INSTITUT DE STATISTIQUE

UNIVERSITÉ CATHOLIQUE DE LOUVAIN

DISCUSSION PAPER

0725

ANALYSING TOTAL FACTOR PRODUCTIVITY THROUGH A BOOTSTRAP-BASED NONPARAMETRIC APPROACH: THE CASE OF ITALIAN REGIONS

BONACCORSI, A. , CURI, C., DARAIO, C., MANCUSO, P. and L. SIMAR

Analysing Total Factor Productivity through a bootstrap-based nonparametric approach: the case of Italian regions

Andrea Bonaccorsi*[1], Claudia Curi**[2], Cinzia Daraio*[3], Paolo Mancuso**[4] Léopold Simar***[5]

University of Pisa, **University of Rome Tor Vergata,*Université Catholique de Louvain*

Abstract

This paper analyzes the Total Factor Productivity (TFP) growth of the Italian regions over the period 1980-2001. Malmquist Productivity Index (MPI) and its components (namely Efficiency Change –EffCh- and Technical Change -TechCh) as well as confidence intervals have been estimated by applying the procedure proposed by Simar and Wilson (1999). Including human capital among inputs, we estimated an overall bias-corrected productivity gain of 2.1 percent, an efficiency gain of 0.5 and a technical gain of 1.6 percent. The bootstrap analysis revealed that for most Italian regions efficiency and technical changes did not show a statistically significant change.

1. Introduction

The analysis of Italian economic growth at regional level has been extensively examined in the context of the well known convergence debate between Northern and Southern regions. So far, the convergence process has been widely dealt with using a parametric approach (*e.g.* Barro and Sala-i-Martin,1991; Sala-i-Martin, 1996; Di Liberto, 1994; Paci and Saba, 1998; Carmeci and Mauro, 2002; Aiello and Scoppa, 2005), without identifying its main sources. The only exception, to the best of our knowledge, is Leonida *et al.* (2004) which estimate the sources of Total Factor Productivity (TFP) growth of Italian regions over the period 1970-1995 by means of the Malmquist Productivity Index decomposition (Färe *et al.*, 1994), using a nonparametric approach based on Data Envelopment Analysis (DEA). However, recent theoretical papers highlight several shortcomings of the DEA methodology (see for a detailed analysis Daraio and Simar, 2007), among which there is the bias in the estimates and the difficulty of carrying out statistical inference in a multivariate nonparametric set up.

The main contribution of this paper is to re-estimate and update TFP growth and its determinants (Efficiency Change and Technical Change) of Italian regions using a rigorous procedure, based on the

[1] Andrea Bonaccorsi, Department of Electrical Systems and Automation, School of Engineering, University of Pisa, Italy 56100, Via Diotisalvi, 2, e-mail: a.bonaccorsi@gmail.com

[2] Claudia Curi, corresponding author, Dipartimento Ingegneria dell'Impresa, University of Rome, Tor Vergata, Italy 00133, Via del Politecnico , e-mail:curi@disp.uniroma2.it

[3] Cinzia Daraio, Department of Electrical Systems and Automation, School of Engineering, University of Pisa, Italy 56100, Via Diotisalvi, 2, e-mail: daraio@dsea.unipi.it

[4] Paolo Mancuso, Dipartimento Ingegneria dell'Impresa, University of Rome, Tor Vergata, Italy 00133, Via del Politecnico 1, e-mail:mancuso@disp.uniroma2.it

[5] Léopold Simar, Institute of Statistics, Université Catholique de Louvain, Louvain la Neuve, Belgium Voie du Roman Pays, E-mail: simar@stat.ucl.ac.be

bootstrap method proposed by Simar and Wilson (1999) for DEA estimations[6]. It allows for the estimation of bias-corrected MPI (and its components) and permits the construction of confidence intervals on the estimated measures. Moreover, we also include human capital in the analysis since it is widely recognized that it generally plays a central role in the economic growth (*e.g.* Del Barrio-Castro *et al.*, 2002; Bassanini and Scarpetta, 2002; Maudos *et al.*, 1999) and also for the Italian regions (Ascari and Di Cosmo, 2004).

