Firm heterogeneity in TFP, sectoral innovation and geography. Evidence from Italy

Aiello, Francesco and Pupo, Valeria and Ricotta, Fernanda

Department of Economics, Statistics and Finance - University of Calabria, Department of Economics, Statistics and Finance - University of Calabria, Department of Economics, Statistics and Finance - University of Calabria

23 July 2013

Firm Heterogeneity in TFP, Sectoral Innovation and Geography. Evidence from Italy•

Francesco Aiello, Valeria Pupo, Fernanda Ricotta

University of Calabria
Department of Economics, Statistics and Finance
87036 Arcavacata di Rende (CS) - Italy

Abstract Sectoral and territorial specificities affect the firm's capabilities of being productive. While there is a wide consensus on this, a quantitative measure of the these effects has been lacking. To this end, we combine a dataset of Italian firms with some *meso* regional and sectoral variables and apply a cross-classified model that allows for a clear distinction between firm, region-specific and sector-specific effects. After observing a marked TFP heterogeneity across firms, the paper addresses the issue of understanding how much differences in firms' productivity depend on regional localisation and sector specificities. Results refer to 2004-2006 and are threefold. Firstly, they confirm that the main source of firm variety is mostly due to differences revealed at individual level. Secondly, we find that sector is more important than location in explaining firms' TFP. Lastly, the results show that firm TFP increases when it belongs to more innovative sectors. Similarly, companies get benefits from belonging to sectors where there is a high proportion of firms using R&D public support and a high propensity to collaborate in innovative projects.

Key words: Total Factor Productivity, Firms' Heterogeneity, Sectoral innovation, Geography, Cross-Classified Models
JEL code: L60, L25, O33

1. Introduction

The role of sector membership and location in firms' performance is one of the most common fields of research in empirical economics. This *incipit* implies that there are at least three key levels of analysis to be considered in order to understand the differences in enterprises' performances better. Firstly, firms differ from each other in several ways, whatever the region and the sector. For instance, they have different sizes, specific approaches to production and different innovation strategies which render their innovation process persistent (Teece *et al*, 1997). Again, every learning process (e.g. learning by doing; learning by exporting) is firm-specific because it is driven by factors relating to the individual skills of owners, workers and managers (Bloom and Van Reenen, 2010). It is not surprising that all this heterogeneity in firm-specific behavior translates into heterogeneity in performance. Another source of variability is location: extensive literature shows that firms' behavior depends on the spatially-binded availability of territorial resources devoted to

• This work has benefitted from the financial support of Regione Calabria within the CALCOM project on "Regional Competitiveness and Innovation" (L.R. 10/1998). The authors would like to thank Paola Cardamone for the very useful comments and suggestions on an earlier version of paper. Usual disclaimer applies.

growth. In short, the higher the quantity and quality of endowments of a given area, the higher the benefits accruing to firms located in that territory (on this see, above all, Vernon Henderson *et al,* 2001 and Krugman, 1999). The same thing applies when referring to sectors: firms operate in different sectors which have specific technological regimes (conditions of appropriability, technological opportunity and knowledge cumulativeness) that influence, *ceteris paribus*, firms' innovation activity and, therefore, performance (among others, Castellacci and Zheng, 2010; Malebra, 2005; Ngai and Samaniego, 2011).

Following this line of reasoning, what we expect is that there will be substantial differences in productivity when comparing individual firms and when grouping them by region or sector. However, even when differences in results are found, some questions still remain unaddressed. For instance, the main issue concerns the evaluation of the importance of firms' heterogeneity compared with that of location and sector. This is a point to be addressed on empirical grounds. Again, location and sector affect firms, but, to our knowledge, there is no evidence quantifying the order of these effects: how much the difference in firm performance is due to individual heterogeneity and how much it is a result of territorial and sectoral influences. This study contributes to the debate by providing new insights regarding the determinants of Italian firms' productivity.

With respect to the literature, the distinguishing features of this study are the following.

Firstly, Italy is an interesting case-study because of its marked productivity variability across regions and sectors. A strand of literature shows that firms located in the North of the country have advantages from operating in highly-endowed regions, while the contrary is true when firms operate in the South of Italy, an area which still suffers from widespread structural deficits. This gap is wide and has been persistent over the last 30 years (ISTAT, 2010; Iuzzolino *et al*, 2011 for a survey). From this perspective, it is clear that location matters in explaining firms' performance. Beside the spatial dimension, the specialisation model of the Italian economy is also of great interest in explaining the variability in firms' productivity. In general terms, we know that current sectoral specialisation does not offer many opportunities for productivity gains because the proportion of high-tech firms is limited and their R&D investments are very low (Bugamelli *et al*, 2012).

Secondly, this study focuses on firms' heterogeneity in TFP. The use of TFP is a novelty in itself when measuring the sources of heterogeneity observed at firm level.[1] The related literature is scant and addresses different research questions from ours. For example, some papers are based on business strategy and compare sectoral regimes and firm-level variety. This literature refers back to contributions made by Schmalensee (1985) and further explored by McGahan and Porter (1997) or Kaniovski and Peneder (2002). Another area of research investigates the role of sector membership in firms' innovation activities and argue how innovation activities made by firms are highly dependent on the economic sector they belong to (see, above all, Nelson and Winter, 1982; Malerba, 2005). Finally, some recent studies support the view that sectors are important in explaining TFP growth, for instance Castellacci and Zheng (2010). These scholars consider a sample of Norwegian firms and decompose the TFP growth into technical progress and technical efficiency for the years 1998-2004. They show that these two components of TFP growth are related to sectoral technological opportunities. In Schumpeter Mark II sectors, the large innovating firms define the frontier and push it further through technical progress, while in Schumpeter Mark I sectors, the enterprises use the technological knowledge in order to increase efficiency. With regard the Italian case, papers by Fazio and Piacentino (2010) and Aiello *et al* (2013) are comparable with our work. While the former investigates the spatial variability in labour productivity of small and

[1] It is also worth noting here that, in the case of Italy, the use of TFP is further motivated by the fact that a great deal of research shows how differences in regional development depend on differences in TFP (Aiello and Scoppa 2000; Ascari and Di Cosmo, 2005; Marrocu and Paci, 2010; Quartaro, 2006).

medium-sized enterprises across Italian provinces (NUTS 3), the latter analyses how firms' characteristics and regional factors affect firms' TFP. Both papers disregard sectors.

The final innovation in the analysis comes from the method applied in the empirical setting. In order to handle data identified at different levels of analysis (firms, regions and sectors), the paper refers to a cross-classified model, which allows us to disentangle the impact exerted by firm-specific factors, location and sector. This is possible because multilevel analysis relates the different levels of data aggregation in ways that render the simultaneous existence of distinct level-one (*firms*) and level-two (*regions* and *sectors*) equations explicit. With respect to single-equation models, multilevel equations exploit the structures of data and properly address the issue of error correlation across firms which operate in the same sector and in the same region. More importantly, the inference phase is made by distinguishing between sample size at the different levels of data aggregation. This limits the high risk of type I errors because the variance in any level higher than the first is correctly estimated (whereas single-equation models underestimate the variance because they use the entire sample size, without differentiating according to level).

The empirical analysis refers to the period 2004-2006 and uses micro-data from the "Survey on Manufacturing Firms" carried out by UniCredit-Capitalia (2008). The main results are as follows. After having found that TFP heterogeneity at firm level is mainly due to differences in firms' internal characteristics, we show that sector plays a more prominent role than location. To be more precise, 86% of the variance in firms' TFP is due to firms' characteristics and 3.4% is ascribable to geography, while the proportion of firms' TFP heterogeneity due to sector is 10%. When detailing the role of industry membership, we find that several proxies of sector innovativeness play a crucial role.

The paper is organised as follows. Section 2 presents the micro-data used in the analysis and points out how relevant the regional and sectoral differences in TFP are. Section 3 illustrates the empirical strategy followed in the estimations. Section 4 discusses the results and section 5 concludes.

2. TFP at firm level in Italy. A brief overview

2.1 The sample

Microdata used in this paper come from the X*th* UniCredit-Capitalia survey (2008) of Italian manufacturing firms. The survey design followed by UniCredit-Capitalia includes all firms with a minimum of 500 employees and a sample of firms with between 11 to 500 employees selected according to three stratifications: geographical area, Pavitt sector and firm size. Although the original dataset comprises 5,100 firms, we use a sample of 3,006 firms which is obtained after carrying out a data cleaning procedure.[2] The survey covers the 2004-2006 three-year period.[3]

Table 1 presents the distribution of firms by area, economic activity and size. Firm distribution by area in 2006 reveals a predominance of enterprises located in the North of Italy (two thirds of the sample), while 15.5% of firms were based in the Centre and around 10% in the South. As far as Pavitt sectors are concerned, the distribution shows a concentration of enterprises in traditional sectors (49% of the total) and in highly specialised sectors (27.6%). The incidence of high-tech firms is residual (only 4.5% in the entire sample). Again, table 1 indicates that the firms in the sample were mainly small and medium firms (about 56% and 35% respectively). This picture is representative of Italian manufacturing industry, which exhibits a predominance of firms located in the North and belonging to traditional sectors. A commonly stylised fact is that there is a very high proportion of small-sized firms in Italy, whatever the area and the economic activity (see, e.g., Bank of Italy, 2009).

