"[11].

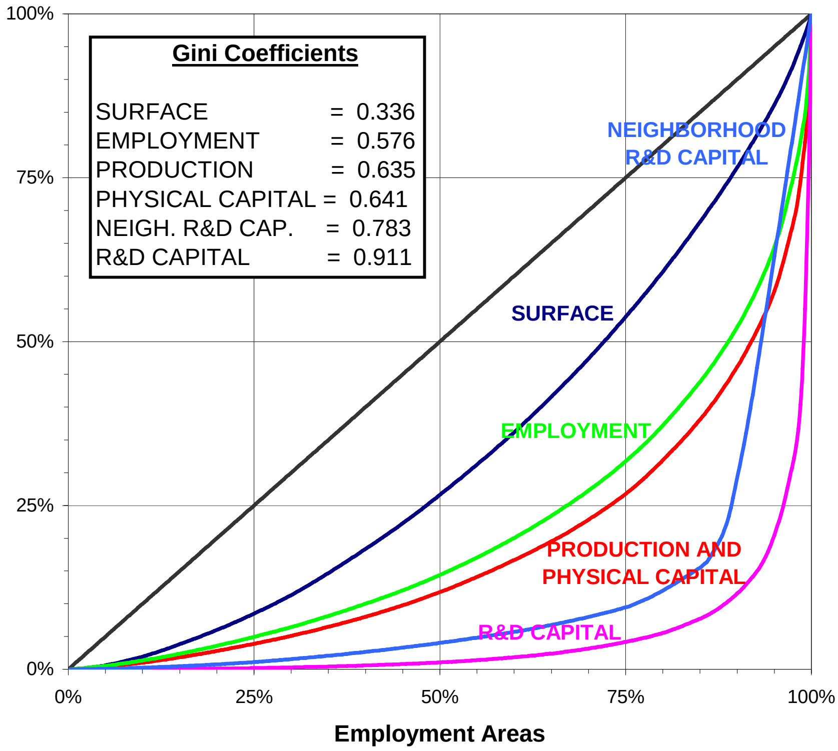


Figure 1 : Concentration Curves and Gini Coefficients

Figure 2 shows the localization and importance of R&D activities in the 312 employment zones in France. These activities are mainly concentrated in the Paris region, and to a lesser extent in the Rhône-Alpes region with Lyon and Grenoble, and the Toulouse region, and they are quite modest or negligible in most other parts of France. In fact, as we just already indicated, about 90 % of the R&D employees (researchers and technicians) employed by firms are located in the 40 employment areas largest in terms of R&D employment, and 70 % in the 10 first of them: 7 in Paris Region and 3 in the province: by decreasing order: Nanterre, Versailles, Boulogne-Billancourt, Paris, Toulouse, Lyon, Les Mureaux, Grenoble, Saint-Denis, Vitry-sur-Seine.

[10] The Lorenz curves for physical capital (*C*) and value-added (*Y*) practically coincide and their Gini coefficients are nearly equal. Note that the Lorenz curve for the neighborhood R&D capital *K*100 appears less concentrated than that for local R&D capital *K*, because of the fact that the different neighborhood areas are by construction greatly overlapping , and the fact that local R&D capital stocks *K* are very small for most employment areas.

[11] Note that because of the high concentration of R&D capital *K* in few employment areas in Paris, Lyon, Toulouse, and Grenoble, and their neighborhood, and because of the fact that the different neighborhood areas are by construction greatly overlapping, the mean neighborhood R&D capital *K*100 appears much larger (by a factor of nearly 30!) than the mean R&D capital stock *K*.

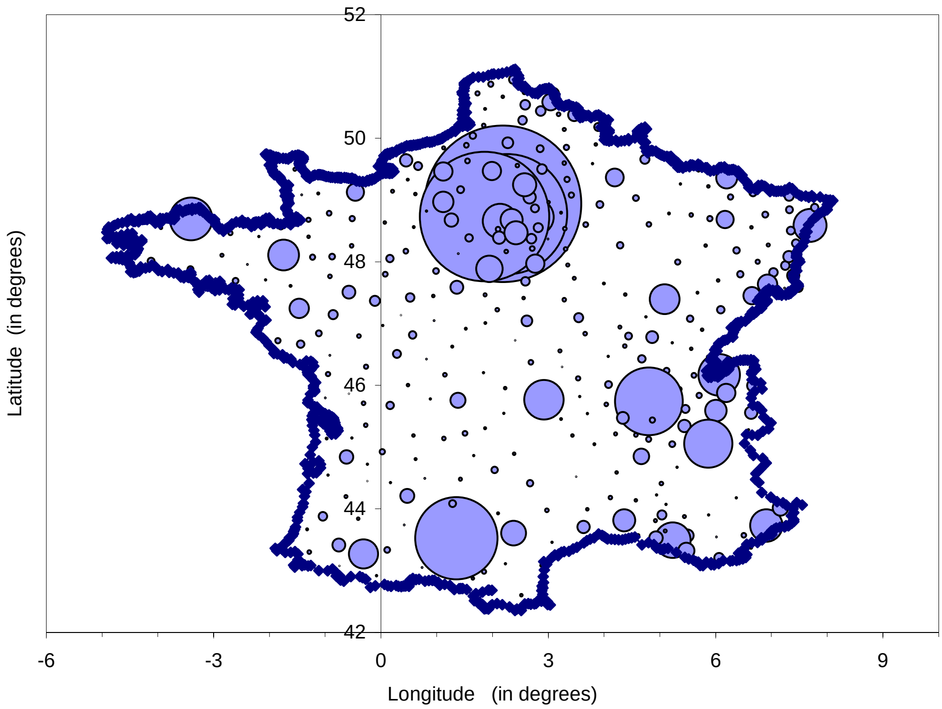


Figure 2 : Geographic Concentration of R&D Employment in France

Finally, Table 2 gives the Moran's coefficients of spatial autocorrelation for our main variables (in logs) using four different contiguity matrices[12]. We can see that the spatial autocorrelation coefficients are statistically significant (at the 1 % confidence interval) for all variables, and whatever the contiguity matrix used in computation. They also tend to be somewhat higher when more weight is given to close proximity, that is when they are computed with the first contiguity matrix (W1) based on the neighboring areas, or the fourth one (W4) based on the inverse of the squared distance. We note also that they are generally close enough for all the variables, in the range of 0.15 to 0.25, with few exceptions. This is a relatively modest order of magnitude, which is high enough, however, to warrant the use of spatial econometric techniques.

[12] See MORAN [1950] or CLIFF – ORD [1980].