2. Bootstrapping MPI

The basic idea of the approach proposed by Simar and Wilson (1999) is to estimate bias corrected estimates and confidence intervals at desidered level of significance for MPI and its components. It allows assessing whether productivity growth (or decline) is statistically significant or not. This means to determinate if the changes are really determined by the analyzed process or merely inducted by sampling variations. The same holds for the significance of both efficiency and technical changes.

The measurement of productivity is done using the decomposition of MPI proposed by Färe *et al.* (1994), with an output orientation and assuming Constant Returns to Scale (CRS) technology, according to the standard neoclassical growth model. Hence, the MPI between period *t* and *t+1* is defined as the product of Efficiency Change (EffCh) and Technical Change (TechCh), as follows:

$$M_i^t(x^{t+1}, y^{t+1}, x^t, y^t) = \frac{D_i^{t+1}(x^{t+1}, y^{t+1})}{D_i^t(x^t, y^t)} \left[\left(\frac{D_i^t(x^{t+1}, y^{t+1})}{D_i^{t+1}(x^{t+1}, y^{t+1})} \right) \left(\frac{D_i^t(x^t, y^t)}{D_i^{t+1}(x^t, y^t)} \right) \right]^{1/2} \quad i = 1, \dots, N \qquad (1)$$

where $D_i^t(x^t, y^t)$ is the Shepard (1970) output distance function for a generic region $i-th$ at time *t*, relative to the technology existing at time t ; it is defined as:

$$D_i^t(x^t, y^t) = \inf\{\theta : (x^t, y^t/\theta) \in S^t\} = \sup\{\theta : (x^t, \theta y^t) \in S^t\} \qquad i = 1, \dots, N \qquad (2)$$

where S^t is the technology corresponding at time *t* which models the transformation of inputs $x^t \in \Re_+^N$ into the outputs $y^t \in \Re_+^M$ and θ is the technical efficiency score; and $D_i^{t+1}(x^t, y^t) = \inf\{\theta : (x^t, y^t/\theta) \in S^{t+1}\} = \sup\{\theta : (x^t, \theta y^t) \in S^{t+1}\}$ is the Shepard output distance function for a region *i* at time *t* relative to the technology existing at time *t+1* (S^{t+1}).

For both MPI and its components, values higher than 1 indicate improvements from period *t* to *t+1*, values less than 1 indicate regress and values equal to 1 indicate no change.

Following the bootstrap-based nonparametric approach by Simar and Wilson (1999, 2000) the MPI bias is estimated by:

$$BIAS\left[\hat{M}_i(t, t+1)\right] = B^{-1} \sum_{b=1}^{B} \hat{M}_{i,b}^*(t, t+1) - \hat{M}_i(t, t+1) \qquad i = 1, \dots, N \qquad (3)$$

where B is the number of bootstrap replications, set to 2000, $\hat{M}_{i,b}^*(t, t+1)$ is the MPI bootstrapped estimate and $\hat{M}_i(t, t+1)$ the estimate for the generic region $i-th$. The bias correction of the estimate is obtained by

[6] For an overview on the bootstrap in nonparametric frontier models and recent developments see Daraio and Simar (2007).

subtracting the bias (3) from the original estimates. Instead, the main idea behind the construction of bootstrap confidence intervals is to compute the quantiles from the empirical bootstrap distribution of the pseudo estimates, approximating the unknown distribution of $\hat{M}_i(t,t+1)-M_i(t,t+1)$ by the known distribution of $\hat{M}_i^*(t,t+1)-\hat{M}_i(t,t+1)$, conditioned on the original Data Generating Process (DGP) where, for a generic region $i-th$, $\hat{M}_i^*(t,t+1)$ is the MPI bootstrapped estimate, $\hat{M}_i\,(t,t+1)$ is the MPI estimate and $M_i\,(t,t+1)$ is the real value of MPI which is unknown. Some algebraic transformations (see more details in Simar and Wilson, 1999), yield an estimated (1-α)–percent confidence interval of the real value of MPI for the generic region i as:

$$\hat{M}_i(t,t+1)+c_{1-\alpha/2} \leq M_i(t,t+1) \leq \hat{M}_i(t,t+1)+c_{\alpha/2} \tag{4}$$

where c_a is the $\alpha*100$-quantile of the distribution of $\hat{M}_i^*(t,t+1)-\hat{M}_i(t,t+1)$. It is straightforward to extend this procedure to the MPI's components EffCh and TechCh: just substitute $\hat{M}_i(t,t+1)$ in (3) and (4) with EffCh (and TechCh).

