Chapter 13
Are Small Firms More Dependent on the Local Environment than Larger Firms? Evidence from Portuguese Manufacturing Firms

Carlos Carreira and Luís Lopes

Abstract This paper analyzes the impact on firm-level total factor productivity of both agglomeration economies and regional knowledge base, using an unbalanced panel of Portuguese manufacturing firms covering the period 1996–2004. Controlling for the endogeneity using the difference generalized method of moments estimator, we found that both localization and urbanization economies have a significant and positive effect on firm productivity, with the latter playing the most important role. Sectoral specialization economies are important for small and medium firms, but not for large firms. However, larger firms, therefore those with higher absorptive capacity, profit more from regional knowledge than smaller ones.

Keywords Agglomeration economies • Regional knowledge • Total factor productivity (TFP) • Small firms • Firm-level studies

13.1 Introduction

The study of spatial agglomeration of both production activities and knowledge base is important to understand their contribution for local and, consequently, national economic growth. Notwithstanding the tendency to reducing transaction costs, there has been observed an increasing propensity for firms to agglomerate their activities in certain regions with economic impact on employment level, wages, knowledge, productivity, and economic growth.

The theories of the location of economic activities are microeconomic in their essence, which means that the empirical studies should use firm-level data. However, the unavailability of large microeconomic datasets has favored empirical investigations at the aggregate rather than micro-level. Even in the cases of micro-level

C. Carreira • L. Lopes (✉)
Faculdade de Economia/GEMF, Universidade de Coimbra, Coimbra, Portugal
e-mail: ccarreir@fe.uc.pt; perlopes@fe.uc.pt

R. Baptista, J. Leitão (eds.), *Entrepreneurship, Human Capital, and Regional Development*, International Studies in Entrepreneurship 31,
DOI 10.1007/978-3-319-12871-9_13

researches, given that productivity growth at firm- or plant-level is generally not available, most of earlier studies use proxies such as employment and wage growth—under the assumptions that there is a national labor market and that labor is homogeneous, then productivity growth will result in proportional employment gains through shifts in labor demand (see, for example, Glaeser et al. 1992; Henderson et al. 1995; Combes 2000).

In this paper, we implement a micro-level analysis in order to shed further light on the extent to which the local environment, namely agglomeration economies and regional knowledge base, has an effect on firms' productivity. Additionally, we also investigate whether smaller firms are more dependent of local environment than larger ones. To conduct the analysis, we will use an unbalanced panel of Portuguese manufacturing firms covering the period 1996–2004.

This paper makes two main contributions to the economic literature. Even though agglomeration economies and regional knowledge base encompass a large number of studies, to our awareness, there has been no research that assesses the role of these two productivity sources together. Furthermore, there is scarce evidence on the effect of local environment on firms' total factor productivity (TFP), especially across firms' size.

The paper proceeds as follows. After a brief review of the background literature in the next section, Sect. 13.3 presents the empirical model and the dataset. Section 13.4 evaluates the effects of agglomeration economies and regional knowledge base on firm productivity through firm size. Section 13.5 offers some brief concluding remarks.

13.2 Theory and Selected Empirical Findings

The location of economic activity within the models of the new economic geography is endogenously determined through the interaction between two forces: the "centripetal" forces that attract economic agents to the same location and the "centrifugal" forces that push them apart (Krugman 1998). Externalities, a key concept developed by Marshall, are the most important centripetal force, as they are central to explain why production activities tend to agglomerate in certain regions.[1] The rationale is that, in the process of choosing its spatial location, a firm looks for the proximity of other firms due to the benefits they can get. Glaeser et al. (1992) identifies three sources of externalities:

- Marshall-Arrow-Romer (MAR)—after the three pioneering contributions of Marshall (1890/1961), Arrow (1962), and Romer (1986)—or localization externalities, which are related to intra-industry economies arising from the regional concentration of firms in the same industry (i.e., sectoral specialization). Firms have advantages in being located near others belonging to the same industry because the geographical concentration of an industry can increase the variety of

[1] Krugman (1998) identifies as the main centrifugal forces the immobile factors (e.g., certain land and natural resources), the high land rents and the external diseconomies (such as congestion).

intermediate goods available (at lower prices) as well as the dimension of final goods demand, can attract a large labor force with the skills demanded by that industry and can spread a great specialized knowledge level (namely via informal channels).

- Jacobs or urbanization externalities, which are connected to inter-industry economies arising from the variety of regional economic activities (Jacobs 1969). A sectoral diversity in a given region can stimulate a more diverse client base protecting firms from volatile demand, can create a vast spectrum of locally available inputs easing their switching in case of scarcity or a rise in prices and can disseminate a more assorted knowledge base increasing the possibility of discovering new products or production processes.
- Porter or competition externalities, which are related with competition intensity within a region. Competition stimulates both production and adoption of innovations and, consequently, improves firms' performance (Porter 1990). Porter externalities are similar to MAR externalities, but unlike earlier, it is local competition and not local monopoly that stimulates a faster search and adoption of innovations.

As it is possible to see, the theories that underlie externalities are microeconomic in essence, which means that empirical studies should use firm-level data. Given that until recently data on firms' productivity was generally not available, most of the studies used proxies. Glaeser et al. (1992), for example, using a dataset of 170 USA cities, between 1956 and 1987, find that MAR externalities have a negative impact on employment growth, while Jacobs and Porter economies positively affect it. Glaeser et al. (1992) approach has been replicated by other authors using both employment and wage growth as a dependent variable (see Cingano and Schivardi 2004, for a brief survey). However, the results of these researches are to some extent puzzling. Using 1991 Italian census data, Cingano and Schivardi (2004) show that, taking local employment growth as the dependent variable, the specialization effect is negative and variety effect has a significant and positive impact, in line with Glaeser et al.'s results, while using firm-level-based TFP indicators, the specialization effect is reversed and becomes positive, and neither sectoral variety nor the degree of local competition has any effect. Cingano and Schivardi (2004) question the conclusions of previous empirical works arguing that they suffer from serious "identification problems" when interpreted as evidence of dynamic externalities, since the chain of causality from agglomeration economies to employment growth could be reversed—the use of employment or wages growth at firm-level as dependent variable is based on the (unlikely) assumption that productivity growth will result in proportional employment gains through shifts in labor demand (see, for example, Glaeser et al. 1992; Henderson et al. 1995; Combes 2000).

