Efficiency And Productivity Growth Across The Italian Regions: The Regional Divide Revisited

Pompei, Fabrizio

Department of Economics, University of Perugia

9 December 2013

EFFICIENCY AND PRODUCTIVITY GROWTH ACROSS THE ITALIAN REGIONS: THE REGIONAL DIVIDE REVISITED

*Fabrizio Pompei**

Abstract
In this paper we study the efficiency and total factor productivity growth of Italian regions by implementing a bootstrap Data Envelopment Analysis method. This approach allows us to perform a sensitivity analysis of the efficiency scores at regional level, in which human capital is included besides traditional inputs. Higher levels of average years of schooling were important for efficiency and TFP growth in the Northern and Central regions. Conversely, the overall scarce human capital accumulation in Southern regions negatively affected their performances. However, both DEA and analysis of decomposition of productivity growth, conducted by means of Malmquist's index, highlighted that also in Southern regions, in which the growth rate of human capital and TFP was remarkable, the contribution of the improvement in pure efficiency to economic growth was totally nonessential.

JEL Classification: O30; O40; R11.

Keywords: Italian regions, Productivity, Data Envelopment Analysis.

*Fabrizio Pompei, Department of Economics, Finance and Statistics, the University of Perugia, Via Pascoli 20, 06123 Perugia, Italy,email: fabrizio.pompei@unipg.it.

1 INTRODUCTION

The regional divide that characterises Italy is still at the centre of empirical research. Besides the profound standard living inequalities and the dramatic social problems of the South of Italy, these persisting economic disparities among Italian regions have undoubtedly contributed to the bad performance of the Italian labour productivity growth over the last decades.

This regional divide has been largely studied in terms of determinants of economic growth (Bronzini and Piselli, 2006; Maffezzoli, 2006; Aiello and Scoppa, 2005; Ascari and Di Cosmo, 2005, Destefanis and Sena, 2005; Paci and Pigliaru, 1999).

More precisely, some authors explored the role played by innovative activities (specifically R&D), human capital and public infrastructures on productivity by performing canonical parametric methods (Bronzini and Piselli, 2006; Aiello and Scoppa, 2005; Ascari and Di Cosmo, 2005; Paci and Pigliaru, 1999). Other authors, performed data envelopment analysis (DEA) upon a standard production function Y=A f(K, L), (Maffezzoli, 2006). Destefanis and Sena (2005) bridge the gap between these two groups: indeed, they studied the role of regional public capital stock as determinant of total factor productivity (TFP) growth by comparing parametric and non-parametric methods.

The advantage in using nonparametric DEA-like methods is that we do not need to specify a particular functional form for the aggregate production function, but only to impose an assumption about returns to scale, as well as input and output disposability. Since in principle all regions have access to the same technology, the production function itself relates to the whole sample of regions; however, regions are allowed to operate at different degrees of technical and allocative efficiency, i.e. different regions with similar levels of capital and labour may produce different amounts of output. On the other hand, this method suffers some severe problems concerning the large sensitivity of the outcomes to the sample variations (Simar and Wilson, 1998) and the fact that DEA efficiency estimates are, by their nature, serially correlated (Simar and Wilson, 2003). Therefore, the inference of the efficiency scores and their use as dependent variables in parametric methods, in order to implement a two-stage approach, is problematic.

This paper aims to take a step forward in the analysis of the efficiency and productivity of the Italian regions by implementing DEA and a two-stage approach that takes into account the problems mentioned above. More precisely, we analyzed efficiency and productivity growth of Italian regions from 1971 to 2003, by means of a DEA-like method[1]. Although linear programming is nowadays considered as a quite classical method, also for regional studies, the original aspect of this article, in regards to similar studies concerning Italian regions, consisted in applying a bootstrap procedure to calculate efficiency scores. Thus, in the first part, after discussing data sources and variables (section 2) and presenting the estimation strategy (section 3), we computed efficiency scores for both standard two-inputs model of the production process and the human capital augmented model (section 4). The bootstrap procedure allowed us to perform a sensitivity analysis of the efficiency scores (section 4.1). Afterwards, in section 4.2, we studied the TFP growth and its decomposition by computing the Malmquist productivity index. Lastly, in section 5 we presented some concluding remarks.

2 DATA SOURCES AND VARIABLES

Two different databases have been used to carry out the empirical analysis: 1) regional accounting statistics (GDP and value added, labour, investments) stem from CRENOS-REGIO.IT[2] database; 2) human capital indicator and the proxy of external economies rely on Census of Population and Census of Industry respectively, that are published by National Institute of Statistics (ISTAT).

In the CRENOS-REGIO.IT database, regional economic aggregates are time series concerning the period 1970-2004. Unfortunately, capital stock is not available in this source, hence we had to estimated it. According to Picci and Bonaglia (2000), we performed a Permanent Inventory Method (PIM) to calculate net capital stock. Our starting points were: a) fixed gross investment series (1970-2003) at 1995 prices and the sector-region level; b) the 1970 gross capital stock at 1995 prices and sector-region level, used as benchmark for our calculations[3]. In the second step we were able to break down investments at the sector-region level into two categories, *i*) dwellings and no-residential buildings, *ii*) machinery and other assets, thanks to specific coefficients that we drew from ISTAT national series. In the third step we implemented the PIM by assuming fixed expected service lives (15 years for machinery and 35 years for buildings), simultaneous exit mortality patterns and linear depreciation. Following the authors mentioned above, we re-constructed investment series before 1970 by splitting up the capital stock of 1970 over years 1936-1969.

---

[1] We were not able to extend the period of the analysis beyond 2003 because of the well-known changes in the European System of National Account (SEC95), that made the series stemming from the CRENOS Database not coherent with the new series developed by the Italian National Institute of Statistics (ISTAT).

[2] CRENOS is the Centre for North South Economic Research.

[3] The data concerning capital stock stem from another CRENOS database, in which time series partially overlap the period we considered, even though they stop at 1994. In order to avoid incoherences in the data, we preferred using only the 1970 data as benchmark and calculating capital stock of following years by cumulating investments of REGIO.IT database.

As far as human capital is concerned, we work out the average years of schooling (Barro and Lee, 1993) as follows

$$\sum_j YR_j \cdot HS_j$$

where j is the schooling level (primary, secondary, tertiary);

YR is the number of years of schooling represented by level j;

HS is the fraction of population (+ 6 years) for which the jth level is the highest value attained.

3 DATA ENVELOPMENT ANALYSIS, BOOTSTRAPPING PROCEDURE AND ESTIMATION STRATEGY

Efficiency, TFP growth and its decomposition at the region-level have been measured by means of non parametric DEA-like methods. These methods develop Farrell's ideas (1957) and employ linear programming techniques to measure efficiency as the distance of each statistical unit from a non parametric production frontier, constructed from convex combinations of observed input–output pairs.

Although DEA is mainly used in those management and business studies focusing on firm efficiency, important upgrading of this method have been performed and tested by analysing economy at the country-level. For example, Fare et al. (1994) suggested a technique to decompose the Constant Return to Scale (CRS) measure of efficiency in two components based on the Variable Returns to Scale (VRS) measure of efficiency and the Scale Efficiency (SE) measure, for a panel of OECD countries. Kumar and Russell (2002) applied a similar method to decompose labour productivity growth for 57 countries in the period 1965-1990.

According to these authors, firstly we assumed in our case that Italian regions have convex technology sets. Furthermore, we applied an output orientation DEA: this means that regions are allowed to operate at different degrees of technical and allocative efficiency, i.e different regions with similar level of capital, labour and human capital may produce different amounts of output.

In the empirical analysis concerning whole regional economies we considered two models: the first one includes the canonical two inputs (capital and labour), whereas in the second one, human capital is added to the latter, hence we performed a one output-three input model.

Thus, the approximated technology set, or *Farrel cone*, for the model including human capital is the following:

$$\Psi_t \equiv \left\{ (Y,K,L,HK) \in \Re_+^4 \mid Y \leq \sum_{i=1}^{n} \lambda_i Y_{it}; K \geq \sum_{i=1}^{n} \lambda_i K_{it}; L \geq \sum_{i=1}^{n} \lambda_i L_{it}; HK \geq \sum_{i=1}^{n} \lambda_i HK_{it}; \lambda_i \geq 0 \forall_i \right\} \quad (1)$$

Each observation is interpreted as a unit operation of a linear process operating at the λ_i level.