The main purpose of this approach is to see whether the Bias-Corrected (BC) estimates, namely BC-EffCh, BC-TechCh and BC-MPI, are significantly different from 1. If the interval expressed by (4) does not cover 1, then there is a significant change in productivity that can be either a progress (if the lower bound is greater than 1) or regress (if the upper bound is less than 1). A statistically significant BC-EffCh (or BC-TechCh) with bounds greater than 1 is interpreted as evidence of a *Catching Up Process*[7] towards the frontier (or *Improvements in Innovation*) while with bounds lower than 1 is interpreted as evidence of shifting away from the frontier (regress in innovation).

3. Data

The sample used consists of 19 Italian regions over the period 1980-2001, using the database provided by Bank of Italy (Bronzini and Piselli, 2006). The regions are geographically classified into three main areas (North, Centre and South), according to the Italian Statistical Institute (Istat)[8]. One region, Valle d'Aosta, is excluded due to lack of data. The variables used in the analysis are: 1) aggregate output measured by the Value Added (VA), expressed at constant price; 2) aggregate labour measured by total employment; 3) Private Capital Stock calculated using the method introduced by Piselli (2001), which breaks down national capital stock by region, sector and type of capital good; and 4) human capital stock calculated by Bronzini and Piselli (2006).

[7] As reported by Färe *et al.*, 1994, this definition of catching up is not the same of that one given by Abramovitz (1986) and other authors, according to which the catching up is based on a inverse correlation between low initial levels of TFP or income and TFP growth.

[8] The Northern area is made up of Piemonte, Lombardia, Trentino-Alto-Adige, Veneto, Friuli-Venezia-Giulia, Liguria, Emilia Romagna, the Central area of Toscana, Umbria, Marche, Lazio and finally the Southern area of Abruzzo, Molise, Campania, Sicilia, Sardegna, Basilicata, Calabria and Puglia.

4. Results

Table 1 shows the annual geometric means of Bias Corrected (BC)-MPI and its components (BC-EffCh and BC-TechCh), for each pair of consecutive years in the sample. In the following we report only the bias corrected estimates, since according to Efron and Tibshirani (1993) we have to correct for the bias when the bias in absolute value is higher than ¼ of the standard deviation of the estimates; and indeed in all our analysis this is the case.

Table 1

BC-MPI and its components. Annual Average Change 1980-2001

Region	BC-EffCh	BC-TechCh	BC-MPI
Piemonte (PIE)	1.005	1.022	1.026
Lombardia (LOM)	1.000	1.013	1.007
Trentino-Alto Adige (TAA)	1.002	1.017	1.019
Veneto (VEN)	1.002	1.024	1.025
Friuli-Venezia Giulia (FVG)	1.006	1.019	1.025
Liguria (LIG)	1.006	0.988	0.992
Emilia Romagna (ERO)	1.001	1.027	1.027
Toscana (TOS)	0.998	0.999	0.995
Umbria (UMB)	1.005	1.023	1.027
Marche (MAR)	1.009	1.027	1.034
Lazio (LAZ)	1.000	1.011	1.009
Abruzzo (ABR)	1.006	1.018	1.024
Molise (MOL)	1.013	1.017	1.030
Campania (CAM)	1.010	1.019	1.028
Puglia (PUG)	1.005	1.020	1.024
Basilicata (BAS)	1.015	1.017	1.032
Calabria (CAL)	1.015	1.017	1.032
Sicilia (SIC)	1.006	1.017	1.022
Sardegna (SAR)	1.000	1.017	1.016
Overall (geometric mean)	*1.005*	*1.016*	*1.021*

The results show an overall annual productivity gain (measured by the MPI) of 2.1 percent over the analyzed period since most regions (except Liguria and Toscana) exhibit progress in both components of MPI[9]. Deeply, the results suggest that improvement in productivity is achieved mainly by the innovation driving force (BC-TechCh) instead of catching up (BC-EffCh). Indeed, the technological progress (upward shift of the frontier) shows an overall annual growth rate of 1.6 percent while the efficiency gain (movement towards the frontier) only of 0.05 percent.