Therefore, since externalities imply a change in output not fully accounted for by a change in inputs, TFP would be a better measure of performance. Martin et al. (2011) show that French plants from 1996 to 2004 benefit in terms of TFP growth from localization economies, but not from urbanization economies. They do not find any consistent pattern for local competition. An explanation can be that competition incentives firms to invest in R&D, but if the succession of innovations is rapid, the returns from R&D are low, which will reduce the R&D investment and, as a consequence, the innovations. In the case of the USA plants, over the period 1972–1992, Henderson (2003) finds that localization economies only have strong positive effects on TFP in high-tech not in mechanical industries. He also finds little evidence of urbanization economies.

Another interesting strand of economic geography research, favored by the flourishing endogenous growth theories, has pointed out that localized knowledge and technology spillovers matter for innovative activity, which is consequently shaped by space and concentrated in certain areas (see, for example, Scott 1988; Feldman 1994; Acs 2002; Johansson and Lööf 2008; Bronzini and Piselli 2009). In particular, it is argued that proximity to the knowledge base can encourage the circulation of ideas and the transmission of knowledge, thanks to face-to-face contacts and social interaction, which in turn facilitates innovation (Storper and Venables 2004; see Audretsch and Feldman 2005, for a review of theoretical and empirical studies). The knowledge-transfer environment in which a firm is embedded can also play a key role in explaining productivity differential between firms located in different geographic areas (Amesse and Cohendet 2001)—for example, *knowledge intensive business services* (KIBS) are crucial to disseminate knowledge across the region and to support firms' innovative activity (Muller and Zenker 2001).

Looking at the firm size, in general small firms could be expected to be more dependent on the local environment than larger firms (Henderson 2003; Andersson and Lööf 2011). Indeed, they are less able than large firms to internalize innovative inputs and to provide complementary activities that may facilitate innovation (Feldman 1994).

On the whole, despite the fact that the literature on agglomeration economies and regional knowledge base encompass a large body of studies, to our awareness, there has been no empirical research that assesses the role of these two productivity sources together. In fact, if both factors affect productivity and interact with each other and if one these factors is omitted, estimations of elasticity can be biased. Moreover, there is scarce evidence on effect of local environment across firms' size. We will try to fill this gap by assessing the role of both agglomeration economies and regional knowledge base effects in enhancing the TFP by firm size.

13.3 Empirical Methodology

13.3.1 The Dataset

To conduct our empirical analysis, we use an unbalanced panel of Portuguese manufacturing firms covering the period 1996–2004. The raw data is drawn from the combination of two statistical data sources, both run by the Portuguese Statistical Office (INE): *Inquérito às Empresas Harmonizado* (IEH), an annual business survey with information on both the input requirements and the output level; and *Ficheiro de Unidades Estatísticas* (FUE) which contains a variety of firm characteristics (activity, number of employees, age, and location) of all Portuguese firms, critical to compute spatial agglomeration variables. The longitudinal dimension of the panel, required for our analysis, was constructed using firm's unique identification code.

The unit of production considered is thus the firm. Each firm is assigned to a given region (at NUTS3 level, definition of 2002) through a spatial identification code. Thus, the first drawback of the data is that multi-plant firms may affect our results if their different plants are located in different regions. We note, however, that the different plants of corporations are often registered as distinct legal entities, thus the multi-plant phenomena impact on results may be small.

The IEH survey comprises all firms operating in Portugal with more than 100 employees, plus a representative random sample of firms with less than 100 employees.[2] For the purpose of this paper, the following filters were applied: firstly, due to lack of good data, firms with less than 20 employees were eliminated from the estimation sample[3]; secondly, firms located in the island regions (i.e., Madeira and Azores) were excluded; thirdly, given the number of observations, those firms operating in the manufacture of tobacco products (CAE 16) and manufacture of coke, refined petroleum products, and nuclear fuel (CAE 23) were also excluded; finally, firms with missing observations or unreasonable values (negative values and outliers) were dropped from the estimation sample. For each industry, we define as an outlier a firm for which the log difference between an input and the output is in the top and bottom one percentile of the respective distribution. As a result of all these procedures, we have, for the period 1996–2004, an unbalanced panel of 8,074 firms and a total of 32,003 (year-firm) observations.

13.3.2 *Empirical Model and Variables*

The main purpose of our analysis is thus to shed further light on the extent to which the local environment has an impact on productivity. In the past few years the study of this issue has greatly shifted from aggregated regional level towards the understanding of the operation of micro units (Stephan 2011; Ottaviano 2011). Accordingly, the general model that we use for our empirical analysis is a firm-level Cobb–Douglas production function—we assume that each firm is located in a given region r and operates in a given industry j[4]

$$Y_{it} = A_{it} K_{it}^{\alpha_j} L_{it}^{\beta_j} M_{it}^{\theta_j} \tag{13.1}$$

[2] The sample is representative of the Portuguese sector disaggregation (at three-digit level), both in terms of employment size and sales.

[3] We note that firms with less than 20 employees represent about 71 % of Portuguese manufacturing firms, but only 16 % of total employment (average over the period; source: OECD database).