More formally, the technically efficiency scores can be calculated by solving the following linear program for each observation:

$$\max_{\{\theta,\lambda_1,\ldots\lambda_n\}} \theta^{-1}$$

$$\text{s.t.} \quad \begin{cases} Y_{it} / \theta \leq \sum_{i=1}^{n} \lambda_i Y_{it} \\ K \geq \sum_{i=1}^{n} \lambda_i K_{it} \\ L \geq \sum_{i=1}^{n} \lambda_i L_{it} \\ HK \geq \sum_{i=1}^{n} \lambda_i HK_{it} \\ \sum_{i=1}^{n} \lambda_i = 1 \\ \lambda_i \geq 0 \quad \forall_i \end{cases} \tag{2}$$

It must be remarked that the output-orientated Farrel index is $1 \leq \theta \leq \infty$ and $\theta - 1$ is the proportional increase that could be achieved by the *i-th* region, with input quantities held constant.
θ^{-1} is $0 \leq \theta^{-1} \leq 1$ the output-orientated Shepard index (1970) of technical efficiency, that is the reciprocal of the Farrel index. Thus, our efficiency scores are less or equal to one, and equal unity only if the production process is efficient.
Constraints ensure that the projection points cannot lie outside the feasible set; $\sum_{i=1}^{n} \lambda_i = 1$ allows Variable Returns to Scale (VRS) specification.
Simar and Wilson (1998) stressed that the statistical estimators of the frontier are obtained from finite samples, hence the corresponding measures of efficiency are sensitive to the sampling variations of the obtained frontier. These authors showed how the bootstrap procedure, that simulates a Data Generating Process (DGP) can approximate the sampling variation of the estimated frontier, allowing us to analyse the sensitivity of the efficiency score of a given production unit[4]. Therefore we consider a bias-corrected estimator of θ_i :

$$\tilde{\theta}^*_{i,b} = \hat{\theta}^*_{i,b} - 2bias_i$$

where the empirical density function is $\hat{\theta}^*_{i,b}, b = 1,\ldots B$, and the usual percentile confidence interval for θ_i with intended coverage $(1 - 2\alpha)$ is given by:

$$(\hat{\theta}_{i,low}, \hat{\theta}_{i,up}) = (\tilde{\theta}_i^{*(\alpha)}, \tilde{\theta}_i^{*(1-\alpha)})$$

where $\tilde{\theta}_i^{*(\alpha)}$ is the $100 \cdot \alpha th$ percentile of the empirical density function $\hat{\theta}^*_{i,b}, b = 1,\ldots B$.

[4] A more detailed discussion of the bootstrap procedure is reported in the Appendix.

In the following section we use this bootstrap procedure to perform a sensitivity analysis on the efficiency scores for Italian regions, hence we calculated correction for bias and percentile confidence interval.
As regards the study of TFP growth and its decomposition, we followed Fare *et al.* (1994) and Coelli *et. al* (2005) in defining productivity change as the geometric mean of two output-based Malmquist productiviy indexes:

$$M_i(\mathbf{X}_i^{t+1}, y_i^{t+1}, \mathbf{X}_i^t, y_i^t) = \frac{D_i^{t+1}(\mathbf{X}_i^{t+1}, y_i^{t+1})}{D_i^t(\mathbf{X}_i^t, y_i^t)} \times \left[\left(\frac{D_i^t(\mathbf{X}_i^{t+1}, y_i^{t+1})}{D_i^{t+1}(\mathbf{X}_i^{t+1}, y_i^{t+1})} \right) \left(\frac{D_i^t(\mathbf{X}_i^t, y_i^t)}{D_i^{t+1}(\mathbf{X}_i^t, y_i^t)} \right) \right]^{1/2} \quad (21)$$

where $\mathbf{X}_i^t$ and $\mathbf{X}_i^{t+1}$ are the vectors of inputs (K, L, HK) in years t and t+1;

y_i^t and y_i^{t+1} are the outputs in years t and t+1;

$\frac{D_i^{t+1}(\mathbf{X}_i^{t+1}, y_i^{t+1})}{D_i^t(\mathbf{X}_i^t, y_i^t)}$ is the ratio of distance functions measuring efficiency change;

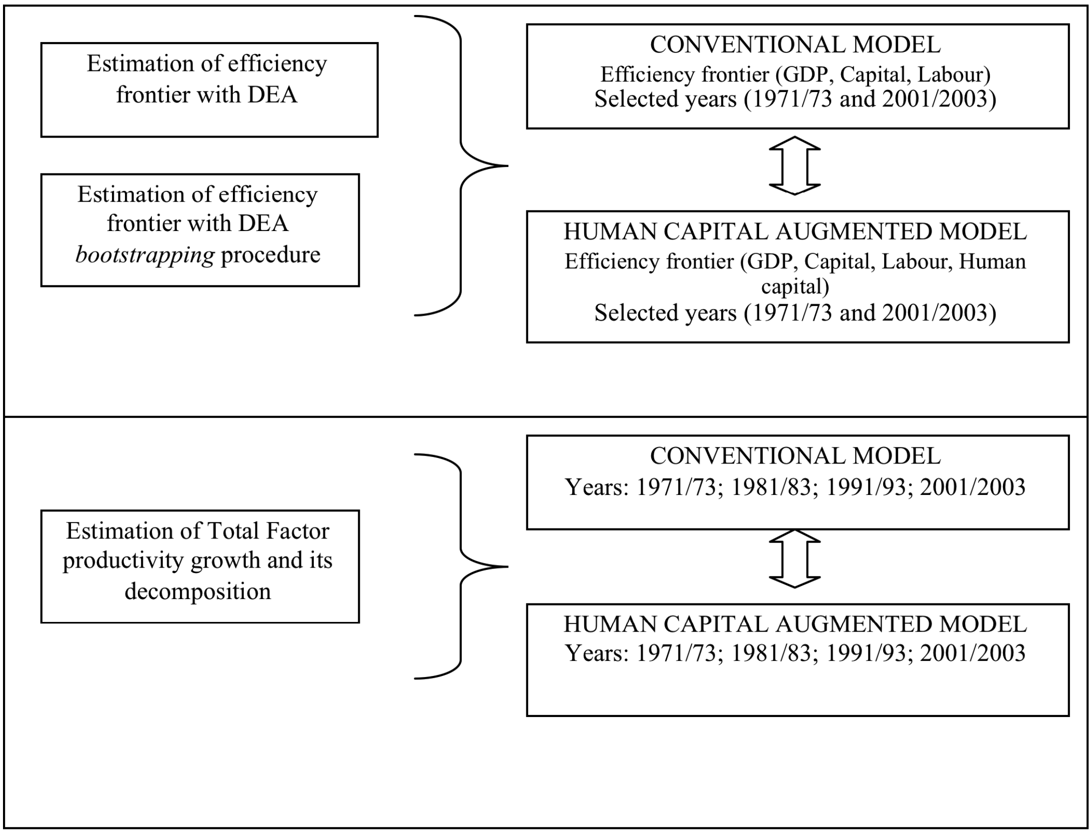
$\left[\left(\frac{D_i^t(\mathbf{X}_i^{t+1}, y_i^{t+1})}{D_i^{t+1}(\mathbf{X}_i^{t+1}, y_i^{t+1})} \right) \left(\frac{D_i^t(\mathbf{X}_i^t, y_i^t)}{D_i^{t+1}(\mathbf{X}_i^t, y_i^t)} \right) \right]^{1/2}$ is the term measuring technological change.

We also relaxed the CRS assumption, hence the Malmquist index can be decomposed in

$M_i(\mathbf{X}_i^{t+1}, y_i^{t+1}, \mathbf{X}_i^t, y_i^t)$ = (Pure eff. change) x (Scale eff. change) x (Tech. change)

Figure 1 sums up the strategy of the empirical analysis. In the first part we calculated efficiency scores for the whole regional economy in both the standard model of production process and the human capital augmented model. Moreover we performed these estimations by using both the canonical DEA-model and the bootstrap procedure model. It must be remarked that the human capital measure and the proxy of external economies are based on ISTAT Census data, available at ten-year intervals. For this reason we performed DEA considering four ten-year-intervals (1971/73, 1981/83, 1991/93, 2001/2003). More precisely for annual data such as GDP, labour and capital stock we took the average of the first three years of the decade.
In the second part we estimated TFP growth and its decomposition over three decades (1970s; 1980s; 1990s) by means of the Malmquist index model (Coelli et al. 2006).

Figure 1 Strategy of the empirical analysis



4 RESULTS OF THE EMPIRICAL ANALYSIS

Before showing the results of non-parametric estimations, we provide some statistics concerning the variables that enter the regional production process. Afterwards, we estimated efficiency, TFP growth and its decomposition both in a standard production process, where capital and labour are the only inputs, and in an augmented human capital model of production.

4.1 A preliminary descriptive analysis

Table 1 shows the well-known heterogeneity between the Northern and Central regions on the one hand, and Southern regions on the other. In this table regions are ranked according to the labour productivity level (GDP/L) achieved at the end of the period (2001/2003). Firstly, we can see almost all regions at the top of ranking in the beginning of period maintaining their position in 2001/03. It is also worth noting that regions such as Liguria and Friuli remarkably gained ground in terms of labour productivity levels. Conversely, the group of regions placed under the Italian average, both at the beginning and at the end of the period, includes all the Southern regions and some regions of the Central Italy, such as Marche and Umbria.
At first glance, we can also see that the higher productivity levels are, the higher are human capital levels, measured as average years of schooling of the population aged 6 and over.

