Explaining TFP at firm level in Italy. Does location matter?•

Francesco Aiello, Valeria Pupo, Fernanda Ricotta
[f.aiello@unical.it; v.pupo@unical.it; f.ricotta@unical.it]
University of Calabria, Department of Economics and Statistics, Italy
87030 Arcavacata di Rende (CS)

Abstract This study analyses how firms' internal variables and regional factors affect Total Factor Productivity (TFP) of Italian manufacturing firms. Due to the hierarchical structure of our data, we employ a multilevel model that allows for a clear distinction between firm and region-specific effects. Results refer to 2004-2006 and show, as expected, the importance of firm-specific determinants of TFP. At the same time, they indicate that location matters, in the sense that the context where firms operate plays a crucial role in determining the level of TFP. In more detail, we find that the regional endowment of infrastructure, the efficiency of local administration and the investments in R&D exert a positive effect on firms' performance. We also argue that regional gaps in the endowment of these factors help to understand the dualistic nature of the Italian economy, where a wealthy North coexists with a less developed South.

Key words: Manufacturing Firms; Total Factor Productivity, Italian Regional Divide; Multilevel Models.
JEL code: L60, R11, C31

1. Introduction

Italy is an interesting case-study in the field of economic development because of the high and persistent disparity between the South and the rest of the country. The level of per capita income in Southern regions was 17,324 euro in 2009, a much lower value than that (29,399 euro) observed in the Centre-North. This is a substantial gap which is also persistent, given that it has not varied significantly over the last 30 years (ISTAT, 2010; and Iuzzolino *et al*. 2011, for a survey).

As a result of these wide disparities, it has been necessary to adopt policies aimed at overcoming the dualist nature of the Italian economy. For instance, through the

• The authors gratefully acknowledge comments and suggestions from Giovanni Anania and Antonio Aquino on an earlier version of the paper. This work benefits of the financial support from Regione Calabria within the CALCOM project on "Regional Competitiveness and Innovation" (L.R. 10/1998).

Intervento Straordinario per il Mezzogiorno (Special Funding Plan for the Development on the South), Southern regions received a large amount of resources from 1951 to 1984. The effectiveness of public expenditure is debatable: while it may have been effective in fostering convergence up to the mid-‘70s, it has not modified the structural conditions needed to ensure growth in the long term (Carey and Galbraith, 1955; Iuzzolino *et al*, 2011). Similar questions have arisen with regards more recent EU structural policy interventions, which have only slightly tackled the Mezzogiorno issue (Aiello and Pupo, 2011). Several factors have contributed to bringing about the failure of these policies. Although the investigation of these aspects goes beyond the purpose of this paper, it is useful to remember here that their impact on TFP has been really weak and that TFP is the main determinant of income differences in Italy (Aiello and Scoppa, 2000; Ascari and Di Cosmo, 2005; Byrne *et al*, 2009; Di Giacinto and Nuzzo, 2006; Piacentino, 2008).

This paper provides further evidence for the debate on the dualistic nature of Italian economy. To this end, it uses data at firm level retrieved from the survey carried out by UniCredit-Capitalia (2008). The main focus is on the determinants of the differences existing in TFP between firms located in Italian Southern regions and those operating in the Northern area of the country. A distinguishing feature of the work is the underlying belief that the environment in which firms operate matters (Krugman, 1991) and thus, from our perspective, TFP is meant to depend not only on firms' internal factors - like size, type of economic activity and internal R&D - but also on external variables beyond firms control. In other words, we are interested in distinguishing between the impact on TFP brought about by internal variables and the role played by territorial factors (e.g. availability of infrastructure, quality of public institutions, propensity to innovate) which the related literature suggests might influence firms' performance. The key question addressed in the paper is whether location matters in terms of firms' performance.

In order to provide an answer to this question we proceed as follows. While TFP is estimated at firm level by employing the Levinshon and Petrin's (2003) approach, the empirical setting we propose is consistent with the type of analysis we carry out. Indeed, in order to explain firms' TFP, we use data at firm and regional levels and consider the multilevel approach. This model - giving proper attention to nesting – allows us to evaluate whether space matters in determining firms' performance. In fact, multilevel regressions combine different levels of data aggregation and relate them in ways that render the simultaneous existence of distinct level-one (*firms*) and level-two (*regions*) equations explicit. This represents a methodological advantage with respect to single-equation models, which are too limited to handle hierarchical structures of data, because they lead to correlated errors for firms belonging to the same region, so violating the assumption of independence among errors. Furthermore, in a single-equation model, the statistical inference is based on the entire sample size and this yields a high risk of type I errors because the variance of the level-two variables is underestimated.

By using data at firm level and following the multilevel approach, we aim at explaining differences in TFP by providing a clear distinction between firm and region-specific effects. Previous works generally used regions as the unit of analysis (Ascari and Di Cosmo, 2005; Destefanis and Sena, 2005; Marrocu and Paci, 2010; Quartaro, 2006). However, finding a correlation at the regional level does not necessarily mean that it also holds when individual firm level data are used. In the literature this is known as the ecological fallacy. To the best of our knowledge, the multilevel analysis has only been applied to the case of Italian dualism by Fazio and Piacentino (2011), although their focus was to investigate the variability of labour productivity across Italian provinces (NUTS 3).