[4] We omit subscripts j and r to simplify the notation except when it causes ambiguity.

where Y_{it} is the real gross output of the ith firm and year t (located in region r and operating in industry j), and K_{it}, L_{it}, and M_{it} are capital, labor, and material (intermediate) inputs, respectively; A_{it} is the TFP. We allow for the coefficients α_j, β_j, and θ_j to vary across industries. Given the regulation of the Portuguese labor market, we cannot assume perfect competition hypothesis, so neither constant returns to scale. The advantage is that, we disentangle TFP changes from production-scale effects, otherwise attributed to TFP.

The *gross output* is given by the sum of total revenues from sales, services rendered, and production subsidies. It is deflated by the producer price index at the three-digit level. The *labor* input is a 12-month employment average. *Materials* include the cost of materials and services purchased and were deflated by the GDP deflator. *Capital stock* is measured as the book value of total net assets (excluding financial investments and cash stock).

We assume that TFP of firm i is driven not only by firm's knowledge, but also by both agglomeration economies and regional knowledge base

$$A_{it} = \left(R_{it}\right)^{\gamma} \left(S_{it}^{jr}\right)^{\varphi} \left(Z_{it}^{jr}\right)^{\phi} \tag{13.2}$$

where R_{it} is the firm's knowledge stock in year t, S_{it}^{jr} is a vector of covariates that reflects the potential for spatial agglomeration economies of industry j in region r, and Z_{it}^{jr} is a vector of covariates that proxies regional knowledge base.

We assume as a proxy for firm's stock of knowledge the inverse of firm's size times its age

$$\text{FKNOW}_{it} = \frac{1}{L_{it} \cdot \text{Age}_{it}} \tag{13.3}$$

The rationale is that older and larger firms often command more resources and have higher managerial experience (Jovanovic 1982). The firm's knowledge returns are assumed nonlinear and decreasing. The index (13.3) ranges between close to zero (high level of knowledge), when firm is very large and old, and one (low level of knowledge), if it had only one employee and 1 year old—in our case, since we have imposed a censoring level of 20 employees, the maximum value is 0.05.

As discussed in Sect. 13.2, three kinds of advantages of the proximity for economic agents (agglomeration economies) can be distinguished: localization, urbanization, and competition economies. The localization (or sectoral specialization) economies are measured, for each firm, as the share of other employees working in the same industry (at the two-digit level) within a region (Combes 2000)[5]

$$\text{LOC}_{it}^{jr} = \frac{L_t^{jr} - L_{it}}{L_t^{r} - L_{it}} \tag{13.4}$$

with $L_t^{jr} = \sum_{i \in J^{jr}} L_{it}$ and $L_t^{r} = \sum_{i \in I^{r}} L_{it}$, where J^{jr} and I^r are the set of firms belonging to industry j in region r and whole region r, respectively, in year t.

[5] Since we subtract ith firm's employment, LOC are firm-specific.

The urbanization (or sectoral diversity) economies are proxied by the inverse of the Herfindahl–Hirschman index of industry concentration based on the employment share of the different industries (at the two-digit level), except the respective industry j, in a region (Henderson et al. 1995; Combes 2000)

$$\mathrm{URB}_t^{jr} = \frac{1}{\mathrm{HR}_t^{jr}} \tag{13.5}$$

with $\mathrm{HR}_t^{jr} = \sum_{g \neq j \wedge g \in G^r} \left[L_t^{gr} / \left(L_t^r - L_t^{jr} \right) \right]^2$, where G^r is the set of industries in region r. The measure of industrial diversity (13.5) ranges between 1 (minimum value), when all other manufacturing employment in the region is concentrated in a single industry, and $J^r - 1$ (maximum value) if it is uniformly distributed across all (other) industries. As pointed out by Combes (2000), the value of this indicator is not directly linked with the previous one of industrial specialization. In fact, if the regional employment is highly concentrated in a given industry and the several remaining industries have approximately the same size, the values of both indexes (concentration and diversity) for this industry are high.

To measure the degree of competition inside each industry at local level (competition externalities), we use the inverse of the Herfindahl–Hirschman index of regional employment concentration

$$\mathrm{COMP}_t^{jr} = \frac{1}{\mathrm{HJ}_t^{jr}} \tag{13.6}$$

with $\mathrm{HJ}_t^{jr} = \sum_{i \in J^{jr}} \left(L_{it} / L_t^{jr} \right)^2$. The higher is the employment share of firm i, therefore lesser uniform distribution of employment across firms, the lower is COMP_t^{jr}. The index also tends to increase with the number of firms.

Taking into account the theories of innovation and technological diffusion outlined in Sect. 13.2, we consider two kinds of factors through which regional innovative environment might impact on firm's productivity: knowledge transfer and knowledge base. Some economic agents such as those that operating in KIBS play a crucial role in disseminating knowledge through the region and supporting firms' innovative activity. We represent the capacity of transfer knowledge as the number of employees working in KIBS sector in the region.[6] In order to capture the effect of knowledge base, we distinguish two sources: regional R&D employment (RD) and the number of higher degree establishments in a region (UNIV)—the role of universities in innovation has been highlighted by various studies, such as Fritsch and Slavtchev (2007) and Cassia et al. (2009).

[6] According to *European Monitoring Centre on Change*, KIBS comprises the following CAE-rev2.1 divisions: (CAE 72) computer and related activities, (CAE 73) research and experimental development, and (CAE 74) other business activities.