Figure 1 shows the trends of annual geometric means of the estimated MPI and its components for Northern, Central and Southern areas. As far as the efficiency change is concerned, the Southern regions are the best performers of the sample with 0.86 percent average annual change while Northern and Central regions exhibit average annual change, respectively, of 0.33 and 0.34 percent. As far as the technical change is concerned, again Southern regions achieve better results with an annual average rate of 1.80 percent (only in 1994-1995 and 1998-1999, they experienced a regress), followed by Northern regions with 1.60 percent

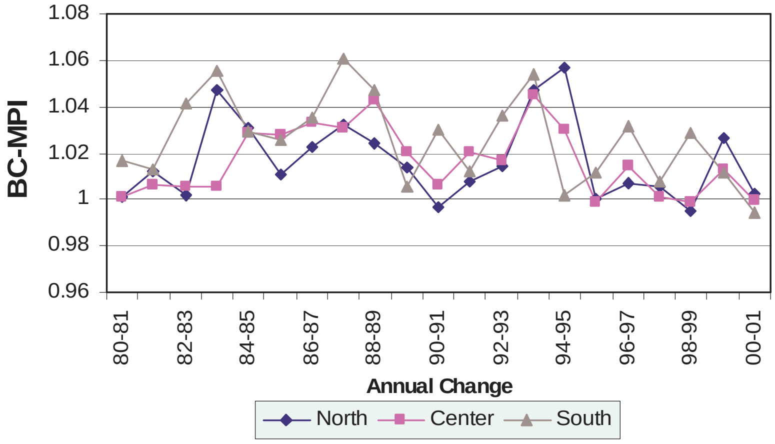
[9] The confidence intervals on all annual changes of MPI and its components are available from the authors upon request.

and Central regions with 1.50 percent. Finally, as far as the Malmquist Productivity Index is concerned, all Italian regions experience improvements in productivity and their ranking is unchanged compared with the former ones. Over the period, Southern regions exhibit productivity gains of 2.61 percent, whereas Northern and Central regions of 1.69 and 1.64 percent, respectively. It is worth to note the complementarity between the evolutions of BC-EffCh and BC-TechCh. Indeed, higher gains in BC-EffCh (*i.e.* between 1987-1988 and 1994-1995) are achieved in correspondence to highest lost in innovation (for the same pairs of consecutive years) and vice versa. This may suggest that Italian regions have experienced new technologies, without improving their efficiency by using them.

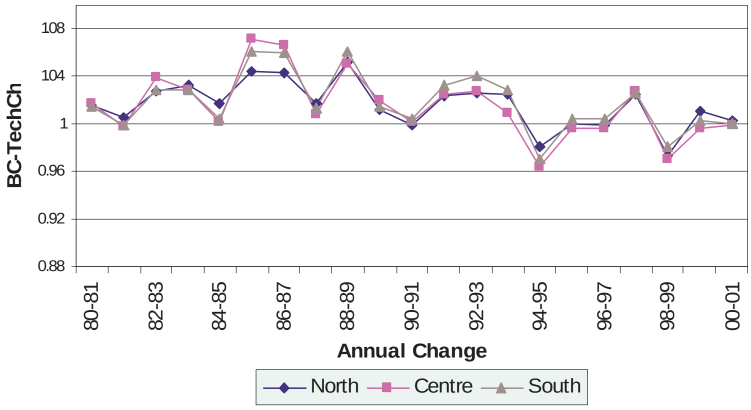
Figure.1.

Annual growth evolution over the period 1980-2001. Geometric means by areas reported.

