

Order-M Frontier and Performance of Territorial Areas: an Example for Italian Regions

Frontiera di Ordine-M e Performance di Aree Territoriali: un Esempio per le Regioni Italiane

Erasmus Vassallo

Dipartimento di Contabilità Nazionale e Analisi dei Processi Sociali, Università degli Studi di Palermo, Viale delle Scienze, Edificio 2, Parco d'Orlèans, 90128 Palermo
erasmo.vassallo@unipa.it

Riassunto: La misura dell'efficienza tecnica tramite tecniche non parametriche tradizionali è influenzata da valori anomali. La presenza di tali casi può distorcere fortemente i risultati e pertanto rendere le conclusioni di scarso interesse interpretativo. La frontiera robusta di ordine-m che tra l'altro evita l'appiattimento su valori unitari, ed un'opportuna scomposizione del corrispondente indice ordine-m di Malmquist, rappresentano una valida soluzione. L'approccio è qui illustrato con un'applicazione su alcuni dati macroeconomici delle 20 regioni italiane tra gli anni 1970 e 2002.

Keywords: efficiency, productivity, order-m frontier, Malmquist index.

1. Introduction

Statistical estimation of technical efficiency uses a model to link up inputs and outputs. The parametric approach utilizes a function, usually a translog form with a two-sided noise term and an one-sided inefficiency term. The nonparametric approach, instead, does not use parametric functions and requires nonparametric methods like Data Envelopment Analysis (DEA) or Free Disposal Hull (FDH). In this case, an estimate of efficiency (inefficiency) for an unit (or DMU) is based on comparison with best practice. Nonparametric estimators do not need a specification of a functional form between inputs and outputs of production process but, as they are built on the envelopment of observed data, are strongly influenced by outliers; accordingly, very often efficiency scores are flattened to maximum (Førsund and Sarafoglou, 2005). This last characteristic, in some real applications, can twist results to the point that the final economic interpretation appears devoid of interest (De Stefanis and Storti, 2004). To solve this problem, Cazals *et al.* (2002) suggest the order-m approach, a robust method to calculate the efficiency frontier and a corresponding efficiency estimator that resists to extreme values and noise and allows a better graduation of units. Moreover, the order-m frontier generally demands less observations than classic nonparametric frontiers to gain meaningful results. The order-m distance function consents also to examine dynamic effects when it is utilized in a Malmquist index framework to analyse productivity changes (Wheelock and Wilson, 1999). In this paper, the order-m approach is applied to some italian regional macroeconomic data; the comparison with traditional frontiers highlights its peculiarity. At last, efficiency scores and Malmquist index are computed for the 20 regions in some years from 1970 to 2002.

2. Order-M Frontier and Malmquist Index Decomposition

Cazals *et al.* (2002) introduce the concept of “expected minimum input frontier of order m ” (input orientation) and “expected maximum output frontier of order m ” (output orientation); the convexity assumption can be relaxed and also noise with zero expected value can be present. With output orientation, for a given level of input $\mathbf{x}_0$, the method

draws m i.i.d. random variables $\{\mathbf{u}_j\}_{j=1}^m, \mathbf{u}_j \in \mathbb{R}_+^q$ from the conditional distribution F ,

$$F_{y|x}^t(\mathbf{y}_0 | \mathbf{x}_0) = \Pr(\mathbf{y} \leq \mathbf{y}_0 | \mathbf{x} \leq \mathbf{x}_0). \text{ Then, } C_m^t(\mathbf{x}_0) = \left\{ (\mathbf{x}, \mathbf{y}) \in \mathbb{R}_+^{p+q} \mid \mathbf{x} \leq \mathbf{x}_0, \bigcup_{j=1}^m \mathbf{y} \leq \mathbf{u}_j \right\}$$

and, like a Shephard output distance function, the method obtains:

$$D(\mathbf{x}, \mathbf{y} | C_m^t(\mathbf{x})) \equiv \inf \{ \theta > 0 \mid (\mathbf{x}, \mathbf{y} / \theta) \in C_m^t(\mathbf{x}) \} \quad (1)$$