13.3.3 Estimation Strategy

We adopt the so-called two-step approach. We firstly estimate the factor elasticity parameters of the following (log) Cobb–Douglas production function for each two-digit industry

$$y_{it} = a + \alpha^j k_{it} + \beta^j l_{it} + \theta^j m_{it} + u_{it} \tag{13.7}$$

where lower-case letters denote the log upper-case variables of Eq. (13.1), to compute firm-level (log) TFP

$$\hat{a}_{it} = y_{it} - \hat{\alpha}^j k_{it} - \hat{\beta}^j l_{it} - \hat{\theta}^j m_{it} \tag{13.8}$$

In the estimation of Eq. (13.1), we control for macroeconomic shocks by including year dummy variables. Additionally, we assume $u_{it} = \omega_{it} + \eta_{it}$, with ω_{it} denoting a firm-specific unobserved component and η_{it} a residual term uncorrelated with input choices. Ordinary least-squares (OLS) estimation of Eq. (13.7) produces inconsistent estimates due to the likely presence of simultaneity and selection bias: the simultaneity bias arises because input demands are also determined by firm's knowledge of its productivity level, which makes ω_{it} correlated with the observed inputs; the selection bias is generated by endogenous exit, as smaller firms, with lower capital intensity, are more likely to exit. Assuming that ω_{it} is time invariant, Eq. (13.7) can be estimated using the least square dummy variable approach or the within transformation.[7] Consistency of the fixed effect model requires, however, strictly exogeneity of the included regressors, a nonrealistic assumption (Griliches and Mairesse 1998). To overcome this problem, we estimate Eq. (13.7) using the generalized method of moments (GMM) methodology for 20 separate industries (at two-digit level). In particular, we employ the Arellano and Bond (1991) one-step difference GMM (GMM-DIF) estimator, which transforms the panel data model in first differences to remove the individual effects and then uses lagged levels of the dependent variable and the predetermined variables as instruments for the endogenous differences.[8]

We then estimate (in the log form) the model (13.2)

$$\begin{aligned} a_{it} = {} & \gamma_0 + \gamma_1 \text{fknow}_{it} + \varphi_1 \text{loc}_{it}^{jr} + \varphi_2 \text{urb}_t^{jr} + \varphi_3 \text{comp}_t^{jr} \\ & + \phi_1 \text{kibs}_t^r + \phi_2 \text{univ}_t^r + \phi_3 \text{rd}_t^r + \upsilon_{it} \end{aligned} \tag{13.9}$$

where the residual term is given by $\upsilon_{it} = \mu_i + \varepsilon_{it}$. We cannot disentangle firm and regional fixed effects with this formulation, but that does not affect the estimation. Since all covariates are expressed in logarithms, the estimated coefficients can be interpreted as elasticity parameters.

[7] The random effects model is rejected in favor of the presence of fixed effects by both Hausman and robust Hausman tests at the 1 % significance level (see Wooldrige 2002).

[8] Regressions were performed using the Stata, *xtabond2* procedure (Roodman 2009). The results presented in the paper are robust to fixed-effects (Olley and Pakes 1996; Levinsohn and Petri 2003) and GMM-System methods. These results are available from the authors upon request.

Regarding Eq. (13.9), we note that it is subject to two main sources of endogeneity: unobserved heterogeneity and simultaneity bias. In fact, some regional characteristics (e.g., public infrastructures, local climate, natural resources, etc.) that are not taken into account in this econometric model can affect the propensity to agglomerate, while at the same time agglomeration influences these regional characteristics—in other words, υ_{it} is correlated with the independent variables. Additionally, self-selection of the more productive firms also creates a simultaneity problem. Higher productivity in larger markets (or denser areas) may not be due to agglomeration economies (learning effect); it might instead be due to the fact that high-productivity firms are more likely to be attracted to these advantageous markets (selection effect).[9] In other words, because more productive firms are likely located in larger/denser regions, average firm productivity in these regions should be higher even if there are negligible agglomeration economies, which means that OLS estimates might be biased (Baldwin and Okubo 2006; Melitz and Ottaviano 2008; Andersson and Lööf 2011; Saito and Gopinath 2009). To deal with the endogeneity problem, we estimate the model using again the GMM-DIF procedure. Industry and regional dummies were also included in the estimation.

As discussed in Sect. 13.2, it can be expected that the role of local environment can be different across firms of different sizes. In order to investigate this, we will split the sample into three size classes: firms with 20–100, 100–250, and 251 or more employees (small, medium, and large firms, respectively). The thresholds are those used by the OECD, except for large firms—in Portugal, there are only a few firms with more than 500 employees, the OECD threshold.

13.3.4 Summary Statistics

Tables 13.1 and 13.2 report the summary statistics and the correlations matrix, respectively, of the main variables used in our estimations. Most variables exhibit strong variability, as shown by the large values of standard deviations respective to their mean (Table 13.1). Even if between variations account for a large part of this heterogeneity, within standard-deviation has a nonnegligible role in its explanation. The mean manufacturing firm in the estimation sample has 122 employees and produce 9,812,000€.

The correlation matrix reveals that, as expected, there is a statistically significant (at 5 %) and negative correlation between TFP and FKNOW—recall that lower values of variable mean higher level of knowledge—and a statistically significant and positive correlation between TFP and both spatial agglomeration and regional knowledge covariates, except in the case of URB (Table 13.2). The correlation between the regional knowledge covariates (i.e., KIBS, RD, and UNIV) is rather high, which should cause multicollinearity problems in the regressions. Given that, the two explanatory variables that measures the knowledge input available in the region, RD and UNIV, are replaced by their product (i.e., RKNOW = RD × UNIV).

[9] In the Portuguese case, larger markets and denser areas are highly correlated.