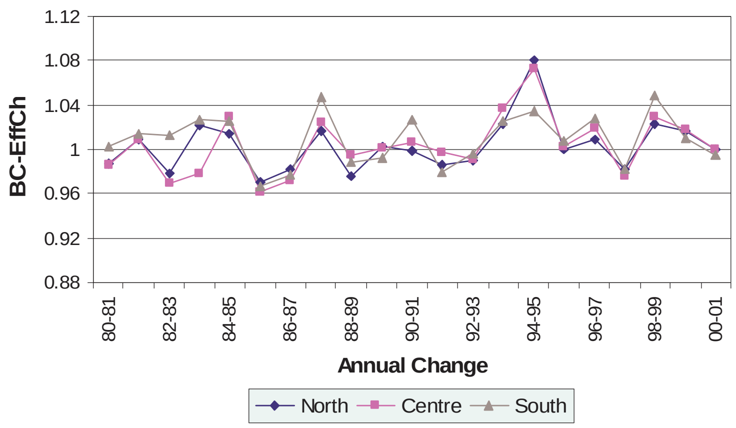
A) Bias-corrected MPI by geographic areas.



B) Bias-corrected TechCh by geographic areas.



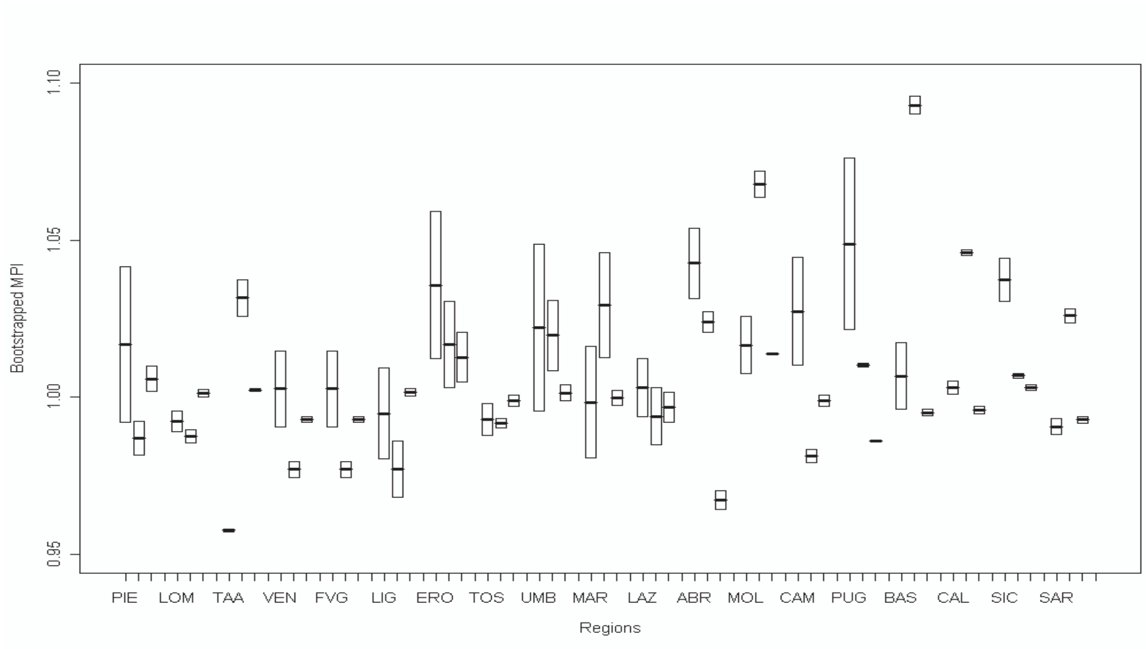
C) Bias-corrected EffCh by geographic areas.