Table 1 GDP per worker, capital/labour ratio and human capital (HK) in selected years

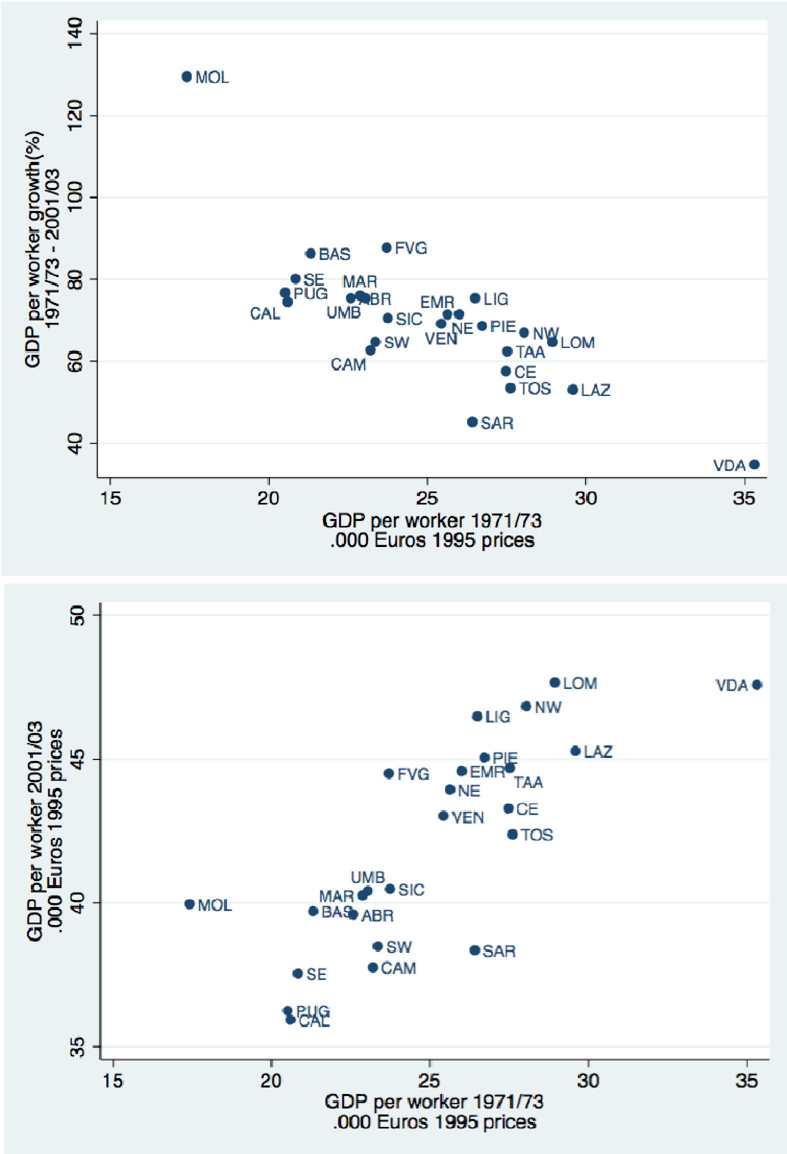
	1971/73			2001/03		
Regions	**GDP/L**	**K/L**	**HK**	**GDP/L**	**K/L**	**HK**
Lombardy (LOM)	28.93	119.61	5.95	47.67	157.61	6.38
V.d'Aosta (VDA)	35.30	204.52	5.67	47.59	214.33	6.50
Liguria (LIG)	26.50	141.79	5.93	46.48	162.63	6.33
Latium (LAZ)	29.58	103.04	5.26	45.28	126.87	5.99
Piedmont (PIE)	26.72	100.56	5.91	45.06	160.03	6.41
Trentino (TAA)	27.52	120.91	6.66	44.70	169.71	6.47
Emilia Romagna (EMR)	26.00	96.63	5.29	44.59	130.41	6.17
Friuli (FVG)	23.71	124.34	6.03	44.51	157.41	6.38
Veneto (VEN)	25.43	129.94	5.85	43.03	147.36	6.33
Tuscany (TOS)	27.61	98.50	5.33	42.38	121.34	6.24
Sicily (SIC)	23.74	121.87	4.09	40.49	161.45	5.81
Umbria (UMB)	23.04	145.22	4.64	40.42	150.58	5.93
Marche (MAR)	22.87	106.28	4.56	40.26	130.77	6.02
Molise (MOL)	17.42	99.71	4.18	39.96	160.82	5.73
Basilicata (BAS)	21.31	134.27	3.88	39.71	198.58	5.56
Abruzzo (ABR)	22.58	121.00	4.34	39.60	143.35	5.83
Sardinia (SAR)	26.41	165.35	4.40	38.36	181.44	6.08
Campania (CAM)	23.20	121.69	4.40	37.76	152.73	5.89
Abulia (PUG)	20.51	104.15	4.22	36.25	128.37	5.88
Calabria (CAL)	20.59	115.70	3.85	35.94	162.05	5.59
Sample mean	**24.949**	**123.754**	**5.02**	**42.001**	**155.892**	**6.08**

Of course in the period 1971/73 – 2001/03 a sort of β-convergence process occurred, both in labour productivity terms and in human capital terms. The top panels of figures 2 and 3 clearly reveal that the Southern regions, with lower levels of labour productivity and human capital, showed better performances. In particular, the South-East (SE in the figure is the mean of Abruzzo, Molise, Apulia and Basilicata) was at the top of the growth rates in productivity, whereas the South-West (SW in the figure is the mean of Campania, Calabria, Sicily and Sardinia) did so in the case of human capital. Nonetheless, the panels at the bottom of the same figures, that show the scatter diagram between productivity levels at the beginning and at the end of the period respectively, also show that convergence is not concluded.

Indeed, in these panels the regional divide is still evident since all the Mezzogiorno regions are located in the lower-left corner of the figures.

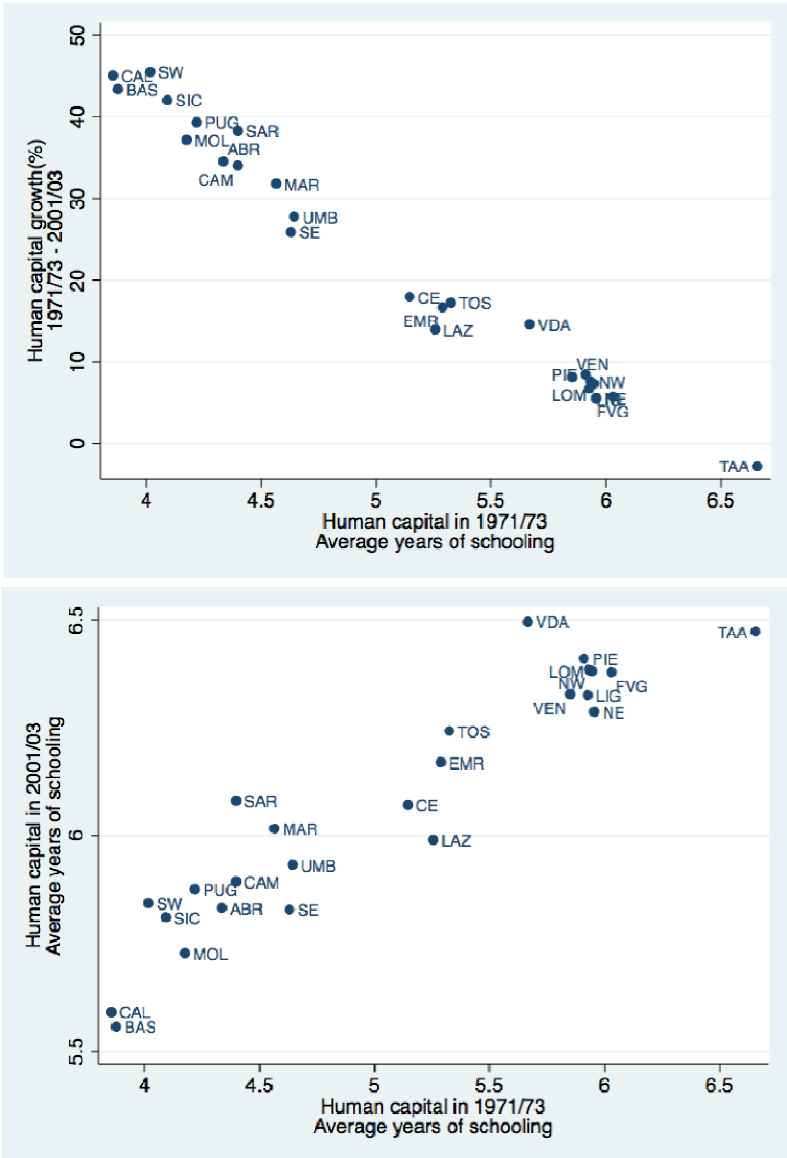
These findings are perfectly coherent with that ones discussed by Maffezzoli (2006).