While table 1 presents data for macro-areas and Pavitt sectors, the A1 and A2 appendix tables refer to the distribution of firms in each individual Italian region and across sectors as classified according to Ateco. When considering location, firm distribution is very uneven. For example, in some regions (Valle d'Aosta, Basilicata, Molise and Calabria) there are few firms. Lombardia, on the other hand, has the largest number of firms (979). The same high variability occurs when one considers the sample composition at sectoral level. The proportion of firms is high in sectors producing "metal products" (386 companies) and "machinery" (458), and low in the production of "coke and refined petroleum products" (16) and "office machinery" and "computers" (14). These details regarding the composition of sample by region and sector will be used in selecting the econometric model to be estimated (*cfr* § 3).

2.2. TFP by sector and region

This section aims at verifying whether firm productivity differs across geographical area and economic sectors. To this end, we use TFP as it is the key variable in explaining the differences in levels of economic activity across countries or regions (on this see, amongst others, Aiello *et al*,

[2] The cleaning procedure was carried out as follows. Firms with negative values of value added were eliminated from the original archive. Moreover, in order to eliminate *outliers*, firms with a growth rate of value added and of employees below the first or above the ninety-ninth percentile of the distribution were also eliminated. Finally, when building the sample used in estimating TFP, we excluded firms for which at least 7 years data regarding the number of employees was not available. It is important to point out that the distribution of our sample overlaps with the distribution of the original dataset when aggregating firms by macro-area, region and sector. On the other hand, the proportion of medium-large firms in the sample is higher than that observed in the original dataset, although small firms still make up more than 50% of the sample.

[3] In this respect, some clarification is required. The Xth Capitalia-UniCredit survey questionnaire refers to 2004-2006 and includes information on firm structure, ownership, work force and investments in physical and technological capital, as well as the degree of internationalisation. Although the survey covers the 2004-2006 period, some parts of the questionnaire only refer to 2006. Finally, the survey is complemented with balance sheet data for the 1998-2006 period.

2012a; Brandolini and Cipollone, 2001; Daveri and Jonia-Lasinio, 2005; ISTAT, 2007; OECD, 2007; Van Ark *et al*, 2007).[4]

We estimate TFP by using the Levinsohn and Petrin (2003) approach, which is detailed in appendix B. Data for 2006 are summarised in table 1, where firms are grouped by area, sector and size. As can be seen from the last row of table 1, TFP of the entire sample of firms is, on average, 867 and its variability is high: when considering the entire sample of firms, TFP ranges from 84 to 5,766. The sample also confirms the existence of a sharp economic divide between the South and the rest of the country: TFP in the South is lower than in other areas, underlining that it is technologically lagging, something that has long been debated in the literature (Ascari and Di Cosmo, 2005; Byrne *et al*, 2009; Ladu, 2010). We also find evidence regarding TFP heterogeneity within each area.[5] In brief, data highlight significant TFP differences across macro-regions and significant heterogeneity within each area.

High differences in TFP are confirmed at sector level. On average, the best performing firms are those operating in high-tech sectors, followed by scale intensive and specialised sectors. At one extreme, the lowest TFP is observed for firms in supplier dominated sectors. However, TFP is highly variable even within each individual sector. Indeed, it should be underlined that firms register not only the highest TFP average, but also a narrow range of variability: the minimum and maximum TFP values are 348 and 3,838 respectively, although standard deviation is high just as it is elsewhere.

Other interesting insights come from figures 1 and 2, where regional and sectoral TFP are plotted for the 1998-2006 period. When considering regions, we find that differences in TFP were not uniform over time: they were wide at both the beginning and at the end of the period, whereas a convergence occurred in 2002-2003. This was, though, not so much due to the performance of southern firms, but mainly the result of what happened in the rest of Italy.[6] What is important for the specific scope of this paper is the evidence on empirical regularity regarding the persistence of the TFP divide across region. For instance, net of cyclical tendency, TFP of southern firms is always lower than that observed in the rest of the country. This also applies when referring to figure 2, which provides a picture of persistent and significant TFP differences across Pavitt sectors. The most important evidence is that the TFP gap in traditional sectors revealed in 2006 (see table 1) is persistent over the years 1998-2006. This means that, other things being fixed, firms working in supplier dominated industries always perform worse than those active in other sectors. Figure 2 shows that a sort of sector-fixed-effect exists when analysing the productivity of Italian manufacturing firms.

The main conclusion that can be drawn from this paragraph is that TFP differs greatly across sectors and areas. The following paragraphs will look at some reasons behind the TFP heterogeneity revealed in the data.

---

[4] Some basic elaborations of data confirm that there is a strong correlation between firms' labour productivity and TFP: e.g. in the period 2001-2006 the correlation index was, on average, 0.96.

[5] TFP ranges from 87 to 5,766 in the North-West of Italy and, at the opposite extreme, varies from 84 to 3,275 in the South of the country.

[6] As figure 1 shows, there was a decline in the TFP gap in Italy in 1999-2001. This was mostly due to the dynamics of northern regions, while an improvement in the efficiency of southern firms only took place subsequently. Again, it is important to emphasise that this recovery in the South was short-lived and much more limited than that registered elsewhere (figure 1).

Table 1 Descriptive statistics of firms in the sample (2006)

	Number of firms	% of firms	TFP			
			Mean	Standard deviation	Min	Max
By territoral area						
North West	1333	44,3%	906	549	87	5766
North East	916	30,5%	867	463	198	4350
Center	466	15,5%	827	519	170	5167
South	291	9,7%	751	513	84	3275
By Pavitt Sector						
Supplier dominated	1463	48,7%	769	426	84	5167
Scale intensive	577	19,2%	969	652	120	5766
Specialised suppliers	831	27,6%	925	492	170	5728
Science based	135	4,5%	1130	683	348	3838
By size (employees)						
Small (11-50)	1708	56,8%	721	406	87	5766
Medium (50-250)	1059	35,2%	959	502	84	5016
Large (>250)	239	8,0%	1501	700	342	5167
Total	3006	100,0%	867	518	84	5766

Figure 1 TFP by area from 1998 to 2006

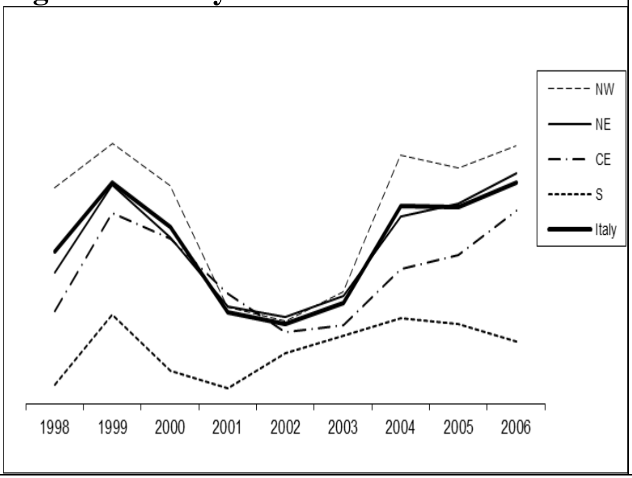
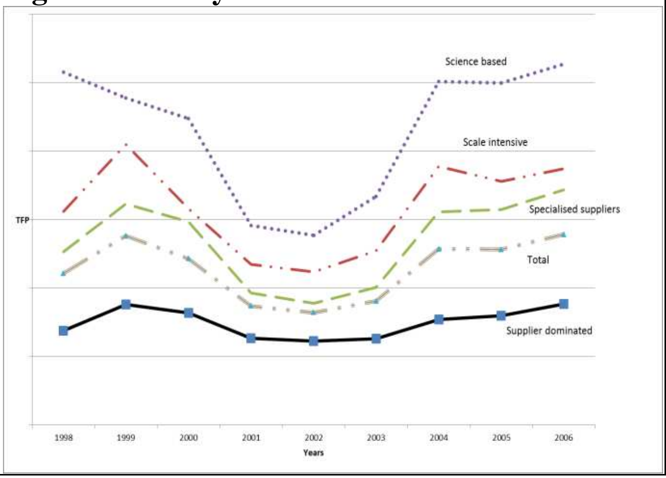


Figure 2 TFP by sector from 1998 to 2006



3 Empirical setting

3.1 Methodology: Cross-Classified Model

Our objective is to take simultaneously into account differences across both sectors and regions which influence the TFP heterogeneity observed at firm-level. One approach that is appropriate for relating a dependent variable to predictor variables at more than one level is multilevel modelling. In a multilevel model, variables at different levels do not simply operate within the one single-level equation, but are linked together in ways that make the simultaneous existence of distinct level-one and level-two equations explicit. In such a way, level-two factors are used not just as independent variables to explain variability in a level-one dependent variable, but also to explain variability in random intercepts and random slopes (Bickel, 2007).