Table 13.1 Descriptive statistics, 1996–2004

Variable	Obs	Mean	Std. dev.			Min	Max
		Overall	Overall	Between	Within	Overall	Overall
(a) Firm-specific							
Y (10^3€)	32,003	9,812	39,822	29,437	8,382	118	2,076,602
K (10^3€)	32,003	9,927	35,903	29,689	10,685	28	2,019,021
L	32,003	122	236	185	47	20	7,455
M (10^3€)	32,003	6,934	32,795	23,938	6,002	11	1,699,340
TFP	32,003	40.9	29.6	29.2	6.8224	7.0	702.8
FKNOW	31,960	0.0015	0.0023	0.0030	0.0009	0.00	0.048
(b) Regional level							
LOC	32,003	0.1428	0.1406	0.1427	0.0213	0.00	0.805
URB	32,003	7.0256	3.0239	3.0009	0.6114	1.23	13.316
COMP	32,003	37.4	46.3	52.0	8.7	1.0	272.5
KIBS	32,003	17,239	35,589	32,548	9,774	24	143,322
RD	32,003	2,166	3381	3,163	688	0	11,991
UNIV	32,003	25.3	32.7	31.4	5.2	0	97

Figure 13.1a displays the distribution of sample firms across the 28 NUTS3 regions. The map shows a high concentration of firms in the North, mainly not only in the regions of Grande Porto, Ave, and Baixo Vouga, but also in the region of Grande Lisboa. Figure 13.1b highlights the spatial distribution of the weighted average of the TFP level. As can be seen, the regions of Minho-Lima, Ave, Cova da Beira, Pinhal Interior Norte, and Pinhal Litoral show the highest values of TFP.[10]

13.4 How Large Are the Local Environment Effects Across Size Classes?

The key results of GMM-DIF estimation of model (13.9) are presented in Table 13.3—the factor elasticity estimates for each industry, used in the second-step to compute firm-level TFP, are in Appendix Table 13.4. The Appendix Table 13.5 summarizes the key coefficient estimates of model (13.9) using ordinary least-squares estimators. Column (1) of Table 13.3 summarizes the main coefficient estimates for the overall sample, while columns (2)–(4) show the results by size classes. The validity of GMM-DIF estimates depends on the absence of second-order serial autocorrelation and on the choice of the appropriate set of instruments. This is indeed the case, since, as expected, the Arellano–Bond AR(1) test shows a negative first-order serial correlation, while the AR(2) test indicates that residuals are seemingly free from second-order serial correlation. Moreover, the null hypothesis of the Hansen test that the overall instruments are valid is not rejected in all four regressions. We note that the Hansen and Sargan tests for over-identifying restrictions show opposite results; however, the Sargan test should be interpreted with care, since the model allows for heteroskedasticity rendering the test baseless.

[10] See NUTS3 regions in Fig. 13.2.

Table 13.2 Correlation across covariates (pooled yearly values), 1996–2004

	Y	*K*	*L*	*M*	TFP	FKNOW	LOC	URB	COMP	KIBS	RD
K	0.83*	1									
L	0.63*	0.57*	1								
M	0.93*	0.79*	0.58*	1							
TFP	0.05*	0.06*	0.09*	0.03*	1						
FKNOW	–0.11*	–0.12*	–0.20*	–0.09*	–0.06*	1					
LOC	–0.04*	–0.04*	0.02*	–0.04*	0.06*	–0.02*	1				
URB	0.09*	0.09*	0.05*	0.08*	–0.01	–0.05*	–0.35*	1			
COMP	–0.05*	–0.06*	–0.00	–0.05*	0.29*	0.00	0.49*	–0.27*	1		
KIBS	0.08*	0.08*	0.05*	0.06*	0.04*	–0.06*	–0.19*	0.53*	–0.05*	1	
RD	0.08*	0.09*	0.05*	0.06*	0.05*	–0.06*	–0.22*	0.62*	–0.05*	0.98*	1
UNIV	0.07*	0.08*	0.05*	0.06*	0.04*	–0.06*	–0.27*	0.71*	–0.06*	0.88*	0.93*

Note: * denotes statistical significance at the 0.05 level

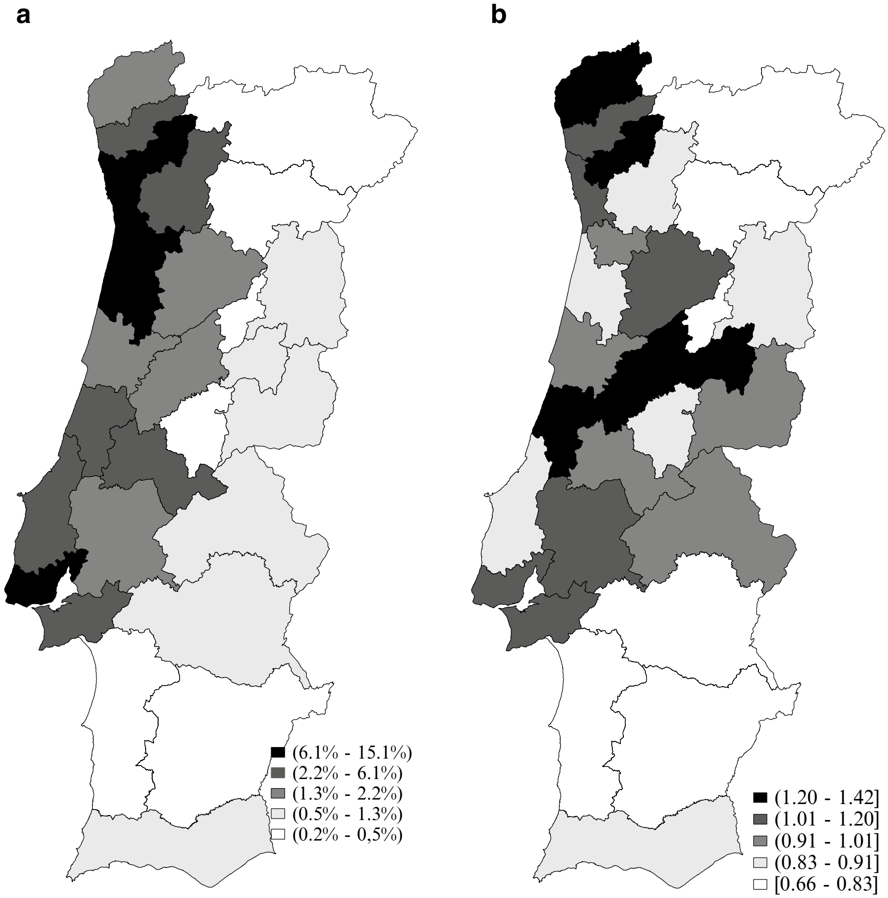


Fig. 13.1 Number of firms and TFP by NUTS3 regions. (**a**) Number of firms (percentage of total). (**b**) Total factor productivity (quintiles)