Figure 2 Labour productivity convergence 1971/73 – 2001/03



* NW=North-West; NE=North-East; CE=Centre ;SE=South-East; SW=South- West

Figure 3 Human capital convergence 1971/73 – 2001/03



* NW=North-West; NE=North-East; CE=Centre ;SE=South-East; SW=South- West

Additional information is obtained by considering the cumulated growth rates of variables whose levels have been used in DEA to calculate the efficiency frontier (GDP, Capital, Labour and Human Capital). It is worth noting, among the other statistics of table 2, that only the GDP (not labour productivity) more than doubled in the North-Eastern and Central regions between 1971/73 and 2001/03, but increased less than the national average in almost all Southern regions (Molise and Abruzzo are the sole exceptions). It means that the findings discussed above, and concerning the improvement of labour productivity in the South, are probably affected by the sluggish growth in employment that has been plaguing the Mezzogiorno for decades.

The same table 2 also reveals that higher cumulated growth rates of average years of schooling (human capital) have not been sufficient to improve GDP growth.

Table 2 Growth rates from 1971/73 to 2001/2003
Gross Domestic Product (GDP), Capital (K), Labour (L), Human Capital (HK)

Regions	Δ GDP	Δ K	Δ L	Δ HK
Veneto	134.64	57.25	38.67	8.14
Molise	131.12	62.44	0.71	37.17
Trentino	127.25	96.36	39.89	-2.78
Umbria	122.60	31.59	26.90	27.77
Emilia Romagna	121.12	74.01	28.94	16.67
Abruzzo	121.09	49.32	26.05	34.52
Marche	114.98	50.29	22.14	31.79
Friuli	112.01	42.97	12.94	5.78
Latium	106.92	66.43	35.17	13.97
Lombardy	101.84	61.45	22.52	7.33
Apulia	98.40	38.37	12.26	39.30
Campania	95.89	51.07	20.37	34.02
Tuscany	89.06	51.75	23.18	17.22
Sicily	87.13	45.38	9.74	42.01
Sardinia	81.14	36.87	24.73	38.27
Calabria	79.21	43.83	2.69	45.04
Basilicata	78.09	41.37	-4.41	43.36
Piedmont	74.12	64.29	3.24	8.45
Liguria	72.39	12.72	-1.73	6.73
V.d'Aosta	57.68	22.60	16.99	14.60
Sample mean	**100.33**	**50.02**	**18.05**	**23.47**

4.2 Efficiency scores with DEA

Tables 3 and 4 show the efficiency scores of regions calculated with the standard output-orientated DEA procedure[5] (Coelli *et al.* 2005). It must be remarked that output-orientation distance function measures the radial expansion of output when input levels are held fixed. In other words, scores equal to one mean that the region is on the efficiency frontier, whereas scores that are less than one signal the size of technical inefficiency: how much output could be expanded without increasing use of inputs. For example, if we consider the case of Tuscany and Calabria, with constant returns to scale (see table 3), we can see that the former, by improving an efficient use of the same level of capital and labour, could increase GDP by 3% in 1971/73 and by 1% in 2001/03, whereas Calabria to be efficient should increase its GDP by 32% in 1971/73 and 25% in 2001/03.
By maintaining our focus on the baseline model of table 3, namely the one-output (GDP) two-inputs (Capital and Labour) model, it also worth noting that only three regions emerged as efficient: Latium and Val d'Aosta in 1971/73, Lombardy and Latium in 2001/03. However, other regions such as Tuscany and Emilia Romagna are located very close to the frontier. Conversely, all Southern regions improved their

[5] We used DEAP as computer program to conduct this DEA procedure (see Coelli *et al.* 2005)

efficiency scores (with the exception of Sardinia) but remained remarkably inefficient.
These results are coherent with that ones concerning similar analyses and reported in the literature (Maffezzoli, 2006; Bollino and Polinori, 2007).
We also assumed that regions are not always operating at the optimal scale, hence a Variable Returns to Scale (VRS) specification is needed. Through this assumption we want to avoid confounding scale inefficiencies with technical inefficiencies, namely that there is a problem of different scale in the use of input levels.
In the baseline model of table 3, scale inefficiencies are not important; if we consider all sample they accounted for about 4% in 1971/73 (it is the complement to the sample mean 0.96, in column SE) and 1% in 2001/03 (it is the complement to the sample mean 0.99).

Table 3 Efficiency scores in selected years calculated with standard DEA Output distance functions (baseline model)

	1971/73			2001/2003		
	CRS	**VRS**	**SE**	**CRS**	**VRS**	**SE**
Lombardy	0.95	1.00	0.95	**1.00**	**1.00**	**1.00**
Latium	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
V.d'Aosta	**1.00**	**1.00**	**1.00**	0.99	1.00	0.99
Tuscany	0.97	0.98	0.99	0.99	1.00	0.99
Emilia Romagna	0.92	0.93	1.00	0.98	0.98	1.00
Liguria	0.84	0.89	0.95	0.97	0.97	1.00
Piedmont	0.92	0.94	0.98	0.94	0.94	1.00
Trentino	0.90	0.90	0.99	0.93	0.93	1.00
Veneto	0.82	0.86	0.96	0.92	0.92	1.00
Friuli	0.76	0.78	0.98	0.92	0.93	0.99
Marche	0.76	0.77	0.98	0.89	0.91	0.97
Umbria	0.72	0.75	0.96	0.85	0.87	0.98
Abruzzo	0.73	0.74	0.99	0.85	0.86	0.99
Sicily	0.78	0.80	0.98	0.83	0.83	1.00
Molise	0.63	1.00	0.63	0.82	0.91	0.90
Basilicata	0.67	0.68	0.99	0.82	0.82	1.00
Apulia	0.70	0.71	0.99	0.80	0.80	0.99
Campania	0.78	0.80	0.97	0.79	0.79	1.00
Sardinia	0.84	0.90	0.93	0.79	0.79	1.00
Calabria	0.68	0.69	0.99	0.75	0.75	1.00
Sample mean	**0.82**	**0.86**	**0.96**	**0.89**	**0.90**	**0.99**

CRS = Constant Return to Scale technical efficiency
VRS = Variable Return to Scale technical efficiency
SE= Scale efficiency (CRS/VRS)

By contrast with the baseline model, scale inefficiencies seem to be important in the model that includes human capital among inputs (see table 4). In this case things change little if we take efficiency scores into the CRS specification (Lombardy, Latium and Val d'Aosta maintain their position as efficient regions), but become notably different in the VRS specification. We not only find the three regions mentioned above on the VRS efficiency frontier, but also Southern regions such as Molise, Basilicata and Calabria. This means that if we introduce a third input as human capital and allow regions with different scales (different levels of used inputs) to be benchmarked only with other regions of similar size, these last three Southern

regions appear as efficient. In other words, given the low level of human capital that they have accumulated, they efficiently use their inputs. Since in the baseline two input model (labour and capital only), Molise, Basilicata and Calabria were not on the VRS efficiency frontier, we can deduce that for these regions a problem of human capital accumulation is evident.
This line of reasoning could be plausible but relies upon non robust results. Indeed, we need to implement at least two kind of tests: the first one is aimed to prove that VRS scores are significantly different from the CRS ones; the second group of tests should prove statistical difference between the baseline model and the human capital augmented model.

Table 4 Efficiency scores in selected years calculated with standard DEA Output distance functions (HK augmented model)

	1971/73			2001/2003		
	CRS	**VRS**	**SE**	**CRS**	**VRS**	**SE**
Latium	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
Lombardy	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
V.d'Aosta	**1.00**	**1.00**	**1.00**	**0.99**	**1.00**	**0.99**
Tuscany	0.97	0.98	0.99	0.99	1.00	0.99
Molise	0.63	1.00	0.63	0.82	1.00	0.82
Basilicata	0.69	1.00	0.69	0.82	1.00	0.82
Calabria	0.69	1.00	0.69	0.75	1.00	0.75
Emilia Romagna	0.92	0.93	1.00	0.98	0.98	1.00
Liguria	0.89	0.89	1.00	0.97	0.98	0.99
Piedmont	0.92	0.94	0.98	0.94	0.95	1.00
Friuli	0.78	0.78	1.00	0.92	0.93	0.99
Trentino	0.90	0.90	1.00	0.93	0.93	1.00
Veneto	0.86	0.86	1.00	0.92	0.92	1.00
Marche	0.76	0.81	0.94	0.89	0.92	0.96
Umbria	0.75	0.86	0.88	0.85	0.89	0.95
Abruzzo	0.74	0.83	0.89	0.85	0.88	0.96
Sicily	0.80	1.00	0.80	0.83	0.86	0.97
Apulia	0.70	0.91	0.77	0.80	0.81	0.99
Campania	0.80	0.91	0.88	0.79	0.80	0.99
Sardinia	0.91	1.00	0.91	0.79	0.81	0.98
Sample mean	**0.84**	**0.93**	**0.90**	**0.89**	**0.93**	**0.96**

CRS = Constant Return to Scale technical efficiency
VRS = Variable Return to Scale technical efficiency
SE= Scale efficiency (CRS/VRS)

In order to improve the statistical properties of the estimators of production frontier and to perform reliable hypothesis tests we re-ran the same DEA analysis by implementing the bootstrap procedure (Simar and Wilson, 1998)[6]. This method is also important because estimators of the frontier are obtained from finite samples, the corresponding measures of efficiency are sensitive to the sampling variations of the efficiency frontier. Therefore, a sensitive analysis of efficiency scores allows us to obtain more accurate information about differences and similarities with statistical significance.

