Labor Market and Industrial Location: Evidence from French Data

Carl Gaigné, Jean-Pierre Huiban and Bertrand Schmitt
Cesaer-Inra, France

June 16, 2003

Abstract

This paper focuses on the respective roles of labor costs, input-output linkages between firms, and the market size effect to explain the location of industries. Empirical estimations are based on an Economic Geography model in which we disaggregate the labor factor to take into account the adjustment between the labor demand structure and local wages based on skill. We use French data at the Labor Market Area level which are assembled from an comprehensive database of plants for 1997 and aggregated in order to obtain a cross-regional and cross-industry database. We found evidence that low local labor costs when they are associated with a large local (final or intermediate) goods markets may explain the industrial location, according to the nature of the industry.

Keywords: Economic Geography, Industrial Location, Labor Costs, Vertical Linkages, Final Demand Size.

JEL classification: R12, F12, J31

[#]Corresponding author: Carl Gaigne, 26 Bld Petitjean, BP 87999, F-21069 Dijon, France. Tel: 33 (0)3 80 77 26 69. Fax: 33 (0)3 80 77 25 71. Email address: gaigne@enesad.inra.fr.

The authors wish to thank P. Blanchard, E. Caroli, S. Chandra, P.-Ph Combes, M. Lafourcade, P. Sevestre and J. Thisse for their valuable comments. They would also like to thank participants at the RSAI North American, CESAER-ERUDITE and CESAER-CORE meetings and seminars at LEA in Paris, as well as the INSEE for providing data and Virginie Piguet for preparing data.

1 Introduction

Recent developments in the empirics of economic geography (see Head and Mayer, 2003 for a survey) have focused their analysis on the mechanisms that generate economic agglomerations, such as the market size effect and the decline of transport costs in different contexts (Combes and Lafourcade, 2001; Davis and Weinstein, 1999; Hanson, 1999; Mion, 2002; Redding and Venables, 2002). Although the firm's location choice is the result of some interplay between centripetal and centrifugal forces (Fujita *et al.*, 1999; Fujita and Thisse, 2002), the empirical literature has given less attention to the mechanisms that could favor the dispersion of activities, especially labor cost.[1] Indeed, firms have incentives to agglomerate in order to take advantage of increasing returns associated with the local size of intermediate and final goods markets. However, this agglomerating effect may be opposed by a dispersing effect, because the spatial concentration of firms favors a rise in wages due to competition to attract workers, which encourages firm dispersion to take advantage of lower labor costs.

This paper aims to take into account the role of regional wage differentials in industrial location, in parallel with the input-output relationship and final demand. The existence of such inter-regional wage differentials based on skill is empirically attested in Europe (Combes and Overman, 2003), particularly in France (as shown by our French micro-geographic data in Appendix A). This fact is related to the low rate of migration between regional labor markets. As now recognized, the spatial mobility of workers is relatively low in Europe (see Faini, 1999, for a survey). While the workforce seems to be relatively indifferent to spatial differentials in wages, we assume that the firm location decision takes this dimension into account. However, the sectoral dimension is important because job structure varies greatly between industries. Indeed, there is less incentive for low-skilled industries than for high-skilled industries to set up in a region where the wages of unskilled workers are high, and vice-versa. This leads us to consider the heterogeneity of industrial sectors as well as to disaggregate the labor factor according to different skill levels and to introduce the corresponding regional wage level.

Therefore, in this paper, we investigate the effect of spatial frictions between labor demand structure and local wages based on skill on the regional

[1] This is the case even though some studies have focused on the role of land rents as a dispersion force (Brakman *et al.*, 2002; Tabuchi and Yoshida, 1999).

industrial specialization. In addition, we consider the effect of the local intensity of vertical linkages and the size of local final demand. Our empirical analysis is driven by a theoretical model based on that of Krugman and Venables (1995), with R regions and S industrial sectors, in which dispersion force is due to labor cost and the agglomeration force is due to the vertical linkages between firms. We develop this model in order to introduce the final goods market size effect by considering an uneven distribution of population between regions. As suggested above, the labor factor is subdivided into several skill categories so as to introduce a vertical differentiation among workers and to allow for a more refined analysis of local adjustments on the labor market. In short, access to both intermediate and final goods markets constitutes the agglomeration forces, while local labor costs, reflecting tension between supply and demand for labor on the local labor markets, act as the dispersion force. Hence, finally we obtain an equation system. To estimate it, we use a French comprehensive database, the 1997 'Déclarations annuelles de données sociales' ('DADS'), aggregated at the sectoral and regional level so that the resulting database is composed of the conjunction of the 341 labor market areas and 67 industrial sectors, forming a 'spatial panel'. As suggested by the model, we pay attention to two econometric issues, which are the endogeneity of labor costs and the sectoral heterogeneity. They leads us to use the methods of panel data and instrumental variables in the same regression and to perform estimations sector by sector.

By controlling for heterogeneity of sectors and endogeneity of labor cost, our analysis reveals that labor cost favors the dispersion of industrial activities while the vertical linkages favor the agglomeration, and the size of the final demand associated with the specific area seems to have a weaker effect. In addition, we show that the sector-based characteristics have a strong effect on the industrial location process. More precisely, the highly competitive sectors tend to spread their firms and jobs whereas high productivity sectors tend to be agglomerated. Moreover, estimations sector by sector show that the proximity of firms demanding or supplying intermediate goods is a powerful agglomeration factor for a very large majority of sectors. Whatever their activities, industrial firms seem to concentrate where input demand or supply is higher. By contrast, the effect of labor costs on the spatial distribution of activities varies from one sector to another. The spatial differentiation of labor costs acts as a dispersion force in the case of activities that are labor intensive and/or where competition on the goods markets is fierce. Other sectors are less sensitive to such differentiation of labor costs, probably because they require a labor force with specific skills. Finally, the final demand affects the location process of only a few sectors. In summary, our findings suggest that, in a large set of sectors, *industrial firms face a trade-off between agglomerating to take advantage of increasing returns due to input-output linkages and dispersing to enjoy lower labor costs.*

These results lead us to question the spatial scale at which the theoretical framework of economic geography is consistent. The Krugman and Venables (1995) model to which we refer is often considered as a relevant set-up to analyze the international distribution of activities, both theoretically and empirically. Our results suggest that the agglomeration and dispersion forces encapsulated in this model remain valid at a micro-geographic level, in labor market areas where the mobility of workers between such areas is relatively low (as in Europe). Indeed, the interregional distribution of industrial activities seems to be driven by the proximity of intermediate goods suppliers and purchasers and by the spatial differentiation of the labor costs for many sectors. Redding and Venables (2002) show the agglomeration force due to the vertical linkages remains influential at an international scale. Conversely, the role of the local final demand, which is a key variable in the Krugman model (1991), is weak at the selected intra-regional level. This result is consistent with Davis and Weinstein (1998), who show economic activities are little sensitive to the fine-scale location of demand (at the Japanese prefectures level). However, they seem to be more sensitive to the distribution of final demand at a large geographic level (OECD countries).

Finally, by focusing on the role of labor cost in industrial location, our analysis completes some existing studies in different ways. Indeed, several approaches that test a reduced form of economic geography models do not consider the impact of wages on the industrial location. First, some studies test the opposite effect, so that the dependent variable is the regional wage (Brakman *et al.*, 2002; Hanson, 1999; Mion, 2002; Redding and Venables, 2000). In this way, the home market effect is captured: the nominal wage rate in a region tends to be higher if the demand addressed to this region is high. However, if the agglomeration of industrial activities raises nominal wages, this wage increase may, in turn, favor dispersion. Second, although Combes and Lafourcade (2001) focus on the impact of the fall in inter-regional transport costs on industrial location, they assume that the wage does not vary in space.[2] Our analysis reveals that the labor costs are significant in

[2] Note also that although the observation level in Combes and Lafourcade (2001) is the

the location process of industrial activities. Finally, we note that Head and Mayer (2002) estimate the impact of the local wage on the location of foreign direct investment, but they do not find a significant effect. However, they do not account for the differences based on skill and the heterogeneity of sectors. Our results suggest that the sectoral specificities must be controlled. As suggested by Davis and Weinstein (1999) and Midelfart-Knarvik *et al.* (2000), firm behavior in the location process depends on the specificities of the production technology.

The paper is organized as follows. The next section describes the model. Section 3 sets out the data and the econometric methods used while the results are reported and analyzed in section 4. The final section provides a summary and some concluding remarks.

2 The model

Our model is based on Krugman and Venables (1995). The economy has R regions. In each region $r = 1, .., R$ two kinds of goods can be produced: an agricultural good and S industrial goods. The agricultural sector produces a homogenous good under constant returns. Each industrial sector $s = 1, .., S$ produces a continuum of varieties of a horizontally differentiated product. The number and the spatial distribution of the firms in each industrial sector are endogenous. Each industrial good is both sold to consumers/workers and used as an intermediate good in other industries. In addition, we pay special attention to the labor input, which we disaggregate into several categories $\left(L^1, ..., L^i, ..., L^I\right)$ that corresponds to a vertical (*i.e.* skill) differentiation of the labor force. Furthermore, we introduce an uneven regional distribution of workers/consumers who are geographically immobile but sectorally mobile.

