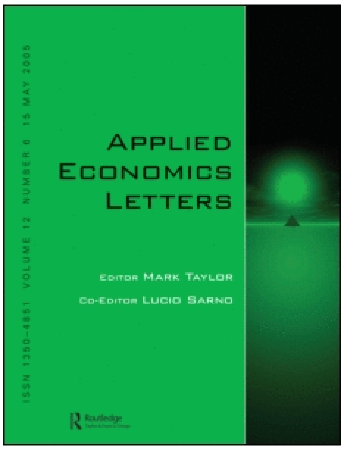

Applied Economics Letters
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713684190

On the approximation of production functions: a comparison of artificial neural networks frontiers and efficiency techniques

Daniel Santin [a]
[a] Department of Applied Economics VI, Complutense University of Madrid, Madrid, Spain

First Published:June2008

To cite this Article Santin, Daniel(2008)'On the approximation of production functions: a comparison of artificial neural networks frontiers and efficiency techniques',Applied Economics Letters,15:8,597 — 600

To link to this Article: DOI: 10.1080/13504850600721973

URL: http://dx.doi.org/10.1080/13504850600721973

Applied Economics Letters, 2008, **15**, 597–600

Routledge
Taylor & Francis Group

On the approximation of production functions: a comparison of artificial neural networks frontiers and efficiency techniques

Daniel Santin

Department of Applied Economics VI, Complutense University of Madrid, Campus de Somosaguas 28223, Pozuelo de Alarcon, Madrid, Spain
E-mail: dsantin@ccee.ucm.es

The aim of this article is to show how Artificial Neural Networks (ANN) is a valid semi-parametric alternative for fitting empirical production functions and measuring technical efficiency. To do this a Monte-Carlo experiment is carried out on a simulated smooth production technology for assessing efficiency results of ANN compared with Data Envelopment Analysis (DEA) and Stochastic Frontier Analysis (SFA). As ANNs provides average production function estimations this article proposes a so-called *thick frontier* strategy for transform average estimations into a productive frontier. Main advantages of ANN are in contexts where the production function is smooth, completely unknown, contains nonlinear relationships among variables and the quantity of noise and efficiency in data is moderate. Under this scenario, the results display that an ANNs algorithm can detect, better than traditional tools, the underlying shape of the production function from observed data.

I. Introduction

Traditional approaches for estimating empirical production functions with the final aim of measuring technical efficiency can be divided in two types. First, stochastic frontier analysis (SFA) (see Kumbhakar and Lovell (2000) for a review) imposes a parametric model where the aim is to adjust the parameters through the empirical data. Secondly, nonparametric approaches like data envelopment analysis (DEA) (see Cooper *et al.* (2004) for a review) are more flexible and do not assume any functional form. This approach draws up a linear piecewise convex production frontier through the efficient decision making units (DMU).

Both approaches, parametric and nonparametric, present different limitations derived from its econometric or mathematical nature. In order to overcome these problems recent research (Athanassopoulos and Curram, 1996; Costa and Markellos, 1997; Papadas and Hutchinson, 2002; Pendharkar and Rodger, 2003; Santin *et al.*, 2004; Pendharkar, 2005) has proposed artificial neural networks (ANN) algorithms (see Bishop (1995) for a review) as a third semi-parametric way.

The objective of this article is to provide additional evidence about the potential benefits of ANN as tool for estimating production functions and efficiency. To fulfil with this purpose I compare through a Monte-Carlo experiment the results obtained by

Applied Economics Letters ISSN 1350–4851 print/ISSN 1466–4291 online © 2008 Taylor & Francis
http://www.tandf.co.uk/journals
DOI: 10.1080/13504850600721973

DEA, SFA and ANN in a smooth nonlinear production function.

II. The Monte-Carlo Experiment

Experimental design

In order to examine the performance of efficiency techniques, let $G(x)$ be the following smooth production function (Equation 1):

$$G(x) = 2\left[\sin\left(\frac{x}{2} - \frac{\pi}{2}\right) + 1\right] + 1 \tag{1}$$

where $G(x)$ is the output, and x is a controllable input. This production function fulfils all well-known smooth properties traditionally pointed out in a microeconomics textbook (Mas Colell *et al*., 1995). Note that a sin production function is only interesting as illustration of the infinite smooth production function alternatives. In this sense this regular production function is more similar to real economic processes that the production function depicted in Santin *et al*. (2004).

Input data (X) was pseudo-randomly generated according with a uniform distribution $X \sim U(0, 2\pi]$ for 50 DMUs. This input vector yields output quantities $G(x)$ on the frontier. This output is modified by fluctuations due to two components: inefficiency and random noise. An inefficiency value is calculated for each DMU with a normal distribution of $u \sim |N(0.7; 0.1)|$ and random noise normally distributed $v \sim N(0; 0.01)$ is also generated. A number of DMUs are allowed to remain efficient belonging to the true frontier. To do this the Bernoulli (0.2) distribution is used to decide what DMUs are efficient. After this process the synthetically generated observed output is obtained. A graphical representation of the cloud of points obtained and the production frontier is shown in Fig. 1.