13.4.1 Overall Sample Analysis

Looking at the estimated parameters in column (1) of Table 13.3, firm's stock of knowledge (FKNOW) has a statistically significant (at 5 %) and virtual impact on firm's productivity—an increase in knowledge implies that the corresponding index reduces, then increasing the productivity—but it is far to explain all productivity gains. Localization (LOC) and urbanization (URB) economies also positively impact (at the 1 % significance level) on the firm's productivity, while no effects of the degree of local competition (COMP) are found at conventional significance levels. In particular, increasing by 1 % the share of other employees working in the same industry region, *ceteris paribus*, increases the TFP of a firm by 0.0068 %. In the case of the employment share of the other industries in the region, the corresponding increment in the TFP is 0.0751 %. These results seem to point out a superiority of sectoral diversity (urbanization) economies.

Table 13.3 Results of GMM-DIF regression

Firm size Variable	Overall (1)	Small (2)	Medium (3)	Large (4)
FKNOW	–0.0690** (0.0271)	0.0017 (0.0470)	–0.2943*** (0.0474)	–0.0991*** (0.0352)
LOC	0.0068*** (0.0025)	0.0065* (0.0038)	0.0087** (0.0042)	0.0006 (0.0036)
URB	0.0751*** (0.0191)	0.0773*** (0.0248)	0.0083 (0.0341)	0.0744** (0.0361)
COMP	0.0023 (0.0066)	0.0026 (0.0091)	0.0092 (0.0096)	0.0008 (0.0131)
KIBS	0.0078** (0.0033)	0.0127** (0.0050)	–0.0034 (0.0055)	0.0184*** (0.0064)
RKNOW	0.0241*** (0.0041)	0.0195*** (0.0054)	0.0141** (0.0061)	0.0318*** (0.0111)
No. of observations	11,015	5,368	3,958	2,107
No. of firms	2,922	1,827	1,046	478
No. of instruments	49	42	31	44
i. AR(1) and Prob(z)	–6.33 0.000	–2.52 0.012	–6.53 0.000	–6.02 0.000
ii. AR(2) and Prob(z)	1.95 0.051	1.64 0.101	0.07 0.948	0.73 0.468
iii. Sargan test and Prob(z)	274.89 0.000	81.26 0.000	0.48 0.785	115.97 0.000
iv. Hansen test and Prob(z)	4.93 0.177	0.00 1.000	0.86 0.651	7.02 0.319

Notes: The table summarizes the key coefficient estimates for four different regressions of model (13.9). GMM-DIF denotes the Arellano–Bond one-step difference GMM estimator. All regressions include industry and regional dummies. Variables are in logarithmic form (except in the case of the dummy variables). Robust standard errors are given in parentheses. ***, **, and * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively

For its part, regional knowledge also seems to play a key role on firms' TFP gains. In fact, both the number of employees working in KIBS sector-region and regional knowledge base have a positive impact (significance at 5 % and 1 %, respectively) on the productivity—increasing KIBS (RKNOW) by 1 %, all else equal, increases the TFP by 0.0078 (0.0241) %.

13.4.2 Differences Across Firms' Size

We now refine our analysis splitting the sample into three size classes—small, medium, and large firms, respectively, columns (2), (3), and (4) of Table 13.3—, considering that agglomeration economies and regional knowledge might have heterogeneous impact across firms. In related works, Martin et al. (2011) and Henderson (2003) find that small firms benefit more from agglomeration economies than larger ones.

Looking at the estimated parameters of agglomeration variables, our first finding is that small and medium firms benefit from localization economies, while at the same time large firms do not benefit from this sectoral specialization. However, we also find that localization economies are stronger for medium than small firms, contrary to the expected. A second finding is that there is a significant and positive relationship between sectoral diversity and productivity for small and large firms, but stronger for the smaller ones. Finally, the impact of regional knowledge (KIBS and RKNOW) seems to be higher for large firms than small firms. One explanation for this unexpected finding can be that small firms have not accumulated enough knowledge to absorb external (regional) knowledge ("absorptive capacity of firms," after Cohen and Levinthal 1989).

Also surprisingly, while firm's internal knowledge has a significant (at 1 %) expected effect on the productivity level of medium and large firms, it does not seem to impact on the productivity of small firms. A possible explanation for this unexpected finding can be that sample partition created a homogeneous group of (small) firms which have not yet accumulated enough internal knowledge to impact on productivity.

13.5 Conclusion

This study focuses on the extent to which the local environment has an impact on productivity across firms' size, using an unbalanced panel of Portuguese manufacturing firms covering the period 1996–2004. We assume that both agglomeration economies and regional knowledge have a positive impact on firms' TFP. Additionally, smaller firms are more dependent of local environment than larger firms.

Our econometric estimates confirm the conjecture that the agglomeration economies and regional knowledge base seem to be important to explain productivity gains at firm-level. In particular, we found that both localization and urbanization economies have a significant and positive effect on firms' TFP, with the latter playing the most important role. Sectoral specialization economies are important for small and medium firms, but not for large firms. However, larger firms, consequently, those with higher absorptive capacity, profit more from regional knowledge than smaller ones.