---

[6] We used FEAR, a software package that uses the R environment (see Wilson, 2006).

Tables 5 and 6 show efficiency scores calculated by the bootstrap procedure. First of all, we can notice that the things change little. If we focus on CRS specification, again Umbria, Marche and all Southern regions are placed under the national average in both the standard and human capital model. Conversely, in the VRS specification of the human capital model (see table 6) we find, as in the previous analysis, Molise, Basilicata and Calabria on the frontier.
However, in this case we could perform hypothesis tests concerning returns to scale and model specification.

Table 5 Efficiency scores in selected years calculated with bootstrap procedure (Simar and Wilson 1998)
Output distance functions (baseline model)

	1971/73			2001/2003		
	CRS	**VRS**	**SE**	**CRS**	**VRS**	**SE**
Lombardy	**0.95**	**1.00**	**0.95**	**1.00**	**1.00**	**1.00**
Latium	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
V.d'Aosta	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
Emilia Romagna	0.94	0.94	1.00	0.98	0.98	1.00
Tuscany	0.98	0.98	0.99	0.98	0.99	0.99
Liguria	0.83	0.89	0.94	0.98	0.98	1.00
Piedmont	0.93	0.94	0.98	0.95	0.95	1.00
Trentino	0.90	0.91	0.99	0.94	0.94	1.00
Friuli	0.77	0.79	0.98	0.93	0.94	0.99
Veneto	0.82	0.86	0.95	0.92	0.92	1.00
Marche	0.77	0.79	0.97	0.88	0.90	0.98
Umbria	0.72	0.76	0.95	0.86	0.87	0.99
Abruzzo	0.74	0.75	0.99	0.85	0.86	0.99
Sicily	0.77	0.80	0.97	0.85	0.85	1.00
Molise	0.61	0.96	0.64	0.84	0.94	0.89
Basilicata	0.68	0.69	0.99	0.83	0.83	1.00
Sardinia	0.80	0.88	0.91	0.80	0.80	1.00
Campania	0.76	0.78	0.97	0.80	0.80	1.00
Apulia	0.69	0.69	1.00	0.80	0.80	1.00
Calabria	0.68	0.69	0.99	0.75	0.75	1.00
Sample mean	**0.82**	**0.86**	**0.96**	**0.90**	**0.90**	**0.99**

CRS = Constant Return to Scale technical efficiency
VRS = Variable Return to Scale technical efficiency
SE= Scale efficiency (CRS/VRS)

Table 7 displays two tests for returns to scale proposed by Simar and Wilson (2002). The null hypothesis to be tested is that the production process exhibits globally constant returns to scale. The *p-value* is the probability to reject the null hypothesis when it is true (that is the probability to make Type 1 error). According to these tests, we should accept the null hypothesis of globally CRS. However, Simar and Wilson warned that these test statistics do not perform well with low observation numbers (less than 40). For this reason we also carried out a Spearman rank correlation, as suggested by Parkin and Hollingsworth (1997). Indeed, table 7 also shows that the Spearman correlation between CRS and VRS ranking is not so high in the case of human capital augmented model. Therefore, we should pay attention to both specifications concerning returns to scale.

This result led us to perform tests for model specification on VRS efficiency scores. Table 8 shows two tests proposed by Banker et al. (1996). The null hypothesis is that an additional input (human capital in our case) does not influence the production correspondence, namely it does not contribute significantly to the production process. In three cases out of four the null hypothesis cannot be rejected, even though the Spearman rank correlation tells us that both at the beginning and at the end of the period the correlation between VRS efficiency score rankings were not particularly high.

Table 6 Efficiency score in selected years calculated with bootstrap procedure (Simar and Wilson 1998)
Output distance functions (HK augmented model)

	1971/73			2001/2003		
	CRS	**VRS**	**SE**	**CRS**	**VRS**	**SE**
Lombardy	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
Latium	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
V.d'Aosta	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
Emilia Romagna	0.94	0.94	1.00	0.98	0.98	1.00
Tuscany	0.98	0.98	0.99	0.98	0.99	0.99
Liguria	0.89	0.89	1.00	0.98	0.98	0.99
Piedmont	0.93	0.94	0.98	0.95	0.95	1.00
Trentino	0.90	0.91	0.99	0.94	0.94	1.00
Friuli	0.79	0.79	1.00	0.93	0.95	0.99
Veneto	0.86	0.86	1.00	0.92	0.92	1.00
Marche	0.77	0.82	0.94	0.88	0.90	0.98
Umbria	0.76	0.86	0.88	0.86	0.89	0.96
Abruzzo	0.75	0.85	0.89	0.85	0.88	0.97
Sicily	0.80	1.00	0.80	0.85	0.87	0.97
Molise	0.61	1.00	0.61	0.84	1.00	0.84
Basilicata	0.70	1.00	0.70	0.83	1.00	0.83
Sardinia	0.88	0.97	0.91	0.80	0.82	0.98
Campania	0.79	0.89	0.89	0.80	0.81	0.99
Apulia	0.69	0.91	0.76	0.80	0.80	0.99
Calabria	0.69	1.00	0.69	0.75	1.00	0.75
Sample mean	**0.83**	**0.93**	**0.89**	**0.89**	**0.93**	**0.96**

CRS = Constant Return to Scale technical efficiency

VRS = Variable Return to Scale technical efficiency

SE= Scale efficiency (CRS/VRS)

Table 7 Tests of returns to scale concerning efficiency scores calculated with bootstrap methodology

	H_0: production frontier is globally CRS p-values for H_0 with bootstrapping procedure (Simar and Wilson, 2002)		Spearman Rank Correlation
	Mean of ratios (S_{1n}^{crs})	Ratio of means (S_{2n}^{crs})	
Baseline model			
1971/73	0.96	0.95	0.80 (p=0.00)
2001/03	0.99	0.99	0.96 (p=0.00)
Human Capital augmented model			
1971/73	0.90	0.89	0.22 (p=0.33)
2001/03	0.96	0.96	0.48 (p=0.03)