Results of our paper are twofold. On one hand, we confirm that firms' specific characteristics greatly affect firms TFP. On the other hand, results support the hypothesis that local environment conditions exert an influence upon firms' TFP. Since firms are clustered within regions, we find that operating in a high R&D-oriented region or in an area with good infrastructure and/or with efficient public services affects private performance. This is an important policy issue in Italy, as firms in Southern Italy suffer from being located in regions which are still poorly endowed in terms of pro-growth local resources.

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

2. The economic divide in Italy and the role of TFP: what firm level data highlight

The firm level data used in this paper come from the X*th* UniCredit-Capitalia survey (2008) of Italian manufacturing firms.[1] Table 1 presents some descriptive statistics of the sample of firms used in the empirical analysis.[2] In particular, it presents the distribution of firms by area and, for 2006, the labour productivity of firms grouped in terms of size and economic activity.[3]

The sample is comprised of 3,019 firms which are concentrated in traditional sectors (49% in the entire sample, 62% when just considering the South) and in highly specialised sectors. The incidence of high-tech firms is residual (only 5% in the whole of Italy and 2% in the South). From a regional perspective, two thirds of the sample is comprised of firms located in the North of Italy, 16% in the Centre and 10% in the South. The proportion of small firms is high (about 56%) and uniform across the area. This picture is representative of the Italian manufacturing industry, which is characterised by a predominance of firms located in the North and belonging to traditional sectors. Again, the share of small-sized firms is very high in Italy, whatever the area and the economic activity (Bank of Italy, 2009; Onida, 2004)**.**

Table 1 further confirms the dualist nature of Italian economy. We find that labour productivity of Southern firms is lower than that recorded in Northern Italy. More importantly, this gap holds whatever the subgroup of firms we consider (table 1).[4]

What clearly emerges from table 1 is a sharp economic divide between firms operating in the South and those located in the rest of the country, something which has been long debated in the literature. While there are complex reasons explaining this

[1] The survey covers a sample of firms with 11-500 employees and all firms with more than 500 employees. The Xth Capitalia-UniCredit survey questionnaire refers to 2004-2006 and contains information on firm structure, ownership, work force and investments in physical and technological capital, as well as the degree of internationalization. Data from balance sheet refer, instead, to 1998-2006.

[2] Although the original data refer to 5,100 firms, a sample of 3,000 firms obtained after carrying out a data cleaning procedure is used in the empirical analysis. The firms which presented negative values of value added have been 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 have also been 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.

[3] Labour productivity is calculated as the weighted average of firms' productivity, using as weights the firm's value added with respect to the group of reference (the whole sample or the value added of the area in the case of averages relative to the territory).

[4] The results for science-based firms operating in the South are not really interpretable because there are only seven of such firms.

phenomenon, here we simply refer to the strand of literature explaining how regional differences in labour productivity may be mainly attributed to differences in TFP (amongst others Aiello *et al*, 2012; Brandolini and Cipollone, 2001; Daveri and Jonia-Lasinio, 2005; ISTAT, 2007; OECD, 2007; Van Ark *et al*, 2007). In this sense, our analysis confirms the role of TFP. Indeed, after retrieving TFP from the Levinshon and Petrin's estimator (see Appendix A), we find that the correlation between firms' labour productivity and TFP from 1998 to 2006 is, on average, 0.86% and 0.96% over the 2001-2006 period. At regional level, this correlation ranges between 0.97 in the South of Italy and 0,82 in the North-West.

Table 1. Main characteristics of the sample (2006)

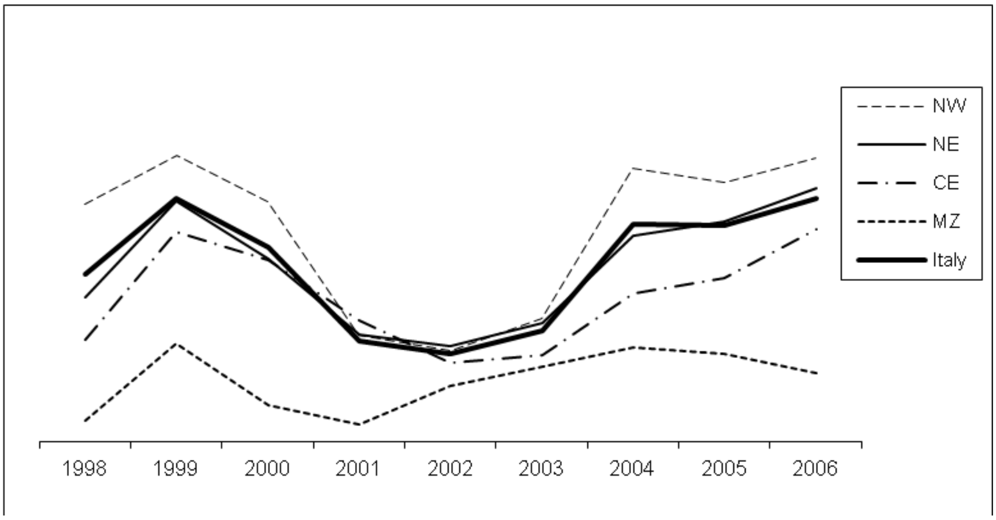
	North West		North East		Centre		South		Italy	
	Labour productivity	% of firms	Labour productivity	% of firms	Labour productivity	% of firms	Labour productivity	% of firms	Labour productivity	% of firms
By Sector										
Supplier dominated	54203	*42,9*	50683	*49,9*	52753	*53,7*	47220	*62,2*	51976	*48,6*
Scale intensive	64377	*19,8*	60887	*16,0*	62846	*23,0*	59862	*21,4*	62740	*19,3*
Specialised suppliers	59935	*31,2*	58379	*30,7*	52542	*19,6*	51003	*13,9*	58168	*27,6*
Science based	66798	*6,1*	54048	*3,4*	61812	*3,6*	80604	*2,4*	64159	*4,5*
By Size										
Small (11- 50 empl.)	58328	*57,92*	53484	*52,07*	54899	*55,1*	49878	*55,6*	55611	*56,08*
Medium (50-250 empl.)	58605	*33,0*	55178	*38,0*	53704	*35,4*	52023	*36,1*	56059	*35,2*
Large (>250 empl.)	63124	*9,0*	60001	*9,9*	65972	*8,5*	58675	*8,8*	62087	*9,2*
Total	58824		54775		55376		51395		56338	
Observations	N.	%	N.	%	N.	%	N.	%	N.	%
	1338	*44,3*	918	*30,4*	469	*15,5*	294	*9,7*	3109	*100,0*