Contiguity matrix	W1	W2	W3	W4
Log(Y)	0.228	0.086	0.141	0.189
	(6.78)	*(4.24)*	*(6.36)*	*(6.87)*
Log(L)	0.198	0.056	0.108	0.156
	(5.88)	*(2.83)*	*(4.91)*	*(5.68)*
Log(C)	0.187	0.068	0.115	0.159
	(5.57)	*(3.37)*	*(5.22)*	*(5.80)*
Log(K)	0.211	0.125	0.172	0.211
	(6.27)	*(6.07)*	*(7.70)*	*(7.64)*
Log(K100)	0.764	0.695	0.728	0.759
	(22.48)	*(33.22)*	*(32.16)*	*(27.20)*
Log(Y/L)	0.278	0.211	0.249	0.282
	(8.23)	*(10.18)*	*(11.11)*	*(10.17)*
Log(C/L)	0.097	0.107	0.118	0.134
	(2.95)	*(5.22)*	*(5.32)*	*(4.91)*
Log(K/L)	0.168	0.137	0.168	0.195
	(5.01)	*(6.67)*	*(7.51)*	*(7.07)*
Log(K100/L)	0.474	0.393	0.417	0.448
	(13.99)	*(18.86)*	*(18.48)*	*(16.11)*

Moran'I statistics, z-test of no spatial autocorrelation in italic (distribted as standard normal under the null of no spatial autocorrelation).

The expected value of Moran'I statistics is always -1/N = -0.003 (N : the number of observations).

The standard error of Moran'I statistics depends on the contiguity matrix. They are respectively 0.034 , 0.021 , 0.023 and 0.028 .

W1 = Contiguity Matrix based on Immediatly Neighboring Employment Areas

W2 = Contiguity Matrix based on Neighboring Employment Areas in a Circle of 100 km

W3 = Continguity Matrix based on the inverse of geographical distance

W4 = Contiguity Matrix based on the inverse of the square of geographical distance

Table 2 : Spatial Autocorrelation Coefficients and Tests

3 Local R&D Spillovers

In order to assess the existence and magnitude of local and neighborhood R&D capital intensities on local productivity, we estimate the following extended simple Cobb-Douglas production function (1):

$$\log\left(\frac{Y_i}{L_i}\right) = \alpha + \nu_1 \log(L_i) + \beta_1 \log\left(\frac{C_i}{L_i}\right) + \gamma_1 \log\left(\frac{K_i}{L_i}\right) + \eta_1 \log\left(\frac{K100_i}{L_i}\right) + \varepsilon_i \qquad (1)$$

and the more general extended Translog production function (2):

$$\log\left(\frac{Y_i}{L_i}\right) = \alpha + \nu_1 \log(L_i) + \nu_2 (\log(L_i))^2$$
$$+ \beta_1 \log\left(\frac{C_i}{L_i}\right) + \beta_2 \left(\log\left(\frac{C_i}{L_i}\right)\right)^2 + \beta_3 \log(L_i) \log\left(\frac{C_i}{L_i}\right)$$
$$+ \gamma_1 \log\left(\frac{K_i}{L_i}\right) + \gamma_2 \left(\log\left(\frac{K_i}{L_i}\right)\right)^2 + \gamma_3 \log(L_i) \log\left(\frac{K_i}{L_i}\right) + \gamma_4 \log\left(\frac{C_i}{L_i}\right) \log\left(\frac{K_i}{L_i}\right) \qquad (2)$$
$$+ \eta_1 \log\left(\frac{K100_i}{L_i}\right) + \eta_2 \left(\log\left(\frac{K100_i}{L_i}\right)\right)^2 + \varepsilon_i$$

where *i* denotes the employment area *i* (i = 1 to 312), and where our main parameters of interest are γ_1 and η_1 for the Cobb-Douglas specification, together with γ_2 and η_2 (and possibly γ_3 and γ_4) for the Translog specification. Note that all capital stocks (*C*, *K* and *K*100) measured at the beginning of the year 1999. Note also that all the squared and cross product log terms in the Translog specification are taken in deviations from the corresponding means, which implies for example that the estimated γ_1 and η_1 in the Translog specification directly measure the local and neighborhood R&D capital elasticities at the mean values of the variables, and that they should be not too different from the constant elasticities γ_1 and η_1 as estimated in the Cobb Douglas specification[13]. Note finally that in order to take into account the different industry structure of the employment areas, we have included in the two Cobb-Douglas and Translog productivity equations eleven control variables measuring the value added shares of the different industries (at the NES16 classification level) in the employment areas.

As we have seen in the previous section, our main variables Log(*Y*), log(*L*), log(*C*), log(*K*) and Log(*K*100) are not only extremely dispersed but they also exhibit spatial autocorrelation patterns, and we can thus expect that the error terms ε in the productivity Cobb-Douglas and Translog equations (1) and (2) are also spatially autocorrelated. To take account of such a spatial autocorrelation, we rely on the spatial econometrics methods as developed in ANSELIN [1988], LESAGE [2000] or LE GALLO [2002]. After various experimentations, we have focused on the Spatial Autoregressive Regression (*SAR*) estimated by maximum likelihood. The *SAR* specification performs better than the usual regression as estimated by Ordinary Least Squares (*OLS*), that is when tested against the null hypothesis of no spatial autocorrelation ($\rho = 0$). It is also performs better when tested against the Spatial Error Model (*SEM*) in the framework the Spatial General Model (*SGM*) encompassing both the *SAR* and *SEM* specifications. It also does well when tested with the Spatial Durbin Model (*SDM*). (See Appendix B for detailed explanations.)

Tables 3 and 4 give the results of the estimation by maximum likelihood of the spatial autoregressive model for the Cobb-Douglas and Translog equations respectively.

[13] Note also that we have not included in the Translog specification the three cross-terms involving Log(*K*100/*L*), that is Log(*L*)* Log(*K*100/*L*), Log(*C*/*L*) * Log(*K*100/*L*) and Log(*K*/*L*) * Log(*K*100/*L*), since these three variables are very strongly collinear.