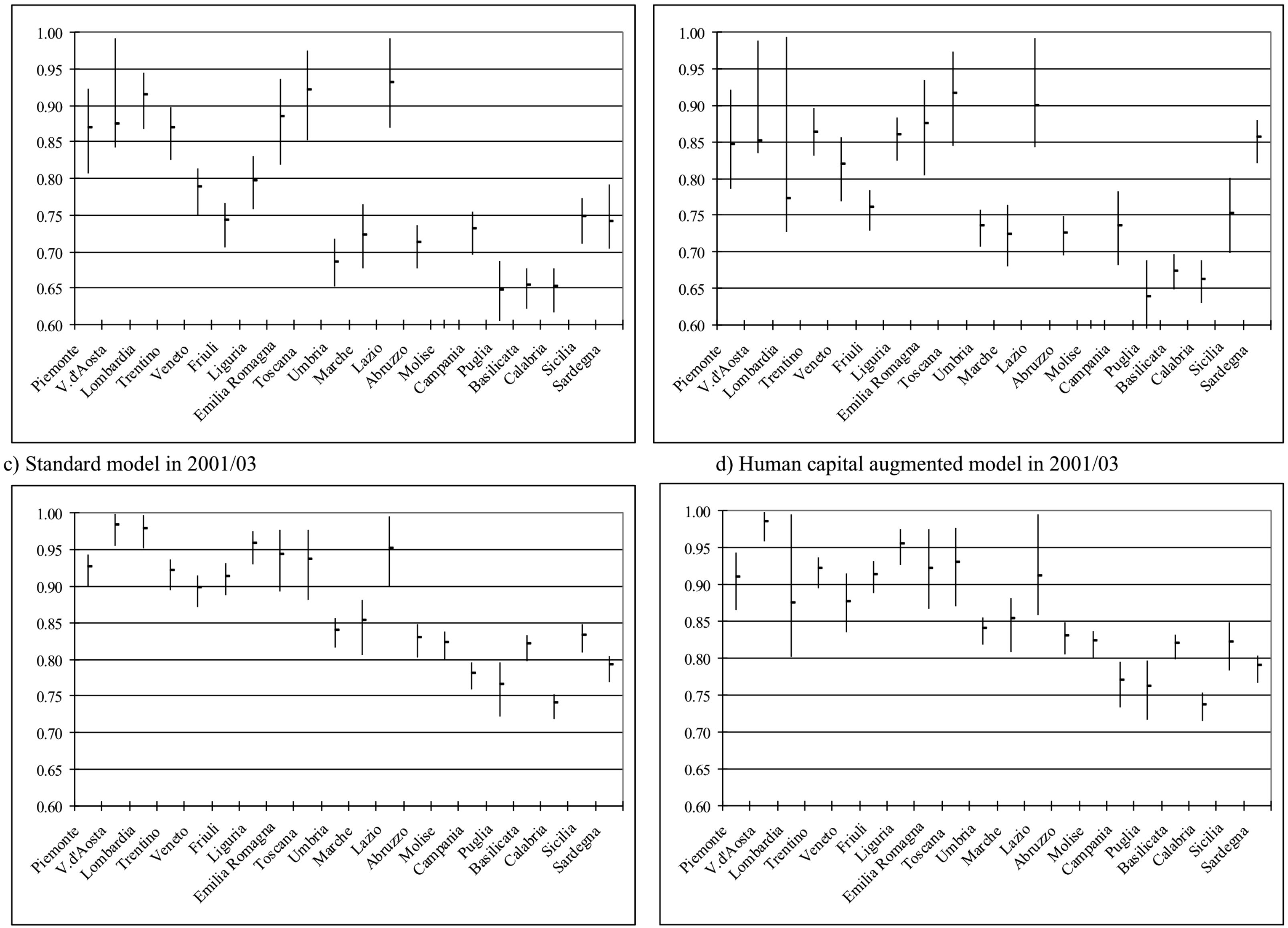
In summary, we can say that DEA analysis conducted with the bootstrap procedure confirm the findings of standard DEA (tables 3 and 4), in which the introduction of human capital in the regional production process somehow provides us with new information. The problem of scarce accumulation of human capital, as underlined by Faini and Sapir (2005), concerns all Southern regions in general terms but it is particularly prominent in regions such as Molise, Basilicata and Calabria. Our results highlight that in these contexts an increase in the average years of schooling could improve, more than elsewhere, an efficient use of standard inputs (capital and labour), given that these regions are located on the VRS frontier. However, we also have to take into account that in other Southern regions, in which human capital levels resulted higher at the end of period (see figure 2), the efficient use of inputs worsened (see for example the case of Sicily, Sardinia, Campania and Apulia in the tables 4 and 6).

Table 8 Tests for Model Specification on VRS efficiency scores calculated with bootstrap procedure

	Output	**Input1**	**Input 2**	**Input3**
Baseline Model (1)	GDP	Capital	Labour	
Human capital Model (2)	GDP	Capital	Labour	Human Capital
H0: Human Capital do not influence prod. correspondence	**Exponential distribution Banker et al. Test (1995)**		**Half-normal distribution Banker et al. Test (1995)**	
	1971/73	**2001/03**	**1971/73**	**2001/03**
Test. Stat.	1.94	1.44	2.92	1.67
Critical F.	2.12	2.12	1.69	1.69
Spearman Rank Correlations				
	1971/73		**2001/03**	
Model 1 *vs* Model 2	0.40 *(p=0.07)*		0.58 *(p=0.00)*	

To conclude this section we present the sensitivity analysis conducted on the bias-corrected estimators discussed in section 3. The interest of figure 4 not only relies on the clear picture of regional divide that it discloses, but it is also grounded on the different information that we can obtain when we take into account the bias of the efficiency score estimator and its confidence interval. First of all, we can see in the human capital augmented model (panels *b* and *d*) that regions on the efficiency frontier, such as Lombardy and Latium are very sensible to sampling variations; in fact, their confidence intervals are very large. Conversely, Valle d'Aosta confirms its good position, accompanied by Liguria, that shows a bias-corrected efficiency above 0.95 and a very small confidence interval. It is worth noting that between 1971 and 2003, the relative position of this region, compared to the ones of other efficient regions such as Emilia Romagna and Tuscany, remarkably improved. Other important information is that the confidence intervals for the efficiency of Latium, Lombardy and other Northern and Central regions (with exception of Marche and Umbria) overlap to a large degree. Therefore, we can say that the Northern and Central regions are significantly more efficient than Southern regions, while Marche and Umbria seem to be a bridge between these two macro-areas. As regards the South, in 1971/73 we could identify three different groups in terms of technical efficiency. Sardinia was at the top, in a position not significantly different from the one of the North-Centre; Abruzzo, Campania and Sicily were in an intermediate position, whereas Apulia, Basilicata and Calabria were located at the bottom of the ranking. In 2001/03 we maintain three groups of Southern regions, but the composition is remarkably changed: Calabria remained alone at the bottom of ranking, Campania, Apulia and Sardinia are located at an intermediate position, while Abruzzo, Molise, Basilicata and Sicily rose to the top of the Southern regions'ranking, even though their efficiency resulted significantly lower that of the Northern and Central regions.

Figure 4 Bias-corrected estimates and confidence intervals



a) Standard model in 1971/73

b) Human capital augmented model in 1971/73

c) Standard model in 2001/03

d) Human capital augmented model in 2001/03

4.3 Total Factor Productivity Growth and its decomposition

After we discussed what occurred at the beginning and at the end of the 1971/73-2001/03 period, we can explore the regional TFP growth and its components. As mentioned in section 3, TFP growth can be decomposed first into: *1) technological change* (namely, how much the shift of the efficiency frontier between *t* and *t+1* contributed to the growth of productivity) and *2) efficiency change* (namely, how much the change in the relative distance from the efficiency frontier between *t* and *t+1* contributed to productivity growth). In turn, by relaxing the CRS assumption, this last measure of general technical efficiency change can be decomposed into *a) scale efficiency change* (i.e., productivity changed because a region improved or worsened its distance from an optimal scale point) and *b) pure efficiency change* (i.e., productivity changed because a region effectively improved or worsened the efficient use of inputs).

Table 9 reports the measures mentioned above for the three decades under examination (1970s; 1980s and 1990s), the bottom-right panel sums up these values over the whole period 1971/73-2001/03. First of all, we can see that the average national TFP growth was down by half, from 26.8% in the 1970s to 13.7% in the 1990s. In the bottom-right panel reporting the whole time interval, we can observe that Southern regions (with the exception of Sardinia) on average show a TFP growth above the national average. This result confirms that a convergence process occurred. However, if we look at the three panels we observe that for Sicily, Apulia, Abruzzo and Molise the convergence process is remarkable only in the 1970s and partially in the 1980s; Campania and Basilicata start to converge in the 1980s, whereas Calabria and Sardinia did so in the 1990s. Its worth noting that in all these cases the major contribution to TFP growth stems from technological change. For example, the value of 20.6% representing the ten-year average growth of Sicily (see the bottom-right panel) was mainly composed of technological change (19%), whereas the general efficiency change contribution of 1.40%, resulted from an improvement of scale efficiency 6.60% (probably it is the human capital accumulation) and from a negative variation of pure efficiency (-4.90%). Indeed the bottom-right panel shows that all the Southern regions registered null or negative variation of the pure efficiency change. This means that the problem relies not only on the scarce accumulation of human capital in Southern regions, but involves the same quality of this human capital. Our findings tell us that human capital accumulated in the South of Italy did not improve an efficient use of inputs.