All variables computed for 2006. Labour Productivity is deflated and expressed in Euro
Italian regions are grouped as follows. North-West: Liguria, Lombardia, Piemonte and Valle d'Aosta. North-East: Emilia-Romagna, Trentino Alto-Adige, Friuli Venezia-Giulia and Veneto, . Centre: Marche, Lazio, Toscana, and Umbria. South: Abruzzo, Basilicata, Calabria, Campania, Molise, Puglia, Sardegna and Sicilia.

Bearing in mind the above-mentioned evidence, in what follows we briefly present the dynamics of TFP over the period under scrutiny and highlight the differences between one area and another. In particular, it can be seen how TFP in the South was lower than in other areas for the whole period, underlining the technological gap in Italy which has already been discussed in the literature (Ascari and Di Cosmo, 2005; Byrne *et al*, 2009; Ladu, 2010). Results also show how this disparity was not uniform over time; although wide at both the beginning and at the end of the period, regional TFP converged in 2002-2003 (fig. 1). 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. 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 the 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).

While previous results bear out the dualistic nature of Italian economy, they leave the question about the reasons underlying the regional gap in TFP open. The next paragraph looks at this issue.

Figure 1. Average TFP by area from 1998 to 2006



Source: elaborations on data from UniCredit – Capitalia (2008)

NW=North West; NE=North East; CE=Centre; MZ=Mezzogiorno (South)

3 Empirical setting

3.1 Methodology: the multilevel analysis

The understanding of how localisation in different regions affects firms' performance is a typical issue with hierarchically structured data, in the sense that the units (firms) refer to different levels of aggregation (regions) (Goldstein, 2003). If a nested structure of data exists, single-level methodologies will suffer from the following potential estimation problems. First, as a result of regionally specific factors, firms operating in a given region are likely to be more similar than firms located in differing regions. This similarity means that the assumption of independence of errors is violated. On the other hand, the multilevel approach addresses this issue and ensures more efficient estimates. Second, single-level regressions yield an inflated significance of level-two coefficients because tests are made by using the number of level-one observations instead of the number of regional units. It is likely that the significant relationships found in OLS regressions will turn out not to be significant in multilevel regressions. In other words, the multilevel model controls for spatial dependence and corrects the measurement of standard errors, so reducing the risk of type I errors. Finally, apart from the statistical improvements, another advantage of the multilevel model is that variables at different levels are not simply add-ons to the same single-level equation, but they 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 intercept and random slopes (Bickel, 2007).

In what follows we present the multilevel regression model. In order to limit complexity, we consider a two-level model where firms are the first-level units and regions those at the second-level.

The dependent variable ω refers to firms and depends on a set *X* of variables measured at firm level and on a set *Z* of variables defined at regional level. The variable ω may be predicted by just considering *X* as explanatory variables:

$$\omega_{ij} = \beta_{0j} + \beta_{1j} X_{ij} + e_{ij} \qquad [1]$$

where β_{0j} is the intercept, β_{1j} are the slope coefficients and e_{ij} is the random error term with zero mean and σ_e^2 as variance, *j* is for regions (*j=1…r*) and *i* for firms (*i=1…N_j*). In eq. 1 the regression parameters β_j vary across level-2 units. This may be modelled as follows:

$$\beta_{0j} = \gamma_{00} + \gamma_{01} Z_j + u_{0j} \qquad [2]$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11} Z_j + u_{1j} \qquad [3]$$

In so doing, β_{0j} and β_{1j} differ across regions and depend on *Zj*, u_{0j} e u_{1j} are random error terms defined at regional level with zero mean and assumed to be independent from e_{ij}. γ denote the fixed level-two parameters.

Combining the micro (eq. 1) and the macro models (eq. 2 and 3) produces a two level mixed model:

$$\omega_{ij} = \gamma_{00} + \gamma_{10} X_{ij} + \gamma_{01} Z_j + \gamma_{11} Z_j X_{ij} + (u_{1j} X_{ij} + u_{0j} + e_{ij}) \qquad [4]$$

The deterministic part of the model, $\gamma_{00} + \gamma_{10} X_{ij} + \gamma_{01} Z_j + \gamma_{11} Z_j X_{ij}$, contains all the fixed coefficients, while the stochastic component is in brackets. The added complexity of the error term stems from the fact that it captures residual variance, in the same way as OLS regression does, as well as group-to-group variability in the random intercept relative to the overall intercept, and group-to-group variability in the random slope relative to the overall slope. It is clear that the error term displayed in eq. 4 is not independently distributed. Indeed, as data are nested at different levels of analysis, firms operating in the same region tend to have correlated residuals, violating the assumption of independence.