Multilevel models are commonly used to model hierarchically structured data, such as firms within a region or firms operating in the same sector. However, firms may belong to more than one grouping within a hierarchy and each grouping can be a source of random variation. For instance, firms' performances may be affected by both the territorial conditions of the regions where they are located and the specificities of sectors in which they operate. Firms from different sectors may be located in the same region and firms from different regions may operate in one given sector. In this sense, sectors are not nested in regions and regions are not nested in sectors, but, rather, the regions and sectors are crossed with one another. There are two separate two-level hierarchies which cross one another: a firms-within-regions hierarchy and a firms-within-sectors hierarchy. In such a condition, data have a cross-classified structure.[7] To sum up, in models for cross-classified data, a lower-level unit belongs uniquely to one higher-level unit of the first type (e.g. a region) and also uniquely to one higher-level unit of the second type (e.g. a sector), but the two types of unit are not nested in either way.

The consequences of ignoring an important cross-classification are similar to those of ignoring an important hierarchical classification: as a result of cluster specific factors, firms operating in a given region/sector are likely to be more similar than firms located in differing regions/sectors and, thus, the assumption of independence of errors is violated. Moreover, assuming a pure nested structure with a hierarchical linear model as opposed to modeling the cross-classified data structure is particularly problematic if one is interested in evaluating the effect of variables at the ignored level. Standard error estimates of the non-ignored cross-classified factor are overestimated and explanatory variables relating to the ignored cross-classified factor will be identified as having a greater impact than they really do (Meyers and Beretvas, 2006). For example, fitting a firms-within-regions two-level model of TFP while ignoring that firms are simultaneously, but separately, nested within sectors will likely lead to an overstating of the importance of regions as a source of variation in firm TFP. Some of the variation that could be attributed to regions may be better categorised as sector-to-sector differences in TFP. The analysis would, therefore, overstate the importance of regions in firm TFP and ignore the role of sectors. Furthermore, including sector-level predictor variables in our firms-within-regions two-level model, but ignoring sector will typically lead to severe underestimation of the standard errors on these sector-level variables.

A simple cross-classified model can be written as:

$$y_{i(sj)} = \alpha_{0sj} + e_{i(sj)} \qquad \text{with } i = \text{firms} \;\; s\text{=sectors} \;\; j\text{=regions} \qquad [1]$$

where there are two indices at the second level, *s* and *j*, denoting simultaneous nesting in sector *s* and in region *j*. The dependent variable, $y_{i(sj)}$, refers to the *i-th* firm from the *(sj)-th* sector/region combination. In eq. [1], the regression parameters α_{0sj} vary across sectors and regions:

[7] In practice, clusters are not always completely crossed and many combinations of sector and region do not occur.

$$\alpha_{0sj} = \gamma_{000} + u_s + u_j \qquad [2]$$

where γ_{000} is the overall mean, u_s and u_j are random intercepts for sector *s* and region *j*.

Combining eq. [1] and eq. [2] yields the "empty" model:

$$y_{i(sj)} = \gamma_{000} + u_s + u_j + e_{i(sj)} \qquad [3]$$

in eq. [3] the variable $y_{i(sj)}$ is equal to the overall mean γ_{000}, plus a random departure u_s due to sector *s*, a random departure u_j due to region *j*, and an individual-level random departure $e_{i(sj)}$, with $e_{i(sj)} \sim N(0, \sigma_e)$, $u_s \sim N(0, \sigma_s)$ and $u_j \sim N(0, \sigma_j)$.

The random intercept for sector u_s is shared across all regions for a given sector, whereas the random intercept for region u_j is shared by all sectors for a given region. The random intercepts are independent of each other, across sectors and regions and are also uncorrelated with $e_{i(sj)}$.[8]

A useful way to interpret the relative magnitude of the variance components is to compute variance partition coefficients (VPCs). VPC statistics reports the proportion of the response variance that lies at each level of the model hierarchy. The region VPC*j* is calculated as the ratio of the region variance to the total variance, that is:

$$VPC_j = \frac{\sigma_j^2}{\sigma_j^2 + \sigma_s^2 + \sigma_e^2} \qquad [4]$$

For sectoral VPC*s*, we have:

$$VPC_s = \frac{\sigma_s^2}{\sigma_j^2 + \sigma_s^2 + \sigma_e^2} \qquad [5]$$

Finally, the firm VPC*i* is written as:

$$VPC_i = \frac{\sigma_i^2}{\sigma_j^2 + \sigma_s^2 + \sigma_e^2} \qquad [6]$$

The model [3] can be extended to identify which region, sector and firm characteristics might explain parts of these components of variance. In this case, the eq. [1] can be expressed as:

$$y_{i(sj)} = \alpha_{0sj} + \beta_{1(sj)} X_{isj} + e_{i(sj)} \qquad [7]$$

where *X* is a set of firm-level variables.

At level 2, the intercept and slope are modelled such that:

$$\alpha_{0sj} = \gamma_{000} + \gamma_{010} S_s + \gamma_{001} Z_j + u_s + u_j \qquad [8]$$

$$\beta_{1(sj)} = \gamma_{100}$$

[8] In model [3], a random interaction term between sector and region, u_{sj} could be added. However, since we have rather few observations for many of the combinations of sector and region, we are not able to consider this possibility, although it would clearly be the preferred specification.

From eq. [8], the intercept variability may be explained by sectoral predictors (S_s) and a set of regional variables (Z_j). The effect of firm-level variables is supposed to be constant across regions and sectors.

Substituting eq. [8] into eq. [7] we obtain:

$$y_{i(sj)} = \gamma_{000} + \gamma_{100} X_{isj} + \gamma_{010} S_s + \gamma_{001} Z_j + u_s + u_j + e_{i(sj)} \quad [9]$$

Eq. [9] is a random intercept model. The variable $y_{i(sj)}$ depends on a set, *X,* of variables measured at firm level, a set, *S,* of sectoral variables and a set, *Z,* of variables defined at regional level. After introducing these variables into the model, we estimate the residual components of variance. This allows us to have a measurement of how much of the variation in outcomes might be attributable to unobserved factors operating at each level of the model. We can use VPCs to interpret estimated variance parameters in models which include covariates. VPCs, in this case, measure the proportion of outcome variation unexplained by the predictor variables that lie at each level of the model hierarchy (Leckie, 2013).

3.2 Econometric specification

Sample size at any level of analysis is an important issue to be addressed when using multilevel approach. This is because in order to get precise measurement of between-group variance, the estimation of random coefficient variability requires a "sufficient" number of clusters. Although there are some, albeit very different from each other, rules of thumb, a clear result does not exist in this respect (Richter, 2006). Some authors suggest that 20 is a sufficient number of groups (Heck and Thomas, 2000; Rabe-Hasketh and Skondal, 2008), others 30 (Hox, 2002) or 50 (Mass and Hox, 2004). In addition, it is worth noting that in random-effects models the clusters must be sized with at least two observations. The alternative is a fixed-effects approach in which the number of groups is not important, although their dimension becomes crucial as the estimated group-effect is unreliable for small-sized groups.

In our case, there are 20 clusters for regions and 21 for sectors. We decided to use a random effect for both regions and sectors for two reasons. With regards group-size, in some regions the number of firms is small but always higher than (or equal to) two (Valle d'Aosta), while the number of firms is high in other regions (956 in Lombardia). Something very similar is true for sectors, where the number of firms varies from 14 in the "manufacturing of office machinery and computers" to 447 in the "machinery" sectors. This heterogeneity in cluster-size supports our decision to use a random effects model at the level of regions and sectors (see Aslam and Corrado, 2012). Moreover, since we wish to test how sectors and location influence TFP, the random coefficient model should be used. The fixed effects model already explains all of the differences between groups and there is no unexplained between-group variability remaining that can be explained by group-level variables (Snijders and Bosker, 2012).

Estimating the empty cross-classified model (eq. [3]) allows us to understand how much of the variation in outcomes might be attributable to differences between regions, between sectors and between firms after controlling for region and sector effects (Fielding-Goldstein, 2006). Table 2 displays results obtained when estimating eq. [3]. A first result to be discussed refers to the likelihood-ratio test. This compares eq. [3] with the standard linear regression and, being highly significant, supports the use of a multilevel model.[9]

[9] The null hypothesis is that $u_s = u_j = 0$. If the null hypothesis is true, an ordinary regression can be used instead of a cross-classified model.

Table 2 also presents the VPC values attributable to the different sources of variability. The calculations show that 3.4% of the unexplained variation in TFP lies at regional level and 10% at sectoral level, while the remaining variability (86.5%) is explained by internal firm characteristics (column 3). When region alone is incorporated (column 1), this factor explains a relatively high percentage of variance (4.7%). The same thing happens when only sector is considered (11% in column 2). This result indicates that when only one of the two levels of analysis is incorporated into the model as a random effect, it will drew to itself part of the other random effect (Snijders and Bosker, 2012). LR tests which compare the cross-classified model (column 3) to a simpler two-level firms-within-regions model (column 1) and a two-level firms-within-sectors model (column 2) demonstrate that the cross-classified model offers a significantly better fit to the data.

As already said (*cfr* § 3.1), after estimating eq. [3], we extend the model by introducing some region, sector and firm variables that might be important in explaining TFP variance. After introducing these variables, we estimate the residual components of variance. This allows us to evaluate how much of the variation in outcomes might still be attributable to unobserved factors operating at each level of the model.