For any $\mathbf{y} \in \mathbb{R}_+^q$, it is possible to define the “expected maximum output level of order m ” ($\hat{\partial}$ denotes an upper boundary) $\mathbf{y}_m^{\hat{\partial}}(\mathbf{x}) \equiv \mathbf{y} / E[D(\mathbf{x}, \mathbf{y} | C_m^t(\mathbf{x}_0))]$ for all $\mathbf{x}$ such that is positive the ratio between $f^t(\mathbf{x}, \mathbf{y})$ and $f^t(\mathbf{y} | \mathbf{x})$, where $f^t(\cdot)$ represents an appropriate probability density function over the production set. At the end, the order- m of production set S^t can be written as $S_m^t \equiv \{ (\mathbf{x}, \mathbf{y}) \mid (\mathbf{x}, \mathbf{y}) \in S^t, \mathbf{y} \leq \mathbf{y}_m^{\hat{\partial}}(\mathbf{x}) \}$ to say the “expected production set of order m ”; therefore, $S_m^{\hat{\partial}t}$ indicates the “order- m frontier”, frontier now more robust to outliers than traditional frontiers since it does not envelop necessarily all data. A Monte Carlo procedure (with k repetitions) is applied to obtain a nonparametric estimate of $E[D(\mathbf{x}, \mathbf{y} | C_m^t(\mathbf{x}))]$. The order- m estimator is used to estimate the order- m frontier $S_m^{\hat{\partial}t}$ but not $S^{\hat{\partial}t}$ so that the root- n consistency is preserved; it is also interesting to underline that if $m \rightarrow \infty$, the order- m estimator converges to the FDH estimator (Simar, 2003). As the traditional way, this robust approach permits to define a Malmquist index useful to measure productivity change pertinent to order- m frontiers. Following Simar and Wilson (1998), this peculiar order- m Malmquist index can be broken up into four factors where t_1, t_2 denote two years and $H(\cdot)$ the convex conical hull of the production set (obtained by dropping in S_m the constraint $\sum(\text{inputs and outputs weights})=1$ as the classic DEA procedure for CRS case):

$$\begin{aligned} M(\mathbf{x}^{t_1}, \mathbf{y}^{t_1}, \mathbf{x}^{t_2}, \mathbf{y}^{t_2} | S_m^{t_1}, S_m^{t_2}) &= \left(\frac{D(\mathbf{x}^{t_2}, \mathbf{y}^{t_2} | H(S_m^{t_1}))}{D(\mathbf{x}^{t_1}, \mathbf{y}^{t_1} | H(S_m^{t_1}))} \frac{D(\mathbf{x}^{t_2}, \mathbf{y}^{t_2} | H(S_m^{t_2}))}{D(\mathbf{x}^{t_1}, \mathbf{y}^{t_1} | H(S_m^{t_2}))} \right)^{1/2} = \\ &= \left(\frac{D(\mathbf{x}^{t_2}, \mathbf{y}^{t_2} | S_m^{t_2})}{D(\mathbf{x}^{t_1}, \mathbf{y}^{t_1} | S_m^{t_1})} \right) \left(\frac{D(\mathbf{x}^{t_2}, \mathbf{y}^{t_2} | H(S_m^{t_2})) / D(\mathbf{x}^{t_2}, \mathbf{y}^{t_2} | S_m^{t_2})}{D(\mathbf{x}^{t_1}, \mathbf{y}^{t_1} | H(S_m^{t_1})) / D(\mathbf{x}^{t_1}, \mathbf{y}^{t_1} | S_m^{t_1})} \right) \\ &\quad \left(\frac{D(\mathbf{x}^{t_1}, \mathbf{y}^{t_1} | S_m^{t_1})}{D(\mathbf{x}^{t_1}, \mathbf{y}^{t_1} | S_m^{t_2})} \frac{D(\mathbf{x}^{t_2}, \mathbf{y}^{t_2} | S_m^{t_1})}{D(\mathbf{x}^{t_2}, \mathbf{y}^{t_2} | S_m^{t_2})} \right)^{1/2} \\ &\quad \left(\frac{D(\mathbf{x}^{t_1}, \mathbf{y}^{t_1} | H(S_m^{t_1})) / D(\mathbf{x}^{t_1}, \mathbf{y}^{t_1} | S_m^{t_1})}{D(\mathbf{x}^{t_1}, \mathbf{y}^{t_1} | H(S_m^{t_2})) / D(\mathbf{x}^{t_1}, \mathbf{y}^{t_1} | S_m^{t_2})} \frac{D(\mathbf{x}^{t_2}, \mathbf{y}^{t_2} | H(S_m^{t_1})) / D(\mathbf{x}^{t_2}, \mathbf{y}^{t_2} | S_m^{t_1})}{D(\mathbf{x}^{t_2}, \mathbf{y}^{t_2} | H(S_m^{t_2})) / D(\mathbf{x}^{t_2}, \mathbf{y}^{t_2} | S_m^{t_2})} \right)^{1/2} \end{aligned} \quad (2)$$