2.1 Assumptions

Let us start by describing the industry. The number of firms in industry s operating in region r is denoted by n_r^s, which is an endogenous variable. Let us assume that the firms belonging to the industry $s = 1, .., S$ located in region r produce one differentiated good in quantity (q_r^s) in a monopolistically competitive good market (*à la* Dixit and Stiglitz, 1977). The output of each

same that we use (the French 'Labor Market Areas', LMAs), our database is different and is more complete.

industry's products can be aggregated *via* a CES function into S aggregates. These aggregates may be used as a consumption good and as an intermediate input. Each firm uses S intermediate goods aggregates and labor factors corresponding to different quality levels l_i, $i = 1, ..., I$. Furthermore, each firm belonging to industry s needs a fixed part α^s and a variable one β^s of those inputs. Therefore, the cost function for a firm in s set up in region r is given by:

$$C_r^s = (\alpha^s + \beta^s q_r^s) \prod^{I} \left(w_{r,s}^i\right)^{\phi_s^i} \prod^{S} (Pm_r^s)^{\mu^{s',s}} \tag{1}$$

where $\mu^{s',s}$ is the intensity of vertical linkages between the two industries s and s' (where s and $s' \in [1, S]$), ϕ_s^i is the intensity of quality i labor in the sector s, $w_{r,s}^i$ is the corresponding local wage and Pm_r^s is the price index of a composite of the manufactured goods s in region r, which is expressed as follows:

$$Pm_r^s = [n_r^s (p_r^s)^{1-\sigma^s} + \sum^{r'} (\tau^s)^{1-\sigma^s} n_{r'}^s (p_{r'}^s)^{1-\sigma^s}]^{1/(1-\sigma^s)} \tag{2}$$

with $\sigma^s > 1$ is the constant elasticity of substitution in industry s and τ^s is the 'iceberg' transportation cost for shipping one unit of good s from the region r to another region r'. Therefore, we assume that the trade cost of a unit of a good is the same regardless of the origin and destination regions. This assumption may be justified by the fact that distance-related shipping costs have become very low, while costs unrelated to distance, such as insurance and loading and unloading costs, are still relatively high.[3]

The variable p_r^s corresponds to the price of goods in the industry s located in region r, which is obtained by solving the first-order condition under demand constraint (Benassy, 1991), and is given by:

$$p_r^s = \frac{\sigma^s}{\sigma^s - 1} \beta^s \prod^{I} \left(w_{r,s}^i\right)^{\phi_s^i} \prod^{S} (Pm_r^s)^{\mu^{s',s}} \tag{3}$$

so that p_r^s is the marginal cost multiplied by a constant mark-up $\frac{\sigma^s}{\sigma^s-1}$. The equilibrium output $(q_r^s)^o$, under a free entry assumption (*i.e.* profit equals zero) is a constant, given by:

$$(q_r^s)^o = [(\sigma^s - 1)\alpha^s]/\beta^s \tag{4}$$

[3] Davis and Weinstein (1999) made a similar assumption.

We now consider the labor demand. Let $(l^i_{r,s})^d$ denote the labor demand of a firm belonging to industry s in region r for the skill level i, which can be expressed as $(l^i_{r,s})^d = \phi^i_s p^s_r q^s_r / w^i_{r,s}$. Therefore, because firms of the same industry located in the same region are symmetric, the labor demand of the industry s in region r may be written as:

$$(L^i_{r,s})^d = \phi^i_s n^s_r p^s_r q^s_r / w^i_{r,s} \tag{5}$$

The population of workers L is unevenly located across the R regions and L_r is the population employed in region r so that $L_r = \lambda_r L$ with $0 < \lambda_r < 1$ and $\Sigma_r \lambda_r = 1$. Moreover, we assume that workers are spatially immobile, although they can shift from one industry to another. Let θ^i_r denote the share of the labor force L_r with skill i living in the region r. Because of labor immobility, θ^i_r is exogenous. Thus, the local labor supply is:

$$(L^i_r)^o = \theta^i_r L_r \quad \text{with } \sum_{i=1}^{I} \theta^i_r = 1 \tag{6}$$

For a given labor skill, we assume that each worker chooses to work in the industry that locally offers the highest wage. Hence, there is competition between industries to attract the labor force. In this case, there is no difference in wages between sectors within a given region. Thus, for each skill i labor factor, we thus have:

$$w^i_{rs} = w^i_r$$

As usual in the Dixit-Stiglitz framework, workers assign a share $1 - \Sigma_s \gamma^s > 0$ of their income to the consumption of the agricultural good and a share $\gamma^s \in (0,1)$ to the consumption of S aggregates of industrial goods (q^s_j) with constant elasticity of substitution (σ^s). The assumption of a Cobb-Douglas utility function combined with the CES utility function between industrial goods gives:

$$U = \prod_{s=1}^{S} (M_s)^{\gamma^s} A^{1-\Sigma\gamma^s} \quad \text{with } M_s \equiv \left(\int_{j=0}^{n^s} \left(q^s_j\right)^{(1-1/\sigma^s)} dj \right)^{\sigma^s/(\sigma^s-1)}$$

The demand for the good s produced in region r is:

$$(q^s_r)^d = [(Pm^s_r)^{\sigma^s-1} E^s_r + \tau^{s(1-\sigma^s)} \sum^{r'} (Pm^s_{r'})^{\sigma^s-1} E^s_{r'}] (p^s_r)^{-\sigma^s} \tag{7}$$

where E_r^s is the share of the regional income spent on industrial goods, given by:

$$E_r^s = \gamma^s \sum_{i=1}^{I} w_r^i \theta_r^i \lambda_r L + \sum_{s'=1}^{S} \mu^{s',s} n_r^{s'} p_r^s q_r^s \tag{8}$$

Finally, we describe the agricultural sector. This sector is perfectly competitive. We assume that it produces a homogenous good using only labor as an input, with constant returns to scale. This output is the *numéraire* ($p_A = 1$). If we assume that one unit of labor produces one unit of output, the wage rate is such that $w_A = 1$. Hence, in order to attract workers industrial firms must be able to pay wages such as $w_{r,s}^i \geq 1$.

2.2 Equilibrium

We first determine the equilibrium on the goods markets, then the equilibrium on the labor markets.

1. *Equilibrium on the goods markets.* We solve $(q_r^s)^o = (q_r^s)^d$ (see 4 and 7) to obtain the equilibrium on the goods market. The linearization of this last equality (see Appendix B for details) gives:

$$\log n_r^s = \Omega_r^s + \Lambda^s + \Gamma_{r,r'} \tag{9}$$

where:

$$\Omega_r^s \equiv -\frac{1}{B^s} \sum_{i=1}^{I} \left(\phi_s^i \log w_r^i\right) + C^s \sum_{s'=1}^{S} \left(\mu^{s',s} \log n_r^{s'}\right) - \frac{1}{B^s} \sum_{s'=1}^{S} \left(\mu^{s',s} \log \varphi_r^{s'}\right) + \log\left(\gamma^s + \Psi^s\right) w_r L_r$$

$$\Lambda^s \equiv -\log \alpha^s - \frac{1}{B^s} \log \sigma^s + \frac{1 - B^s}{B^s} \log\left((\sigma^s - 1)/\beta^s\right)$$

with

$$C^s = 1/[B^s(\sigma^s - 1)] \quad \text{and} \quad B^s = 1 - \sum_{s'=1}^{S} \mu^{s',s} > 0$$

Finally, the term $\Gamma_{r,r'}$ in (9) is defined in Appendix B (equation B.7).

Although the equation (9) is an industrial firm location equation, it is related to a regional labor demand equation for each sector, taking account of equilibrium on the intermediate and final goods markets. The endogenous variable is the number of production units belonging to sector s and located in region r. This variable depends on three groups of variables.

The first group (Ω_r^s) is made up of explanatory variables that measure interactions between geographical and sectoral features. It involves three main terms:

(i). the *Local Labor Cost* is the product of local wages based on skill level and the skill requirements of the sector $\Sigma_i \left(\phi_s^i \log w_r^i\right)$. The lower the local labor cost, the higher the number of firms set up in the region.

(ii). the *Local Vertical Linkages* is defined as the sum of the products of the number of local firms in a sector s' by the intensity of vertical linkages between firms in sector s and firms in sector s' $(\Sigma_{s'} \mu^{s',s} \log n_r^{s'})$. We remove from this term the sum of the products of the intensity of vertical linkages $\mu^{s',s}$ by the level of relative prices of intermediate goods purchased in the region r $(\Sigma_{s'} \mu^{s',s} \log \varphi_r^{s'})$. The more suppliers of industrial goods there are in a region and the more frequent their purchases are, the more firms cluster in this region. However, the higher the prices charged by suppliers in a region, the fewer firms located in the region.

(iii). The *Local Demand* is given by the product of the income of the region's workers by the coefficient of household spending assigned to a sector, plus an indicator of demand for intermediate goods $(\log (\gamma^s + \Psi^s) w_r L_r)$. This final point shows that the agglomeration of sector s firms in region r depends positively on the local demand.