Régression	**(1)**		**(2)**		**(3)**		**(4)**	
Constant	1.758**	*(.238)*	2.107**	*(.254)*	1.776**	*(.239)*	2.161**	*(.258)*
Log(L)	0.030**	*(.007)*	0.045**	*(.008)*	0.035**	*(.009)*	0.039**	*(.010)*
Log(C/L)	0.268**	*(.024)*	0.256**	*(.024)*	0.266**	*(.024)*	0.257**	*(.024)*
Log(K/L)	0.031**	*(.005)*	0.030**	*(.005)*	0.031**	*(.005)*	0.030**	*(.005)*
log(K100/L)			0.014**	*(.004)*			0.013**	*(.004)*
log(K200/L)					0.004	*(.004)*		
log((K200-K100)/L)							-0.005	*(.004)*
W * Log(Y/L)	0.314*	*(.139)*	0.269*	*(.137)*	0.298*	*(.140)*	0.289*	*(.137)*
s	0.0814		0.0801		0.0814		0.0798	
R^2-adjusted	0.7628		0.7700		0.7624		0.7706	
Log. Likelihood	338.09		344.63		338.55		345.55	
LM Test OLS vs. SAR	27.953	*[0.000]*	5.128	*[0.029]*	20.601	*[0.000]*	5.352	*[0.021]*
LM Test SAR vs. SGM	2.015	*[0.156]*	0.979	*[0.323]*	1.561	*[0.212]*	1.568	*[0.211]*

Maximum Likelihood Estimates with standard errors in parenthesis, 312 Observations.
All regressions include 11 industry shares (NES 16 level).
** : significant at 5% level; ** : significant at 1% level.*
LM Test OLS vs. SAR : Lagrange multiplier test of Autoregressive model vs. no spatial model (distributed as khi-squared (1) under the null)
LM Test SAR vs. SGM : Lagrange multiplier test of spatial generalized model vs. spatial autoregressive model (distributed as khi-squared (1) under the null)

Table 3: Estimates of Cobb-Douglas production function with local R&D spillovers

Making first a few general observations, we see that for all eight different regressions that we thought useful to document in these tables, the absence of spatial autocorrelation is rejected at 5% level, while the Spatial Autoregressive Regression (*SAR*) is accepted against the Spatial General Model (*SGM*). The spatial autocorrelation parameter, i.e. the coefficient of W*Log(*Y/L*), is statistically significant of the order of 0.3 to 0.4 depending on the regressions. This can be interpreted as indirect evidence of local spillovers effects, other than the direct evidence provided by the estimates of the R&D capital stocks elasticities.

We also observe that the general fit of the regressions are strongly improved when we move from the Cobb-Douglas to the Translog specifications. This is mainly accounted by the inclusion in the equations of the squared log-variables (and not by the cross-product terms), as indicated by the likelihood ratio tests. Following the interpretation of such a result proposed by CREPON – MAIRESSE [1993] in the context of firm panel data, we can view it as strong evidence of the heterogeneity of the production function across individual units: that is for us here across employment areas. The Translog equation takes explicitly into account such heterogeneity by including squares and cross-product of the log-variables, contrary to the more parsimonious Cobb-Douglas equation. Note, however, that, as could be expected, the estimates of average elasticities (i.e., when computed at the sample means of the variables for the Translog specification) are all practically the same for both type of equations.

In the Translog specification, the average elasticities are not constant. They varies according to the magnitude of the variables. For example, the average elasticity of physical capital stock is :

$$\hat{\beta}_A = \beta_1 + 2\beta_2\left(\log\left(\frac{C}{L}\right) - \overline{\log\left(\frac{C}{L}\right)}\right) + \beta_3\left(\log(L) - \overline{\log(L)}\right) + \gamma_4\left(\log\left(\frac{K}{L}\right) - \overline{\log\left(\frac{K}{L}\right)}\right)$$

Let-us note that the squared and cross-product variables are taken in deviation from their means in the Translog specification. The last two terms corresponding to the cross-products are negligible in the estimation. Therefore they are dropped in the following comments.

We find in all eight regressions estimates of the <u>average</u> elasticity of physical capital stock $\hat{\beta}_A$, which are both statistically very significant and of a reasonable order of magnitude of 0.25. The Translog estimates show, however, that the elasticity β is far from being constant across employment areas, increasing strongly with physical capital intensity : $\hat{\beta} = 0.23 + 0.28\left(\log(C/L) - \overline{\log(C/L)}\right)$. We also find small but significant increasing returns to scale ν of about 3 to 5 %, which appear to be practically constant across employment areas (contrary to β).

Turning now to our parameters of main interest: the elasticities of R&D capital, we see clearly that in all eight regressions the <u>average</u> elasticity of local R&D capital $\hat{\gamma}_A$ is as statistically significant as the average elasticity of physical capital $\hat{\beta}_A$, and about equal to 0.03. Such an order of magnitude which may seem small is in fact on the high side of what could be expected. The similar cross-sectional estimates of R&D capital elasticity performed at the firm level for samples of R&D doing firms in manufacturing industries are in the range of 0.05 to 0.10[14]. Considering that only a minority of firms do R&D, a simplistic guess would be that at the aggregate level of employment areas, the estimated elasticity of local R&D capital would be a great deal smaller. Finding that it is actually of about 0.03 is clear evidence for the existence of sizeable R&D spillovers within employment areas. The Translog estimates show that the elasticity γ, like β, is not constant across employment areas but is strongly increasing with the intensity of local R&D capital : $\hat{\gamma} = 0.03 + 0.02\left(\log(K/L) - \overline{\log(K/L)}\right)$.

[14] See for example CREPON – MAIRESSE [1993] for such cross-sectional estimates for French manufacturing industries. See also MAIRESSE – SASSENOU [1991] for a survey of both cross-sectional and time series types of estimates for other countries, which remains representative of the results that can be found in recent studies.

Regression	(5)		(6)		(7)		(8)	
Constant	1.787**	*(0.223)*	1.926**	*(0.228)*	2.171**	*(0.236)*	2.321**	*(0.240)*
Log(L)	0.035**	*(0.007)*	0.030**	*(0.007)*	0.051**	*(0.008)*	0.045**	*(0.008)*
Log(L)²	-0.002	*(0.004)*	-0.008	*(0.005)*	-0.002	*(0.004)*	-0.010*	*(0.005)*
Log(C/L)	0.244**	*(0.022)*	0.236**	*(0.023)*	0.232	*(0.022)*	0.226**	*(0.022)*
Log(C/L)²	0.137**	*(0.026)*	0.139**	*(0.025)*	0.139**	*(0.025)*	0.141**	*(0.025)*
Log(K/L)	0.030**	*(0.005)*	0.031**	*(0.005)*	0.029**	*(0.004)*	0.030**	*(0.004)*
Log(K/L)²	0.010**	*(0.002)*	0.008**	*(0.002)*	0.010**	*(0.002)*	0.008**	*(0.002)*
Log(K100/L)					0.013**	*(0.003)*	0.013**	*(0.003)*
Log(K100/L)²					0,002*	*(0.001)*	0.002*	*(0.001)*
Log(L) * Log(C/L)			-0.009	*(0.017)*			-0.003	*(0.017)*
Log(L) * Log(K/L)			0.009	*(0.005)*			0.011*	*(0.005)*
Log(C/L) * Log(K/L)			0.016	*(0.011)*			0.017	*(0.011)*
W * Log(Y/L)	0.459**	*(0.128)*	0.431**	*(0.128)*	0.417**	*(0.125)*	0.389**	*(0.124)*
s	0.0739		0.0733		0.0719		0.0710	
R² - adj.	0.8024		0.8038		0.8120		0.8149	
Log. Likelihood	368.6088		371.4119		378.4054		382.4365	
LM Test SAR vs. OLS	22.537	*[0.000]*	19.887	*[0.000]*	2.432	*[0.119]*	1.677	*[0.195]*
LM Test SGM vs. SAR	1.144	*[0.285]*	0.821	*[0.365]*	0.578	*[0.447]*	0.663	*[0.416]*
LR Squared Variables	61.043 (3)	*[0.000]*			67.543 (4)	*[0.000]*		
LR Cross-Product Var.			5.606 (3)	*[0.132]*			8.062 (3)	*[0.045]*
LR Translog Variables			66.649 (6)	*[0.000]*			75.605 (7)	*[0.000]*
LR Neighborood Var.					19.593 (2)	*[0.000]*	22.049 (2)	*[0.000]*

Maximum Likelihood Estimates with standard errors in parenthesis, 312 Observations.