Finally the objective of the experiment is to fit this theoretical production function and measuring efficiency with DEA under variable returns to scale (DEAvrs) using Banker *et al*. (1984) model, SFA (Battese and Coelli, 1988) and a MLP trained with a backpropagation algorithm. SFA method was performed through the functional form $\ln(y) = \beta_0 + \beta_1 \ln(x) + \beta_{11}(\ln x)^2 + v - u$ where $u \sim |N(0, \sigma_u^2)|$ and $v \sim N(0, \sigma_v^2)$.

The computation of thick frontiers with artificial neural networks

Since MLP average production function estimation is not a frontier, it is necessary to perform a second stage to bound efficiency scores between one and zero in order to assess efficiency measures. To do this two main strategies are followed based on Athanassopoulos and Curram (1996, pp. 1003–1004) work. The first one consists of adding up the maximum residual term to the average output predicted by the MLP for each DMU. This methodology will be named as MLPMAX and efficiency calculated according to the following equation.

$$TE^{\mathrm{MLPMAX}} = \frac{y_i}{\hat{y}_i + \max(R_j)} \tag{2}$$

where y_i denotes the observed output for DMU i, $\hat{y}_i$ is the predicted output by the MLP for DMU i and R_j is the maximum residual value observed in DMU j. In order to alleviate extreme maximum residual terms the second methodology propose applying (2) to different segments of the distribution of the dependent variables which leads to the concept of the so-called *thick frontiers*. This methodology will be developed in this article. To do this the frontier will be drawn up from the least average computed output value to the large one in the following way.

(1) Order the N DMUs from the least average fitted output to the large one. $\mathrm{DMU} \in [1; N] \,|\, \mathrm{DMU}_1, \ldots, \mathrm{DMU}_N$ where DMU_1 has the least estimated output and DMU_N has the highest one.
(2) Following this ranking detect the first positive error ε_1 belonging to DMU i.

$$\varepsilon_1 = y_i - \hat{y}_i > 0$$

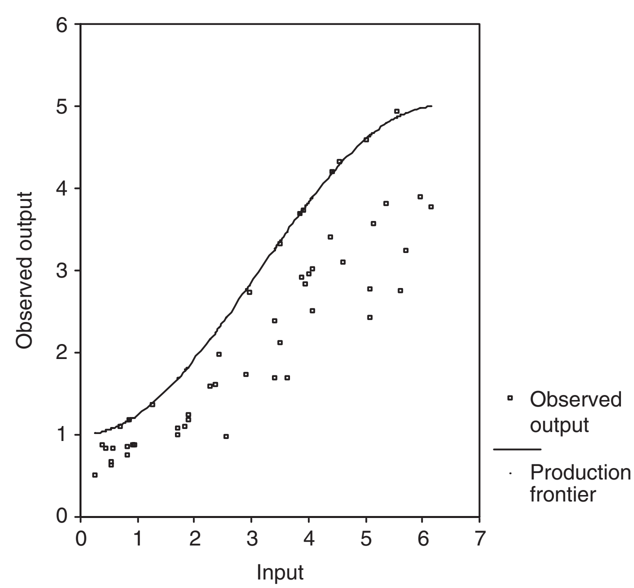

Fig. 1. An example of the input observed-output production function problem

(3) Once a positive error ε_1 is found, fitted output for all observations with a computed output less than DMU i, from i to 1, are shifted upwards adding up ε_1 to each average computed output.
(4) Detect the second positive error ε_2 belonging to DMU j with $\varepsilon_2 > \varepsilon_1 > 0$.
(5) Add up ε_2 to the fitted output for all DMUs between DMU j and DMU i.
(6) Repeat the stages described above until find DMU_k with the highest error $\varepsilon_k > \varepsilon_{k-1} > \varepsilon_2 > \varepsilon_1$.
(7) Add ε_k to the fitted output for all DMUs between DMU_k and DMU with error ε_{k-1}.
(8) The process can finish in these two ways:

- SMOOTHMLP: Adding up ε_k for remaining DMU from DMU_k to DMU_N.
- DEAMLP: If between DMU_k and DMU_N a DMU_h is found with the biggest real output y_h, and this output is smaller or equal that fitted output plus ε_k, i.e. $y_h \leq \hat{y}_h + \varepsilon_k$, then assign y_h for remaining DMUs from DMU_h to DMU_N.

When the entire process is over, a *thick frontier* is obtained for measuring efficiency. An illustration of both kinds of thick frontiers is showed in Fig. 2.

Simulation results

A number of 100 samples with size $N = 50$ were generated in the Monte-Carlo experiment to fit the production frontiers with each technique. Previous to training the MLP, data set was split assigning 80% to training and 20% to validation sets. After an exploratory analysis, it was tested that error differences for training and validation patterns were almost identical. Thus, in-sample (training set) and out-of-sample (validation set) estimations were joined for computing estimated output. A number of five neurons in one hidden layer were selected with learning coefficient and a weight decay term fixed both at 0.1. A logistic function in all neurons was used as transfer function.