In addition to this first group of variables (Ω_r^s), the sectoral features are also decisive factors in the location process (Λ^s). We have three sectoral variables which are a combination of sectoral characteristics. The first gives a measure of an 'internal economies of scale effect'. The higher are the fixed costs α^s, the fewer firms there are. The second $(-1/B^s \log \sigma^s)$ measures a 'competition effect'. The more competitive is the sector (high σ^s), the more dispersed the firms are. This is a classic mechanism of economic geography. The third variable accounts for a 'productivity effect' $(\frac{1-B^s}{B^s} \log ((\sigma^s - 1)/\beta^s))$. It emphasizes the inputs productivity $1/\beta^s$ weighted by the level of competition on the goods market σ_s. The higher the level of productivity is, the more closely firms agglomerate. This effect is enhanced when industry is not very labor intensive (low B^s).

The third group of variables ($\Gamma_{r,r'}$) emphasizes the cross-regional dimension. It indirectly measures the relative importance of the region in terms of population, firms and prices ($\varphi_D^{r,s}$ and $\varphi_P^{r,s}$). In addition, it incorporates the ambiguous effect of transport costs τ^s on agglomeration: as we know, their influence is both negative and positive.

2. *Equilibrium on labor markets.* The equality of labor supply and demand (see 6 and 5) implies the equilibrium wage of each skill, given by:

$$w_{r,s}^{i} = \frac{\phi_{s}^{i}}{\theta_{r}^{i} L_{r}} n_{r}^{s} p_{r}^{s} q_{r}^{s} \tag{10}$$

All things being equal, this suggests that setting up a new firm in region r tends to increase the level of $w_{r,s}^{i}$. However, this effect on wages depends also on the ratio $\phi_{s}^{i}/\theta_{r}^{i}$, which is exogenous because workers are immobile. It is lower (higher) when the local supply of skill i, θ_{r}^{i}, is relatively high (low) with respect to the industry demand for this skill, ϕ_{s}^{i}. This point greatly influences local competition for labor, reducing (intensifying) the dispersion effect. This effect, which is not taken into account in Krugman and Venables (1995), affects the intensity of competition in the local labor market.

To sum up, we obtain a system of $1+I$ equations. The location equation (9) looks like a regional equation for sectoral labor demand. It takes into account the labor demand behavior of sector s in region r in view of the wage levels based on skill and considers equilibria on the final and intermediate goods markets. This equation also controls the sectoral characteristics and the local characteristics both independently and in their interaction. The i wage equations (10) suggest the existence of endogeneity between the number of firms and the level of wages.

3 Data and econometric issues

Estimation of these kind of structural models is very difficult because they yield a general equilibrium in a spatial economy with non-linear equations. The gap between the structural parameters and the parameters actually estimated depends on two main factors: *the analytical complexity of the structural form* and *the availability of data* (Keane and Wolpin, 1997). We allow for these difficulties in two complementary ways. First, we adopt a form that can be estimated econometrically while keeping as close as possible to the structural model. As will be seen, this form is determined by the possibility of finding analytical solutions to complex forms and by the availability of data providing proxies of the model primitives. Second, we have attempted to use the most appropriate methods of estimation. In order to estimate this reduced form of the previous model, we use panel data crossing economic activities with spatial areas. As a certain number of econometric problems arise, we use several complementary estimators, including the usual panel data estimators as well as the instrumental variable estimator.

3.1 Data used

We used the French 'Labor Market Areas' (LMAs or '*zones d'emploi*') which are at the intra-regional level. This economic rather than administrative delineation was made by the *INSEE* (French National Institute for Statistics and Economic Studies) using the commuting patterns between communes (French municipalities) as recorded in the 1990 population census. France was then divided into 341 labor market areas, the relatively small mean size (1570 km2) of which conceals wide disparity (standard deviation: 987 km2).

Similarly, the choice of a sectoral level of analysis has led to a fine division. The *NES114* level was chosen, which divides the French economy into 114 sectors corresponding roughly to the ISIC three-digit level. The nature of the model implies that we have focused only on private manufacturing sectors and business services. We excluded agriculture, wholesale and retail trade, consumer services and the public sector. Some 67 sectors were therefore included in the analysis.[4] The resulting database is therefore composed of the conjunction of the chosen spatial scale (341 labor market areas) and the chosen sectoral division (67 sectors), forming what might be called a 'spatial panel'. The basic data are assembled from a specific French database the *Déclarations annuelles de données sociales* (*DADS*) or Annual Corporate Data Returns. Each private firm with at least one employee must report the main characteristics of each of its employees annually, in terms of skill level, length of employment contract, earnings, and so on. The 67 sectors selected cover 412,320 plants and a volume of work of 5,916 million full-time equivalents, or 77% of the industrial employment in France.

At this level of analysis, available French data cannot include price data, meaning that the term $\Gamma_{r,r'}$ in the location equation (9), which combines items of the environment outside the area (notably relative prices) with the size of the sector and of the area, cannot be completed. In order not to omit these components completely from the analysis, it is possible to take

[4] The sectors and their main characteristics are listed in Appendix C. For technical reasons, due to the common practice of grouping the returns for the different plants under the head office, a few industrial sectors could not be included, most notably for gas and electricity supply, construction, post and telecommunications, and financial intermediation, among others.

account of the sizes of the area and of the sector in the left-hand side of the equation, by weighting the variable n_r^s by the size of the sector (n^s) and by the relative size of the region (n_r/n), n being the total mass of firms in the economy and n_r the total mass of firms in the region. Our dependent variable will then no longer be an absolute number of firms but will become an index of local specialization. Therefore, *we will seek to explain the under- or over-representation of a sector within a region.*

In addition, there is no available indicator to capture α^s the fixed cost effect included in the term Λ^s in the location equation (9). The mean plant size could be used as a proxy as, when fixed costs are high, plant size tends to grow in order to generate internal scale returns. This is the reason why, in addition to the previous employment index, we use an index of plant location. Our dependent variables will therefore be an index of the local specialization of plants and an index of the local specialization of jobs.

Using the information available in our database for 1997, we first calculated, as follows, both local specialization indexes (one for plants, Zn and one for jobs, ZL), to make up our explained variables:

$$Z_r^s \begin{cases} Zn_r^s = \log \frac{n_r^s/n^s}{n_r/n} \\ ZL_r^s = \log \frac{L_r^s/L^s}{L_r/L} \end{cases}$$

where n_r^s (L_r^s) is the number of plants (jobs) in industry s and region r, n^s (L^s) is the number of plants (jobs) of industry s in the whole country, n_r (L_r) is the number of plants (jobs) in region r, and n (L) is the number of plants (jobs) in the whole country. So, if $Z_r^s > 1$, then region r is considered to be specialized in activity s.

Second, the same database has enabled us to construct the variable capturing the spatial differentiation of labor costs $\Sigma_i \left(\phi_s^i \log w_r^i\right)$. To do this, we identified six categories of workers by skill level : senior managers, middle managers, intermediate categories (such as technicians, supervisors, etc.), clerical workers, skilled manual workers, and unskilled manual workers. For each category, as suggested by equation (9), we weighted the logarithm of local wages, $\log w_r^i$, by L_s^i/L_s, the share of category i in the total employment in the sector, L_s. The variable for the local labor cost is then obtained by summing the results for each of the six categories:

$$LC_r^s = \sum_{i=1}^{6} \left(\frac{L_s^i}{L_s} \log \left(w_r^i\right) \right)$$

Third, the 'Local Demand' variable ($\log(\gamma^s + \Psi^s) w_r L_r$) involves local demand for final goods γ^s and for intermediate goods Ψ^s. As γ^s is available for the *NES114* level and Ψ^s for the *NES16* level (a level between one-digit and two-digit ISIC), we had to either aggregate γ^s at the *NES16* level to obtain a homogenous variable or to disregard the Ψ^s variable to maintain maximum sectoral disaggregation. The second solution is preferred to the first. The 'Local Demand' variable thus becomes a Local Final Demand variable. To calculate it we used the taxable income of local households (Inc_r) from the taxable income records of the *Direction générale des Impôts* (Revenue Service) aggregated at the LMA level. This local income was weighted by the consumer's budget coefficient γ^s, which is available from the Annual Series on French Household Consumption.[5] The variable for final local demand made on sector s is thus obtained:

$$LFD_r^s = \log \gamma^s Inc_r$$

Fourth, the construction of a variable capturing the effects of local vertical linkages was more difficult. Here we have no local prices for intermediate consumption, so that the term $\Sigma_{s'} \mu^{s',s} \log \varphi_r^{s'}$ featured in Ω_r^s (9) cannot be calculated. Conversely, the term $\Sigma_{s'} \mu^{s',s} \log n_r^{s'}$ can be evaluated by using *DADS* data to capture the local occurrence of firms or jobs in sectors that are vertically linked with the sector in question. However, this does not address the question of the availability of technical coefficients at the desired sectoral and geographic level. These are only available at the national level in the *NES16* classification (that is, the nine sectors of interest to us here). Therefore, we calculated an intensity indicator for local vertical linkages at this level. We weighted the local employment coefficients in all the other sectors ($L_r^{s''}$) by the technical coefficient of the intersectoral relation in question recorded in the national input-output matrix ($\mu_{ID}^{s,s''}$ for demand for intermediate goods made on sector s, and $\mu_{IS}^{s,s''}$ for the supply of intermediate goods from sector s). The results of this calculation were then applied at the *NES114* level and the indicators of backward and forward linkages were summed as a single indicator. The local vertical linkages indicator (*LIO*) is then obtained from:

[5] Obviously, some sectors have no direct relationships to the final consumers. There are eight sectors, where the final demand variable equals zero: metal components of construction, boiler-making, mechanical equipment, engines & transformers, casting of metals, metal working services activities, firm administration, research & development.

$$LIO = \sum_{s''=1}^{9} \left(\left(\mu_{ID}^{s,s''} + \mu_{IS}^{s,s''} \right) \log(L_r^{s''}) \right)$$

Finally, calculating the sectoral characteristics featured in the term Λ^s of the equation (9), $\frac{1-B^s}{B^s} \log \left(P^s(\sigma^s - 1) \right)$ and $\frac{1}{B^s} \log \sigma^s$, requires us to evaluate the intensity of labor B^s, the level of competition σ^s, and the productivity of inputs $P^s = 1/\beta$. Using national sectoral information from the INSEE, we chose to represent B^s by the inverse of capital intensity, P^s by labor productivity, and σ^s by the inverse of a concentration index evaluated as the weight of the top ten firms in the sector's added value. Thus, the 'productivity effect' (PE^s) could be calculated as:

$$PE^s = \frac{1 - B^s}{B^s} \log(P^s \left(\sigma^s - 1 \right))$$

and the 'competition effect' (CE^s) as :

$$CE^s = \frac{1}{B^s} \log \sigma^s$$

3.2 Econometrics

In relation to our previous choices, we obtain the following econometric model:

$$Z_r^s = \delta_0 + \delta_{LC}.LC_r^s + \delta_{FD}.LFD_r^s + \delta_{IO}.LIO_r^s + \delta_{PE}.PE^s + \delta_{CE}.CE^s + \varepsilon_r^s \quad (11)$$

where δ_{LC}, δ_{FD}, δ_{IO}, δ_{PE}, δ_{CE} are the coefficients to be estimated, δ_0 is a constant and ε_r^s is the error term. Appendix D gives some descriptive statistics on the variables of equation (11): the means and standard deviations of the dependent and explanatory variables of the region/sector couples and the correlation matrix between variables. *There is no major problem of colinearity between explanatory variables.* The highest correlation coefficient is 0.16 and concerns the Local Vertical Linkages and Local Final Demand variables. Finally, as expected, the two specialization indexes are closely correlated with one another.

However, estimating the econometric model (11) raises two sorts of problems. The first relates to the nature of our data, which reveals the heterogeneity of both sector and region. The second is related to the possible endogeneity of the labor costs as suggested by the structural model. We present both econometric problems separately, but we correct both in the same regression using the Breusch-Mizon-Schmidt (1989) method.

1. The occurrence of sectoral variables in the reduced form may raise hopes that any sector effects will be brought under control. However, that does not prevent the need to resort to panel data econometrics (Matyas and Sevestre, 1996). The sector dimension is then likened to individuals and the geographical dimension to time and a two-component error model is evaluated, such that:

$$\varepsilon_r^s = \kappa^s + \nu_r^s$$

where ν_r^s is a random error and κ^s is a sector-specific component. Note that κ^s captures the sectoral characteristics of firms in each region that are unobservable or omitted from the equation, but do not vary over regions. In our case, the *Hausman test* applied to the choice between both models leads us to opt for the fixed-effect model, that is, the *within* estimator.

2. The second problem is that of the endogeneity of the local labor cost variable, which may introduce a new source of correlation between this variable and the error term. In fact, equation (10) for labor market equilibrium clearly shows that local wages are directly related to the number of firms established locally. Our local labor cost variable is based explicitly on local wages and our dependent variables roughly include the number of local firms. Thus, endogeneity is suggested directly by the theoretical model. This problem requires the use of the instrumental variable method (IV) to obtain convergent estimates. The choice of instruments other than the exogenous variables is driven by the equations (10) describing equilibrium on the labor market. Thus, we need to capture the local skill structure and the region's level of wealth. A set of instruments to capture the first variable can be built in using the 1990 population census. A second set could come from the 1991 DADS data, that is, from the same database that we used, but for an earlier date.[6] An exogeneity test based on the augmented regression method is conducted to check if there is a need for correction of endogeneity. The choice of instruments was validated by the *Sargan test* and by the regression of the

[6] In addition to the exogenous variables and their 'spatial lags' (obtained by weighting them with a spatial weight matrix based on the inverse of the distance between the centroides of the labor market areas), the instruments used in the different estimations are: local stock of 'managerial' employees in 1990, the unemployment rate within the labor market area in 1992, the average wage per worker in the local labor area observed in 1991 and the share of tax-exempt households within the total households in 1990.

residuals of the augmented regression on the instruments (Robin, 1999) for checking whether our instruments are not correlated with the error terms.

4 Results

The results of our estimations are presented in two stages. First, we estimate the model with all of the available data, that is, the 'spatial panel' made up of the 67 sectors and the 341 labor market areas. Then, the differences between sectors are investigated by performing 67 different estimations, one for each industry.

4.1 Estimation for the spatial panel

Table 1 shows the results of several estimations of the econometric model (11). The estimations were performed with both endogenous variables: the local specialization indexes in jobs and in firms. Four different estimators were used: the OLS estimator, the instrument variable estimator (IV), the Within estimator for controlling the sectoral fixed effects and a fourth (Within-IV) obtained in applying the instrument variable method to the Within estimator in order to control both heterogeneity between sectors and endogeneity of the labor cost. The exogeneity tests lead us to conclude that the local labor cost variable lacks exogeneity. Therefore, we use instrumental variable method. The tests of validity of our selected instruments confirm they are valid instruments.

Insert Table 1 about here.

The results reported in Table 1 show the clear convergence between the results obtained with the local employment specialization index and those obtained with the local firm specialization index. They highlight the importance of the local vertical linkages and of the local final demand in the industrial location process. Whichever estimator is used, the coefficient of the local intensity indicator for input-output relations of the sector (LIO) is always highly significant and positive. The estimates of the local final demand (LFD) are often significant and positive: with the OLS and the IV estimators, final local demand is highly significant; with the Within-IV estimator it is poorly significant; and with the Within estimator it is not significant. Because the LFD and LIO variables have similar dimensions (see descriptive statistics in Appendix D), we can compare both estimates. While the parameter value is clearly lower than that of the LIO parameter with the OLS and IV estimators, both parameter values become very close when we control the sectoral fixed effects and allow for the labor cost endogeneity. Thus, at the geographical level involved here (that is, at an intra-regional scale), firms tend to agglomerate mainly in order to reduce the costs of their intermediate consumption or access to their industrial customers, but also to be close to the final demand. The lower impact of the final demand suggests that the market areas for final goods and services overlap with the market areas of intermediate goods. Therefore, to capture the final demand effect requires working on a higher geographical scale.[7]

The results are a little different with regard to the local labor cost. The parameters obtained with the different estimators seem to be somewhat contradictory. Focusing on the OLS and IV estimators leads to the conclusion that the effect of the labor cost on local economic specialization is significantly positive, which is not consistent with the expectation based on the theoretical model. However, this effect becomes significantly negative, as expected, when we control the sectoral fixed effects and allow for the labor cost endogeneity (Within-IV estimator). Note that the set of instruments is composed of regional variables, so that we control for some geographical effects, such as the existence of geographic differentials in productivity, which may offset the geographic differentials in wages. Thus, for the spatial panel analyzed as a whole, the dispersion force of activities exerted by the labor cost differentials is masked by the variability in wage costs between sectors. It appears only when the major distinctive features of the sectors are taken into account. The differences between the estimation results obtained for this variable (LC_r^s) suggest that the spatial behavior of the firms depends

[7]Some additional estimations were made in order to capture also the role of the surrounding market areas. The final demand and the vertical linkages variables were recalculated as Market Potentials (see Head and Mayer, 2002) by adding the values of the nearby LMAs, weighted by a 'spatial lag', to the LFD and LIO variables . Thus we constructed our new explanatory variables as $(W_d + I) * LIO$ and $(W_d + I) * LFD$ where I is the identity matrix and W_d is a spatial weight matrix, which links each labor market area to the others in relation to the inverse of the distance between their centroides. The elements of this matrix W_d are as $a_{ij} = (1/d_{ij})$. The results of these estimations are close to those obtained with the local values only. The impact of the final demand could be lower than we have seen before and may even be negative.

on the sector to which they belong. This point requires further investigation, which we will undertake in the sector-by-sector estimations.

The sectoral characteristics (PE^s and CE^s) included in (11), which are introduced only when we do not control the sectoral fixed effects, play significant roles in the location of industrial firms. As expected, the parameter of the variable that we have interpreted as capturing a 'competition effect' (CE^s) is significantly negative, while that of the 'productivity effect' (PE^s) is significantly positive. Thus, the sectors where there is fierce competition between firms tend to be geographically dispersed, whereas those sectors with high productivity levels tend to be geographically concentrated. We find here the usual theoretical conclusions of economic geography, but these effects are now measured within a multi-sectoral economy rather than with a static comparative approach. These two sector-based characteristics, however, do not seem to be the only ones involved here. This is what is suggested by the sign change in the LC parameter when allowing for fixed sectoral effects. Indeed, the local labor cost, which seems to be an agglomeration force for firms when only these two sector-based variables are taken into account, becomes a dispersion force after all of the sector-based characteristics (whether unobservable or omitted) are introduced.