Eq. 4 also allows identification of the errors due to differences across firms or regions. To this end, it is necessary to use an "empty" model, i.e. a model without any explanatory variables:

$$\omega_{ij} = \gamma_{00} + u_{0j} + e_{ij} \qquad [5]$$

which allows decomposition of the variance of ω into two independent components, which are the variance of e_{ij} (σ_e^2), the so-called within-group variance, and the variance of u_{0j} (σ_{u0}^2), also known as between-group variance. Hence, one can calculate the proportion of total variance "explained" by the grouping structure, i.e. the intra-class correlation ICC:

$$ICC = \frac{\sigma_{\mu 0}^2}{\sigma_{\mu 0}^2 + \sigma_e^2} \qquad [6]$$

Furthermore, eq. 4 is a general formulation and represents different reduced specifications. For example, it incorporates the so-called "random intercept model", where the intercept only is a function of level-two predictors, without considering the cross-level fixed effects. Since there are 20 regions in Italy, there is a relatively limited number of groups. Such a constraint limits the number of parameters to be estimated and, for this reason, we proceed by considering a two-level random intercept model with firms (lowest level) and regions (highest level).[5] The specification used in this paper is given by:

$$\beta_{0j} = \gamma_{00} + \gamma_{01} Z_j + u_{0j} \qquad [2']$$

$$\beta_{1j} = \gamma_{10} \qquad [3']$$

$$\omega_{ij} = \gamma_{00} + \gamma_{10} X_{ij} + \gamma_{01} Z_j + u_{0j} + e_{ij} \qquad [4']$$

The random component u_{0j} captures variability in the intercept across regions, while the fixed component γ_{00} is a weighted average of the intercept across all regions. Finally, equation 4' allows us to account for the variability in the random component by introducing one or more levels of analysis through choosing appropriate contextual variables.[6]

3.2 Econometric specification and data

In line with eq. 4', the model used in the empirical analysis is specified as follows

$$\omega_{ij} = \beta_0 + \sum_{s=1}^{k} \beta_s X_{sij} + \sum_{z=1}^{v} \eta_z Z_{zj} + \sum_{p=1}^{3} \lambda_p S_{pi} + \sum_{q=1}^{2} \lambda_p D_{qi} + \lambda South + u_{0j} + e_{ij} \qquad [7]$$

where ω is the TFP of the *i-th* firm (in logarithm) operating in region *j*, *X* is a vector of firm-level variables which are meant to be important drivers of TFP and *Z* are variables

[5] As we have already pointed out, in multilevel models when level-one coefficients are permitted to vary across groups, the number of groups, not the number of level-one observations, is used to test the significance of level-one slopes. Consequently, cross-level interaction terms are likely to have unstable coefficients and uncertain inferential properties, unless there is a comparatively large number of cases at both level one and two (Bickel, 2007).

[6] The possibility to employ contextual factors *(Z_j)* and, in the general specification (equ. [4]), cross-level interaction terms *($Z_j X_{jj}$)* to explain variability in random components is the main difference between the multilevel model and random coefficient regression.

at regional level. TFP determinants defined at firm level include the 2004 R&D intensity (R&D investments/sales), the white collar workers as share of total workers in 2006 and export intensity in 2006 (exports/sales). One of the basic empirical facts related to productivity is a strong positive association between productivity and exporting activity and, therefore, we include the share of exports in total sales among firms' characteristics (Melitz, 2003; ISGEP, 2008). Similarly, it is widely argued that a firm's performance improves as a result of its innovative behaviour and in the presence of skilled workers (see, i.e., Krueger and Lindahl, 2001; Sveikauskas, 2007).

As far as the regional level is concerned, we selected three variables to be included in the analysis: the R&D intensity of the private sector, an index of the endowment of infrastructure and the efficiency of the public sector. This is in line with the literature dealing with the reasons for the economic divide in Italy (an exhaustive survey can be found in Iuzzolino *et al*, 2011). In this paper, R&D intensity is measured as a 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[7] and summarises the availability of different kinds of infrastructure directly relevant for production, such as roads, railways, telecommunications, ports, water, electricity, airports). The index refers to the year 2004. Following 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 of infrastructure and the physical inventory of public capital which has effectively been built, after controlling for regional differences in cost of public construction. This measure refers to 1997 and has also been used as corruption index (Golden and Picci, 2005). All regional indicators are linked with firms' data using the location of the company headquarters. Table 2 displays the values of the regional variables used in the paper.

In estimating the multilevel equation, we also control for sector, size and "South" effects by adding a set of dummy variables. In order to control for sectoral heterogeneity in the production process, we include three sector dummies (*S*) according to the Pavitt taxonomy (S_2 is unity for firms belonging to scale intensive sectors, S_3 is unity for firms operating in specialised suppliers, while S_4 is unity when firms operate in

[7] 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 the EU countries (http://www.cnel.it/cnelstats/index.asp).

science-based sectors. Firms in supplied-dominated sectors form the control group). Furthermore, regressions also include two dummy variables to control for size effect (D_M refers to medium-sized firms and D_L is for large-sized firms, whereas the control group is comprised of small firms). Finally, in some specifications we also use the variable South which is a binary variable equal to unity if the firm is located in the South of Italy and zero otherwise. The variable South is supposed to capture the non-observable differences between the Centre-North and the South of Italy and is used when regressions include only data at firm level. This has been done because in the more extended models we control for location through the three selected regional determinants of TFP.