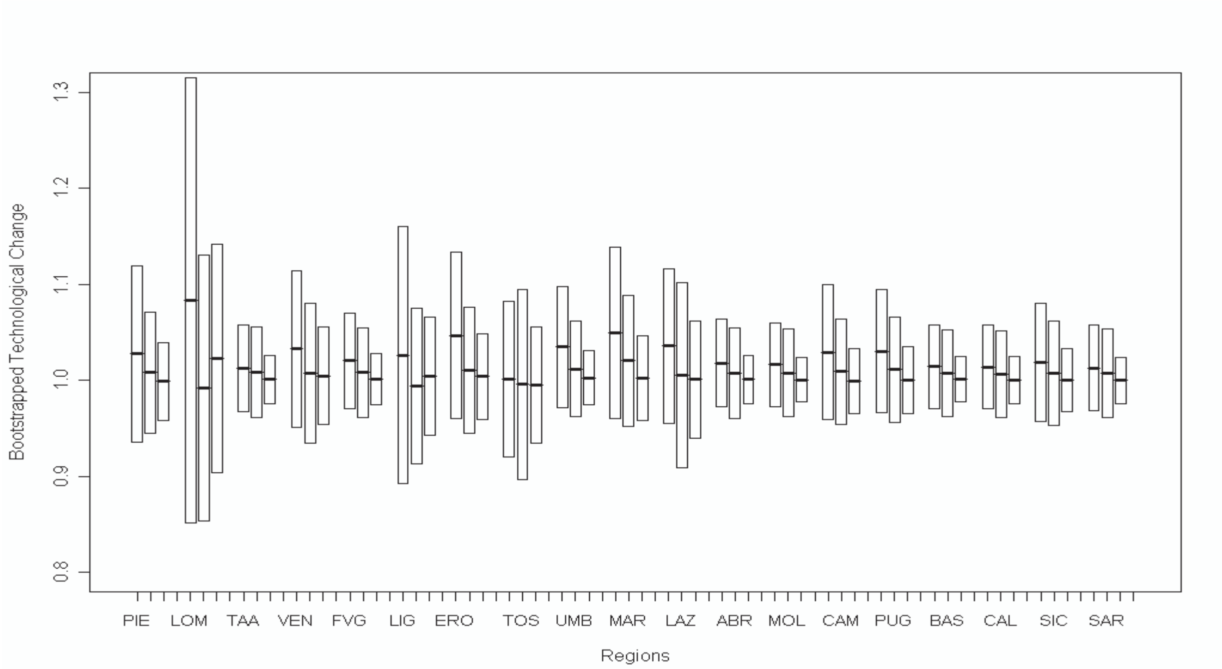
More accurate statistical insights are obtained for each region from the distribution of the bootstrapped estimates of MPI and its components. The distribution of bootstrap estimates, in fact, is used to construct confidence intervals on MPI and its components according to equation (4) described above. Hence, it allows the assessment of whether efficiency, technical and productivity change are statistically different from unity. For the sake of brevity, Figure 2 discloses information on the bootstrapped distribution of regional estimates for three annual changes: 1980-1981, 1990-1991 and 2000-2001.

Figure.2.
Boxplots of bootstrapped MPI and its components. For each Italian region, the following three pairs of years are reported: 1980-1981 (first box), 1990-1991 (second box) and 2000-2001 (third box).

