Forecasting the role of public expenditure in economic growth Using DEA-neural network approach

Arshia Amiri and Bruno Ventelou

GREQAM CNRS – ORS PACA – INSERM U912, 23 rue Stanislas Torrents, F-13006 Marseille, France, Department of Agricultural economics, College of Agriculture, Shiraz University, Shiraz, Iran

13. September 2011

Forecasting the role of public expenditure in economic growth Using DEA-neural network approach

Arshia Amiri [a, b],*, and Bruno Ventelou [b],

[a] ***Department of Agricultural economics, College of Agriculture, Shiraz University, Shiraz, Iran***
[b] ***GREQAM CNRS – ORS PACA – INSERM U912, 23 rue Stanislas Torrents, F-13006 Marseille, France***

Abstract
This paper integrates data envelopment analysis (DEA) and artificial neural networks (ANN) to forecast the role of public expenditure in economic growth in OCDE countries. The results show that this approach is a powerful and appropriate method to forecast this role. DEA method allows us to develop a neutral evaluation, unbiased a priori by any type of criteria, of the proportions in which the goal of productive spending is pursued, for any expenditure. Then we apply ANN to forecast economic growth by using input data taken at frontier. At the end of the DEA-ANN chain, prediction-power tests appear positive: best structures of multiple hidden layers indicate more ability to forecast according to best structures of single hidden layer but the difference between those is not much.

Keywords: DEA method; Economic growth; Public expenditure; Artificial neural network; OCDE countries

*** Corresponding author. Tel.: +98 917 308 57 22; fax: +98 711 822 42 22.**

E-mail addresses: arshiaamirishiraz@gmail.com (A. Amiri).

Introduction

Forecasting the role of public expenditure in economic growth has a special importance in politics of macroeconomic, for example, appointment the optimum size of governments and effects of expansionary and contractionnary policies of governments in economic growth in short run.

This research uses an explicit endogenous growth model due to Barro (1990), in which public expenditure is considered as an input of the macroeconomic production function. For y, the GDP per unit of labour, we have: $y = f(k, d)$ with k, the private capital by unit of labour, and d, a "productive public expenditure". As first shown by Barro (1990), this enrichment of the Solow model allows generating positive and permanent growth rate for the economy: the law of decreasing returns (valid for the private capital) could be offset by a continuous flow of public expenditure, counterbalancing period after period the "falling tendency of the rate of profit".

On the empirical side, we have no research like this that forecast the role of public expenditure in economic growth, but nearly we have some good attempts that found the relationship between public expenditure and economic growth, like Bleaney et al. (2001), and Ventelou and Bry (2006). Ventelou, and Bry applied DEA method to correct the public expenditure data by a factor of "productive efficiency", supposing that inefficiency is not "pure inefficiency" associated with wasted resources but could be "hidden output" (Ventelou and Bry, 2006)[1].

We also call attention to studies that used DEA- Neural network approach to their works, The idea of combination of neural networks and DEA for classification and/or prediction was first introduced by Athanassopoulos and Curram (1996). They treated DEA as a preprocessing methodology to screen training cases in a study of forecasting the number of employees in the health care industry. After selecting samples, the ANNs are then trained as tool to learn a nonlinear forecasting model. In Costa and Markellos (1997), ANNs are compared with corrected ordinary least squares (COLS) and DEA in the application to the London underground efficiency analysis. Their findings reveal that ANNs perform better in regard of the decision making, the impact of constant vs. variable returns to scale or congestion areas. Fleissig et al. (2000) estimate the cost functions using neutral networks. They also show some techniques to treat convergence problems in the training of ANNs. Pendharkar and Rodger (2003) used DEA as a data screening approach to create a sub sample training data set that is 'approximately' monotonic, which is a key property assumed in certain forecasting problems. Their results indicate that the predictive power of an ANN that is trained on the 'efficient' training data subset is stronger than the predictive performance of an ANN that is trained on the 'inefficient' training data subset. With base these results, first we use DEA method to correct the public expenditure data, then we apply neural network to forecast.