As regards technical change, according to Fare *et al.* (1994), we can examine the single components of distance functions in order to identify those regions that caused the frontier shift. More precisely, for each region *i*, we can say that this region caused a frontier shift if

$$D_i^t(\mathbf{X}_i^{t+1}, y_i^{t+1}) > 1$$

and

$$D_i^{t+1}(\mathbf{X}_i^{t+1}, y_i^{t+1}) = 1$$

Table 9 Scores of Malmquist analysis in HK model

	change between 1981/83 and 1971/73					change between 1991/93 and 1981/83				
	EFFch	**TECHch**	**PEch**	**SEch**	**TFPch**	**EFFch**	**TECHch**	**PEch**	**SEch**	**TFPch**
Calabria	-1.60	24.50	-9.00	8.20	22.50	4.10	17.30	9.90	-5.30	22.20
Friuli	3.60	23.90	3.40	0.10	28.30	7.70	17.10	8.10	-0.30	26.20
Piedmont	0.00	25.40	-1.80	1.90	25.40	2.50	15.40	2.40	0.10	18.30
Abruzzo	8.90	24.00	2.60	6.20	35.10	3.00	17.10	0.70	2.30	20.60
Marche	-0.70	24.80	-1.00	0.30	23.90	4.70	10.90	0.40	4.30	16.10
Trentino	3.60	23.30	3.20	0.40	27.70	0.30	16.20	0.20	0.10	16.50
V.d'Aosta	0.00	15.30	0.00	0.00	15.30	0.00	14.20	0.00	0.00	14.20
Lombardy	0.00	25.90	0.00	0.00	25.90	0.00	22.50	0.00	0.00	22.50
Liguria	8.70	24.30	8.60	0.10	35.10	-0.40	17.40	0.00	-0.40	16.90
Latium	0.00	25.30	0.00	0.00	25.30	0.00	10.50	0.00	0.00	10.50
Molise	14.20	23.90	0.00	14.20	41.50	12.80	15.60	0.00	12.80	30.40
Basilicata	-3.10	22.80	0.00	-3.10	19.00	11.80	16.80	0.00	11.80	30.70
Veneto	2.80	24.80	2.80	0.00	28.40	-1.30	16.50	-1.20	-0.10	14.90
Tuscany	-2.20	25.90	-2.10	-0.10	23.20	-1.30	7.10	-1.20	-0.10	5.70
Umbria	9.50	23.40	2.50	6.80	35.00	-0.20	17.10	-2.50	2.40	16.80
Sicily	8.90	24.90	0.00	8.90	36.00	-2.80	17.50	-3.30	0.40	14.20
Emilia Romagna	1.20	25.90	2.10	-0.90	27.40	-2.40	10.50	-3.60	1.20	7.90
Campania	-6.50	24.80	-7.50	1.10	16.80	3.00	16.40	-4.60	8.00	19.80
Apulia	4.00	25.30	-8.80	14.00	30.30	1.60	15.80	-9.00	11.70	17.70
Sardinia	-5.40	24.10	-9.80	4.90	17.40	-10.70	17.30	-12.70	2.30	4.80
Sample mean	***2.20***	***24.10***	***-0.90***	***3.00***	***26.80***	***1.50***	***15.40***	***-0.90***	***2.50***	***17.10***
	change between 2001/03 and 1991/93					**Summary of region means over the three decades**				
	EFFch	**TECHch**	**PEch**	**SEch**	**TFPch**	**EFFch**	**TECHch**	**PEch**	**SEch**	**TFPch**
Friuli	5.30	13.20	6.50	-1.10	19.20	**5.50**	**18.00**	**6.00**	**-0.40**	**24.50**
Marche	12.60	4.90	14.30	-1.40	18.20	**5.40**	**13.20**	**4.30**	**1.00**	**19.30**
Liguria	1.00	13.10	1.30	-0.30	14.20	**3.00**	**18.10**	**3.20**	**-0.20**	**21.70**
Veneto	5.40	9.40	5.30	0.10	15.40	**2.30**	**16.70**	**2.30**	**0.00**	**19.40**
Emilia Romagna	7.20	3.90	7.20	0.10	11.50	**1.90**	**13.10**	**1.80**	**0.10**	**15.30**
Abruzzo	2.00	10.80	2.20	-0.20	13.10	**4.60**	**17.20**	**1.80**	**2.70**	**22.60**
Umbria	3.20	11.60	3.90	-0.70	15.20	**4.10**	**17.20**	**1.20**	**2.80**	**22.00**
Trentino	-1.00	12.40	-0.80	-0.20	11.30	**0.90**	**17.20**	**0.90**	**0.10**	**18.30**
Tuscany	5.80	-0.20	5.40	0.40	5.60	**0.70**	**10.40**	**0.60**	**0.10**	**11.20**
Piedmont	0.10	12.40	0.10	0.00	12.60	**0.90**	**17.60**	**0.20**	**0.70**	**18.60**
V.d'Aosta	-1.00	10.40	0.00	-1.00	9.40	**-0.30**	**13.30**	**0.00**	**-0.30**	**12.90**
Lombardy	0.00	7.60	0.00	0.00	7.60	**0.00**	**18.40**	**0.00**	**0.00**	**18.40**
Latium	0.00	2.40	0.00	0.00	2.40	**0.00**	**12.30**	**0.00**	**0.00**	**12.30**
Molise	0.80	12.90	0.00	0.80	13.80	**9.10**	**17.40**	**0.00**	**9.10**	**28.10**
Basilicata	10.60	13.70	0.00	10.60	25.70	**6.20**	**17.70**	**0.00**	**6.20**	**25.00**
Calabria	5.90	13.80	0.00	5.90	20.50	**2.80**	**18.40**	**0.00**	**2.80**	**21.70**
Apulia	7.60	5.90	6.60	0.90	13.90	**4.40**	**15.40**	**-4.00**	**8.70**	**20.50**
Campania	2.00	13.70	-0.60	2.60	15.90	**-0.60**	**18.20**	**-4.30**	**3.80**	**17.50**
Sicily	-1.50	14.70	-11.00	10.80	13.00	**1.40**	**19.00**	**-4.90**	**6.60**	**20.60**
Sardinia	3.10	14.50	2.80	0.30	18.00	**-4.50**	**18.60**	**-6.80**	**2.50**	**13.30**
Sample mean	***3.40***	***10.00***	***2.00***	***1.30***	***13.70***	***2.30***	***16.40***	***0.10***	***2.30***	***19.10***

Note:
EFFch *Efficiency change*
TECHch *Technological change*
PEch *Pure Efficiency ch'ange (it measures the Catching up)*
SEch *Scale efficiency change*
TFPch *Total Factor Productivity change*
TFPch = TECHch + EFFch of which EFFch = Pech + Sech

Following this method, we found that only the efficient regions identified in the DEA analysis caused a shift in the frontier (Valle d'Aosta, Lombardy and Latium). If we also take into account the findings of the bootstrap procedure, we can encompass most of the Northern and Central regions in this club. In any case, the Southern regions did not contributed to the frontier shift, hence the short-lived convergence process depends on these explanations: it was not grounded on the improvement of an efficiency use of inputs and was not driven by endogenous technological change.

5 CONCLUSIONS

In this paper we studied the efficiency and TFP growth of Italian regions by paying attention to both sectoral and territorial dimensions. First of all, we aimed to take a step forward in the field of empirical research, by implementing a bootstrap procedure. The bootstrap method applied to DEA allowed us to perform a sensitivity analysis of the efficiency scores in the first part of this work, and also permitted us to set up a two-stage approach in order to evaluate the influence of environmental variables such as innovative activities and external economies on efficiency.

The results coming out of the first part of this empirical analysis somehow contribute to the debate concerning the role that human capital accumulation plays on the economic growth of Italian regions. Undoubtedly, higher levels of average years of schooling were important for efficiency and TFP growth in the Northern and Central regions. Conversely, the overall scarce human capital accumulation in Southern regions negatively affected their performances. However, both DEA and analysis of decomposition of productivity growth, conducted by means of Malmquist's index, highlighted that also in Southern regions, in which the growth rate of human capital and TFP was remarkable, the contribution of the improvement in pure efficiency to economic growth was totally nonessential. This means that the accumulation of human capital in those regions is not crucial to increase an efficient use of capital and labour in the production process. Therefore, a problem concerning the specific quality of human capital that accumulates in the South of Italy emerges.

REFERENCES

Aiello, F. Scoppa, V. (2005) Convergence and Regional Productivity Divide in Italy: Evidence from Panel Data, *mimeo*, Department of Economics and Statistics, University of Calabria.

Ascari, G. Di Cosmo V. (2005) Determinants of Total Factor Productivity in the Italian Regions. *Italian Journal of Regional Science*, 4: 27-49.

Arrow, K. J. (1962) The Economic Implications of Learning by Doing. *Review of Economic Studies* 29:155-173.

Banker, R.D. (1996) Hypothesis tests using data envelopment analysis. *Journal of Productivity Analysis* , 7:139–159.

Barro, R.J., Lee, J.W. (1993), International comparisons of educational attainments, *NBER Working Paper* n.4349.

Becattini, G., Musotti, F. (2003) Measuring the District Effect. Reflections on the Literature. *Banca Nazionale del Lavoro Quarterly Review* 56: 259-290.

Bonaglia, F., Picci, L. (2000), Lo stock di capitale nelle regioni italiane. *Mimeo*.

Bollino, C.A., Polinori, P. (2007) *Efficienza dei sistemi regionali e politica per le imprese*. Roma: Donzelli Editore.

Bracalente, B., Perugini C., Pompei F. (2008) What Sorts of Agglomerations Really Matter to Productivity? A Regional Analysis of Europe's Manufacturing Sector . The Review of Regional Studies, 38:145-172.

Bronzini, R., Piselli, P. (2006) Determinants of long-run regional productivity: The role of R&D, human capital and public infrastructure. *Temi di discussione*,597, Banca d'Italia.