TFP is expressed as the average of the three-year period 2004-2006. Although TFP at firm level is available for a longer period (see fig. 1), we restrict estimations to a cross-section analysis averaging TFP data from 2004 to 2006. This is done because of data constraints.[8, 9]

[8] We average TFP over the three-year period 2004-2006 in order to control for the influence of shocks and measurement errors in a specific year and to limit the extent of missing data.

[9] Equation [7] probably suffers from omitted variable problems since unit heterogeneity is not considered. One way to allow for unobserved heterogeneity is the fixed effects model. However, panel data analysis cannot be performed, due to the lack of time series in variables such as white collar share and exports. Moreover, endogeneity has been addressed by lagging variables at firm level. Furthermore, for each firm the variables defined at regional level act as exogenous factors and this limits the endogeneity issue.

Table 2 Infrastructure, R&D Intensity and efficency of public administration by region (Italy=100)

Regions	Infrastructure in 2004	Private R&D Intensity in 2004	Efficiency of Public Administration in 1997
North			
Emilia-Romagna	109,8	125,7	161,1
Friuli-Venezia Giulia	123,9	98,0	107,7
Lazio	146,2	83,2	81,7
Liguria	191,2	120,2	66,8
Lombardia	123,9	153,5	116,1
Marche	88,6	49,9	131,2
Piemonte	88,3	244,1	163,8
Toscana	111,4	64,7	161,3
Trentino-Alto Adige	60,2	46,2	123,5
Umbria	86,7	29,6	178,3
Valle d'Aosta	44,4	48,1	85,5
Veneto	117,3	51,8	122,0
South			
Abruzzo	77,8	86,9	95,6
Basilicata	38,6	37,0	53,3
Calabria	74	3,7	40,8
Campania	95,7	77,7	36,2
Molise	50,6	11,1	58,2
Puglia	79	29,6	72,2
Sardegna	55,5	5,5	83,8
Sicilia	84,2	42,5	60,7

Source: National Council of the Economy and Labour for infrastructure, National Institute of Statistics for R&D intensity and Golden and Picci (2005) for efficiency of public administration.

4 Econometric results

Results are displayed in tables 3 and 4. While table 3 refers to OLS estimates of eq. 7, table 4 presents the evidence provided by the multilevel approach.

The OLS estimator is only used for reference and to verify the bias when data at different levels of aggregation are evaluated within a single-equation model. However, because OLS regressions are performed using micro and regional data, we control for the potential downward bias in the estimated errors by clustering firms at regional level.[10] In brief, two key results emerge from table 3. On one hand, we find that location

[10] Clustering data at regional level relaxes the assumption of independence and, therefore, increases the error term to accommodate the lack of independence of firms within regions. However, while clustered

matters in determining firms' TFP. This can be seen from the negative and significant coefficient of the South dummy (table 3, Model 2). This implies that, *ceteris paribus*, the level of TFP in firms located in southern regions is lower than that in firms located in the North of Italy. To some extent, similar evidence comes from Model 3, where the parameters associated with regional variables (R&S intensity, infrastructure index and the efficiency of public administration) are all positive and highly significant. By referring to these results it is possible to argue that high regional R&D intensity, good infrastructure and high efficiency of public administration do help firms to improve performance. On the other hand, by comparing data in tables 3 and 4 we obtain evidence that using a single-equation model when data are available in a hierarchical structure yields deflated standard errors. In this sense, the interpretation of OLS results is bounded by the actual statistical significance of the estimated coefficients.

The results of the multilevel models are presented in table 4, where each column of data refers to different specifications of eq. 7 according to the set of explanatory variables included in the model.[11] Model 1 is the empty model, i.e. a model without regressors (eq. 5), while Model 2 only includes level-1 predictors. With respect to Model 2, in Model 3 we add the South dummy, while Model 4 includes the level-2 regressors.

An initial result comes from the likelihood-ratio test, which compares the empty model (eq. 5) with the standard linear regression. This test, which is highly significant, supports the use of multilevel methodology[12] and indicates that the intercept should be considered as a region-by-region variant coefficient. Moreover, the ICC value (*cfr* eq. 6) indicates that 4.6% of firms' TFP can be explained by their mere spatial location (Model 1, table 4) while internal firm characteristics explain 95% of firms' TFP.[13]

A further interesting aspect of the approach refers to the possibility of using the variance at the different levels of analysis to calculate the coefficient of determination

---

OLS leaves both the noise associated with difference between firms and noise associated with differences between regions in the error term, the multilevel model goes further by allowing these two error components (see equ. [4]) to be separated.

[11] The multilevel analysis was implemented in Stata using the "xtmixed" subroutine. All models were estimated using restricted maximum likelihood (REML) over maximum likelihood (ML) since the latter is more sensitive to loss of degrees of freedom when dealing with a small number of groups (Bickel, 2007).

[12] The null hypothesis is that $u_{0j} = 0$ or that there is no random intercept in the model. If the null hypothesis is true, an ordinary regression can be used instead of a variance-components model.

[13] For Italy our findings are in line with Fazio and Piacentino's (2010) results at the provincial level. For the Netherlands, Raspe and van Oort (2011) find that 2.3% of firm productivity can be related to location and that more than 97% to internal characteristics.

and, in such a way, to obtain a proportional reduction in the estimated total residual variance. This is done by comparing the "empty model" with an extended specification of the model (Rabe-Hesketh and Skrondal, 2008).[14]

For instance, we see in table 4 that variables at firm level as a whole are able to explain 27% of TFP firm variance (table 4, model 2). In Raspe and van Oort (2011) the selected set of firm level variables only explains 7.7% of the within-group variance. When including the region-level predictors we find that the variance of regional intercepts decreases by 89%. A large proportion of region-by-region variability in the intercepts has been accounted for by the regional variables included in the analysis (private R&D intensity, infrastructure and public administration efficiency). This evidence ensures that the selected regional factors of TFP capture a great deal of intercept variability, which we attribute to unobserved TFP heterogeneity when considering the empty model.