3. Efficiency and Productivity Change: Application for Italian Regions

Traditional and order-m approaches are used here to measure the technical efficiency and the productivity change in Italian regions in the wide period 1970-2002. Matlab code in Simar (2003) and aggregate data in Felli *et al.* (2004) are used. A simple production function is assumed, where output is added value and inputs are the number of total workers in full-time standard measure and the stock of physical private capital. It is just an exemplification helpful in the context of this paper; furthermore, the number of units (20 regions) are not inadequate to extract conclusions full of interest (1 output plus 2 inputs represent a small number of dimension and, in any case, the root-n consistency of order-m frontiers reduces quite a lot the problem of dimensionality). Efficiency scores with robust and traditional methods (and bias and bootstrap estimates of confidence intervals, Simar and Wilson, 2000) are computed in 1970, 1978, 1986, 1994 and 2002; for the order-m frontier, the draws are $m=5$, 10 and 15 repeated $k=200$ times. Then, the Malmquist index is computed in 1970-1978, 1978-1986, 1986-1994, 1994-2002 and 1970-2002. The Table 1 presents some main results for order-m frontiers with $m=10$ (higher values can affect the quality of the estimates).

Table 1: *Efficiency Scores and Malmquist Index Decomposition in Italian Regions*

Order-M(10)	Years					Order-M(10)	Malmquist index and its decomposition, 1970-2002				
Regions	1970	1978	1986	1994	2002	Regions	Total	T.1	T.2	T.3	T.4
Abruzzo	0,9066	0,8883	0,9426	1,0324	1,0248	Abruzzo	0,9634	1,1209	1,0289	0,7080	1,1798
Basilicata	1,0074	1,0065	1,0087	1,0187	1,0124	Basilicata	1,5124	1,0050	0,9909	1,5187	1,0000
Calabria	0,9728	0,9496	0,9359	0,8928	0,9509	Calabria	1,0276	0,9774	0,9915	0,9543	1,1111
Campania	1,0640	1,0445	1,0750	1,1054	1,0839	Campania	0,9811	1,0329	0,9892	0,9327	1,0294
Emilia	1,1881	1,4433	1,2208	1,2553	1,2342	Emilia	0,8751	1,0063	1,0031	0,8496	1,0204
Friuli	1,0346	1,0171	1,0103	1,0178	1,0297	Friuli	1,1095	0,9827	1,0270	0,9792	1,1226
Lazio	1,3150	1,3164	1,1912	1,1469	1,1651	Lazio	0,9169	0,8974	0,9988	1,0130	1,0098
Liguria	1,2661	1,1750	1,1886	1,1062	1,0914	Liguria	0,8406	0,8459	0,9948	0,9320	1,0717
Lombardia	1,7795	1,7637	1,6962	1,7922	1,7788	Lombardia	1,0926	1,0638	1,0812	1,1507	0,8256
Marche	1,0421	1,0377	1,0303	1,0200	1,0320	Marche	1,1070	0,9885	0,9984	0,9754	1,1499
Molise	1,0000	1,0000	1,0000	1,0000	1,0000	Molise	1,8614	1,0000	1,4380	1,5649	0,8272
Piemonte	1,1388	1,0666	1,1363	1,2535	1,2103	Piemonte	0,9255	1,0765	1,0125	0,8404	1,0103
Puglia	1,1740	1,2026	1,3333	1,2882	1,2806	Puglia	1,2289	1,0759	0,9971	1,0852	1,0556
Sardegna	1,0613	1,0400	1,0212	0,9734	0,9194	Sardegna	0,8961	0,8634	0,9952	0,9439	1,1048
Sicilia	1,2954	1,3696	1,3841	1,3134	1,2422	Sicilia	0,9600	0,9632	0,9977	0,9614	1,0392
Toscana	1,3772	1,3389	1,3468	1,3436	1,4075	Toscana	0,9224	1,0461	0,9997	0,8445	1,0445
Trentino	1,0679	1,0561	1,0476	1,0509	1,0578	Trentino	0,9461	0,9739	1,0415	0,7749	1,2038
Umbria	1,0660	1,0389	1,0665	1,0588	1,0520	Umbria	0,9995	1,0231	0,9000	0,8327	1,3037
Valle d'Aosta	1,0000	1,0000	1,0000	1,0000	1,0000	Valle d'Aosta	1,3497	1,0000	0,9499	1,4076	1,0095
Veneto	0,9255	1,1450	1,0977	1,1035	1,1159	Veneto	1,3145	1,1785	1,0070	1,1052	1,0022