Overall, this paper contributes to a better understanding of the economic mechanisms and, consequently, may contribute to the implementation of the adequate regional policies to enhance economic growth. Our findings imply that fostering productivity could require different instruments across firms' size. Regional specialization seems to be a worthwhile policy to promote productivity gains of small firms. To help small firms to benefit from regional knowledge base, policy makers could promote the creation of internal knowledge inside of these firms' type.

Several issues remain in question, which should deserve our attention in the future, namely the unexpected results for the localization economies and firm's internal knowledge within the small firms.

13.6 Appendix

Table 13.4 Production function elasticities by industry

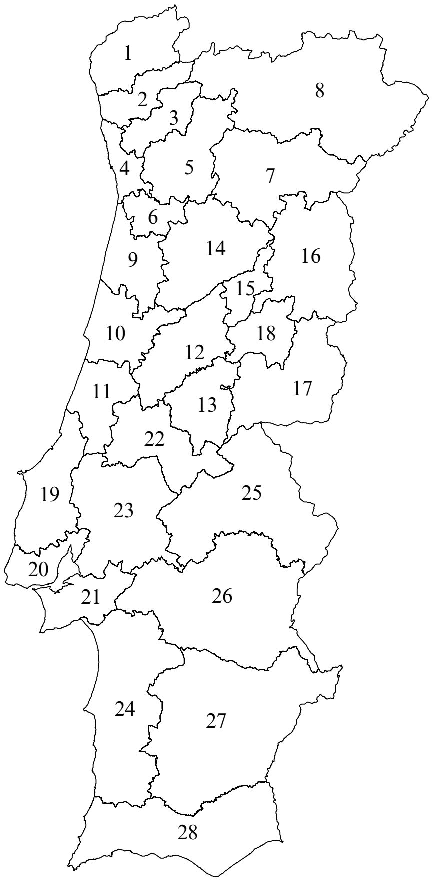
Industry	α	β	θ
Food products and beverages	0.031* (0.017)	0.053* (0.029)	0.759*** (0.031)
Textiles	0.026 (0.024)	0.156*** (0.035)	0.712*** (0.022)
Wearing apparel	0.146*** (0.021)	0.421*** (0.087)	0.457*** (0.023)
Leather and leather products	0.079*** (0.027)	0.202*** (0.055)	0.714*** (0.032)
Wood and wood products	0.011 (0.021)	0.100*** (0.039)	0.720*** (0.023)
Pulp, paper, and paper products	0.099*** (0.035)	0.140* (0.072)	0.676*** (0.046)
Publishing and printing	0.058*** (0.020)	0.143*** (0.048)	0.656*** (0.028)
Chemical and chemical products	0.030 (0.021)	0.124*** (0.029)	0.770*** (0.025)
Rubber products	0.003 (0.057)	0.107 (0.097)	0.636*** (0.064)
Plastics products	0.001 (0.025)	0.103* (0.056)	0.710*** (0.037)
Other nonmetallic mineral products	0.027 (0.022)	0.116*** (0.034)	0.736*** (0.025)
Basic metals	0.023 (0.029)	0.270*** (0.063)	0.731*** (0.029)
Fabricated metal products	0.091** (0.036)	0.216*** (0.036)	0.627*** (0.038)
Machinery and equipment	0.079*** (0.029)	0.300*** (0.053)	0.632*** (0.024)
Electrical and optical equipment	0.068** (0.031)	0.104*** (0.040)	0.742*** (0.028)
Motor vehicles, trailers, and semi-trailers	0.025 (0.027)	0.147*** (0.042)	0.736*** (0.031)
Other transport equipment	0.067 (0.085)	0.220* (0.119)	0.590*** (0.079)
Furniture, manufacturing n.e.c., and recycling	0.124*** (0.038)	0.075** (0.035)	0.698*** (0.029)

Notes: Arellano and Bond (1991) one-step difference GMM estimates of Eq. (13.9). α, β, and θ denote capital, labor, and material elasticities, respectively. All regressions include year dummies. Robust standard errors are given in parentheses. ***, **, and * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively

Table 13.5 Results of OLS and FE

Variable	Pooled OLS				FE			
	Overall (1)	Small (2)	Medium (3)	Large (4)	Overall (5)	Small (6)	Medium (7)	Large (8)
FKNOW	–0.019*** (0.002)	–0.006*** (0.002)	0.006 (0.003)	–0.009*** (0.003)	–0.051*** (0.007)	–0.046*** (0.012)	–0.081*** (0.014)	–0.054*** (0.017)
LOC	0.001 (0.001)	0.001 (0.002)	0.000 (0.002)	0.002 (0.002)	0.002 (0.002)	0.006* (0.003)	0.006 (0.005)	–0.003 (0.005)
URB	0.005 (0.008)	0.003 (0.013)	0.018 (0.013)	0.029 (0.023)	0.059*** (0.016)	0.045** (0.023)	0.067** (0.032)	0.057* (0.034)
COMP	0.004 (0.002)	0.007** (0.003)	0.006 (0.004)	–0.015*** (0.005)	0.000 (0.006)	–0.013 (0.008)	0.008 (0.010)	–0.009 (0.01)
KIBS	0.010*** (0.003)	0.011*** (0.004)	0.008** (0.004)	0.016** (0.007)	0.011*** (0.003)	0.007 (0.005)	0.006 (0.005)	0.020*** (0.006)
RKNOW	0.015*** (0.004)	0.010* (0.005)	0.026*** (0.006)	0.029*** (0.010)	0.025*** (0.004)	0.014** (0.006)	0.030*** (0.007)	0.041*** (0.011)
No. of observations	14,882	7,803	4,972	2,107	14,882	7,803	4,972	2,107
R^2	0.950	0.943	0.956	0.962	0.847	0.672	0.480	0.308