Moving on to discuss the results of estimated coefficients, table 4 shows that, at firm level, the parameter associated with internal R&D has the expected positive sign and is highly significant. Firms investing more in R&D obtain higher TFP levels than firms with limited innovative activities. Again, an important role is played by the human capital employed by firms. The result is that TFP increases with human capital. These findings are in line with the literature showing that R&D and human capital induce higher firm TFP because they directly affect the possibility to introduce and use more productive processes and, hence, translate innovation efforts into profitable opportunities (Griliches 2000; Parisi *et al*, 2006). Furthermore, and consistent with existing literature, we find that TFP tends to increase with exports. Many studies explain the positive relationship between export activity and productivity by self-selection of more efficient Italian firms into the export markets (see, for instance

---

[14] The coefficient of determination for two-level model is given by:

$$R^2 = \frac{(\sigma^2_{\mu 0N} + \sigma^2_{eN}) - (\sigma^2_{\mu 0M} + \sigma^2_{eM})}{\sigma^2_{\mu 0N} + \sigma^2_{eN}}$$

where N stands for the null model and M for the model of interest.
The proportional reduction in each of the variance components can be calculated separately. The proportion of the level-2 variance explained by the covariates is:

$$R^2_2 = \frac{(\sigma^2_{\mu 0N} - \sigma^2_{\mu 0M})}{\sigma^2_{\mu 0N}}$$

and the proportion of the level-1 variance explained is:

$$R^2_1 = \frac{(\sigma^2_{eN} - \sigma^2_{eM})}{\sigma^2_{eN}}.$$

Benfratello and Razzolini, 2008; ISGEP, 2008; Serti and Tomasi, 2008 among many others), while few studies also find support for the "learning by exporting" hypothesis (ISGEP, 2008; Serti and Tomasi, 2008, and, then just for exporters with a high share of export intensity, Castellani, 2002).[15]

In addition, the positive coefficients associated with the D_M and D_L dummies highlight the role of size for TFP. Medium-sized firms perform better than small firms, but less well than large enterprises. In short, in the case of Italian manufacturing firms, TFP increases with firm size, indicating that economies of scale are at work. Another influential factor is the type of economic activity. It is widely accepted that TFP differs across sectors and it is found that firms in high-tech sectors perform better than others, followed by firms operating in scale intensive and specialised sectors. The lowest value of TFP is obtained for firms belonging to traditional sectors (our group of control). This result indicates that sectoral characteristics in producing innovative products allow high-tech firms to perform better than those operating in other sectors.

When considering the first level of the analysis, results indicate how firms' internal factors are relevant in determining the level of TFP. However, the main interest lies in the role of variables defined at regional level. An initial finding regards the role of infrastructure. We find that TFP at firm level is positively affected by the endowment of regional infrastructure, in the sense that firms' located in regions with an adequate provision of infrastructure benefit more than firms operating in under-endowed regions. Due to the sharp differences in regional endowment of infrastructure (see table 2), this result indicates that, other things being equal, TFP of Southern firms will be lower (fig. 1) because they operate in areas suffering from a lack of public capital. This is in line with the conclusions drawn, for instance, by Aiello *et al* (2010), Destefanis and Sena (2005) and Marrucu and Paci (2010). With regards regional private R&D activity, we find a positive impact on firms' TFP. This is consistent with the literature (e.g, Camagni 1991; Ciccone and Hall, 1993). It is an indication of the spillover effects as a product of innovations, in the sense that being located in a region with high innovation-creating

[15] In the literature two hypotheses about the positive correlation between export activity and productivity are investigated. The first hypothesis is that the most productive firms self-select into foreign markets because they can overcome sunk costs associated with foreign sales (Melitz, 2003). The second hypothesis raises the possibility of "learning by exporting". Firms participating in international markets acquire knowledge and technology with positive feedback effects on firms' knowledge and technology accumulation. Furthermore, firms which are active in world markets are exposed to more intensive competition than firms which only sell their products domestically. In summarizing the results achieved in this field of research it can be said that the more productive firms self-select into export market (ISGEP, 2008; Melitz, 2003).

potential makes individual firms perform better. In this respect, it appears clear that the TFP divide between Northern and Southern firms is also due to differences in regional innovativeness: data used in this paper indicate that the level of innovative efforts made by the private sector in Southern regions is very low and far less than in the rest of the country (table 2). Therefore, stimulating R&D investments in the South of Italy has to be a priority in policy agendas because this might help to build a R&D environment from which firms may acquire innovative opportunities that can be translated into internal efficiency. Finally, the efficient provision of public services is an important factor for firms' productivity. In Italy, public administration is most successful in providing services in the Northern regions (Bank of Italy, 2009; Tabellini, 2010). This fact contributes to explain why TFP of firms operating in that part of the country is higher than the TFP levels observed in the South. Firms operating in regions with efficient public institutions benefit from a reduction in transaction costs they face when introducing more productive activities and creating an environment which is conducive to growth.