In line with eq. [9], the empirical analysis is based on the following model[10]

$$\omega_{i(sj)} = \gamma_{000} + \sum_{f=1}^{3} \gamma_f X_{fi(sj)} + \sum_{h=1}^{3} \beta_h Z_{hj} + \sum_{p=1}^{k} \lambda_p S_{pis} + \sum_{q=1}^{2} \eta_q D_{qi} + u_s + u_j + e_{i(sj)} \qquad [10]$$

where ω is the TFP of the *i-th* firm (in logs) operating in sector s and located in region j, X is a vector of firm-level variables which are considered to be important drivers of TFP, Z represents the variables at regional level, S is the set of sectoral variables and D stands for two dummy variables to control for size effect. The variables used in estimating eq. [10] are described in the next paragraph.

[10] The analysis was implemented in Stata using the "xtmixed" subroutine following the suggestions of Rabe-Hasketh and Skondal (2008). All models were estimated by employing the restricted maximum likelihood (REML) over maximum likelihood (ML). This is because the latter is more sensitive to loss of degrees of freedom when dealing with a small number of groups (Bickel, 2007).

Table 2 TFP heterogeneity at firm level. Estimations from empty models

		(1)	(2)	(3)
	Constant	6.588***	6.665***	6.618***
		(0.028)	(0.036)	(0.042)
Variance				
	Regions	0.0105		0,008
		(0.005)		(0.004)
	Sectors		0.024	0,022
			(0.009)	(0.008)
	Firms	0.211	0.197	0,193
		(0.005)	(0.005)	(0.005)
Variance Partition Coefficient (VPC)				
	Regions	4,7%		3,4%
	Sectors		10,9%	10,0%
	Firms	95,3%	89,1%	86,5%
	LR test	61,14	239,11	271,40
	Log likelihood	-1882,9	-1793,9	-1777,7
	LR test model 3 versus model	210,25	32,28	
	Number of observations	2907	2907	2907
		Regions	Sectors	
	N. of groups	20	21	
	N. of firms			
	Min	2	14	
	Max	956	447	
	Average	145,3	138,4	

Dependent variable: log of TFP (average values for the 2004-2006 period). Level of significance: *** 1%. In parentheses standard deviation.

3.3 Data and variables

The data used in this paper are from various sources. As already said (cfr § 2), the firm level data come from the Xth UniCredit-Capitalia survey (2008) of Italian manufacturing firms. With reference to this dataset, we estimate TFP by using the Levinshon and Petrin (2003) approach (see Appendix B for details). In all estimations, TFP is expressed as the average for the 2004-2006 three-year period. This is done to control for annual variability of TFP resulting from the influence of shocks and measurement errors. Furthermore, the average over the 2004-2006 three-year period limits the extent of missing data for TFP due to the estimation of the production function. The analysis regresses firms' TFP on R&D investment in 2004, the number of white collar workers as a share of total workers and export intensity (exports/sales) in 2006.[11] Furthermore, regressions also include two dummy variables to control for size effects (DM refers to medium-sized firms and DL is for large-sized firms, whereas the control group is comprised of small firms).

As far as the regional variables are concerned, we included the R&D intensity of the private sector, an index of infrastructure endowment and the public sector efficiency. R&D intensity is measured as the share of private R&D expenditure in regional GDP in 2004 (data are from ISTAT). The index of total infrastructure (Italy = 100) is from the CNEL (National Council of the Economy and Labour) database Cnelstats[12] and summarises the availability of different kinds of infrastructure which are directly relevant to production (such as roads, railways, telecommunications, ports, water, electricity and airports). The index refers to the year 2004. Following the method used by Golden and Picci (2005), the index of the efficiency of public institutions is retrieved by using the difference between the total amount of financial resources allocated to endow regions with infrastructure and the physical inventory of public capital which has effectively been built, after controlling for regional differences in the cost of public construction. This measure refers to 1997 and has also been used as a corruption index (Golden and Picci, 2005).

The sectoral level data refer to the years 2002-2004 and are retrieved from the 4th Community Innovation Survey (CIS4), except for Researchers and R&D intensity, which come from ISTAT. The CIS-dataset covers innovation activities enterprises with at least ten employees. In particular, CIS collects information on the type of innovation introduced, innovation outputs and inputs, as well as sources, effects, cooperation and hampering factors of innovation.

Three groups of indicators have been considered. The first group is related to the "innovativeness" of the sector and includes six variables: two indicators of innovative activity (researchers as a share of total employment and R&D *intra-muros* expenditure as a share of total turnover) and three indicators that constitute an attempt to gauge the direction of technical change followed by manufacturing industries. The first measures the number of firms that have at least one type of innovation (as a percentage of the number of firms in the sector). The second is the number of new product innovators in any sector (this is expressed as "number of innovative firms"/"total firms in that sector"). The same thing applies for process innovators and for any other type of innovation.

The second group refers to public funding for innovation and considers the share of enterprises which receive R&D public funding with respect to the total number of enterprises working in the

[11] One of the basic empirical facts relating to productivity is a strong positive association between productivity and exporting activity and, therefore, we include the share of exports in total sales (Melitz, 2003; ISGEP, 2008). Similarly, it is widely argued that a firm's performance improves as a result of its innovative behavior and the presence of skilled workers (see, i.e., Krueger and Lindahl, 2001; Sveikauskas, 2007). In the survey carried out by Unicredit-Capitalia, white collar workers and exports only refer to 2006.

[12] The Cnelstats database, built in cooperation with the Guglielmo Tagliacarne Institute, provides both information and statistical indicators on economic trends, the productive network and social situation for Italy and other EU countries (http://www.cnel.it/cnelstats/index.asp).

sector. The four variables of public subsidies are labelled according to the institution which manages the support. Thus, we distinguish between subsidies from (a) local Government, (b) national Government, (c) EU (public funding from regular EU support) and (d) the EU's 5th or 6th Framework Programme for research and technological development. The *INNO support* variable combines all of the previously named sources of funding and is equal to the number of firms using at least one type of public funding for innovation as a proportion of the firms operating in the sector.

The last group of sectoral variables shows the degree to which firms collaborate on innovative projects with any kind of institution (public research institutions or other enterprises).[13]

Table 3 summarises the variables that were used in our study and provides information on their description and sources. Instead, table 4, presents the summary statistics for all variables included in the regression models.

[13] The location of company headquarters is used to link regional indicators to firms. It is worth noting that the unit of analysis in the Capitalia-Unicredit survey is the firm and no information is reported on the number of each firm's establishments. While this limits us to having to make the most of the heterogeneity within the micro-data, there is no possibility of using plant-data level when the source of data is the Capitalia-Unicredit survey. Thus, results have to be interpreted cautiously, although it is also important to bear in mind that more than 50% of our dataset is formed by small-sized firms which are probably single-plant firms (see table 1).

Table 3 Variables included in the study. Description and source

Variable	Description	Data source
Firm level		
TFP	Average of the three-year period 2004-2006	Own Computations
R&D_firm	R&D investments in 2004	UniCredit-Capitalia
Human capital	White collar workers as a share of total workers in 2006	UniCredit-Capitalia
Export intensity	Exports/sales in 2006	UniCredit-Capitalia
Size	DM: medium-sized firms; DL: large-sized firms	UniCredit-Capitalia
Regional level		
INFRA	Economic Infrastructure in 2004	CNEL
R&D_region	Private R&D expenditure over GDP in 2004	ISTAT
Eff_PA	Efficiency of Public Administration in 2004	Golden and Picci (2005)
Sector level		
Innovativeness		
Researchers	Researchers as share of total employment in 2005	ISTAT
R&D intensity	R&D intra-muros expenditure as a share of total turnover in 2004	ISTAT
INNO	Number of firms that have at least one type of innovation as percentage of sector's number of firms (2002-2004)	CIS(4)
Product innovators	Number of new product innovators as percentage of sector's number of firms (2002-2004)	CIS(4)
Process innovators	Number of process innovators as percentage of sector's number of firms (2002-2004)	CIS(4)
Other innovators	Number of other innovators as percentage of sector's number of firms (2002-2004)	CIS(4)
Public financial support for innovation (2002-2004)		
INNO support	A least one type of public funding for innovation	CIS(4)
Local	Firms benefitting from local or regional support as percentage of sector enterprises	CIS(4)
National	Firms benefitting from national support as percentage of sector enterprises	CIS(4)
EU	Firms benefitting from EU support as percentage of sector enterprises	CIS(4)
RTD	Firms benefitting from 5th or 6th RTD support as percentage of sector enterprises	CIS(4)
Cooperation (2002-2004)		
INNO cooperation	Joint innovation activities with others	CIS(4)
University	Joint innovation activities with universities	CIS(4)

Table 4 Descriptive statistics for variables included in the regression model

Variable		Obs	Mean	Std. Dev.	Min	Max
Firm level						
	TFP	2907	866.414	513.293	83.805	5766
	R&D_firms		186.263	1739	0	50000
	Human capital		0,4	0.265	0	1
	Export intensity		29.357	30.499	0	100
Regional level		20				
	INFRA		92,365	37,093	38,6	191,2
	R&D_region		70,45	57,645	3,698	244,078
	Eff_PA		100	43,383	36,235	178,31
Sectorial level		21				
	Innovativeness					
	Researchers		0.264	0.142	0.06	0.51
	R&D intensity		0.01	0.018	0.00012	0.069
	INNO		0.406	0.121	0.152	0.638
	Product innovators		0.239	0.145	0.035	0.494
	Process innovators		0.25	0.582	0.893	0.328
	Other innovators		0.175	0.054	0.08	0.313
	Public financial support					
	INNO support		0.204	0.083	0.057	0.349
	Local		0.120	0.057	0.024	0.253
	National		0.098	0.046	0.033	0.212
	EU		0.027	0.022	0	0.084
	RTD		0.0152	0.0163	0	0.0579
	Cooperation					
	INNO cooperation		0.081	0.063	0.013	0.217
	University		0.04	0.036	0	0.131

See table 3 for *legenda* and sources.