All regressions include 11 Industry Shares (NES 16 Level).

The square and cross-product are computed in deviation from the mean of the variable.

** : significant at 5 % level ; ** : significant at 1 % level.*

LM Test SAR vs. OLS : Test of Spatial Autoregressive Model vs. No Spatial Model (distributed as $X^2(1)$ under the null)

LM Test SGM vs. SAR : Test of Spatial Generalized Model (with autoregressive error) vs. Spatial Autoregressive Model (distributed as $X^2(1)$ under the null)

LR Tests : Likelihood ratio test of squared, cross product and all translog variables, with degrees of freedom in parenthesis and p-value in squared brackets.

Table 4: Estimates of Translog production function with local R&D spillovers

Looking finally at the estimates of the average elasticity $\hat{\eta}_A$ of neighborhood R&D capital in regressions (2), (4), (7) and (8) where we used our preferred measure *K*100, we see that they are statistically significant and of nearly 0.015, half of the average elasticity $\hat{\gamma}_A$ [15]. The Translog estimates show again that the elasticity η is not constant across employment areas but appears to increase moderately with the intensity of the neighborhood R&D capital. In regressions (3) and (4), we present two among the different regressions we did in order to assess approximately the spatial range of R&D spillovers beyond employment areas, using different measures of neighborhood R&D capital stocks *K*80, *K*150, *K*200 and *K*250 in a circle of respectively 80 km, 150 km, 200 km and 250 km. We see in regression (3) that the average elasticity $\hat{\eta}_A$ becomes not statistically different from zero if we use the broader definition of neighborhood R&D capital *K*200 instead of our preferred one *K*100. Equivalently in regression (4) where we include in addition to Log(*K*100/*L*) the variable Log((*K*200-*K*100)/*L*) measuring the intensity of R&D capital stocks in the neighboring employment areas centered in the 100 km to 200 km ring, we see that this variable is also not statistically different from zero.

4 Industry R&D Spillovers

In this section, we attempt both to confirm and be more specific about our findings on local R&D spillovers by pursuing our analysis at the level of five large manufacturing industries and by differentiating between own-industry and other-industry R&D spillovers. We have been able to partition our employment area database according to the French one-digit industry classification NES 16, and we can focus on five large manufacturing industries, leaving aside trade, transport, services, and other industries which typically invest very little in R&D. These five broad manufacturing industries are the following: (B) food and beverage industries; (C) consumption good industries; (D) motor vehicles industries; (E) equipment good industries; and (F) intermediate good industries.

We are thus now considering a much larger sample of 1 538 "industry-employment area" observations for which we computed, as we did previously for the whole economy, both an "own-industry" local R&D capital stock (*K*) and an "own-industry" neighborhood R&D stock (*K*100)[16]. To test whether we could find evidence of R&D spillovers across different industries, we also defined an "other-industry" local R&D capital stock (*Kdif*), simply computed for all industry-employment area observations as the sum of the own-industry local R&D capital for the four other industries[17].

Table 5 reports the estimates of the R&D capital stocks elasticities of interest for three regressions of the Translog productivity equation. All three regressions include fixed

[15] Taking for *K* and *K*100 their median values (in Table 1) that only differ by 20 %, this implies that the corresponding gross rate of return of neighborhood R&D capital would be about 60 % of that of local R&D capital, which is quite high, still plausible enough.
[16] We deleted 22 observations (out of 5x312=1560) because of zero own industry local and neighborhood R&D capital stocks *K* and *K*100.
[17] We also computed an "other-industry" neighborhood R&D capital stock (*Kdif*100); however the regression estimates of the corresponding elasticities were very small and non significant, and not worthwhile to be reported here. The same applies for the estimates we obtained when we tried to include in the regressions separately the logs of the "own-industry" local R&D capital stocks for the other industries, as four additional separate variables instead of the log of their sum Log(*Kdif*).

industry effects, and the results shown are the usual within-industry OLS estimates, since we do not find anymore significant evidence in favor of the (*SAR*) specification in our larger sample, once we control for industry effects. Regression (9) assumes that all parameters are equal across industry (except for the industry fixed effects), while regression (10) only restricts the R&D capital elasticities to be equal across industry, and regression (11) also allows the R&D capital elasticities to be industry specific. The within-industry OLS estimates of regression (11) are thus the same as the OLS ones, when estimating it separately for each five industries. The complete estimates for regression (11), including the elasticities for physical capital, are recorded in Table C1 in Appendix C.

Looking first at the χ^2 tests of equality of the R&D capital stocks elasticities in regression (11), as well as the Likelihood ratio tests of the fully pooled and semi-pooled regressions (9) and (10) against the more general regression(11), clearly conclude in favor of the latter specification. However, it also appears that the specification of regression (10) is mainly rejected because of very significant industry differences in the estimated elasticities of the local R&D capital Log(K/L). Actually, the estimates of the five other R&D capital elasticities in the Translog equation, that is for Log($Kdif/L$), Log($K100/L$), Log$(K/L)^2$, Log$(Kdif/L)^2$ and Log$(K100/L)^2$ are not statistically different across industry at the 5 % (or more) confidence level.

Focusing now on the magnitude of the estimates, we see that the <u>average</u> elasticity $\hat{\gamma}_A$ of the local R&D capital, as estimated overall five manufacturing industries in regressions (9) an (10), is again statistically very significant (as when estimated for the French business non agricultural economy as a whole in the previous section), but that it is of a much higher order of magnitude of about 0.09 (as against 0.03 before). This important difference in size is largely explained by the fact that we are now considering manufacturing industries only[18]. We also see that the elasticity γ is not constant across industry and employment areas and that it is increasing as before, but even more strongly, with the intensity of local R&D capital: $\hat{\gamma} \;\; = 0.09 + 0.05\left(\log(K/L) - \overline{\log(K/L)}\right)$. We find, however, when considering regression (11), that the estimated average elasticity $\hat{\gamma}_A$ can be quite different across industries. It is significantly higher, of about 0.21, in the Motor vehicles industries, but falls in the range of 0.05 to 0.10 in the other industries. It is noteworthy that the average elasticity $\hat{\gamma}_A$ remains statistically different from zero at the 1 % confidence level, except in the Food industries where it is only significant at the 10 % confidence level. It is also interesting to observe that the elasticity γ tends to be significantly increasing with local R&D capital intensity Log(K/L) even within industry.