Notes: The table summarizes the key coefficient estimates of Eq. (13.9). OLS and FE denote ordinary least-squares and fixed effects (within) estimators. All regressions include industry and regional dummies. Variables are in logarithmic form (except in the case of the dummy variables). Robust standard errors are given in parentheses. ***, **, and * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively



Region:
1-Minho-Lima
2-Cávado
3-Ave
4-Grande Porto
5-Tâmega
6-Entre Douro e Vouga
7-Douro
8-Alto Trás-os-Montes
9-Baixo Vouga
10-Baixo Mondego
11-Pinhal Litoral
12-Pinhal Interior Norte
13-Pinhal Interior Sul
14-Dão-Lafões
15-Serra da Estrela
16-Beira Interior Norte
17-Beira Interior Sul
18-Cova da Beira
19-Oeste
20-Grande Lisboa
21-Península de Setúbal
22-Médio Tejo
23-Lezíria do Tejo
24-Alentejo Litoral
25-Alto Alentejo
26-Alentejo Central
27-Baixo Alentejo
28-Algarve

Fig. 13.2 NUTS3 regions of mainland Portugal

References

Acs Z (2002) Innovation and the growth of cities. Edward Elgar, Cheltenham

Amesse F, Cohendet P (2001) Technology transfer revisited from the perspective of the knowledge-based economy. Res Policy 30:1459–1478

Andersson M, Lööf H (2011) Agglomeration and productivity: evidence from firm-level data. Ann Reg Sci 46(3):601–620

Arellano M, Bond S (1991) Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. Rev Econ Stud 58:277–297

Arrow KJ (1962) The economic implications of learning by doing. Rev Econ Stud 29:155–173

Audretsch DB, Feldman MP (2005) Knowledge spillovers and the geography of innovation. In: Henderson V, Thisse JF (eds) Handbook of urban and regional economics: cities and geography, vol 4. North-Holland, Amsterdam, pp 2713–2739

Baldwin RE, Okubo T (2006) Heterogeneous firms, agglomeration and economic geography: spatial selection and sorting. J Econ Geogr 6:323–346

Bronzini R, Piselli P (2009) Determinants of long-run regional productivity with geographical spillovers: the role of R&D, human capital and public infrastructure. Reg Sci Urban Econ 39:187–199

Cassia L, Colombell A, Paleari S (2009) Firms' growth: does the innovations system matter? Struct Chang Econ Dyn 20(3):211–220

Cingano F, Schivardi F (2004) Identifying the sources of local productivity growth. J Eur Econ Assoc 2(4):720–742

Cohen WM, Levinthal DA (1989) Innovation and learning: the two faces of R&D. Econ J 99(397):569–596

Combes P (2000) Economic structure and local growth: France 1984-1993. J Urban Econ 47: 329–355

Feldman MP (1994) The geography of innovation. Kluwer Academic, Boston

Fritsch M, Slavtchev V (2007) Universities and innovation in space. Ind Innov 14:201–218

Glaeser E, Kallal H, Sheikman J et al (1992) Growth in cities. J Polit Econ 100(6):1126–1152

Griliches Z, Mairesse J (1998) Production functions: the search for identification. In: Strøm S (ed) Econometrics and economic theory in the 20th century: the Ragnar Frisch centennial symposium. Cambridge University Press, Cambridge, pp 169–203

Henderson JV (2003) Marshall's scale economies. J Urban Econ 53:1–28

Henderson JV, Kuncoro A, Turner M (1995) Industrial development of cities. J Polit Econ 103:1067–1090

Jacobs J (1969) The economy of cities. Vintage, New York

Johansson B, Lööf H (2008) Innovation activities explained by firm attributes and location. Econ Innov New Technol 17(6):533–552

Jovanovic B (1982) Selection and the evolution of industry. Econometrica 50(3):649–670

Krugman P (1998) What's new about the new economic geography? Oxf Rev Econ Policy 14(2):7–17

Levinsohn J, Petri A (2003) Estimating production functions using inputs to control for unobservables. Rev Econ Stud 70:317–342

Marshall A (1890/1961) Principles of economics, 9th edn. Macmillan, London

Martin P, Mayer T, Mayneris F (2011) Spatial concentration and plant-level productivity in France. J Urban Econ 69:182–195

Melitz MJ, Ottaviano GIP (2008) Market size, trade, and productivity. Rev Econ Stud 75(1): 295–316

Muller E, Zenker A (2001) Business services as actors of knowledge transformation: the role of KIBS in regional and national innovation systems. Res Policy 30:1501–1516

Olley GS, Pakes A (1996) The dynamics of productivity in the telecommunications equipment industry. Econometrica 64(6):1263–1297

Ottaviano G (2011) 'New' new economic geography: firm heterogeneity and agglomeration economies. J Econ Geogr 11(2):231–240

Porter ME (1990) The competitive advantage of nations. Macmillan, London

Romer PM (1986) Increasing returns and long-run growth. J Polit Econ 94:1002–1037

Roodman D (2009) How to do xtabond2: an introduction to difference and system GMM in Stata. Stata J 9(1):86–136

Saito H, Gopinath M (2009) Plants' self-selection, agglomeration economies and regional productivity in Chile. J Econ Geogr 9:539–558

Scott AJ (1988) New industrial spaces. Pion, London

Stephan A (2011) Locational conditions and firm performance: introduction to the special issue. Ann Reg Sci 46:487–494

Storper M, Venables AJ (2004) Buzz: face-to-face contact and the urban economy. J Econ Geogr 4:351–370

Wooldrige JM (2002) Econometric analysis of cross section and panel data. MIT, Cambridge