4. Results

Estimations are displayed in tables 5 and 6. The first evidence regards the impact exerted by firms' characteristics as explanatory factors of TFP: the estimated parameters are invariant in all models and not substantially different from those obtained by other researchers. R&D and human capital lead to higher firm TFP since they directly affect the possibility to introduce and use more productive processes and, hence, translate innovation efforts into profitable opportunities. This is in line with the literature (see, e.g. Griliches 2000; Parisi *et al*, 2006). Secondly, we find that TFP tends to increase with exports, as shown by, among many others, Benfratello and Razzolini (2008), ISGEP (2008) and Serti and Tomasi (2008). Thirdly, the positive coefficients associated with the DM and DL dummies highlight the role of size in TFP. Medium-sized firms perform better than small firms, but less well than large enterprises, indicating that economies of scale are at work.

Firm performance is influenced by the territorial specificities of the regions where they operate: all of the variables defined at regional level have a positive effect on TFP.[14] This is consistent with the results of a previous work (Aiello *et al*, 2013). The endowment of regional infrastructure influences TFP positively and this is in line with the conclusions drawn by, for instance, Aiello *et al* (2012b) and Marrucu and Paci (2010). Furthermore, regional private R&D activity has a positive impact on firms' TFP: being located in a region with high innovation-creating potential helps individual firms perform better (Camagni 1991; Ciccone and Hall, 1993). Finally, firms operating in regions with efficient public institutions benefit from a reduction in the transaction costs they face when introducing more productive activities and creating an environment which is conducive to growth. As far as the random part of the model is concerned, all variance across regions is detected by the regional level predictors used in the analysis.

After considering the role of firm characteristics and geography, we focus on the relationship between firms' TFP and sector-specific technological characteristics. What the innovation literature says is that firms' performances depend on sector-specificities: in every sector there are striking differences among innovation actors and heterogeneous sources of knowledge. All this acts, at firm-level, as a stimulus for large firms to introduce technological improvements or, for others, to render the use of productive factors more efficient for any level of knowledge (Castellacci and Zheng, 2010; Malerba, 2005; Ngai and Samaniego, 2011; Van Dijk, 2000). Therefore, we expect operating in sectors with high innovation-creating potential to make individual firms perform better. Following this, the empirical approach explores the relationship between innovation and productivity by using some sectoral variables which are related to specific aspects of innovation, such as different degrees of innovativeness and different participation in public R&D programs and cooperative R&D agreements.

As expected, the sectoral indicators of innovation are highly correlated (see appendix C) and this precludes the possibility of using them in a single regression. Hence, the potential collinearity bias is addressed by including one sectoral variable at a time in the model. Table 5 refers to the evidence found when considering the indicators capturing innovativeness and industries' direction of technical change, while table 6 displays the results obtained when the regressions focus on public funding for innovation and on the degree to which firms cooperate in innovative activities.

It emerges from table 5 that all sectoral variables except for R&D intensity have a significant effect on firms' TFP.

---

[14] As displayed in Appendix C, the regional indicators are not correlated.

With regards the inputs of innovation, the estimations indicate that firms' TFP is positively related to the share of researchers working in the same sector (model 1), whereas the R&D intensity does not exert a significant effect (model 2). It seems that firms benefit more from people than from the intensity of R&D expenditures. This is confirmed by sectoral VPC, which is at its lowest level (4,1%) in model 1. In interpreting this, one should bear in mind that innovation activity in Italy operates across several dimensions besides R&D investments and this is particularly true for small firms, which are still a relevant part of the fragmented Italian production system (Bugamelli et al 2012; Hall et al, 2009). Italian firms have a knowledge paradigm based on innovation without research, i.e. on the recombination and re-use of knowledge practices. Hence, highly qualified employees allow to produce firms-specific technology but also to widespread knowledge across firms. What the evidence suggests is that the role of researchers in diffusing sectoral knowledge is more significant than formal R&D expenditures.

When referring to the outcome of the innovation process, we find that firms obtain benefits from belonging to sectors with a high proportion of firms which introduce innovation: firm's TFP increases when the number of innovators in the concerned sector is high. This is shown by the coefficient associated with the variable relating to the proportion of firms which introduce at least one type of innovation (product, process or other innovations) into the sector: the estimated parameter of *INNO* is 0.55 and its significance is 1% (table 5, model 3). The same applies as far as the variables *Product Innovators*, *Process Innovators* and *Others* are individually concerned (models 4, 5, 6).[15] To some extent, this evidence indicates the presence of likely external effects upon firms which gain in TFP from operating in an environment where there is a high number of innovations introduced by sector-competitors.

Moreover, when considering the degree of cooperation in an innovating project, estimations indicate that the higher the share is of firms in a sector being involved in innovation cooperation, the higher the TFP is. Participation in innovative projects reduces unexplained industry variability to 5.7%. Put differently, participation in innovative projects explains more than 3% of industry variability (table 6, model 1). Our results are consistent with those of Medda *et al* (2005) who, when considering a sample of Italian manufacturing firms over the period 1995-1997, find that there are positive returns to collaborative research with other companies and research centres. Similar evidence is provided by Belderbos *et al* (2004) for Dutch firms over the 1996–1998 period. Obviously, a detailed comparison with these studies is not possible because they use different firm level data and different methodologies. However, what clearly emerges is the empirical regularity regarding the positive influence on firms' productivity of participation in innovative projects. This also applies when we consider the effect of cooperation with universities (table 2, model 2). In this case, our findings differ from those of Medda *et al* (2005), but are in line with the positive effect identified by Belderbos *et al* (2004) and Harris *et al* (2011) for Dutch and British firms respectively.

Moving on to the relationship between firms' TFP and participation in R&D policies, we find that a positive impact on firms' TFP is exerted by the share of enterprises receiving at least one public funding for innovation. Indeed, the coefficient associated with the variable *Inno Support* is 0.45 and highly significant (table 6, model 3). The importance for a firm of operating in sectors whose competitors widely use R&D policy is reflected in the VPC index, which now decreases to about 7.2% (see model 3 in table 6). In other words, *Inno Support* absorbs much of the unexplained variance due to industry membership: this explains around 3% of industry variability. When splitting *Inno Support* on the basis of which institution manages the R&D program, positive and significant parameters have been found for *National*, *EU* and *RTD* indicators, whereas the impact of

[15] Results are not substantially different when jointly considering the number of product and service innovators or the number of process and logistic innovators (these results available upon request).

regional policies (*Local*), although still positive, is of weak significance. In terms of size effect, we find that the strength of the link between participation in R&D programs and firms' TFP increases as the complexity of research funded projects increases: this can be seen by looking at the coefficient associated with each variable, which is 1.6 for national support (model 5), 2.14 for policies set up by the EU (model 6) and 3.04 for the research initiatives of the EU innovation framework program (model 7). A firm obtains no benefit from participation in regional R&D programs, probably because these location-based policies usually have a low threshold of requirements for access and are widely used, involve R&D projects with cross-sector objectives, mobilise small amounts of finance and, thus, only marginally influence the innovative system. At the other extreme, EU programs are more sector-specific and their complexity entails the introduction of more radical innovations (Cefis and Evangelista, 2007). Some considerations can be drawn from this. Bearing in mind the period studied, the low estimated value for the *Local* variable might be caused by the absence of regional political autonomy in innovation policy until 2002 and by limited financial resources (Muscio *et al*, 2013). Again, Italian regions do not use a unified model to implement innovation policies and coordination between national and regional authorities is difficult. Whilst regional governments now have competence in the area of innovation policy, national government might promote national initiatives that are similar to those promoted by regions (Ciffolilli, 2010). This does not necessarily have a negative impact on innovation, but it certainly creates some extra coordination problems due to overlaps and gaps.