4.2 Sector-by-sector estimations

In the previous estimations we looked for a mean effect of explanatory variables on firm location. We now determine whether that mean effect hides heterogeneity of behavior between sectors. To do this, and go beyond fixed sectoral effects, we consider heterogeneity of slopes and therefore of parameter coefficients by sector.[8] We estimate the model (11) sector by sector. Thus, we omit the two sector variables CE^s and PE^s. Because of the similarity between the results obtained with the specialization indexes for employment and those for firms, we confine the analysis to the employment specialization index.

[8]In addition to the test of fixed effects significance, we performed tests of slope and intercept homogeneity, based on Hsiao (1986). In our case, the homogeneity test of both slopes and intercepts consists of comparison between the results obtained by the OLS estimator for the whole spatial panel and those obtained in estimating the model for each sector. The homogeneity test of the sectoral slopes compares the Within estimators with the OLS estimators for each sector. Both tests lead us to conclude that there is heterogeneity of both intercepts and slopes.

As before, such estimates are likely to suffer from endogeneity bias caused by the local labor cost variable. Again, this can be corrected by using the instrumental variables method. This has been implemented for sectors where the result of the exogeneity test suggests that such a problem arises. Some 34 out of the 67 sectors are involved. Several sets of instruments are used, the choice being guided by the results of the Sargan test for validity of instruments.[9]

The sector-by-sector estimates are set out in detail in Appendix E. We can identify whether the exogeneity test leads toward the OLS estimator or VI estimator. An analysis of the homogeneity or heterogeneity of location behavior across sectors is addressed in Table 2, which summarizes the number of sectors by the sign of their estimated parameters (positive, negative or not significant). The table also compares the uncorrected results for endogeneity (*OLS*) with the results corrected, when necessary, for labor cost endogeneity (33 sectors in *OLS* and 34 sectors in *IV*).

Insert Table 2 about here

Local labor cost

As our previous results suggested, the effect of the labor cost varies from one sector to another. The results obtained in *OLS* show that about half of the industrial sectors (37 out of 67) react to labor cost differentials as if they were a dispersion force, whereas for 29 sectors this variable does not seem to be involved in the location process. First, it should be observed that allowance for labor cost endogeneity (where observed) affects the distribution of sectors according to the impact of local labor costs. The switch from *OLS* to *IV* reduces the number of sectors with negative coefficients (from 37 to 24), and slightly increases the number of sectors that are 'attracted' by high labor costs (from 3 to 16). Therefore, control of the agglomeration feedback effect on the local labor cost reduces the effect of the dispersion force exerted by local tensions related to adjustments between local labor supply and demand, as suggested by the model.

Nonetheless, a significant number of sectors remain sensitive to this dispersion force. Among them are the food industry sectors (except for the beverages industry), many sectors of the consumer goods industry (wearing

[9]See Appendix E for the different sets of instruments used, when necessary.

apparel, shoes and leather, printing and publishing, motor vehicles, sporting equipment and toys) as well as a few sectors in the equipment goods industry (shipbuilding, metal components for construction and agricultural machinery) and in the intermediate goods industry (glass products, fabrics, wood products and metalworking industrial services). As shown in the table in Appendix C , these are activities where the payroll is a large proportion of the value added, where labor is comparatively unskilled and/or where the number of employees per firm is comparatively small (a sign of fiercer competition on the goods markets in these sectors). The reduction of wage bills is therefore an important strategy for these sectors. Their dispersion may serve this end even if it reduces the gain provided by agglomeration economies.

These sector-based characteristics do not apply to the sectors less sensitive to spatial differentiation of labor costs. This is the case for activities such as equipment and intermediate goods and the business service sectors. Some other activities' estimates of the LC variable are even positive, which means that firms tend to locate where labor cost is high. The occurrence of such coefficients in the chemical industry, aeronautical construction and Research & Development could be explained by some of the characteristics of these sectors. The reduction of labor costs is not a major goal in itself when wages do not represent the greatest part of costs. The search for skills could lead such firms to locate where this resource is available, which often means a place where wage levels are high.

Local vertical linkages

Relationships linking firms engaged in trading inputs and outputs always appear, at this scale of analysis, as a powerful agglomeration force, whatever the industrial sector concerned. Thus the coefficient of the LIO variable is significantly positive in 65 of the 67 sectors in OLS and in 59 of them when correcting for any endogeneity bias of labor cost. These results confirm the result obtained in the previous section.

Local final demand

The homogeneous behavior observed with vertical linkages is no longer apparent if we look at the impact of final demand. Unlike the results obtained with the whole 'spatial' panel, local final demand has no effect on the location decisions in many sectors. The estimated LFD coefficient is not significantly different from zero in most instances. This is the case in some 28 sectors in OLS, and the 8 sectors where final demand actually equals zero (see note 5 in

Section 3.1). When we allow for the labor cost endogeneity (when necessary), this figure varies a little. In this case, the corresponding results are 31 and 8. In other words, the final demand does not seem to affect the location process of 60% of the sectors.

However, the sectors where local final demand does have an effect are divided almost equally between those with a positive effect (which was expected) and those with a negative (more unexpected) effect.[10] It should be noted that the 16 sectors (15 in *OLS*) where location of final demand has an apparent dispersion effect include the food industries, intermediate goods industries (such as iron and steel, textiles, water treatment, and rubber and plastics) and transport (rail or road). Here, we find: (i) sectors closely related to a resource that is widely scattered throughout the country (and for which the 'repulsive' character of final demand is only an appearance); (ii) sectors, like transportation, which aim to cover each part of the country and which, therefore, can be expected to be found even in sparsely populated areas; and (iii) polluting or dangerous industries.

Thus, there are few sectors where final demand has the expected effect of an agglomeration force. It even tends to fall when the labor cost endogeneity bias is corrected: it declines from 16 sectors in *OLS* to 12 in *OLS* or *IV*. This includes a few consumer goods industries and some equipment and intermediate goods industries. Therefore, our results suggest that firms are not very sensitive to the precise location of final demand (that is, within an LMA). Recall that we have put aside the wholesale and retail trade sectors as well as the personal services ones, in order to focus on the manufacturing sectors and their services. Omitting the sectors directly related to the final demand obviously reduces the number of sectors involved in this dimension.

5 Summary and concluding remarks

In this paper, we have investigated some factors that explain the spatial distribution of activities as highlighted by economic geography. Such factors include vertical linkages, access to final goods markets and spatial tensions between the labor demand structure and the local skill structure. An econometric model, which was based on an economic geographic model and includes these three factors, was estimated using French data for 1997 aggregated by sector (67 sectors were analyzed) and by a fine delineation of

[10] Davis and Weinstein (1999) obtain a similar result for some sectors.

France (into 341 labor market areas). Several econometric methods were used to take into account the nature of the data (*i.e.* the spatial panel) and the endogeneity of one explanatory variable (*i.e.* the local labor costs). The results obtained for individual sectors are consistent with, and can refine, those obtained by estimating the complete panel. Our results reveal the strong effect of vertical linkages and the weaker role of the size of local final demand at this geographical scale. The impact of local labor costs seems to vary from one sector to another, inducing dispersion of firms in some industries (especially in competitive and/or labor-intensive ones) but exerting no effect on other industries. Therefore, firms in a large array of sectors are faced with a choice between agglomerating to take advantage of increasing returns from input-output linkages and dispersing to take advantage of lower labor costs.

This work may be pursued in several ways. First, we could improve the way the variable measuring the intensity of vertical linkages was constructed. Input-output matrixes at more disaggregated sector-based levels or at finer geographical scales could bring about such an improvement. Second, as we used spatial data, some spatial autocorrelation between error terms could occur. We perform some tests that do not reject this assumption. Thus, we would have to use estimators that take account of either spatial autoregression on the dependant variable or spatial autocorrelation between error terms related to unobservable or omitted characteristics. Finally, the assumption of the geographical immobility of the labor force needs further investigation. Driven by economic geography models, our approach needed to include such a hypothesis to highlight spatial tensions on the labor market. One way to relax this assumption would be to consider an economy with two categories of workers: *(i)* geographically mobile and skilled workers; and *(ii)* geographically immobile and unskilled workers.

References

[1] Benassy J.-P. (1991). Monopolistic competition. In Hildenbrand W. and H. Sonnenschein (eds), *Handbook of Mathematical Economics*, Vol. IV, 1997-2045.

[2] Brakman S., H. Garretsen and M. Schramm (2002). The spatial distribution of wages and employment: testing the Helpman-Hanson model

for Germany. *Mimeo.*

[3] Breusch T., Mizon G., and P. Schmidt (1989). Efficient estimation using panel data. *Econometrica* 57, 695-700.

[4] Combes P.-Ph. and M. Lafourcade (2001). Transportation costs decline and regional employment inequalities: Evidence from France, 1978-1993. *Working Paper CERAS* 01-01, Paris.

[5] Combes P.-Ph. and H. Overman (2003). The spatial distribution of economic activities in the European Union, in Henderson J.-V. and J. Thisse (eds), *Handbook of Regional and Urban Economics* (forthcoming).