[18] It is also explained by the related fact that $\hat{\gamma}_A$ is now measured at a different sample average value $\overline{\log(K/L)}$ of the local R&D capital intensity, which is much higher for manufacturing industries than for the overall business economy.

Regression	**(9)**	**(10)**			**(11)**		
	common	**common**	**B**	**C**	**D**	**E**	**F**
Log(K/L)	0.089** (0.013)	0.093** (0.011)	0.040 (0.021)	0.091** (0.012)	0.210** (0.035)	0.075** (0.011)	0.055** (0.011)
Log(Kdif/L)	0.009* (0.004)	0.009* (0.004)	0.008 (0.006)	0.018* (0.007)	-0.001 (0.013)	0.020** (0.007)	0.003 (0.006)
Log(K100/L)	0.015** (0.004)	0.019** (0.004)	0.023** (0.006)	0.019** (0.007)	0.020 (0.013)	0.015* (0.006)	0.014* (0.006)
$Log(K/L)^2$	0.025** (0.004)	0.025** (0.004)	0.011 (0.007)	0.020** (0.004)	0.045** (0.010)	0.020** (0.005)	0.016** (0.005)
$Log(Kdif/L)^2$	0.000 (0.002)	0.001 (0.002)	0.001 (0.003)	0.005* (0.002)	-0.005 (0.005)	0.006* (0.003)	0.002 (0.002)
$Log(K100/L)^2$	0,003* (0.001)	0.003* (0.001)	0.002 (0.002)	-0.000 (0.002)	0.004 (0.004)	0.003 (0.002)	0.005* (0.002)
s	0.229748	0.220377	0.214060				
R^2 adj.	0.6288	0.6584	0.6777				
Log. Likelihood	91.814	182.738	240.195				
Equality Log(K/L)					22.734 [0.000]		
Equality Log(Kdif/L)					5.577 {0.233]		
Equality Log(K100/L)					1.439 [0.837]		
Equality $Log(K/L)^2$					0.534 [0.970]		
Equality $Log(Kdif/L)^2$					5.190 [0.268]		
Equality $Log(K100/L)^2$					8.378 [0.079]		

OLS Estimation with heteroskedastic-consistent standard-errors. 1 538 Observations.
* : significant at 5% level; ** : significant at 1% level.
Equality : Wald test of equality of parameters across industries (p-value in squared brackets)
Regression (9) is a pooled regression with industry specific effects.
Regression (10) allows all non-R&D capital parameters to vary accross industries.
Regression (11) allows all parameters to vary accross industries.
B=Food industries; C=Comsumption good industries; D=Motor vehicles industries, E=Equipment good industries; F=Intermediate good industries.

Table 5: Estimates of Translog production function with local and industry R&D spillovers

The estimates of the average elasticity $\hat{\eta}_A$ of the neighborhood R&D capital in regressions (9) and (10) remain statistically significant as before, but with the same order of magnitude of 0.015 (or perhaps just slightly higher), contrary to the average elasticity estimate $\hat{\gamma}_A$. We also find a moderate tendency for the elasticity η to increase with the intensity of neighborhood R&D capital. We observe in regression (11) that the industry estimates of η do not statistically differ and are roughly constant across industries, again contrary to the corresponding estimates γ for local R&D capital.

Finally, we only find weak evidence that local spillovers are not only industry specific, but are also significant and sizeable across different industries. The estimated elasticities of other-industry local R&D capital (*Kdif*) in regressions (9) and (10) are just significant at the 5% level of confidence and of about 0.01, that is much smaller by a factor of nearly 10 than the estimated elasticities of own-industry local R&D capital (*K*). In regression (11) we see that the elasticities of other-industry local R&D capital are significant and of about 0.02 for two industries out of the five: consumption goods and equipment goods industries.

This result is in contrast with the results obtained by AUTANT-BERNARD, LESAGE [2008] where they find between-industries spillovers. Our evidence of weak effect of other-industry local R&D capital can be explained by the high level of aggregation of the data in a employment area and in broad industry classification. rather than at the firm level or at a detailed industry level. In fact in our data, they are only a small share of firms in each employment area which undertake R&D projects. The cross-industries spillovers is small in our estimates because the R&D spillovers are disseminated on a small number of firms doing R&D and a large number of firms not doing R&D, and because the spillovers are concentrated within broad industry class rather than between these classes.

5 Conclusion

This note is a contribution to the existing literature on the effects of local R&D spillovers on productivity in their geographic and industrial dimensions. Our estimations of an extended Cobb-Douglas and Translog production function with local and neighborhood R&D capital are performed at the level of some 300 employment areas for the French non agricultural business economy as a whole in 1999. They are also confirmed and generalized on a larger sample of some 1 500 observations for each employment area for five broad manufacturing industries. Even though R&D investments are very highly concentrated in a few employment areas around Paris and other large French cities, we find statistically significant and large but plausible average spillover effects of local R&D capital on productivity. In addition to such more direct effects, we also find statistically significant but smaller effects of neighborhood capital R&D for the neighboring employment areas extending on average as far as 100 km, but not beyond. We also observe that these effects are not constant across employment areas, but increase very significantly with the R&D capital intensity. These findings are strongly confirmed at the industry-employment level, which show that local R&D spillovers tend to be mostly industry specific, and that the evidence for R&D spillovers across different industries is much weaker.

Although surprisingly good and robust, our results should still be considered as exploratory in view of many shortcomings related mainly to the data, and in particular its cross-sectional nature, and its consequences in terms of econometric modeling and estimation. Data on comparable cross-sectional employment area for a few number of years, and least one more recent year than 1999 would be very useful, but the data construction is complex and costly. An analysis at a more detail industry classification level might also be possible, though difficult. Localized data, and preferably panel data, at the establishment level is *a priori* the preferable way to go; it has however, also its own important problems.[19]

Références

Agrawal A., 2002, « Innovation, Growth Theory and the Role of Knowledge Spillovers », *Innovation Analysis Bulletin*, 4(3), p. 3-6.

Anselin L., 1988, *Spatial Econometrics: Methods and Models*, Kluwer Academic Publishers, Doordrecht.

Anselin L., Bera A. K., Florax R., Yoon M. J., 1996, « Simple Diagnostic Tests for Spatial Dependence », *Regional Science and Urban Economics*, 26, p. 77-104.

Audretsch D., Feldman M., 2004, « Knowledge Spillovers and the Geography of Innovation », in Henderson V. and J. F. Thisse, J. F., eds, *Handbook of Regional and Urban Economics, Volume 4,* , Amsterdam: Elsevier.