Table 5 Explaining TFP of Italian manufacturing firms for 2004-06: multilevel regressions. The role of sectoral innovativeness

	(1)	(2)	(3)	(4)	(5)	(6)
Fixed Effects						
Firms characteristics						
R&D_firm	0.00002***	0.00002***	0.00002***	0.00002***	0.00002***	0.00002***
	(4.61)	(4.61)	(4.54)	(4.63)	(4.63)	(4.63)
Human capital	0.200***	0.199***	0.198***	0.201***	0.198***	0.199***
	(7.16)	(7.10)	(7.08)	(7.21)	(7.08)	(7.11)
Export intensity	0.00079***	0.00078***	0.00078***	0.00080***	0.00079***	0.00079***
	(3.18)	(3.12)	(3.12)	(3.21)	(3.17)	(3.15)
DM	0.288***	0.287***	0.288***	0.287***	0.287***	0.287***
	(18.38)	(18.37)	(18.38)	(18.37)	(18.37)	(18.37)
DL	0.672***	0.673***	0.673***	0.673***	0.674***	0.674***
	(24.15)	(24.20)	(24.20)	(24.18)	(24.21)	(24.21)
Regional characteristics						
INFRA	0.00187***	0.00188***	0.00181***	0.00191***	0.00189***	0.00189***
	(4.30)	(4.34)	(4.08)	(4.47)	(4.37)	(4.38)
R&D_region	0.0549*	0.0541*	0.0539*	0.0544*	0.0536*	0.0533*
	(1.78)	(1.77)	(1.68)	(1.82)	(1.75)	(1.75)
Eff_PA	0.00123***	0.00122***	0.00124***	0.00123***	0.00123***	0.00123***
	(4.03)	(4.03)	(3.96)	(4.11)	(4.06)	(4.06)
Sectoral characteristics - *Innovativeness*						
Researchers	0.670***					
	(4.49)					
R&D intensity		1.962				
		(1.20)				
INNO			0.555***			
			(2.64)			
Product innovators				0.413**		
				(2.28)		
Process innovators					0.861*	
					(1.89)	
Other innovators						1.230**
						(2.51)
Constant	5.832***	5.978***	5.771***	5.899***	5.774***	5.781***
	(78.43)	(87.21)	(54.59)	(76.09)	(43.81)	(54.32)
Random effects						
Variance						
Regions	0,000	0,000	0,000	0,000	0,000	0,000
Sectors	0,006	0,014	0,010	0,011	0,012	0,011
Firms	0,148	0,148	0,148	0,148	0,148	0,148
Variance Partition Coefficient (VPC)						
Regions	0,3%	0,2%	0,2%	0,2%	0,2%	0,2%
Sectors	4,1%	8,5%	6,3%	6,9%	7,5%	6,6%
Firms	95,6%	91,3%	93,5%	92,9%	92,3%	93,2%
Log likelihood	-1407,2	-1411,3	-1411,1	-1411,9	-1411,6	-1410,5
Observations	2907	2907	2907	2907	2907	2907

Dependent variable: log of TFP (average values for 2004-2006 period). In parentheses, t-values. Level of significance: *** 1%, ** 5%, * 10%.

Table 6 Explaining TFP of Italian manufacturing firms for 2004-06: multilevel regressions. The role of R&D policy support and firms' collaboration in Innovative Projects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Fixed Effects							
Firms characteristics							
R&D_firm	0.00002***	0.00002***	0.00002***	0.00002***	0.00002***	0.00002***	0.00002***
	(4.63)	(4.61)	(4.59)	(4.59)	(4.61)	(4.60)	(4.59)
Human capital	0.200***	0.198***	0.200***	0.200***	0.199***	0.197***	0.199***
	(7.16)	(7.10)	(7.16)	(7.16)	(7.12)	(7.06)	(7.10)
Export intensity	0.00080***	0.00076***	0.00079***	0.00079***	0.00078***	0.00077***	0.00076***
	(3.20)	(3.06)	(3.16)	(3.15)	(3.14)	(3.09)	(3.06)
DM	0.288***	0.287***	0.287***	0.287***	0.287***	0.287***	0.287***
	(18.37)	(18.35)	(18.37)	(18.37)	(18.36)	(18.36)	(18.36)
DL	0.672***	0.673***	0.673***	0.672***	0.673***	0.674***	0.673***
	(24.16)	(24.19)	(24.17)	(24.16)	(24.19)	(24.21)	(24.18)
Regional characteristics							
INFRA	0.00190***	0.00188***	0.00187***	0.00186***	0.00188***	0.00187***	0.00186***
	(4.40)	(4.37)	(4.30)	(4.26)	(4.34)	(4.27)	(4.26)
R&D_region	0.0550*	0.0531*	0.0544*	0.0545*	0.0539*	0.0538*	0.0537*
	(1.81)	(1.75)	(1.76)	(1.75)	(1.76)	(1.73)	(1.74)
Eff_PA	0.00123***	0.00123***	0.00122***	0.00122***	0.00123***	0.00122***	0.00123***
	(4.07)	(4.09)	(4.01)	(3.99)	(4.04)	(3.99)	(4.02)
Sectoral characteristics							
Cooperation							
INNO cooperation	1.196***						
	(3.08)						
University		2.020***					
		(3.09)					
Pubblic financial support							
INNO support			0.455**				
			(2.05)				
Local				0.391			
				(0.72)			
National					1.604***		
					(3.21)		
EU						2.141*	
						(1.66)	
RTD							3.045*
							(1.79)
Constant	5.906***	5.920***	5.879***	5.946***	5.840***	5.941***	5.954***
	(82.78)	(85.23)	(68.28)	(64.19)	(73.15)	(80.33)	(84.42)
Random effects							
Variance							
Regions	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Sectors	0,009	0,009	0,012	0,014	0,009	0,013	0,013
Firms	0,148	0,148	0,148	0,148	0,148	0,148	0,148
Variance Partition Coefficient (VPC)							
Regions	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%
Sectors	5,7%	5,7%	7,2%	8,8%	5,5%	7,9%	7,8%
Firms	94,1%	94,0%	92,6%	90,9%	94,3%	91,9%	92,0%
Log likelihood	-1409,6	-1409,1	-1412,1	-1412,9	-1409,1	-1410,9	-1410,4
Observations	2907	2907	2907	2907	2907	2907	2907

Dependent variable: log of TFP (average values for 2004-2006 period). In parentheses, t-values. Level of significance: *** 1%, ** 5%, * 10%.

5. Conclusions

This study analyses to what extent sector affect the TFP of Italian manufacturing firms. To this end, we combine a dataset of Italian firms with some *meso* regional and sectoral variables and apply a cross-classified model that allows us to evaluate the relative variability of firm specific *versus* sectoral and geographical factors. Our main results can be summarised as follows.

The first outcome is the measurement of the degree to which the effects of firms' heterogeneity and sectoral and territorial conditions explain firms' performance. In fact, while there is a wide consensus regarding the importance of these aspects in influencing the firm's capabilities of being productive, a quantitative measure of the geographic and sectoral effects on firms' TFP has been lacking. We find that 86% of firms' TFP is explained by internal firm characteristics and 3.4% by location, while the sector-fixed effects explain 10% of firms' productivity. This result confirms that the main source of firm variety are differences at individual level, and on the other hand, suggests that sectors play a more prominent role than territory. Some policy implications emerge from this because a great deal of public support in Italy is activated in order to reduce regional gaps, while less attention has been paid to industrial policy. For instance, a great deal of EU policy effort is based on horizontal intervention which is neutral to sectoral support requirements. Examples of such a horizontal approach include energy, R&D, regional development, SMEs, training and employment, the sum of which constitute almost all of the funds mobilised by EU-25.

Other results regard the impact of specific sectoral variables. We find that many proxies of sectoral innovation play a crucial role in influencing firms' TFP. Overall, firms operating in a sector with a high propensity to innovate obtain significant benefits in terms of TFP. In some ways, it seems that a spillover effect is at work. In particular, we find that the output indicators of sector innovativeness (product innovations, process innovations, other innovations etc.) do affect individual productivity. When considering the inputs of the innovation process, we find that firms' TFP is positively related to the presence of skilled personnel in the sector, whereas the R&D intensity does not exert a significant effect. This result confirms the evidence that innovation in Italy exhibits some peculiar features that R&D investments do not capture: Italian firms have a knowledge paradigm based on innovation without research, i.e. on the recombination and re-use of knowledge practices, rather than on usual R&D expenditures.

When considering the other sectoral variables, an important role is assigned to innovation that takes place in projects which are carried out jointly with other companies or with universities, so confirming that there are positive returns to collaboration in innovative projects. This result has clear policy implications, especially when considering that university-industry collaboration is a recent phenomenon in Italy, while it is long established in other countries (for example, the Centres of Expertise in Finland and the *Poles de Compétitivité* in France).

Finally, our evidence suggests that a high proportion of enterprises receiving public funding for innovation has a positive impact on TFP. When looking at the institutions which manage R&D programs, we find that the impact of National, EU and RTD policies is positive and significant, while a limited role is played by regional strategies. Phrased differently, firms' TFP is neutral to the fact that the sector exhibits a large use of local R&D support. This fact merits further study given the increasing role assigned to regions by the EU in programming and implementing R&D policy.

One conclusion that might be drawn from this paper regards the necessity to apply industrial policies without taking regional economic divide into consideration because sectors matter much more than location. More in general, this essay just suggests the need for a recovery of sectoral policy, which has been left out of the policy agenda for too long. There is a need to implement selective processes and support innovation in high-performing sectors. However, the asymmetries across regions still render the adoption of a place-based policy necessary in order to ensure that firms which operate in the poorer regions will increase their efficiency and approach the technological frontier.