Data resource and modeling

In order to conduct this analysis, we reconsidered public expenditure data categorized by functions for a maximum number of OCDE countries, as well as data concerning their GDP growth rates per person. Although the small size of the sample (only 15 countries), this is the only available source of data allowing international comparisons.

We chose to analyse growth rates of the "per capita GDP" over a 10-year period, 1989–1999, rather than use raw data from any 1 year. There were several reasons for this choice. First, the

---

[1] More precisely, we call attention to studies conducted by Aschauer (1988, 1989), Lynde (1992), Devarajan et al. (1996), Kneller et al. (1999), Hamiltona and Turton (2002), and we base the results of Ventelou and Bry (2006), to apply DEA method for correcting the public expenditure data.

decade chosen allows us to study the majority of the countries in our chosen sample within the framework of a nearly identical (peak to peak) economic cycle so that the differences in growth do not reflect cycle discrepancies but, rather, the (tested) efficiency of public expenditure. From the perspective of a conditional convergence test, it was necessary to enhance this growth measurement by an adjustment factor of the "catching-up effect", which was, in this case, a value for per capita GDP at the beginning of the study period (1990). Then, after a standard OLS-adjustment to GDP/capita in 1990 (in logarithm), we obtained the data series "Index of Adjusted Growth Rates", which indicate by comparison to what degree economic-growth was or was not strong during the given period.
As for the input, in order to work on a precise breakdown of public expenditure by function, we used the OCDE nomenclature. This nomenclature, called classification of the functions of government (COFOG), breaks down spending into eleven items: general administrative services, justice and police, defense, education, housing, transport and communication, other economic services, cultural/worship/leisure activities, social security, and miscellaneous items. We regrouped these in our analysis into five representative items: administrative services, justice/police/defense, education, health/social security, and miscellaneous expenditures. The reported figures are based on PPP in dollars and are deflated with the use of the US consumer price index in order to obtain volume data for the base year, 1987. Finally, we divided this adjusted data by the total population in order to obtain public expenditure per person.
We chose 1991 as the observation base year for public expenditure. That is to say, for each country we observed a level and pattern of public s expenditure at the beginning of the period (the decade defined for peak to peak growth rates) and sought to evaluate to what extent exactly these patterns were able to contribute to the relative growth in these countries during the years which followed.

1. Data Envelopment Analysis method (DEA)
Data envelopment analysis is the non-parametric mathematical programming approach to frontier estimation. The piecewise-linear convex hull approach to frontier estimation, proposed by Farrell (1957), was considered by only a handful of authors in the two decades following Farrell paper. Authors such as Afriat (1972) suggested mathematical programming which could achieve the task, but the method did not receive wide attention until a paper by Charnes et al. (1978) which coined the term Data Envelopment Analysis (DEA). There have since been a large number of papers which have extended and applied the DEA methodology.
Charnes et al. (1978) proposed a model which had an input orientation and assumed constant returns to scale (CRS). Subsequent papers have considered alternative set of assumption, such as Banker et al. (1984) who proposed a variable returns to scale (VRS) model. One form of their (VRS) model is

$$E_r = \max \sum_{k=1}^{t} u_k y_{rk} \Big/ \left(v_0 + \sum_{j=1}^{s} v_j x_{rj} \right)$$

$$\text{s.t.} \quad \sum_{k=1}^{t} u_k y_{rk} \Big/ \left(v_0 + \sum_{j=1}^{s} v_j x_{rj} \right) \leq 1, \quad i = 1,\ldots,n$$