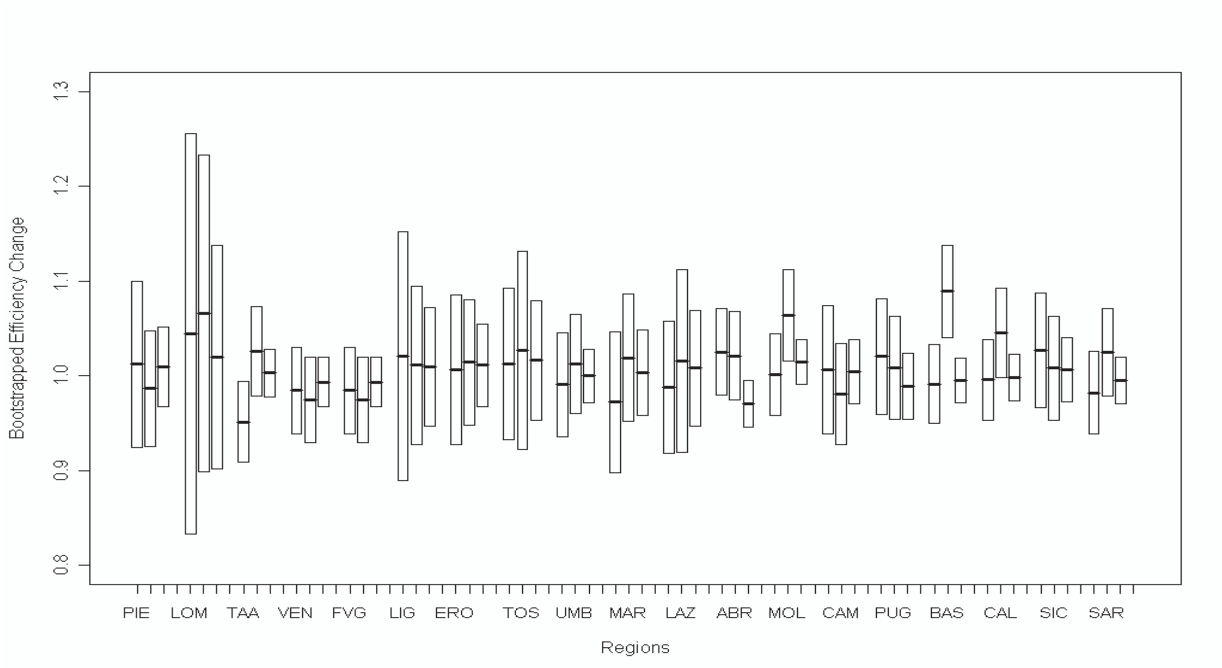
A) Boxplots of bootstrapped MPI



B) Boxplots of bootstrapped TechCh



C) Boxplots of bootstrapped EffCh



Each box represents the interquartile range (IQR), in which the 0.75 and 0.25 quartiles define the upper and lower segments, respectively.

Boxplots of EffCh and TechCh show that IQRs vary around 1, reaviling steady levels and isolated cases of statistical significant changes. Looking at the efficiency changes, Trentino Alto Adige (TAA) and Abruzzo (ABR) show significant regress in 1980-1981 and 2000-2001, respectively, and only Basilicata (BAS) shows a significant improvement in 1990-1991. Most regions reveal a density fairly concentrate, instead Lombardia (LOM) and Liguria (LIG) show high dispersion. Looking at the technical changes, a more concentrate density points out the persistence of the lack of innovation over the period. A different picture emerges from the boxplots of MPI. In 1980-81, 47% of regions show no change in productivity against 32% which experience productivity gain and 21% of regions which experience productivity lost. In 1990-91, most regions (58%) experience a gain in productivity while in 2000-2001 only the 21% of Italian regions shows a productivity gain. These descriptive comments are confirmed by the confidence intervals constructed following equation (4). Moreover, confidence intervals of MPI are narrower than those of EffCh and TechCh, indicating more accuracy in the estimation. For all the considered pairs of years, a statistically significant progress is achieved by Emilia Romagna (ERO) and Molise (MOL) whereas a statistically significant regress is experienced by Toscana (TOS).

5. Concluding remarks

We applied a rigorous statistical procedure based on the bootstrap in a DEA framework to accurately re-estimate and update the measurement of the overall productivity growth of Italian regions, including human capital among inputs. We found a bias corrected improvement of overall productivity of 2.1 percent over the period 1980-2001, mainly driven by technological improvements in production possibilities (overall annual change: 1.6%), whereas efficiency improvements remained fairly constant (overall annual change: 0.5%). The bootstrap procedure (Simar and Wilson, 1999), allows us a more careful analysis at regional level. It has confirmed that only a few Italian regions have a Malmquist Productivity Index significantly different from 1 at a significant level of 95%. The percentages of regions which experienced progress are 32%, 58% and 21%, respectively, for the three annual changes (1980-1981, 1990-1991 and 2000-2001). As far as the technical change is concerned, most Italian regions exhibited a significant persistence of lack of innovation. Concerning the efficiency change, isolated cases of significant regional improvement are identified (*i.e.* Basilicata) as well as of regional regress (*i.e.* Trentino Alto Adige and Abruzzo). According to these results, the inferential approach applied in the paper has provided a more rigorous and accurate insights on the Italian regional TFP than the traditional Data Envelopment Analysis (DEA) estimation carried out by Leonida *et al.*(2004 ,Table 1, p. 2190) in which all the estimated values are interpreted as progress or regress without taking into account the bias of the estimated values and their statistical significance.