While the empirical results of this article do provide original evidence for debate on productivity, it is nevertheless important to emphasise that some issues deserve more attention. For instance, we focus on manufacturing industries, while the service sectors also play an important role by providing knowledge to manufacturing industries and encouraging growth in advanced countries. Secondly, more detailed micro-datasets than ours would be useful in order to relax the static nature of this study. These are left for future research.

Appendix A - Sample composition by region and Sector

Table A1 Sample composition at regional level

Region	N. firms
Abruzzo	61
Basilicata	5
Calabria	9
Campania	80
Emilia-Romagna	363
Friuli Venezia-Giulia	101
Lazio	67
Liguria	30
Lombardia	979
Marche	125
Molise	6
Piemonte	315
Puglia	58
Sardegna	31
Sicilia	34
Toscana	216
Trentino Alto-Adige	40
Umbria	56
Valle D'Aosta	2
Veneto	407
Total	2985

Table A2 Sample composition at sectoral level

Sector Ateco-91 Classification (2 digit)	N. firms
Manufacture of food products, beverages and tobacco	261
Manufacture of textiles	226
Manufacture of wearing apparel	87
Manufacture of leather and leather products	102
Manufacture of wood and of products of wood and cork, except furniture	76
Manufacture of paper and paper products	78
Publishing, printing and reproduction of recorded media	80
Manufacture of coke and refined petroleum products	16
Manufacture of chemicals and chemical products	173
Manufacture of rubber and plastic products	174
Manufacture of other non-metallic mineral products	194
Manufacture of basic metals and fabricated metal products	140
Manufacture of fabricated metal products, except machinery and equipment	386
Manufacture of machinery and equipment n.e.c.	458
Manufacture of office machinery and computers	14
Manufacture of electrical machinery and apparatus n.e.c.	121
Manufacture of radio, television and communication equipment and apparatus	48
Manufacture of medical, precision and optical instruments, watches and clocks	73
Manufacture of motor vehicles, trailers and semi-trailers	46
Manufacture of other transport equipment	33
Manufacture of furniture; manufacturing n.e.c.	199
Total Manufacturing	**2985**

Appendix B - A measure of TFP

TFP at firm level is estimated by using the Levinshon and Petrin (2003) approach. Productivity was estimated using the following log-linear specification of a production function:

$$y_{it} = \beta_0 + \beta_K k_{it} + \beta_l l_{it} + u_{it} \qquad \text{(A1)}$$

with $i = 1, \ldots\ldots N$ firms, t = 1998, ……2006 and where *y* represents the value added, *l* the number of employees, k the stock of physical capital, β_0 measures the average efficiency and u_{it} represents the deviation from this average of firm *i* at time *t*. The error term can be decomposed into two parts:

$$u_{it} = \omega_{it} + \eta_{it} \qquad \text{(A2)}$$

where the term ω_{it} represents the productivity of firm *i* at time *t* and η_{it} is a stochastic term which includes not only the measurement error, but also the shocks which are unobservable to firms and, therefore, do not correlate with inputs.

Productivity ω_{it} is known to firms, which, therefore, in the case of positive shocks to productivity, can decide to increase production by raising the level of inputs. This causes a problem of simultaneity which Levinshon and Petrin (2003) resolved by identifying the demand for intermediate goods (m_{it}) as a proxy for the variations in TFP known to firms.

In order to include m_{it} in the final equation to be estimated, Levinshon and Petrin (2003) assume that m_{it} depends on k_{it} and ω_{it}, namely $m_{it}=f(\omega_{it}+k_{it})$. In addition, if this function is invertible, then ω_{it} may be expressed in terms of observable variables, that is $\omega_{it}=h(m_{it}+k_{it})$. After substitutions, one obtains:

$$\begin{aligned} y_{it} &= \beta_0 + \beta_l l_{it} + \beta_k k_{it} + h(m_{it}, k_{it})_{it} + \eta_{it} \\ &= \beta_l l_{it} + \phi_{it}(m_{it}, k_{it}) + \eta_{it} \qquad \text{(A3)} \end{aligned}$$

with

$$\phi_{it}(m_{it}, k_{it}) = \beta_0 + \beta_k k_{it} + h(m_{it}, k_{it}) \qquad \text{(A4)}$$

The estimates of β_l and β_k are obtained by applying a two-steps procedure as proposed by Levinsohn and Petrin (2003)

From a empirical perspective, this analysis has been carried out by utilising tangible fixed assets as a proxy for the stock of physical capital and demand for intermediate goods has been measured by using operating costs. The value added has been deflated by using the ISTAT production price index available for each ATECO sector. As regards tangible fixed assets, data have been deflated by using the average production price indices of the following sectors: machines and mechanical appliances, electrical machines and electrical equipment, electronics and optics and means of transport. For the operating costs, we adopted the intermediate consumption deflator calculated by using data from ISTAT.

Appendix C – Correlation matrix

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	TFP	1																							
2	R&D_firm	0,16	1																						
3	Human capital	0,11	0,01	1																					
4	Export intensity	0,15	0,06	-0,04	1																				
5	d_m	0,20	-0,01	-0,11	0,13	1																			
6	d_g	0,36	0,19	-0,03	0,13	-0,22	1																		
7	R&D_region	0,11	0,00	0,06	0,06	-0,03	0,02	1																	
8	INFRA	0,08	0,02	0,03	0,04	-0,04	0,00	0,10	1																
9	Eff_PA	0,09	0,00	0,04	0,09	0,00	0,01	0,33	-0,16	1															
10	Researchers	0,23	0,09	0,15	0,01	-0,01	0,04	0,08	0,05	-0,03	1														
11	R&D intensity	0,13	0,05	0,11	0,08	0,00	0,03	0,10	0,10	0,01	0,59	1													
12	INNO	0,21	0,04	0,14	0,13	0,03	0,02	0,18	0,10	0,05	0,62	0,46	1												
13	Product innovators	0,22	0,05	0,15	0,17	0,03	0,03	0,16	0,09	0,07	0,68	0,54	0,93	1											
14	Process innovators	0,14	0,03	0,04	0,02	0,03	0,00	0,16	0,08	0,00	0,33	0,10	0,75	0,48	1										
15	Other innovators	0,18	0,04	0,15	0,10	0,03	0,00	0,15	0,07	0,02	0,66	0,37	0,92	0,84	0,70	1									
16	Product and service innovators	0,22	0,04	0,16	0,17	0,03	0,03	0,17	0,09	0,07	0,66	0,53	0,94	1,00	0,49	0,8	1,00								
17	Process and logistic innovators	0,16	0,03	0,06	0,03	0,03	0,00	0,16	0,08	0,00	0,42	0,15	0,81	0,57	0,99	0,77	0,58	1							
18	INNO support	0,21	0,06	0,15	0,14	0,03	0,03	0,19	0,09	0,05	0,69	0,58	0,89	0,92	0,58	0,82	0,91	0,65	1						
19	Local	0,13	0,03	0,12	0,12	0,04	-0,01	0,16	0,07	0,05	0,49	0,37	0,85	0,83	0,63	0,83	0,84	0,69	0,91	1					
20	National	0,26	0,06	0,15	0,15	0,03	0,06	0,19	0,11	0,04	0,66	0,58	0,86	0,89	0,54	0,68	0,89	0,60	0,91	0,71	1				
21	EU	0,15	0,06	0,13	0,07	0,01	0,02	0,13	0,05	0,01	0,74	0,53	0,57	0,62	0,33	0,64	0,60	0,39	0,79	0,64	0,58	1			
22	RTD	0,17	0,07	0,13	0,11	0,02	0,03	0,15	0,09	0,04	0,69	0,71	0,57	0,67	0,25	0,58	0,66	0,31	0,83	0,66	0,65	0,92	1		
23	INNO cooperation	0,24	0,07	0,17	0,13	0,02	0,04	0,15	0,09	0,04	0,78	0,57	0,87	0,92	0,50	0,80	0,92	0,58	0,90	0,75	0,88	0,71	0,74	1	
24	University	0,25	0,07	0,16	0,15	0,02	0,05	0,17	0,10	0,03	0,73	0,66	0,86	0,90	0,51	0,74	0,89	0,57	0,88	0,67	0,94	0,63	0,69	0,96	1

References

Ascari G., Di Cosmo V. (2005), Determinants of total factor productivity in the Italian regions, Scienze Regionali – Italian Journal of Regional Science, 4, 27–49.

Aiello F., Pupo V., Ricotta F. (2012a), A territorial analysis of TFP in Italy, Scienze Regionali - Italian Journal of Regional Science, 11, 2, 23-46.

Aiello F., Iona A., Leone L. (2012b), Regional Infrastructure and Firm Investment. Theory and Empirical Evidence for Italy, Empirical Economics, 42, 835–862.

Aiello F., Scoppa V. (2000), Uneven regional development in Italy: explaining differences in productivity levels, Giornale degli Economisti e Annali di Economia, 60, 270–298.