[6] Davis D. and D. Weinstein (1998). Market access, economic geography and comparative advantages: an empirical assessment. *NBER Working Paper* N 6787.

[7] Davis D. and D. Weinstein (1999). Economic geography and regional production structure: an empirical investigation. *European Economic Review* 43, 379-407.

[8] Dixit A. and J. Stiglitz (1977). Monopolistic competition and optimum product diversity. *American Economic Review* 67, 297-308.

[9] Faini R. (1999). European migrants: an endangered species? In Baldwin R., D. Cohen, A. Sapir and A. Venables (eds), *Market Integration, Regionalism and the Global Economy.* Cambridge (MA): Cambridge University Press, 228-253.

[10] Fujita M., P. Krugman and A. Venables (1999). *The spatial economy. Cities, regions and international trade.* Cambridge (MA): MIT Press.

[11] Fujita M. and J.-F. Thisse (2002). *Economics of Agglomeration. Cities, industrial location and regional growth.* Cambridge (MA): Cambridge University Press.

[12] Hanson G. (1999). Market potential, increasing returns, and geographic concentration. *Discussion Paper* 439, University of Michigan.

[13] Head K. and T. Mayer (2002). Market potential and the location of Japanese investment in the European Union. *CEPR Working Paper 3455.*

[14] Head K. and T. Mayer (2003). The empirics of trade and agglomeration, in Henderson J.-V. and J. Thisse (eds), *Handbook of Regional and Urban Economics* (forthcoming).

[15] Hsiao C. (1986). *Analysis of Panel Data*, Cambridge (MA): Cambridge University Press.

[16] Keane M. and K. Wolpin (1997). Introduction to the JBES special issue on structural estimation in applied microeconomics. *Journal of Business and Economic Statistics* 15, 111-114.

[17] Krugman P. (1991). Increasing returns and economic geography. *Journal of Political Economy* 99, 483-499.

[18] Krugman P. and A. Venables (1995). Globalization and the inequality of nations. *Quarterly Journal of Economics* 110, 857-880.

[19] Matyas L. and P. Sevestre (1996). *The econometrics of panel data*. Dordrecht-Boston-London: Kluwer Academic Publishers.

[20] Midelfart-Knarvit K., H. Overman and A. Venables (2000). Comparative advantage and economic geography: estimating the location of production in the EU. *CEPR Discussion Paper*, 2618.

[21] Mion G. (2002). *Spatial externalities and empirical analysis: the case of Italy*, processed CORE-UCL.

[22] Overman H., S. Redding and A. Venables (2001). Trade and geography: a survey of empirics. *Mimeo*, London School of Economics.

[23] Redding S. and A. Venables (2002). Economic geography and international inequality. *CEPR Discussion Paper* 2568.

[24] Robin J.-M. (1999). Endogénéité et variables instrumentales dans les sciences sociales. *Working Paper "Méthodologie Statistique"* 99(02), Paris: INSEE.

[25] Tabuchi T. and A. Yoshida (1999). Urban agglomeration economies in consumption and production. *Working paper*, University of Tokyo.

Appendix A: Means and standard deviations of wages by skill across French labor market areas (in Francs 1997)

Skill	Mean	Std-dev
Senior managers	290438	49784
Middle managers	286607	29723
Intermediate categories	171642	13732
Clerical workers	119568	8388
Skilled manuals	123753	10746
Unskilled manuals	106837	6925

Appendix B: Linearization of the equilibrium relation on the goods market

From the equations (4) and (7), the equilibrium relation on the goods market is:

$$\begin{aligned}\frac{(\sigma^s-1)\alpha^s}{\beta^s} &= \left(E_r\left(Pm_r^s\right)^{\sigma^s-1}+(\tau^s)^{1-\sigma^s}(R-1)D^s\right)\left(p_r^s\right)^{-\sigma^s} \quad \text{(B.1)}\\ \text{with } D^s &= \frac{\sum_{r'\neq r}E_{r'}\left(Pm_{r'}^s\right)^{\sigma^s-1}}{R-1}\end{aligned}$$

where D^s stands for the mean of regional expenditures in goods s weighted by the regional price index in sector s. Let $E_r\left(Pm_r^s\right)^{\sigma^s-1}\left(p_r^s\right)^{-\sigma^s}=\varphi_D^{s,r}D^s\left(p_r^s\right)^{-\sigma^s}$ where $\varphi_D^{s,r}$ measures the relative deviation from the mean of demand in region r as an aggregate of the goods s. Thus, (12) becomes:

$$\frac{(\sigma^s-1)\alpha^s}{\beta^s}=E_r^s\left(Pm_r^s\right)^{\sigma^s-1}\left(1+\frac{(\tau^s)^{1-\sigma^s}(R-1)}{\varphi_D^{s,r}}\right)\left(p_r^s\right)^{-\sigma^s} \quad \text{(B.2)}$$

Factorize now the expenditures in goods s made in region r (E_r^s) as well as the index price in region r and sector s (Pm_r^s). Consider first the price index. Let's suppose $n_r^s\left(p_r^s\right)^{1-\sigma^s}=\varphi_p^{r,s}\sum_{r'\neq r}(n_r^s\left(p_r^s\right)^{1-\sigma^s})/(R-1)$ where $\varphi_p^{r,s}$ is the relative deviation from the mean of the product between the number and the price of goods s in region r. We rewritten the price index:

$$Pm_r^s=\left(n_r^s\left(p_r^s\right)^{1-\sigma^s}\left(1+\frac{(\tau^s)^{1-\sigma^s}(R-1)}{\varphi_p^{r,s}}\right)\right)^{1/(1-\sigma^s)} \quad \text{(B.3)}$$

Consider now E_r^s. In summing by labor quality (i), the equation (8) can be written as $n_r^s p_r^s q_r^s=w_r^s L_r^s/\sum^I\phi_s^i$. And E_r^s becomes:

$$E_r^s=\left(\gamma^s w_r L_r+\sum_{s'=1}^{S}\mu^{s',s}n_r^{s'}p_r^{s'}q_r^{s'}\right)=w_r L_r\left(\gamma^s+\Psi^s\right) \quad \text{(B.4)}$$

where $\Psi^s=\sum^{s'}\mu^{s',s}B^{s'}\delta_r^{s'}$ with $\delta_r^{s'}=w_r^{s'}L_r^{s'}/w_r L_r$. In substituting Pm_r^s and E_r^s by expressions (B.3) and (B.4) in equation (B.2), we obtain:

$$\frac{(\sigma^s-1)\alpha^s}{\beta^s}=\left(n_r^s\right)^{-1}\left(p_r^s\right)^{-1}w_r L_r\left(\gamma^s+\Psi^s\right)\left(1+\frac{(\tau^s)^{1-\sigma^s}(R-1)}{\varphi_D^{r,s}}\right)\left(1+\frac{(\tau^s)^{1-\sigma^s}(R-1)}{\varphi_p^{r,s}}\right)^{-1}$$

Under its logarithmic form, the equilibrium relation on the goods market is thus:

$$\log(n_r^s) = -\log\left(\frac{(\sigma^s-1)\alpha^s}{\beta^s}\right) - \log\left(p_r^s\right) + \log\left[\left(\gamma^s + \Psi^s\right) w_r L_r\right] \qquad (B.5)$$
$$+\log\left(1+\frac{(\tau^s)^{1-\sigma^s}(R-1)}{\varphi_D^{r,s}}\right) - \log\left(1+\frac{(\tau^s)^{1-\sigma^s}(R-1)}{\varphi_P^{r,s}}\right)$$

Turn now to the price equation (3) and write it under its logarithmic form:

$$\log(p_r^s) = \log\left(\frac{\sigma^s}{\sigma^s-1}\beta^s\right) + \sum_{i=1}^{I}\phi_s^i \log\left(w_r^i\right) + \sum_{s'=1}^{S}\mu^{s',s}\log\left(Pm_r^{s'}\right)$$

In substituting price indexes by equation (B.3), we obtain:

$$\log(p_r^s) = \log\left(\frac{\sigma^s}{\sigma^s-1}\beta^s\right) + \sum_{i=1}^{I}\phi_s^i \log\left(w_r^i\right) +$$
$$\sum_{s'=1}^{S}\left(\frac{\mu^{s',s}}{1-\sigma^s}\log\left(n_r^{s'}\right) + \mu^{s',s}\log\left(p_r^{s'}\right) + \frac{\mu^{s',s}}{1-\sigma^{s'}}\log(1+\frac{(\tau^{s'})^{1-\sigma^{s'}}(R-1)}{\varphi_P^{r,s'}})\right)$$

Let $p_r^{s'} = \varphi_r^{s'} p_r^s$ where $\varphi_r^{s'}$ is the regional price ratio between goods $s\prime$ and s, the previous equation becomes:

$$\left(1-\sum_{s'=1}^{S}\mu^{s',s}\right)\log(p_r^s) = \log\left(\frac{\sigma^s\beta^s}{\sigma^s-1}\right) + \sum_{i=1}^{I}\phi_s^i\log\left(w_r^i\right) + \sum_{s'=1}^{S}\left(\frac{\mu^{s',s}}{1-\sigma^s}\log\left(n_r^s\right)\right) \qquad (B.6)$$
$$+\sum_{s'=1}^{S}\mu^{s',s}\log\left(\varphi_r^{s'}\right) + \sum_{s'=1}^{S}\left(\frac{\mu^{s',s}}{1-\sigma^s}\log(1+\frac{(\tau^{s'})^{1-\sigma^{s'}}(R-1)}{\varphi_P^{r,s'}})\right)$$

In substituting this expression of p_r^s in equation (12), we obtain the final expression which explains the location of each sector s in the different regions r (see equation 9) where

$$\Gamma_{r,r'} \equiv \log\left(1+\frac{(\tau^s)^{1-\sigma^s}(R-1)}{\varphi_D^{r,s}}\right) - \log\left(1+\frac{(\tau^s)^{1-\sigma^s}(R-1)}{\varphi_P^{r,s}}\right) \qquad (B.7)$$
$$+\frac{1}{C^s}\sum_{s'=1}^{S}\left(\mu^{s',s}\log(1+\frac{(\tau^{s'})^{1-\sigma^{s'}}(R-1)}{\varphi_P^{r,s'}})\right)$$

where $\varphi_D^{r,s}$ measures the relative deviation from the mean of demand in region r as an aggregate of industrial goods s and $\varphi_P^{r,s}$ the relative deviation from the mean of the product between the number and the price of goods of sector s in region r.