Acknowledgements

We would like to thank the participants at the EURO XXII Conference, Prague, 6-9 July 2007 for helpful comments and suggestions. The usual disclaimers apply. Authors 1 and 3 acknowledge the financial support of the Aquameth project, PRIME NoE, funded by the European Commission under the 6th Framework Programme and of the Italian IRIS project (FIRB, The reorganization of the public system of research for the technological transfer: governance, tools and interventions). Authors 2 and 4 acknowledge Drs. P. Piselli and

Dott. R. Bronzini from Bank of Italy for their support and for providing data. Author 5 acknowledges the financial support of the "Interuniversity Attraction Pole", IAP-Phase VI (No. P06/03).

References

Abramovitz, M., 1986, Catching Up, Forgin Ahead, and Falling Behind, Journal of Economic History, 46(2), 385-406.

Aiello F., Scoppa V., 2005, Convergence and Regional Productivity Divide in Italy: Evidence from Panel Data , Impresa, mercato, lealtà territoriale, Atti della XXVII Conferenza italiana di Scienze Regionali

Ascari G., Di Cosmo V., 2004 Determination of Total Factor Productivity in the Italian Regions, Atti della XXV Conferenza Italiana di Scienze Regionali, Novara.

Barro R., Sala-i-Martin X.,1991, Convergence Across States and Regions, Brooking Papers on Economic Activity, 1, 107-158.

Bassanini A., Scarpetta S., 2002, Does human capital matter for the growth in OECD countries? A pooled mean-group approach, Economics Letters, 74, 399-405.

Bronzini R., Piselli P., 2006, Determinants of long-run regional productivity: the role of R&D, human capital and public infrastructure, Temi di discussione del Servizio Studi, Bank of Italy.

Carmeci G., Mauro L. 2002, The Convergence of the Italian Regions and Unemployment. Theory and Evidence, Journal of Regional Science, 42, 3, 509-32.

Daraio C., Simar L., 2007, Advanced robust and nonparametric methods in efficiency analysis. Methodology and Applications, Springer, New-York

Del Barrio-Castro T., Lopez-Bazo E., Serrano-Domingo G., 2002, New evidence on international R&D spillovers, human capital and productivity in the OECD, Economics Letters 77, 41-45.

Di Liberto A., 1994 Convergence Across Italian Regions, Nota di Lavoro 168, Fondazione ENI Enrico Mattei, Rome.

Efron B., Tibshirani R.J. (1993), An introduction to the bootstrap. Chapman and Hall, London

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

Leonida L., Petraglia C., Murillo-Zamorano L.R.,2004, Total Factor Productivity and the convergence hypothesis in the Italian regions, Applied Economics, 36, 2187-2193.

Maudos J., Pastora J.M., Serrano L., 1999, Total factor productivity measurement and human capital in OECD countries, Economics Letters, 63, 39-44

Paci R., Saba F.,1998, The Empirics of Regional Economic Growth in Italy 1951-1993, Rivista Internazionale di Scienze Economiche e Sociali, 45, 515-542.

Piselli P., 2001, Lo stock di capitale nelle regioni italiane: una ricostruzione 1980-1996, Bank of Italy.

Sala-i-Martin X. ,1996, The Classical Approach to Convergence Analysis, The Economic Journal, vol. 106, pp. 1019-1036.

Shepard R.W., 1970, Theory of cost and production functions. Princeton, NJ: Princeton University Press.

Simar, L., Wilson, P.W., 1999, Estimating and bootstrapping Malmquist indices, European Journal of Operational Research, 115, 459-471.

Simar, L., Wilson, P.W., 2000, A general methodology for bootstrapping in non-parametric frontier models, Journal of Applied Statistics, 6, 779-802.