Autant-Bernard C., Lesage J., 2008, « Specialization, Diversity and Geographical Diffusion of Knowledge », *Mimeo* CREUSET.

Autant-Bernard C., Mairesse J., Massard N., 2007, « Spatial Knowledge Diffusion through Collaborative Networks: An Introduction », *Papers in Regional* Science, 86(3), p. 341-350.

Ciccone A., 2002, « Agglomeration Effects in Europe », *European Economic Review*, 46, p. 213-227.

Cliff A. D., Ord K., 1980, *Spatial Process : Models and Applications*, Pion, London.

Cockburn I., Henderson R., 1998, « Absorptive Capacity, Coauthoring Behavior, and the Organization of Research in Drug Discovery », *Journal of Industrial Economics*, 46(2), p. 157-182.

Cohen W. M., Levinthal D. A., 1989, « Innovation and Learning: the Two Faces of R&D », *Economic Journal*, 99(3), p. 569-596.

Crépon B., Mairesse J., 1993, « Productivité, Recherche – Développement et Qualifications », *INSEE-Méthodes*, 37-38, p. 181-220.

Gambardella A., Mariani M., Torrisi S., 2002, « How Provincial is your Region? Effects on Labour Productivity in Europe », Research Memorandum N° 2002-002, MERIT, Maastricht.

Griliches Z., 1979, « Issues in Assessing the Contribution of Research and Development to Productivity Growth », *Bell Journal of Economics*, 10(1), p. 92-116.

Griliches Z., 1992, « The Search for R&D Spillovers », *Scandinavian Journal of Economics*, 94(Supplement), p. S24-S47.

Griliches Z., Mairesse J., 1998, « Production Functions: The Search for Identification », in Ström S., ed., *Econometrics and Economic Theory in the 20th Century: The Ragnar Frish Centennial Symposium*, Cambridge University Press, p. 169-203.

[19] See for example GRLICHES – MAIRESSE [1998] for a survey of the difficulties involved in the identification and estimation of the production function using micro-panel data.

INSEE, 1994, tlas des Zones d'Emploi, DATAR – Ministère de l'Education Nationale – Ministère de l'Industrie et du Commerce Extérieur, Ministère du Travail, de l'Emploi et de la Formation Professionnelle.

Julia J.-L., 2003, « Base de Données Economiques Localisées », *Mimeo,* PSAR Etudes Economiques Régionales, INSEE – DR Midi-Pyrénées.

Krugman P., 1991, *Geography and Trade*, The MIT Press, Cambridge .

Le Gallo J., 2002, « Econométrie spatiale : l'autocorrélation spatiale dans les modèles de régression linéaire », *Economie et Prévision*, 155(4), p. 139-157.

Lesage J., 2000, *Spatial Econometrics*, Mimeo, University of Toledo, available at http://www.spatial-econometrics.com/html/wbook.pdf.

Mairesse J., Sassenou M., 1991, « R-D and Productivity: a Survey of Econometric Studies at the Firm Level », Science-Technology Industry Review, Paris, OECD, 8, p. 9-43.

Moran P. A. P., 1950, « A Test for the Serial Independence of Residuals », *Biometrika*, 37(1/2), p. 178-181.

Ord K., 1975, « Estimation Methods for Models of Spatial Interaction », *Journal of American Statistical Association*, 70(349), p. 120-126.

APPENDIX A : Measurement of local and neighborhood R&D capital variables

The R&D data we use to measure the R&D capital stocks at the employment area level are provided by the annual surveys on firms' R&D expenditures conducted by the statistical office of The French Ministry of Research since the seventies. In these surveys, since 1993, firms which have several laboratories or research centers are asked to report the geographical decomposition of their total internal R&D expenditures and total number of R&D workers by French "departments" (NUTS 3 level). We use this decomposition together with the postal addresses of firms' establishments to determine the localization of their R&D expenditures and number of workers at the very detailed geographical level of the some 36 000 French "communes" or municipalities. These estimates are then summed up to the level of the 341 employment areas which are aggregates of municipalities.

The local R&D capital stocks (K) at the beginning of year 1999 are estimated for each employment area by the permanent inventory method applied on the basis of the past internal R&D expenditures (R) so obtained for the six years 1993 – 1998, after deflation by an overall R&D price index and depreciation assuming with a constant depreciation rate δ of 15 %, that is using the following formula:

$$K_{1999} = \sum_{\tau=1993}^{1998} (1-\delta)^{1998-\tau} \frac{R_\tau}{P_\tau^{RD}} \qquad \text{(A-1)}$$

Note that we did not try to make any adjustment for the unknown initial stock of R&D capital in 1993, since this should not affect noticeably our cross-sectional estimates of the R&D capital elasticities of interest here. With a rate of depreciation δ of 15 %, it is also the case that about 38 % of the R&D capital stock at the beginning of 1992 is not depreciated at the beginning of 1999, which will represent about 28 % of the R&D capital stock at the beginning of year 1999, when assuming that R&D investments have been growing at an average annual growth rate of 5 %.

The neighborhood R&D capital stock ($K100$) for any given employment area is simply computed by summing up the R&D capital stocks (K) in the employment areas which are in a circle of 100 km around this given area. In this procedure, we assume that all the R&D capital of an employment area is localized at its geographical center. Precisely, we have constructed a matrix A_{100} which indicates if the distance between two employment areas i and j is less then 100 km:

$$A_{100} = [a_{i,j}] \quad \text{such that} \quad a_{i,j} = \begin{cases} 1 & \text{if } 0 < dist(i,j) \leq 100km \\ 0 & \text{otherwise} \end{cases} \qquad \text{(A-2)}$$

Denoting by $\underline{K}$ the vector of local capital stock for all employment areas and by $\underline{K100}$ the corresponding vector of neighborhood capital stock, we can compute simply the latter as:

$$\underline{K100} = A_{100}\,\underline{K} \qquad \text{(A-3)}$$

Note that by construction the matrix A_{100} is symmetric and the coefficients of its main diagonal are zeros. Note also that this matrix is not row-standardized as a classical spatial weight matrix since $K100$ is defined as the sum (not the average) of the local R&D capital stocks K for the neighboring areas.

To assess approximately the spatial range of R&D spillovers we have also considered different measures of neighborhood R&D capital stocks, based on alternative choices of distance between the geographical centers of an employment area and its neighboring areas. Besides using $K100$, we have thus experimented with neighborhood R&D capital stock in a circle of 80 km, 150 km, 200 km and 250 km : respectively $K80$, $K150$, $K200$ and $K250$. See Table 3 where we report different estimates of the Cob-Douglas productivity equation using respectively $K100$ and $K200$ alone, and both $K100$ and ($K200$ – $K100$).