Italian regions in Centro-Nord show frequently higher efficiency scores than regions in Mezzogiorno. Problematic outliers don't emerge with different m , even if the value of 1 is exceeded many times. The regional differences of results are similar to DEA, but the classification by order-m frontier is more accurate; instead, with FDH, too many scores equal one. The conclusion is analogous if the frontier is estimated for all years together (the number of units is now 660, 20 regions by 33 years from 1970 to 2002 trying $m=5$, 20, 50, 150 and 300, k always 200). As regards the example in Table 1, the lowest score characterizes regions in Mezzogiorno (Abruzzo for 1970 and 1978, Calabria for 1986 and 1994, Sardegna for 2002); on the contrary, the highest efficiency is always in Lombardia. In the whole period 1970-2002, the Malmquist index underlines an increase in productivity (Total>1) only in 9 regions between Calabria (1,0276) and Molise (1,8614); this is contrary to expectations of a widespread improvement. In particular, the first term of the decomposition (named T.1 in the table) indicates changes in the order-m technical efficiency, the second term (named T.2) represents changes in the order-m

scale efficiency, the third term (named T.3) indicates changes in the order-m frontier over time and, finally, the last term (named T.4) represents changes in the order-m scale efficiency due to changes in the order-m frontier (Wheelock and Wilson, 1999). To illustrate the benefit of this decomposition, it is useful to call attention to two examples: Sardegna e Lombardia (smallest and largest efficiency score in 2002). Lombardia, from 1970 to 2002, shows an improvement in productivity ($1,0926 > 1$) due to an improvement in technical efficiency ($1,0638 > 1$), improvement in scale efficiency ($1,0812 > 1$), scrolling up of the frontier ($1,1507 > 1$) and increment of order-m scale efficiency from shifts or changes in shape of the frontier ($0,8256 < 1$); just the opposite happens in Sardegna (with 0,8961, 0,8634, 0,9952, 0,9439 and 1,1048 respectively). But this is true in the whole period 1970-2002 since the regional trends often are different in an other segment.

4. Conclusion

In this paper the order-m approach is applied to macroeconomic data for Italian regions from 1970 to 2002. The order-m frontier allows to reduce the influence of outliers and enables to graduate efficiency scores better than traditional methods; moreover, an order-m Malmquist index can be computed. The result obtained here is more robust than efficiency scores with classic DEA/FDH methods and allows more precise economic interpretations. In particular, order-m efficiency scores and productivity changes are more heterogeneous than expected among regions and years; this conclusion has important effects about the local development policy.

References

- Cazals C., Florens J.P., Simar L. (2002) Nonparametric frontier estimation: a robust approach, *Journal of Econometrics*, 106-1, 1-25.
- De Stefanis S., Storti G. (2004) Robust estimation of production frontiers, *Atti della XLII Riunione Scientifica SIS*, Bari, 9-11 Giugno.
- Felli E., Gerli M., Piacentino D. (2004) Supply-side and more conventional demand-side shocks in a structural macroeconomic model for Italian regional growth, *17th European Advanced Studies Institute in Regional Science*, Split, Croatia, June 28th – July 4th.
- Førsund F., Sarafoglou N. (2005) The tale of two research communities: the diffusion of research on productive efficiency, *International Journal of Production Economics*, 98, 17-40.
- Simar L. (2003) Detecting outliers in frontier models: a simple approach, *Journal of Productivity Analysis*, 20, 391-424.
- Simar L., Wilson P. (1998) Productivity growth in industrialized countries, *Discussion Paper, Institut de Statistique, Université Catholique de Louvain*, 9810.
- Simar L., Wilson P. (2000) A general methodology for bootstrapping in non parametric frontier models, *Journal of Applied Statistics*, 27, 779-802.
- Wheelock D., Wilson P. (1999) Technical progress, inefficiency and productivity change in U.S. banking, 1984-1993, *Journal of Money, Credit and Banking*, 31, 212-234.