$u_k\,;\ v_j \geq \varepsilon > 0;\ v_0$ unconstrained in sign,

Where X_{ij} and Y_{ik} represent input and output data for the *i*th DMU with *j* ranging from 1 to s and k from 1 to t, and ε is a small non-Archimedean quantity (Charnes and Cooper, 1984; Charnes et al. 1979). Index r indicates the DMU to be rated, and there are n DMUs. When v_0 is set to 0, the assumption of constant returns to scale is imposed, and the model becomes that of Charnes et al. (1979). Note that Model (1) is a linear fractional program which can be transformed to a linear program:

$$E_r = \max \sum_{k=1}^{t} u_k y_{rk}$$

$$\text{s.t.} \quad \left(v_0 + \sum_{j=1}^{s} v_j x_{rj} \right) = 1$$

$$\sum_{k=1}^{t} u_k y_{ik} - \left(v_0 + \sum_{j=1}^{s} v_j x_{ij} \right) \leq 0, \quad i = 1,\ldots,n$$

$$u_k\,;\ v_j \geq \varepsilon > 0;\ v_0 \text{ unconstrained in sign,}$$

Therefore, the conventional LP method can be applied to solve Er. DEA model seeks to determine which of the n DMUs define an envelopment surface that represents best practice, referred to as the empirical production function or the efficient frontier. Units that lie on the surface are deemed. Efficient in DEA while those units that do not, are termed inefficient. DEA provides a comprehensive analysis of relative efficiencies for multiple input-multiple output situations by evaluating each DMU and measuring its performance relative to an envelopment surface composed of other DMUs. Those DMUs are the peer group for the inefficient units known as the efficient reference set. As the inefficient units are projected onto the envelopment surface, the efficient units closest to the projection and whose linear combination comprises this virtual unit form the peer group for that particular DMU. The targets defined by the efficient projections give an indication of how this DMU can improve to be efficient (Wu et al., 2006).

2. Artificial neural network (ANN)

ANN acts according to by many names, for example connectionist models and parallel distributed processing models. The ANN modeling is a capability to solve problems by implementing information acquired from past knowledge to new problems in other words is a computer method that tries to simulate some important features of the human nervous system.

Comparable to a human brain, an ANN employs numerous simple calculative constituents, named artificial neurons, connected by variable weights. Although each neuron, alone, can only perform simple calculations, the hierarchical organization of a network of interconnected neurons makes an ANN capable of performing complicated jobs such as pattern classification and prediction. Recently, artificial neural networks (ANNs) have been proposed as efficient tools for modeling and forecasting. ANNs are supposed to possess the capability to reproduce the unknown relationship existing between a set of input explanatory variables of the system and the output variables.

The multi-layer perceptron structure is commonly applied for prediction. Input layers, hidden layers, and an output layer are layers of multi-layer structure and neurons are organized in these.

A typical configuration for a multi-layer perceptron, a special class of ANN that will be used in this research, is shown in Fig. 1. Before computations by neural network model is effective that Input neurons; get values of an instance of the input parameters that are deliver over to the network after being scaled into a numeric range. Output neurons act instance of output parameters. Hidden neurons link the input neurons to the output neurons and prepare nonlinearity to the network. Every neuron is linked to other neuron in contiguous layers by a relation weight, which determines the strength of the relationship between two connected neurons. In neural network before being introduced as input to the neuron in the subsequent layer, the input from a neuron is multiplied by the relation weight. Actually network for obtain most efficient fit, checks numerous types of relationships. The back-propagation algorithm has been comprehensively utilized for training multi-layer neural networks.

The algorithm uses a gradient to search method of performance to reduce as much as possible a cost function equivalent to the mean square difference between the wanted and the real net outputs. It needs a continuous differentiable nonlinearity to be used as the transfer function by the neurons. Two types transfer function that used in back propagation networks are tan-sigmoid transfer function (tansig) and linear transfer function (purelin). The transfer function relating the inputs to the it h hidden neurons is given by

$$h_i = \tanh\left(\sum_i w_{ij} x_j + \theta_i\right)$$

The relationship between the hidden neurons and the output layer is linear, that is

$$D = \sum_i w_i b_i + \theta$$

The coefficient w and biases b of these equations are determined in such a way as to minimize the energy function. Because the hyperbolic function is a nonlinear function, a non-linear relationship can be predicted using this model.