APPENDIX B : Brief overview of spatial econometric methods

We use the classical spatial econometrics method developed by ANSELIN [1988], LESAGE [2000] or LE GALLO [2002]. We can treat the error terms as a first-order spatial autocorrelation process to give the Spatial Error Model (*SEM*), but we prefer to add the spatial lag of the dependent variable as an additional explanatory variable and consider the Spatial Autoregressive Regression (*SAR*), or spatial lag model. If we write y the vector of observations on the dependent variable, X the matrix of the regressors, and W the contiguity matrix, the (*SAR*) model can be written as:

$$y = \rho W y + X\beta + \varepsilon \tag{B-1}$$

This model cannot be consistently estimated by least-squares because it includes the spatially lagged dependent variable as a regressor. We have instead to rely on the maximum likelihood method. Assuming normality of the error term:

$$(I - \rho W)y - X\beta = \varepsilon \approx N(0, \sigma^2 I)$$

the log likelihood function is the following:

$$\begin{aligned} \log L(\rho, \beta, \sigma^2) = & -\frac{N}{2}\log(2\pi) - \frac{N}{2}\log(\sigma^2) - \log|I - \rho W| \\ & -\frac{1}{2\sigma^2}\left[\left((I - \rho W)y - X\beta\right)'\left((I - \rho W)y - X\beta\right)\right] \end{aligned}$$

Using the transformation proposed by ORD [1975], the log of the determinant $|I - \rho W|$ can be computed as:

$$\log|I - \rho W| = \sum_{i=1}^{N} \log(1 - \rho\omega_i)$$

where ω_i are the eigenvalues of the contiguity matrix W, which can be computed themselves once for all in the iterative maximization procedure. The global solution for maximization of log likelihood function is quite fast using a Matlab estimation program software adapted from the routines provided by James LESAGE on the web site: http://www.spatial-econometrics.com/.

Following ANSELIN [1988] or ANSELIN – BERA – FLORAX – YOON [1996] and using Lagrange multiplier tests, we perform various specification tests of the *SAR* specification against other specifications. First of all, we can test the (*SAR*) regression for the null of no spatial autocorrelation ($\rho = 0$), that is against the usual regression as estimated by Ordinary Least Squares (*OLS*). We can also test the (*SAR*) model against a more general model called the Spatial Generalized Model (*SGM*), which allows spatial autocorrelation in the error term, and thus encompass the *SEM* specification as well as the *SAR* specification. The *SGM* model can be written as:

$$y = \rho Wy + X\beta + u \qquad \text{with } u = \lambda Wu + \varepsilon \tag{B-2}$$

or also as the following second order spatial autoregressive model:

$$y = (\rho + \lambda) Wy + \rho\lambda WWy + X\beta - WX\lambda\beta + \varepsilon \tag{B-3}$$

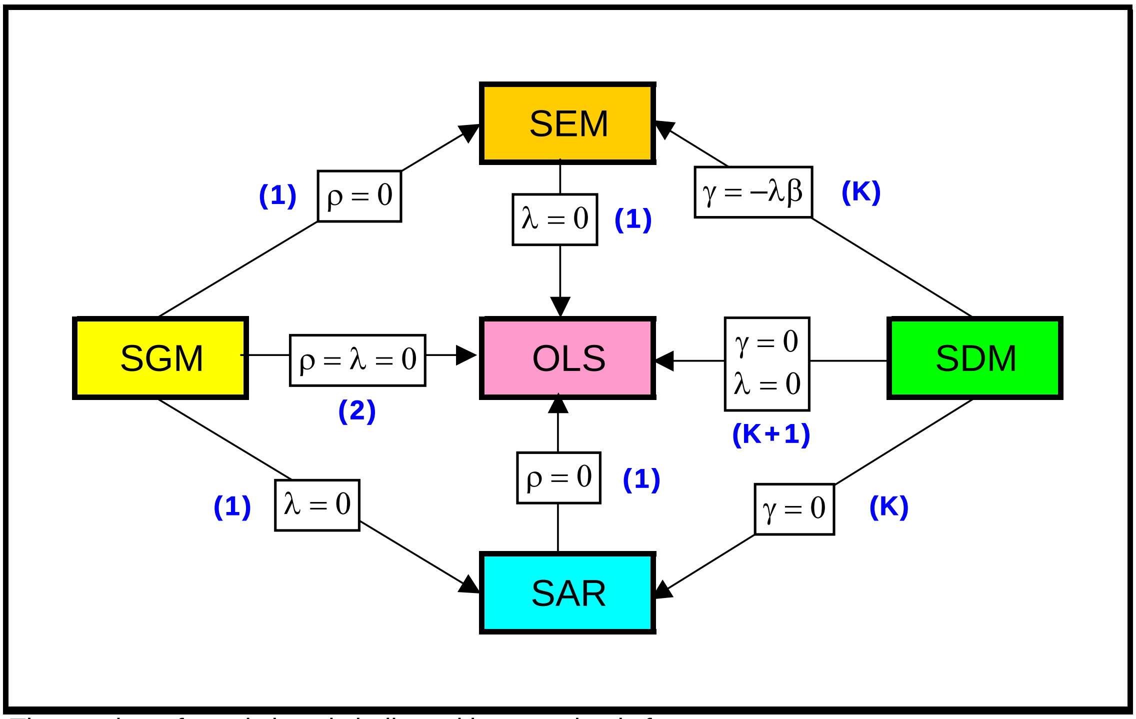
with one cofactor restriction on the parameters of spatial lagged regressors. When there is no lagged dependent variables ($\rho = 0$), the spatial error model (*SEM*) is obtained with such restrictions again. :

$$y = \lambda Wy + X\beta - WX\lambda\beta + \varepsilon \tag{B-4}$$

Finally without this cofactors restriction, we obtain a model proposed by Durbin, called the Spatial Durbin Model (*SDM*), which can be tested against the previous SEM:

$$y = \lambda Wy + X\beta + WX\gamma + \varepsilon \tag{B-5}$$

Figure B1 summarizes the relations between the usual (non spatial) regression (*OLS*) and the four spatial regression models: *SEM, SAR, SGM* and *SDM*. We have tested that the SAR specification was preferable to the *OLS* and *SEM* specifications and was an acceptable restriction to the *SGM* and *SDM* specifications.



The number of restrictions is indicated in parenthesis for K regressors.