Aiello F., Pupo V., Ricotta F. (2013), Explaining TFP at firm level in Italy. Does location matter? Spatial Economic Analysis (forthcoming).

Aslam A., L. Corrado (2012), The Geography of Well-Being, Journal of Economic Geography, 12, 3, pp. 627-649.

Bank of Italy (2009), Rapporto sulle tendenze nel sistema produttivo italiano, Questioni di Economia e Finanza, n. 45, Rome.

Belderbos R., Carree M. , Lokshin B. (2004), Cooperative R&D and firm performance, Research Policy, 33, 1477–92.

Benfratello L., Razzolini T. (2008), Firms' productivity and internationalization choices: evidence for a large sample of manufacturing firms. In Piscitello L., Santangelo G. (eds.), Multinationals and local competitiveness, Franco Angeli, Milan.

Bickel R. (2007), Multilevel Analysis for Applied Research, It's Just Regression!, The Guilford Press, New York.

Bloom N., Van Reenen J. (2010), Why Do Management Practices Differ across Firms and Countries?, Journal of Economic Perspectives, 24, 1, 203–224.

Byrne J. P., Fazio G., Piacentino D. (2009), Total Factor Productivity convergence among Italian Regions: Some evidence from panel unit root tests, Regional Studies, 43,1, 63-76.

Brandolini A., Cipollone P. (2001), Multifactor productivity and Labour Quality in Italy, 1981-2000, in Temi di Discussione, n. 422, Bank of Italy, Rome.

Bugamelli M., Cannari L., Lotti F., Magri S., 2012, Il gap innovativo del sistema produttivo italiano: radici e possibili rimedi, Questioni di Economia e Finanza, Occasional Papers, N. 121, April, Bank of Italy, Rome.

Camagni R. (1991), Innovation Networks: Spatial Perspectives (eds.), Belhaven Press, London.

Castellacci F., Zheng J. (2010), Technological regimes, Schumpeterian patterns of innovation and firm-level productivity growth, Industrial and Corporate Change, 19, 1829–1865.

Ciccone A., Hall R.E. (1993), Productivity and the Density of Economic Activity, American Economic Review, 86, 54-70.

Ciffolilli A. (2010), Expert Evaluation Network Delivering Policy Analysis on the Performance of Cohesion Policy 2007-2013 - Task 1: Policy Paper on Innovation Italy, A Report to the European Commission Directorate-General Regional Policy, August.

Daveri F., Jona Lasinio C. (2005), Italy's decline: getting the facts right, Giornale degli Economisti e Annali di Economia, 64, 4, 365-410.

Cefis E. , Evangelista R. (2007) La valutazione delle politiche per l'innovazione: un confronto tra Italia e Paesi Bassi, L'industria, XXVII, 2, 243-264.

Fazio G., Piacentino D. (2010), A Spatial Multilevel Analysis of Italian SMEs' Productivity, Spatial Economic Analysis, 5, 3, 299-316.

Fielding A., Goldstein H. (2006), Cross-classified and Multiple Membership Structures in Multilevel Models: An Introduction and Review, Department for Education and Skills, Research Report RR791, London.

Golden M., Picci L. (2005), Proposal for a New Measure of Corruption, Illustrated with Italian Data, Economics & Politics, 17, 37–75.

Griliches, Z. (2000), R&D, Education and Productivity A Retrospective. Harvard University Press, Cambridge, Massachusetts.
Hall B. H. , Lotti F., Mairesse J., 2009, Innovation and productivity in SMEs: empirical evidence for Italy, Small Bus Econ, 33, 13–33.
Harris R., Li Q.C., Moffat J. (2011), The impact of higher education institution-firm knowledge links on firm-level productivity in Britain, Applied Economics Letters, 18:13, 1243-1246.
Heck R.H., Thomas, S.L. (2000), An introduction to multilevel modelling techniques. Lawrence Erlbaum Associates, Mahwah, New Jersey.
Hox J.J. (2002), Multilevel analysis: techniques and applications. Lawrence Erlbaum Associates, Mahwah, New Jersey.
ISGEP, International Study Group on Exports and Productivity (2008), Understanding Cross-Country Differences in Exporter Premia: Comparable Evidence for 14 Countries, Review of World Economics, 144, 4, 596 – 635.
ISTAT (2007), Misure di produttività, in Statistiche in breve, October 5, Rome.
ISTAT (2010), Conti Economici Regionali, 1995-2009, Rome.
Iuzzolino G., Pellegrini G. Viesti G. (2011), Convergence among Italian Regions, 1861-2011, Economic History Working Papers, n. 22, Bank of Italy, Rome.
Kaniovski, S., Peneder, M. (2002), On the structural dimension of competitive strategy, Industrial and Corporate Change 11, 3, 257–279.
Krueger A.B., Lindahl M. (2001), Education for Growth; Why and for Whom? Journal of Economic Literature, 39, 4, 1101-1136.
Krugman P. (1999), The Role of Geography in Development, International Regional Science Review, 22, 2, 142-161
Ladu M. G. (2010), Total factor productivity estimates: some evidence from European regions, WIFO Working Papers, 380.
Leckie, G. (2013). Cross-Classified Multilevel Models - Concepts. LEMMA VLE Module 12, 1-60. http://www.bristol.ac.uk/cmm/learning/course.html.
Levinsohn J., Petrin, A. (2003), Estimating production functions using inputs to control for unobservables, Review of Economic Studies, 7, 317-341.
Malerba F. (2005), Sectoral Systems: How and Why Innovation Differs across Sectors. In: J. Fagerberg, D. Mowery, R.R. Nelson (eds) The Oxford Handbook of Innovation, Oxford University Press, Oxford.
Marrocu E., Paci R. (2010), The effects of public capital on the productivity of Italian regions, Applied Economics, 42, 989-1002.
Mass, J.M. and Hox, J.J. (2004), Robustness issues in multilevel regression analysis. Statistica Neederlandica, 58, 127-137.
McGahan A. M., Porter M.E. (1997), How Much Does Industry Matter, Really?, Strategic Management Journal, 18, 15–30
Medda G., Piga C. , Siegel D. S. (2005), University R&D and firm productivity: evidence from Italy, Journal of Technology Transfer, 30, 199–205.
Meyers J.J, Beretvas S.N. (2006), The impact of Inappropriate Modeling of Cross-Classified Data Structures, Multivariate Behavioral research, 41, 4, 473-497.
Melitz M. J. (2003), The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity, Econometrica, 71, 6, 1695–1725.
Muscio A., Quaglione D., Vallanti G. (2013), Does government funding complement or substitute private research funding to universities?, Research Policy, 42, 1, 63-75.
Nelson, R.R., Winter, S.G., 1982. An Evolutionary Theory of Economic Change. Harvard University Press, Cambridge.
Ngai L. R, Samaniego R. M. (2011), Accounting for research and productivity growth across industries, Review of Economic Dynamics, 14, 3, 475–495.
OECD (2007), Economic Survey: Italy, OECD, Paris.

Parisi M. L., Schiantarelli F., Sembenelli A. (2006), Productivity, Innovation and R&D: Micro Evidence for Italy, European Economic Review, 50, 8, 2037-2061.

Quatraro F. (2006), Technological change and productivity growth in Italian regions, 1982-2001, Région et Développement, n.24.

Rabe-Hesketh S., Skrondal A. (2008), Multilevel and Longitudinal Modeling Using Stata. Stata Press, 2nd edition.

Raspe O., van Oort F. (2011), Firm heterogeneity, productivity and spatially bounded knowledge externalities, Socio-Economic Planning Sciences, 45, 38-47.

Richter T. (2006), What is wrong with ANOVA and multiple regression? Analyzing sentence reading times with hierarchical linear models, Discourse Processes, 41, 3, 221-250.

Schmalensee, R. (1985), Do markets differ much? The American Economic Review, 75, 3, 341–351.

Serti F., Tomasi C. (2008), Self-Selection and Post-Entry Effects of Exports: Evidence from Italian Manufacturing Firms, Review of World Economics, 144, 4, 660–694.

Snijders T.A.B., Bosker R.J. (2012), Multilevel Analysis. An Introduction to Basic and Advanced Multilevel Modeling, 2nd Edition, SAGE Publications Ltd, London.

Sveikauskas L. (2007), R&D and Productivity Growth: a Review of the Literature, US Bureau of Labour Statistics, BLS Working Papers n. 208.

Teece D. J., Pisan G., Shuen A. (1997), Dynamic Capabilities and Strategic Management, Strategic Management Journal, 18,7, 509-533.

UniCredit-Capitalia (2008), Decima indagine sulle imprese manifatturiere italiane. Rapporto Corporate, n.1.

Van Ark B., O'Mahony M., Ypma G. (2007), Productivity in the European Union: A comparative Industry Approach. In The EU KLEMS Productivity Report, n. 1.

Van Dijk, M. (2000), Technological regimes and industrial dynamics: The evidence from Dutch manufacturing, Industrial and Corporate Change, 9(2), 173–194.

Vernon Henderson J., Shalizi Z., Venables A.J. (2001), Geography and development, Journal of Economic Geography, 1, 1, 81-105