Both the input and output variables were first normalized within the range 0–1 as follows:

$$\mathrm{X}_N = \frac{X}{X_{\max}}$$

Where X_N is the normalized value of X, X_N is the maximum value of each variable of the original data. This normalization is not essential to the neural network approach, but allows the network to be trained better. Using the normalized data, the coefficients (weights) w and bias b were determined in such a way as to minimize the following energy function

$$M(w) = \beta E_D + \sum_c \alpha_c E_{w(c)}$$

The minimization was implemented using a variable metric optimizer. The gradient of M (w) was computed using a back-propagation algorithm. For the structure of an ANN framework use a feed forward network with one, two and three hidden layer.

Feed forward network

In feed forward network neurons in each layer linked to neurons in the next layer

Fig. 1 shows the structure diagram of a feed forward network. Mathematical descriptive of a feed forward procedure are as is explained in the coming section

$$s_j^J = \sum_i w_{ji}^{JI} x_i^I + b_j^J$$

$$x_j^J = f_J\left(s_j^J\right)$$

$$s_k^k = \sum_j w_{kj}^{KJ} x_j^J + b_k^K$$

$$x_k^K = f_K\left(s_k^K\right)$$

where superscripts I, J and K show the input, hidden and output layers, respectively, subscripts i, j and k mean the nodes of I, J and K layers, respectively, x indicates the nodal value, w indicates the weight between two nodes, b indicates the nodal bias, s indicates the weighted summation of nodal values in the previous layer with a nodal bias, and f is an activation function of every layer.

Back propagation algorithm

Back propagation neural network is the most widely used neural network technique for classification or prediction (Hecht, 1990; Shavlik et al. 1991; Poli and Jones, 1994; Liang and Wu, 2005). Input vectors and the corresponding target vectors are used to train a network until it can approximate a function, associate input vectors with specific output vectors, or classify input vectors in an appropriate way. Networks with biases, a sigmoid layer, and a linear output layer are capable of approximating any function with a finite number of discontinuities.

Network validations

Validation of the trained network as previously mentioned was done using 20% of the available data, which were not considered by the network during the learning process.
Figure 1 Conceptual diagram of a feed forward network with one hidden layer.

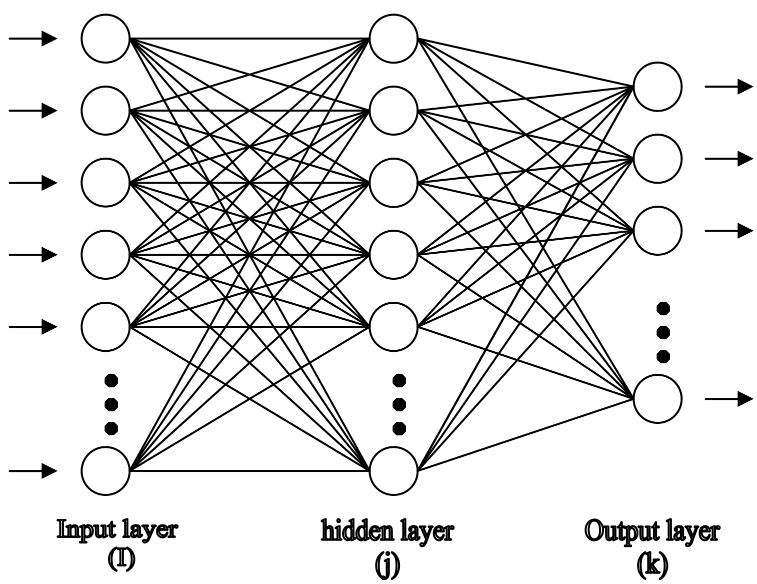

Figure 1: Conceptual diagram of a feed forward network

Result and discussion

DEA results (Effective public expenditure)

Tables 1 illustrate the results of DEA method, adjustment score of DEA efficiency and index of adjusted growth rates. DEA was implemented by using Frontier analyst software package (Frontier analyst version 4 with DEA toolboxes).
After using DEA method, ANN was applied to prediction.