Table 3 Explaining TFP of Italian manufacturing firms in 2004-2006: OLS Results

Explanatory Variables	Model 1	Model 2	Model 3
Firm level covariates			
R&D Investments/Sales	0.00002***	0.00002***	0.00002***
	(7.85)	(7.52)	(9.54)
White Collar Share	0.26421***	0.25228***	0.21599***
	(6.72)	(6.89)	(6.789)
Exports/Sales	0.00087***	0.00066***	0.00059**
	(3.87)	(3.24)	(2.67)
Medium firms (DM)	0.28628***	0.28794***	0.29112***
	(32.25)	(30.62)	(23.27)
Large firms (DL)<	0.70147***	0.70332***	0.69024***
	(23.95)	(25.09)	(24.50)
Scale intensive (S2)			0.16957***
			(8.64)
Specialised suppliers (S3)			0.13288***
			(7.01)
Science based (S4)			0.23032***
			(4.12)
Regional level covariates			
Private R&D over Regional GDP			0.07898***
			(5.18)
Index of Infrastructure			0.00181***
			(4.86)
Efficiency of Public Administration			0.00102***
			(3.33)
South		-0.17450***	
		(-5.03)	
Constant	6.35030***	6.37730***	5.92164***
	(235.46)	(289.67)	(89.12)
R-squared	0.24	0.26	0.29
Observations	2941	2941	2941

Dependent variable: log of TFP (average values for 2004-2006 period). In parentheses, t-values based on standard errors clustered at firm level. Level of significance: *** 1%, ** 5%, * 10%.

Table 4 Explaining TFP of Italian manufacturing firms in 2004-06: multilevel regressions

Explanatory Variables	Model 1	Model 2	Model 3	Model 4
Fixed effects				
Level 1: Firms				
R&D Investments/Sales		0.00002***	0.0000205***	0.00002***
		(4.78)	(4.75)	(4.70)
White Collar Share		0.216***	0.214***	0.215***
		(7.70)	(7.64)	(7.66)
Exports/Sales		0.00064***	0.00060**	0.00060**
		(2.53)	(2.41)	(2.40)
Medium firms (DM)		0.291***	0.290***	0.291***
		(18.26)	(18.23)	(18.29)
Large firms (DL)		0.690***	0.689***	0.690***
		(24.42)	(24.40)	(24.41)
Scale intensive (S2)		0.164***	0.165***	0.166***
		(8.30)	(8.36)	(8.42)
Specialised suppliers (S3)		0.130***	0.130***	0.130***
		(7.31)	(7.31)	(7.32)
Science based (S4)		0.227***	0.227***	0.227***
		(6.16)	(6.18)	(6.15)
Level 2: Regions				
Private R&D over Regional GDP				0.0775**
				(1.92)
Index of Infrastructure				0.0015***
				(2.98)
Efficiency of Public Administration				0.00084**
			-0.130***	(2.32)
South			(-3.30)	
Constant	6.590***	6.257***	6.297***	5.975***
	(239.73)	(226.59)	(242.43)	(88.14)
Random-Effects				
Variance				
Region	0.010	0.007	0.003	0.001
Firms	0.212	0.155	0.155	0.155
Intraclass correlation (ICC)	4.6%			
R		0.27	0.29	0.30
R^2 *level 2*		0.31	0.68	0.89
R^2 *level 1*		0.27	0.27	0.27
LR test	60.37***			
Log restricted-likelihood	-1912.4	-1438.4	-1475.4	-1486.8
Number of observations	2941	2941	2941	2941

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

5. Conclusions

Many recent contributions have investigated the Italian economic divide by focussing on the role played by TFP. This literature comes to the conclusion that TFP is a key variable in explaining regional differences in Italian growth. Following this line of reasoning, this article provides further evidence by analysing the TFP of Italian manufacturing firms. In particular, the key question is to understand the role of location in determining firms' TFP. To this end we employ a multilevel approach which allows to disentangle the role of firms' factors from that played by regional variables. An initial finding of the paper confirms that firm specific characteristics highly affect firms TFP. We also show that location matters in explaining the level of firms' TFP. To be more precise, firms located in the South of Italy are less efficient than those operating in the rest of the country and, in this sense, the analysis supports the hypothesis that Southern regions are technologically lagging behind Northern regions.

This emerges from the basic multilevel model, which points out the incidence of variant regional intercepts in explaining the variance of firms' TFP. Furthermore, the multilevel regressions which include the regional determinants of TFP, i.e. R&D, the efficiency of public administration and the state of infrastructure, explain a large proportion of the average regional TFP variability. This outcome supports our choices because, given the high share of TFP variability explained by these environmental variables, any excluded regional factor of TFP can only be of marginal importance.

Interesting insights come from the estimated impact of regional variables. The selected regional variables, namely R&D intensity, efficiency of public administration and infrastructure, always register the expected positive sign. More importantly, the impact of regional factors remains highly significant in multilevel regressions and this is not a foregone conclusion. Indeed, this approach uses the number of regions, instead of the entire sample of firms, and, therefore, statistical significance may be lost. The fact that they are still significant means that they are important sources of TFP differences across Italian regions. In a nutshell, this paper indicates that operating in R&D oriented regions, which guarantee good quality public services and with an appropriate endowment of infrastructure, ensures firms to achieve a high level of economic performance.