Appendix C: Main characteristics of the 67 selected sectors

Table C.1: Characteristics of sectors

	CI	LP	PE	K/L	SkW
Food products and beverages					
Meat	16.4	218	76.3	279	13.9
Dairy products	31.8	327	65.9	597	22.8
Beverages & alcohol	36.0	582	46.8	1557	32.0
Grain mill products	28.8	399	62.2	992	32.8
Others food products	21.2	288	61.6	395	16.0
Manufacture of consumption goods					
Wearing apparel & dressing	13.8	188	83.0	106	18.8
Shoes & leather	30.1	206	76.3	142	15.8
Printing, press & publishing	12.9	359	73.6	247	39.8
Pharmaceutical Industry	26.6	599	57.6	601	58.9
Soaps, perfumes & cosmetics	37.6	455	63.7	467	45.2
Furniture	21.0	235	77.5	220	19.0
Jewelry & musical instrum.	30.9	264	78.1	144	24.8
Sporting equipment & toys	29.4	254	73.4	304	27.7
Household equipment	79.0	287	72.7	420	28.7
Optical instruments & watches	29.8	272	78.5	230	31.4
Automotive					
Motor vehicles	87.8	435	58.6	1038	32.2
Parts and accessories	47.4	343	68.1	590	25.5
Manufacture of equipment goods					
Shipbuilding	66.4	269	77.7	260	25.3
Aircraft & spacecraft	88.4	492	69.4	597	62.4
Railway & other transport equipment	85.2	282	85.3	290	35.5
Metal components of construction	12.1	239	82.8	177	24.8
Boilermaking industry	24.7	256	85.5	170	24.7
Mechanical equipment	31.6	306	80.6	359	35.1
General purpose machinery	19.4	308	78.2	178	38.2
Agricultural machinery	36.1	293	69.4	248	24.1

CI: Sector concentration index evaluated as the weight of the top ten firms in the sector's added value (in %); LP: the labor productivity expressed in thousands of Francs; PE: the share (in %) of personal expenditures in sector's added value; K/L: the capitalistic intensity expressed in thousands of Francs; SkW: the share (in %) of skill workers in the sector's employment.

Table C.1 (continued)

	CI	LP	PE	K/L	SkW
Manufacture of equipment goods (continued)					
Tools machinery	33.0	314	80.7	258	40.0
Other special purpose machinery	33.0	328	76.0	280	35.7
Office & computing machinery	87.8	546	71.1	436	72.9
Engines & transformers	47.1	298	74.0	258	32.6
Broadcasting transmitters	77.2	402	83.1	306	65.5
Medical appliances & instruments	37.0	305	76.1	242	53.6
Measuring & checking instruments	35.3	335	87.7	188	58.8
Manufacture of intermediate goods					
Glass & glass products	52.0	331	71.2	508	20.5
Construction materials	25.5	332	68.6	676	26.0
Spinning & weaving of textiles	18.8	246	74.7	408	19.2
Textiles	20.6	248	72.7	296	21.1
Fabrics	42.6	207	78.3	224	18.9
Wood products	12.8	232	73.9	320	17.5
Paper & paper products	32.8	356	66.9	849	24.6
Basic chemicals	52.9	614	54.1	1749	46.2
Other chemical products	35.7	434	68.3	668	43.7
Rubber products	70.4	319	71.7	427	26.0
Plastics products	12.3	285	72.0	394	24.2
Basic iron & steel	66.1	397	70.9	1039	37.4
Casting of metals	30.6	246	80.2	311	18.6
Metal working service activities	8.5	263	78.3	251	20.7
Fabricated metal products	19.4	287	73.5	355	22.8
Electrical machinery & apparatus	38.9	308	79.5	372	35.6
Electronic components	58.0	344	69.9	635	39.8

Table C.1 (continued)

	CI	LP	PE	K/L	SkW
Transport					
Transport *via* railway	90.1	274	94.6	640	33.4
Road transport of travelers	27.2	278	80.1	909	15.5
Road transport of goods	5.7	229	80.6	229	10.0
Water transport	65.5	348	83.7	1410	34.4
Air transport	92.0	442	80.0	868	51.0
Supporting & auxiliary transport activities	51.1	565	41.4	613	23.5
Other transport activities	23.9	245	85.3	191	29.4
Business activities					
Computer activities	11.3	381	82.4	115	83.6
Professional services	31.5	356	84.3	100	59.4
Firms administration	38.8	353	103	963	64.8
Advertising & marketing	23.7	330	79.7	121	64.1
Architectural & engineering activities	19.4	321	91.1	220	70.3
Renting without operator	19.2	910	22.2	3702	35.6
Cleaning & security activities	9.3	186	88.5	111	21.3
Sewage, refuse disposal & sanitation	18.9	282	74.4	465	19.2
Research & Development	52.8	344	92.8	594	84.1
Collection, purification & distribution of water	79.8	382	77.5	1815	35.5
Insurance	38.3	183	25.0	1564	61.7

Appendix D: Sample statistics

Table D.1: Summary statistics for both cross-regional and cross-sectorial data

Variable	Mean	Std-dev	Mini.	Maxi.
Local Industrial Specialization (Plant)	-1.88	3.73	-9	4.20
Local Industrial Specialization (Employment)	-2.49	3.65	-9	4.86
Labor Costs	4.98	0.14	4.66	6.02
Local Final Demand	8.57	3.60	0	13.27
Local Input/Output Linkages	11.10	2.39	0.17	18.91

Table D.2:. Correlation matrix

Variable	(1)	(2)	(3)	(4)	(5)
(1) Local Industrial Special. (Plant)	1	0.94	0.03	0.01	0.20
(2) Local Industrial Special. (Empl.)	0.94	1	0.02	0.01	0.17
(3) Labor Costs	0.03	0.02	1	-0.004	0.08
(4) Local Final Demand	0.01	0.01	-0.004	1	0.16
(5) Local Input/Output Linkages	0.20	0.17	0.08	0.16	1

Appendix E: Estimation results by sector

Table E.1 presents the estimation results by sector obtained with the employment specialization index. When necessary (in case of negative result from the exogeneity test on labor cost variable), we use the instrumental variables method (noted IV in the 'Method' column). The different sets of instruments we used are the following:

[1] the exogenous variables + the local stock of 'managerial' personnel in 1990 and the share of tax-exempt households within the households for 1990.

[2] the exogenous variables + the average wage per worker in the LMA observed in 1991 and the share of tax-exempt households within the total households for 1990.

[3] the exogenous variables + the local stock of 'managerial' personnel in 1990 and the unemployment rate within the LMA in 1992

[4] the exogenous variables + the share of tax-exempt households within the local households for 1990 and the spatial lag of the local labor cost variable.

[5] the exogenous variables + the share of of tax-exempt households within the total households for 1990 and the unemployment rate within the LMA in 1992.