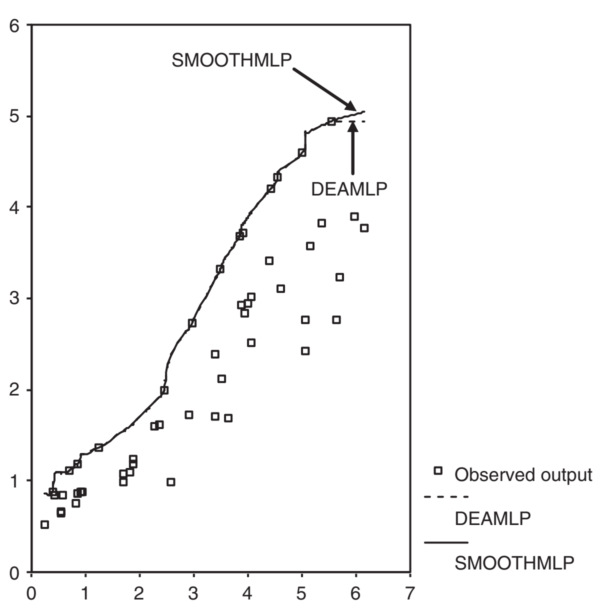

Fig. 2. Two kinds of thick frontiers, DEAMLP and SMOOTHMLP

Average technical efficiency and Pearson's correlation coefficient between real and estimated efficiency scores were computed in order to compare the performance of each approach regarding real efficiency. Simulation results are presented in Table 1.

The comparative results of the differences between the real efficiency and the various approaches used reveal that both ANN thick frontier methodologies obtain the best results. Thick frontier approaches show superior results over DEAvrs mainly in terms of correlation with real for average efficiency estimations. However DEAvrs results are substantially better than the MLPMAX approach. Moreover the results displayed in Table 1 for the parametric tool (SFA) are very similar to MLPMAX but worse than remaining techniques.

III. Conclusions

The results obtained in this article can be summarized as follows. First, main relative advantages of ANN compared with traditional efficiency techniques are for those problems with nonlinear relationships between variables that presents a weak theoretical knowledge about the production technology.

Last, the so-called thick frontiers developed in this article overcome traditional methods to draw up the

Table 1. The Monte-Carlo experiment results

	Average computed efficiency	Pearson's correlation coefficient between real and estimated efficiency
Real Efficiency	0.7506 (0.0222)	–
DEAMLP	0.7324 (0.039)	0.9417 (0.06526)
SMOOTHMLP	0.7291 (0.0376)	0.9434 (0.0687)
DEAvrs	0.7004 (0.028)	0.8501 (0.0418)
MLPMAX	0.6032 (0.0424)	0.6417 (0.1295)
SFA	0.5962 (0.0388)	0.6387 (0.1372)

Note: SD is showed in parenthesis.

production frontier from ANN average production function. However further research is still necessary in order to generalize this result in different scenarios (number of DMUs, signal to noise ratio, average efficiency and so on). This research should also explore the possibilities of integrating several approaches, combining its potential benefits, in order to enhance technical efficiency measurement and production function estimations.

Acknowledgements

This article owes a debt to Francisco Pedraja-Chaparro, Abel Santin and Aurelia Valiño for valuable comments, help and suggestions.

References

Athnassopoulos, A. and Curram, S. (1996) A comparison of data envelopment analysis and artificial neural networks as tools for assessing the efficiency of decision-making units, *Journal of the Operational Research Society*, **47**, 1000–16.

Banker, R. D., Charnes, A. and Cooper, W. W. (1984) Models for the estimation of technical and scale inefficiencies in data envelopment analysis, *Management Science*, **30**, 1078–92.

Battese, G. E. and Coelli, T. J. (1988) Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data, *Journal of Econometrics*, **38**, 387–99.

Bishop, C. M. (1995) *Neural Networks for Pattern Recognition*, Oxford University Press, Oxford.

Cooper, W. W., Seiford, L. M. and Zhu, J. (2004) *Handbook on Data Envelopment Analysis*, Springer, Boston.

Costa, A. and Markellos, R. N. (1997) Evaluating public transport efficiency with neural network models, *Transportation Research C*, **5**, 301–12.

Kumbhakar, S. C. and Lovell, C. A. K. (2000) *Stochastic Frontier Analysis*, Cambridge University Press, New York.

Mas-Colell, A., Whinston, M. and Green, J. (1995) *Microeconomic Theory*, Oxford University Press, London.

Papadas, C. T. and Hutchinson, W. G. (2002) Neural network forecast of input-output technology, *Applied Economics*, **34**, 1607–15.

Pendharkar, P. C. (2005) A data envelopment analysis-based approach for data preprocessing, *IEEE Transactions on Knowledge and data engineering*, **17**, 1379–88.

Pendharkar, P. C. and Rodger, J. A. (2003) Technical efficiency-based selection of learning cases to improve forecasting accuracy of neural networks under monotonicity assumptions, *Decision Support Systems*, **36**, 117–36.

Santin, D., Delgado, F. J. and Valiño, A. (2004) The measurement of technical efficiency: a neural network approach, *Applied Economics*, **36**, 627–35.