Figure B1 : Relations between Spatial Regression Models

APPENDIX C

INDUSTRY	B	C	D	E	F
Industry Dummy	1.278**	2.085**	2.665**	2.813**	2.514**
	(0.165)	*(0.131)*	*(0.171)*	*(0.182)*	*(0.133)*
Log(L)	0.099**	0.074**	0.042	0.072**	0.027*
	(0.017)	*(0.015)*	*(0.024)*	*(0.015)*	*(0.014)*
Log(C/L)	0.403**	0.296**	0.215**	0.119*	0.261**
	(0.021)	*(0.024)*	*(0.029)*	*(0.049)*	*(0.023)*
Log(K/L)	0.040	0.091**	0.210**	0.075**	0.055**
	(0.021)	*(0.012)*	*(0.035)*	*(0.011)*	*(0.011)*
Log(Kdif/L)	0.008	0.018*	-0.001	0.020**	0.003
	(0.006)	*(0.007)*	*(0.013)*	*(0.007)*	*(0.006)*
Log(K100/L)	0.023**	0.019**	0.020	0.015*	0.014*
	(0.006)	*(0.007)*	*(0.013)*	*(0.006)*	*(0.006)*
$Log(L)^2$	-0.276	0.005	0.005	0.001	0.004
	(0.011)	*(0.006)*	*(0.009)*	*(0.008)*	*(0.009)*
$Log(C/L)^2$	0.069*	0.136**	0.050**	0.022	0.121*
	(0.031)	*(0.043)*	*(0.016)*	*(0.086)*	*(0.050)*
$Log(K/L)^2$	0.011	0.020**	0.045**	0.021**	0.016**
	(0.007)	*(0.004)*	*(0.010)*	*(0.005)*	*(0.005)*
$Log(Kdif/L)^2$	0.001	0.005*	-0.005	0.006*	0.002
	(0.003)	*(0.002)*	*(0.005)*	*(0.003)*	*(0.002)*
$Log(K100/L)^2$	0.002	0.000	0.004	0.003	0.005*
	(0.002)	*(0.002)*	*(0.004)*	*(0.002)*	*(0.002)*
Log(L) * Log(C/L)	0.055*	-0.049	-0.030	0.025	-0.044
	(0.023)	*(0.031)*	*(0.018)*	*(0.053)*	*(0.026)*
Log(L) * Log(K/L)	0.006	-0.012	-0.355	-0.505	-0.483
	(0.012)	*(0.007)*	*(0.022)*	*(0.011)*	*(0.011)*
Log(L) * Log(Kdif/L)	0.006	0.022*	-0.506	-0.553	0.019*
	(0.007)	*(0.009)*	*(0.010)*	*(0.007)*	*(0.009)*
Log(C/L) * Log(K/L)	0.018	0.020	-0.054*	0.030	-0.847
	(0.017)	*(0.019)*	*(0.023)*	*(0.024)*	*(0.021)*
Log(C/L) * Log(Kdif/L)	-0.036**	-0.030	-0.015	0.043*	-0.037
	(0.013)	*(0.022)*	*(0.014)*	*(0.018)*	*(0.023)*
Log(K/L) * Log(Kdif/L)	-0.550	0.002	0.016	-0.831	-0.177
	(0.004)	*(0.006)*	*(0.014)*	*(0.005)*	*(0.005)*
Sum of Squared residuals	65.8918				
Stdandard error of residuals	0.214060				
R^2 adusted	0.6777				
LM Test Heteroskedasticity	15.6043 [p-value : 0.000]				
LR Test Pooled Model (9)	296.761 (df = 72) [p-value : 0.000]				
LR Test Semi-Pooled Model (10)	114.914 (df = 24) [p-value : 0.000]				

OLS Estimation with heteroskedastic-consistent standard-error. 1538 observations.
The regression includes three binary indicators for the few observations with respectively missing or zero values of C/L, K/L and Kdif/L.
* : significant at 5% level; ** : significant at 1% level.
B=Food industries; C=Comsumption good industries; D=Motor vehicles industries, E=Equipment good industries; F=Intermediate good industries.

Table C1 : Translog production function with local and industry R&D spillovers:

Cahiers du GRES

Le **GRES (Groupement de Recherche Economiques et Sociales)** est un Groupement d'Intérêt Scientifique entre l'Université Montesquieu-Bordeaux IV et l'Université des Sciences Sociales Toulouse I.

Il regroupe des chercheurs appartenant à plusieurs laboratoires :

- **GREThA - UMR CNRS 5113** (Groupe de Recherche en Economie Théorique et Appliquée), **Université Montesquieu-Bordeaux IV** ;
- **LEREPS - EA 790** (Laboratoire d'Etudes et de Recherche sur l'Economie, les Politiques et les Systèmes Sociaux), **Université de Toulouse 1 Sciences Sociales**;
- L'**UR 023** "Développement local urbain. Dynamiques et régulations"**, IRD** (Institut de Recherches pour le Développement) ;
- Le laboratoire **EGERIE** (Economie et de Gestion des Espaces Ruraux, de l'Information et de l'Entreprise), **ENITAB** (Ecole Nationale des Ingénieurs des Travaux Agricoles de Bordeaux).

www.gres-so.org

Cahiers du GRES (derniers numéros)

2008-03 : CORIS Marie, *Les délocalisations sous l'angle de la coordination : Une lecture par les catégories de la proximité appliquée au logiciel*
2008-04 : DUPUY Claude, LAVIGNE Stéphanie, *Investment behaviors of the key actors in capitalism: when geography matters*
2008-05 : GAY Brigitte, *Competitive intelligence and network mapping of interfirm alliances: strategic implications*
2008-06 : POUYANNE Guillaume, *Théorie économique de l'urbanisation discontinue*
2008-07 : LACOUR Claude, PUISSANT Sylvette, *Medium-sized cities and the dynamics of creative services*
2007-08 : MINDA Alexandre, *The strategies of multilatinas: from the quest for regional leadership to the myth of the global corporation*
2007-09 : NGUYEN Huu Thanh Tam, KECHIDI Med, *Investissement direct étranger d'exportation-plateforme: quels impacts sur la production locale ?*
2008-10 : CLEMENT Matthieu, MEUNIE André, *Economic Growth, inequality and environment quality: An empirical analysis applied to developing and transition countries*
2008-11 : KECHIDI Med, *Constructeur aéronautique versus architecte-intégrateur : le nouveau modèle Airbus*
2008-12 : THOMAS Olivier, *L'acceptation des dérapages de l'intercommunalité en France : une manière détournée d'établir les fondements d'une fusion implicite entre communes ?*
2008-13 : AUVRAY Tristan, BROSSARD Olivier, *Ownership concentration and market discipline in European banking: Good monitoring but bad influence?*
2008-14 : LAYAN Jean-Bernard, LUNG Yannick, *Attractivité et agglomération de l'industrie automobile au Maroc et en Tunisie : une analyse comparative*
2008-15 : MAIRESSE Jacques, MULKAY Benoît, *An exploration of local R&D spillovers in France*

La coordination scientifique des Cahiers du GRES est assurée par Alexandre MINDA (LEREPS) et Vincent FRIGANT (GREThA). La mise en page est assurée par Dominique REBOLLO.