Coelli, T.J, Prasada Rao, D.S., O'Donnel, C.J., Battese, G.E. (2005) *An introduction to efficiency and productivity analysis*, second edition. NY: Springer.

Destefanis, S., Sena, V. (2005) Public Capital and Total Factor Productivity: Evidence from the Italian Regions, 1970-98, *Regional Studies*. 39: 603-617.

Faini, G., Sapir, J. (2005), Un modello obsoleto? Crescita e specializzazione dell'economia Italiana, in Faini, G., Sapir J., *Oltre il declino*. Bologna : Il Mulino.

Fare, R., Grosskopf, S., Norris, M., Zhang, Z. (1994 Productivity Growth, Technical Progress, and Efficiency Change in Industrialized Countries. *American Economic Review, 84:* 66-83.

Farrell, M. J. (1957) The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society, Series A, General,* , 720: 253-82.
Glaeser, E. L., H. D. Kallal, J. A. Scheinkman, and A. Shleifer, (1992) Growth in Cities. *Journal of Political Economy* 100: 1126-1152.
ISTAT, Censimento della popolazione e delle abitazioni, anni 1971, 1981, 1991, 2001, Roma.
ISTAT, Censimento dell'industria e dei servizi, anni 1971, 1981, 1991, 2001, Roma.
Kumar, S., Russell, R.R. (2002) Technological Change, Technological Catch-up, and Capital Deepening: Relative Contributions to Growth and Convergence. *American Economic Review*, 95: 527-548.
Maffezzoli, M. (2006) Convergence Across Italian Regions and the Role of Technological Catch-Up. The BE Journal of Macroeconomics, Topics in macroeconomics, 6: 1-43
Marshall, A. (1920) *Principles of Economics*. London : Macmillan.
Paci, R., Pigliaru F. (1999) Growth and sectoral dynamics in the Italian regions.In *Economic Growth and Change. National and Regional Patterns of Convergence and Divergence*, (Adams J., Pigliaru F., Ed.), Cheltenham (UK): Edward Elgar,
Parkin, D., Hollingsworth, B. (1997) Measuring production efficiency of acute hospitals in Scotland, 1991-1994: validity issues in data envelopment analysis. Applied Economics, 29:1425-1433.
Romer, P. M. (1986) Increasing Returns and Long-Run Growth. *Journal of Political Economy* 94:1002-1037.
Shepard, R., W. (1970) Theory of cost and production function. Princeton, NJ : Princeton University Press,.
Simar, L., Wilson, P., W. (1998) Sensitivity analysis of efficiency scores: how to bootstrap in nonparametric frontier models. *Management Science*, 44: 49-61.
Simar, L., Wilson, P., W. (2002) Non-parametric tests of returns to scale. *European Journal of Operational Research*,139:115–132.
Simar, L., Wilson, P., W. (2001) Testing restrictions in nonparametric efficiency models. *Communication Statistics*, 30:159-184.
Wilson, P., W. (2006),FEAR: A Software Package for Frontier Efficiency Analysis with R. *mimeo.*

APPENDIX

Simar and Wilson (1998), showed how the bootstrap procedure, that simulates a Data Generating Process (DGP) can approximate the sampling variation of the estimated frontier.
More formally, we consider a list of *p* inputs *x* and one output *y*, a production set

$$\Psi = \{(x, y) \in \Re_+^{p+1} \mid x, produce, y\} \tag{1}$$

and an output correspondence set defined for each $x \in \Psi$

$$Y(x) = \{y \in \Re_+ \mid (x, y) \in \Psi\} \tag{2}$$

The Farrell efficiency boundaries are subsets of Y(*x*) denoted by:

$$\partial Y(x) = \{ y \mid y \in Y(x), \theta y \notin Y(x); \forall \theta > 1 \} \quad (3)$$

These may be used to define the Farrell output measure of efficiency for a given point i[7]

$$\theta_i = \max\{\theta \mid \theta y_i \in Y(x_i)\} \quad (4)$$

If $\theta_i = 1$, the unit $((x_i, y_i)$ is considered as being output efficient.
We also denote the efficient level of output corresponding to the input level x_i as

$$y^{\vartheta}(y_i \mid x_i) = \theta_i y_i \quad (5)$$

Thus, $y^{\vartheta}(y_i \mid x_i)$ is the intersection of $\partial Y(x)$ and the ray θy_i.

Tipically, Ψ, $Y(x)$ and $\partial Y(x)$ are unknown, hence for a given unit $(y_i \mid x_i)$, θ_i is also unknown. Now suppose that a DGP, P, generates a random sample $\aleph = \{(x_i, y_i) \mid i = 1,...n\}$ that in turn defines, by some method, the estimators $\hat{\Psi}, \hat{Y}(x), \partial\hat{Y}(x)$. For a given unit we can estimate its efficiency

$$\hat{\theta}_i = \max\{\theta \mid \theta y_i \in \hat{Y}(x_i)\} \quad (6)$$

The sampling properties of $\hat{\Psi}, \hat{Y}(x), \partial\hat{Y}(x)$ and $\hat{\theta}_i$ depend on P, which is unknown.
At this point we can use bootstrap, that repeatedly simulates a DGP through re-sampling, to produce a reasonable estimator $\hat{\mathrm{P}}$ of P from the data $\aleph$.
Consider now the dataset $\aleph^* = \{(x_i^*, y_i^*) \mid i = 1,...n\}$ generated by $\hat{\mathrm{P}}$. This pseudo sample defines the corresponding quantities $\hat{\Psi}^*, \hat{Y}^*(x), \partial\hat{Y}^*(x)$ and for a given unit its measure of efficiency is

$$\hat{\theta}_i^* = \max\{\theta \mid \theta y_i \in \hat{Y}^*(x_i)\} \quad (7)$$

By means of bootstrap, $\hat{\mathrm{P}}$ generates B samples $\aleph_b^*, b = 1,....B.$ In particular for a given unit (x_i^*, y_i^*), we have

$$\{\hat{\theta}_{ib}^*\}_{b=1}^{B}, b = 1,..B; \quad (8)$$

the empirical density function of $\{\hat{\theta}_{ib}^*\}_{b=1}^{B}$ is the Monte Carlo approximation of the distribution of $\hat{\theta}_i^*$, conditional on $\hat{\mathrm{P}}$.
Thus, the known bootstrap distributions will mimic the original unknown sampling distributions of the estimators of interest. For the efficiency measure θ_i, we have

$$(\hat{\theta}_i^* - \hat{\theta}_i) \mid \hat{\mathrm{P}} \sim (\hat{\theta}_i - \theta_i) \mid \mathrm{P} \quad (9)$$

This analogy allows us to estimate the bias of $\hat{\theta}_i$, the original estimator of θ_i, by its bootstrap estimate:

[7] Following Simar and Wilson (1998), we use in this formalisation the Farrell output measure of efficiency, instead of the Shepard one, that is simply the reciprocal.

$$bias_{\hat{P},i} = E_{\hat{P}}(\hat{\theta}_i^*) - \hat{\theta}_i \tag{10}$$

Expression (12) may be approximated by means of the Monte Carlo realizations $\hat{\theta}_{ib}^*$:

$$bias_i = \frac{1}{B}\sum_{b=1}^{B}(\hat{\theta}_{i,b}^*) - \hat{\theta}_i = \overline{\theta}_i^* - \hat{\theta}_i \tag{11}$$

Therefore a bias-corrected estimator of θ_i is

$$\tilde{\theta}_i = \hat{\theta}_i - bias_i = 2\hat{\theta}_i - \overline{\theta}_i^* \tag{12}$$

The empirical distribution of $\hat{\theta}_{i,b}^*, b = 1,...B$, provides after correction for bias, confidence intervals for θ_i, hence we take the corrected empirical density function centered on $\tilde{\theta}_i$, the bias corrected estimator of θ_i.

It must be remarked at this point that the empirical density function of $\hat{\theta}_{i,b}^*$, has to be shifted by $2 \cdot bias_i$ to the left, since a correction of $1 \cdot bias_i$ would center on the biased $\hat{\theta}_i$ rather than $\tilde{\theta}_i$.

Thus, we consider the empirical density function $\hat{\theta}_{i,b}^*, b = 1,...B$, where

$$\tilde{\theta}_{i,b}^* = \hat{\theta}_{i,b}^* - 2bias_i \tag{13}$$

Finally, the usual percentile confidence interval for θ_i with intended coverage $(1-2\alpha)$ is given by:

$$(\hat{\theta}_{i,low}, \hat{\theta}_{i,up}) = (\tilde{\theta}_i^{*(\alpha)}, \tilde{\theta}_i^{*(1-\alpha)}) \tag{14}$$

where $\tilde{\theta}_i^{*(\alpha)}$ is the $100 \cdot \alpha th$ percentile of the empirical density function $\hat{\theta}_{i,b}^*, b = 1,...B$.