From a policy perspective, this evidence implies that there is still room for public intervention aimed at overcoming Italy's North-South dualism. One area of interest is that of the external diseconomies that generate territorial disadvantages such as those produced by the deficit of infrastructure and the inefficiency of public institutions. On one hand, and unlike the past, the infrastructural deficit has to be addressed by using national and European funds efficiently. This might be enough to build an effective stock of public capital and, thus, increase the benefits accruing to firms from better infrastructure. On the other hand, improving the quality of institutions requires complex policies because it involves institutional, social and relational factors which, by their nature, are difficult to change. With regards innovation, public support for private R&D is a good policy option per se¸ because increasing technological potential through sizeable investments should lead to innovation and, ultimately, growth in an economy. This particularly holds true for Italy, where R&D investments are low and concentrated in the richer areas of the country. Therefore, stimulating innovative activities would help lagging southern regions to reduce their distance from the technological frontier and, hence, gain higher returns from investing in R&D.

To sum up, we show that the local context helps in explaining the Italian economic divide. Bearing in mind the weak availability of territorial resources in the South of Italy, the evidence provided by this paper leads to pessimistic conclusions for the future. This is because the impending federalist reform and the current global crisis are likely to exacerbate the economic conditions in Southern Italian regions and this will occur in a country which excludes the "Southern Question" from the national policy agenda (Cannari, Magnani and Pellegrini, 2009). It is hard to be optimistic under these circumstances.

Appendix A– A measure of TFP

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

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

with $i = 1,\dots\dots N$ firms, t = 1998, ……2006 and where *y* represents the value added, *l* the number of employees, k^{MAT} the stock of physical capital, β_0 measures the average efficiency and u_{it} represents the deviation of firm *i* from this average 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 the firm which, therefore, in the case of positive shocks to productivity, can decide to increase production by raising the level of inputs. This determines a problem of simultaneity which Levinshon and Petrin (2003) resolved by identifying the demand for intermediate goods as a proxy for the variations in TFP known to firms.

The equation [A1] was estimated by utilizing the tangible fixed assets as a proxy for the stock of physical capital and the demand for intermediate goods was 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 the 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 adopt the intermediate consumption deflator calculated by using data from ISTAT.

REFERENCES

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

Aiello F., Iona A., Leone L. (2010), Regional Infrastructure and Firm Investment. Theory and Empirical Evidence for Italy, Empirical Economics, DOI 10.1007/s00181-010-0445-9.

Aiello F., Pupo V. (2011), Structural Funds and the Economic Divide in Italy", Journal of Policy Modeling (2011), doi:10.1016/j.jpolmod.2011.10.006.

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

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.

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

Benfratello L., Razzolini T. (2008), Firms' productivity and internationalisation choices: evidence for a large sample of manufacturing firms in Multinationals and local competitiveness, Lucia Piscitello and Grazia Santangelo (Eds.), Franco Angeli

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

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

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.

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

Cannari L., Magnani M., Pellegrini G. (2009) Quali Politiche per il Sud?, in Mezzogiorno e Politiche regionali, pp.673-733, Bank of Italy, Rome.

Carey J.P.C, Galbraith C. A. (1955), The South of Italy and the Cassa per il Mezzogiorno, Political Research Quarterly, 8, 569-588.

Castellani, D. (2002), Export behaviour and productivity growth: evidence from Italian manufacturing firms, Weltwirtschaftliches Archiv, 138, 605–28.

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

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

Destefanis S., Sena V. (2005), Public capital and total factor productivity: New evidence from the Italian regions, 1970-98, Regional Studies, 39, 5, 603-617.

Di Giacinto V., Nuzzo G. (2006), Explaining labour productivity differentials across Italian regions: the role of socio-economic structure and factor endowments, Papers in Regional Science, 85, 2, 299-320.

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

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

Goldstein H. (2003), Multilevel Statistical Models, London, Arnold.

Griliches, Z. (2000), R&D, Education and Productivity A Retrospective. Harvard University Press, Cambridge, Massachusetts.

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, Bank of Italy, Economic History Working Papers, n. 22.
Krueger A.B., Lindahl M. (2001), Education for Growth; Why and for Whom? Journal of Economic Literature, 39, 4, 1101-1136.
Krugman P. (1991), Geography and Trade, Cambridge University Press.
Ladu M. G. (2010), Total factor productivity estimates: some evidence from European regions, WIFO Working Papers, 380.
Levinsohn J., Petrin, A. (2003), Estimating production functions using inputs to control for unobservables, Review of Economic Studies, 7, 317-341.
Luke A. (2004), Multilevel Modelling. London: Sage Publications.
Marrocu E., Paci R. (2010), The effects of public capital on the productivity of Italian regions, Applied Economics, 42, 989-1002.
Melitz M. J. (2003), The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. Econometrica 71, 6, 1695–1725.
OECD (2007), Economic Survey: Italy, OECD, Parigi.
Onida F. (2004), Se il piccolo non cresce. Piccole e medie imprese in affanno, Il Mulino, Bologna.
Parisi M. L., Schiantarelli F., Sembenelli A. (2006), Productivity, Innovation and R&D: Micro Evidence for Italy, European Economic Review, 50, 8, 2037-2061.
Piacentino D. (2008), Productivity, Infrastructures and Convergence: Panel Data Evidence on Italian Regions, Scienze-Regionali/Italian-Journal-of-Regional-Science, 7, 2, 5-26.
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.
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.
Sveikauskas L. (2007), R&D and Productivity Growth: a Review of the Literature, US Bureau of Labour Statistics, BLS Working Papers n. 208.
Tabellini G. (2010), Culture and Institutions: Economic Development in the Regions of Europe, Journal of the European Economic Association, 8, 4, 677-716.
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.