Table E.1: Estimates parameters by sector (specialization index in employment)

Sectors	Method	Intercept	LC	LIO	LFD	r^2	Sargan
Food products and beverages							
Meat	IV[2]	73.5***	-16.9***	0.22***	0.5	0.36	0.09
Dairy products	OLS	73.9***	-7.0*	0.97***	-4.1**	0.10	
Beverages & alcohol	IV[1]	-0.22	10.7	1.0***	-5.3*	0.08	0.02
Grain mill products	IV[2]	95.9***	-31.6***	0.86***	5.0**	0.16	0.21
Others food products	OLS	29.2***	-3.9***	0.08***	-0.8***	0.21	
Manufacture of consumption goods							
Wearing apparel & dressing	OLS	76.7***	-8.0***	0.57***	-3.7***	0.13	
Shoes & leather	OLS	69.1***	-12.1***	0.75***	-1.8	0.09	
Printing, press & publishing	OLS	-5.3	-3.0***	0.11***	1.6***	0.12	
Pharmaceutical industry	OLS	-19.1	-5.2	1.21***	2.3	0.23	
Soaps, perfumes & cosmetics	OLS	-27.8	-7.5**	0.81***	4.5**	0.14	
Furniture	IV[3]	-63.4**	51.3***	-0.10	-17.2***	0.10	1.70
Jewelry & musical instrum.	IV[1]	-48.7**	7.5	0.54**	0.2	0.08	0.42
Sporting equipment & toys	OLS	2.4	-7.2**	0.45***	2.4	0.08	
Household equipment	OLS	-9.0	1.5	1.18***	-1.4	0.21	
Optical instruments & watches	IV[1]	-74.0***	7.7	0.73***	2.2	0.17	0.55
Automotive							
Motor vehicles	IV[1]	73.1***	-16.6***	0.63***	0.0	0.40	2.29
Parts and accessories	IV[3]	-295.8***	185.8***	0.37**	-54.4***	0.53	0.04

Table E.1 (continued)

Sectors	Meth.	Interc.	LC	LIO	LFD	r^2	Sargan
Manufacture of equipment goods							
Shipbuilding	IV[1]	63.2***	-29.7***	0.78***	7.9**	0.09	0.00
Aircraft & spacecraft	IV[1]	-89.1***	16.9*	0.49***	-1.2	0.14	0.72
Railway & other transport equip.	OLS	20.9	6.4*	1.01***	-7.1***	0.16	
Metal components of construction	OLS	25.9***	-5.7***	0.30***	- - -	0.07	
Boilermaking industry	IV[5]	33.3**	8.12**	0.52***	- - -	0.11	1.78
Mechanical equipment	OLS	-19.3	0.7	1.08***	- - -	0.29	
General purpose machinery	IV[2]	-20.2	4.3	0.30***	-0.08	0.13	0.07
Agricultural machinery	IV[2]	187.1***	-44.9***	0.37***	5.1***	0.26	0.76
Tools machinery	IV[1]	-106.5***	26.8**	0.63***	-5.6	0.15	1.16
Other special purpose machinery	IV[2]	40.1**	-13.6***	0.75***	2.6	0.18	0.22
Office & computing machinery	IV[1]	-61.7***	5.8	0.99***	1.7	0.33	0.42
Engines & transformers	OLS	-12.7	-0.3	-0.98***	- - -	0.26	
Broadcasting transmitters	OLS	22.1	-14.2***	1.13***	3.7**	0.31	
Medical appliances & instruments	IV[1]	-6.5	0.9	0.36***	-1.04	0.11	0.44
Measuring & checking instruments	OLS	-17.4	-4.47*	0.69***	5.0***	0.28	
Manufacture of intermediate goods							
Glass & glass products	OLS	-9.4	-8.2**	0.70***	4.0**	0.13	
Construction materials	OLS	15.8***	0.54	0.01	1.9***	0.04	
Spinning & weaving of textiles	OLS	37.5**	-3.2	1.03***	-4.4**	0.15	
Textiles	OLS	36.5**	-5.6*	0.80***	-2.0***	0.13	
Fabrics	OLS	29.4	-8.3**	0.89***	-0.5	0.11	
Wood products	IV[4]	55.1***	-9.6***	-0.07*	-1.08	0.25	0.43
Paper & paper products	OLS	-11.9	0.3	0.63***	0.1	0.10	
Basic chemicals	OLS	-20.5	5.4	0.63***	-2.5	0.09	
Other chemical products	OLS	-26.2	-0.1	0.64***	1.6	0.11	
Rubber products	IV[1]	-28.5	15.5*	0.79	-6.4	0.14	0.01
Plastics products	IV[3]	-158.1*	69.1*	-0.52	-18.2**	0.01	0.26
Basic iron & steel	IV[1]	-141.2***	39.4***	0.52***	-9.7***	0.16	0.05
Casting of metals	OLS	-18.2	1.5	0.54***	- - -	0.05	
Metal working service activities	IV[2]	21.5***	-4.8***	0.17***	- - -	0.05	1.65
Fabricated metal products	IV[3]	-195.6	90.0	-0.27	-24.8	0.01	0.07
Electrical machinery & apparatus	IV[3]	-193.9***	94.2**	-0.17	-25.6**	0.17	0.08
Electronic components	OLS	-42.6***	1.2	0.58***	5.11**	0.15	

Table E.1 (continued)

Sectors	Method	Intercept	LC	LIO	LFD	r^2	Sargan
Transport							
Transport *via* railways	IV[1]	-24.6**	20.5***	0.31***	-8.3***	0.20	0.11
Road transport of travelers	IV[2]	0.24	5.0***	0.11***	-2.3***	0.08	0.28
Road transport of goods	OLS	20.2***	-2.5***	0.09***	-1.1***	0.19	
Water transport	OLS	3.6	-7.9**	0.68***	2.9	0.17	
Air transport	OLS	-36.0***	-1.5	0.79***	2.7**	0.36	
Supporting & auxiliary transport act.	OLS	-25.7*	-0.4	0.76***	1.7	0.31	
Other transport activities	OLS	4.1	-1.9	0.98***	1.8	0.38	
Business activities							
Computer activities	IV[2]	-23.0*	0.9	0.61***	1.1	0.39	2.51
Professional services	IV[2]	7.1*	-0.07	0.08***	-0.4	0.05	2.08
Firms administration	IV[1]	-11.6*	1.7	0.17***	- - -	0.18	1.08
Advertising & marketing	IV[4]	-22.1	6.2	0.56***	-1.9	0.35	1.42
Architectural & engineering act.	OLS	-1.9	-2.2***	0.20***	1.2***	0.36	
Renting without operator	IV[2]	-35.5**	9.9**	0.43***	-2.4	0.22	1.69
Cleaning & security activities	IV[2]	-13.9***	2.9**	0.25***	-0.4	0.46	0.08
Sewage, refuse disposal & sanitation	OLS	1.8	0.6	0.75***	-0.8	0.23	
Research & Development	IV[1]	-70.9***	11.3***	0.82***	- - -	0.32	0.65
Collect., purificat. & distribut. of water	IV[4]	42.5***	6.1	1.31***	-7.7***	0.62	0.66
Insurance	OLS	32.6***	6.5***	0.90***	-5.6	0.68	

- - -: omission of local final demand variable when $\gamma^s = 0$ (that's, for 8 sectors).

Table 1: Estimation results (dependent variable: Plant or Employment Specialization Indexes)

	OLS		Instr. Variable		Within		IV on within	
Variable	Empl	Plant	Empl	Plant	Empl	Plant	Empl	Plant
Constant	-12.2***	-12.4***	-28.5***	-33.5***				
	0.79	0.82	2.28	2.40				
Labor cost	0.50***	0.83***	3.79***	5.09***	-3.46***	-2.28***	-4.98***	-3.38***
	0.15	0.16	0.46	0.48	0.34	0.17	0.81	0.86
Vertical linkages	0.31***	0.27***	0.30***	0.25***	0.62***	0.57***	0.64***	0.58***
	0.009	0.009	0.01	0.01	0.01	0.01	0.01	0.01
Final demand	0.07***	0.07***	0.07***	0.07***	0.03	0.22	0.44*	0.64**
	0.006	0.006	0.006	0.006	0.19	0.20	0.27	0.29
Productivity effect	0.17***	0.16***	0.17***	0.19***				
	0.01	0.01	0.01	0.01				
Competitivity effect	-2.40***	-1.96***	-2.37***	-1.92***				
	0.05	0.06	0.05	0.06				
Sectorial effect					Fix	Fix	Fix	Fix
adjusted r^2	0.16	0.11	0.17	0.12	0.36	0.30	0.58	0.46
Sargan test			1.69	0.25			1.32	2.61

Table 2: Number of sectors (NES114) having the same parameter signs classified by NES16 (with OLS or Instr. Var. methods)

Variables		Food goods	Consumption goods	Equipment goods	Intermediate goods	Transport	Business activities	Total
	-	4 (5)[a]	6 (7)	6 (9)	5 (5)	2 (2)	1 (9)	24 (37)
Labor cost	ns	1 (0)	4 (5)	5 (4)	8 (11)	3 (5)	6 (2)	27 (27)
	+	0 (0)	2 (0)	4 (2)	4 (1)	2 (0)	4 (0)	16 (3)
Local	-	0 (0)	0 (0)	1 (0)	1 (1)	0 (0)	0 (0)	2 (1)
I/O	ns	0 (0)	1 (0)	0 (0)	5 (1)	0 (0)	0 (0)	6 (1)
Linkages	+	5 (5)	11 (12)	14 (15)	11 (15)	7 (7)	11 (11)	59 (65)
Local	-	3 (3)	3 (3)	1 (2)	5 (4)	3 (2)	1 (1)	16 (15)
Final	ns	1 (2)	7 (5)	6 (5)	7 (9)	3 (4)	7 (3)	31 (28)
Demand	+	1 (0)	2 (4)	4 (4)	3 (2)	1 (1)	1 (5)	12 (16)
	none	0 (0)	0 (0)	4 (4)	2 (2)	0 (0)	2 (2)	8 (8)
Total		5	12	15	17	7	11	67

[a]: The number in brackets corresponds to the results obtained in OLS. The first one corresponds to this obtained in OLS or in IV when endogeneity correcting is necessary.