Countries	General Public Services	Defense & Law Enforcing	Education	Health & social security	Other expenditures	Adjustment score (DEA efficience) (%)	Index of adjusted growth rates
Germany	245.85	202.12	23.49	1851.44	659.70	69.45	99.01
Australia	316.77	332.37	253.26	1538.08	1168.66	100.00	100.43
Austria	187.16	166.65	325.15	1987.72	764.32	60.96	100.57
Denmark	288.61	268.11	371.77	1538.42	1329.70	63.72	99.39
Spain	87.88	237.72	146.04	1382.05	1263.05	83.81	101.29
United States	348.87	1067.47	81.78	1885.73	1343.63	100.00	99.39
Finland	153.21	212.82	422.76	1461.06	791.97	62.72	99.20
France	182.26	248.47	240.46	2071.66	648.55	51.73	100.05
Greece	181.32	319.47	231.60	682.61	1545.53	100.00	99.06
Japan	84.00	200.29	213.98	1309.37	1588.44	100.00	100.02
Norway	219.05	438.04	269.91	1835.00	1882.27	71.11	101.21
The Netherlands	280.54	301.87	454.55	2178.31	1076.75	59.60	100.94
United Kingdom	134.13	451.28	106.01	1378.01	1119.01	61.72	99.54
Sweden	146.59	305.94	319.66	1754.43	785.81	55.32	99.80
Switzerland	135.17	260.32	86.02	2006.14	582.96	71.71	99.25

Table 1. Public expenditure by function in PPP $ per person/DEA-adjusted data/Index of adjusted growth rates

Development of ANN method for prediction the role of public expenditure in economic growth

In this study, two types of ANN models were developed (1) single hidden-layer ANN models that consisting of only one hidden layer and (2) multiple hidden-layer ANN models consisting of two and three hidden layers. But as mentioned, the number of neurons in the hidden layer(s) can be determined through the use of trial and error procedure The optimal architecture was determined by varying the number of hidden neurons (from 1 to 20), and the best structure was selected. The training of the ANN models was stopped when either the acceptable level of error was achieved or when the number of iterations exceeded a prescribed maximum of 2500. The learning rate of 0.05 was also used. ANN was implemented by using MATLAB software package (MATLAB version 7.2 with neural network toolboxes).

The performance of all ANN configurations was assessed based on calculating the mean absolute error (MAE), and the root mean square error (RMSE). The coefficient of determination, R2, of linear regression line between the predicted values from either the ANN and the desired output was also used as a measure of performance. The three statistical parameters used to compare the performance of the various ANN configurations are:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|O_i - t_i\right|,$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}\left(O_i - t_i\right)^2}{N}},$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}\left(O_i - t_i\right)^2}{\sum_{i=1}^{N}\left(O_i - \overline{O}_i\right)^2}$$

Where O_i and t_i are observed and predicted for the i th output (observed output that comes from catching-up effect and predicted output that ANN forecasts it), and $\overline{O}_i$ is the average of predicted, and N is the total number of events considered. The models that minimized the two error measures described in the previous section (and optimum R^2) were selected as the optimum. The whole analysis was repeated several times.

Forecasting the role of public expenditure in economic growth

The main reason of executing this part is determine ability of the DEA-ANN chain for prediction of public expenditure in economic growth. The whole data set consisting of 15 data points which was divided into two parts randomly: a training set consisting of 80% data points and a validation or testing set consisting of 20% data points (out of sample).

In single hidden-layer ANN prediction optimal architecture determined by varying the number of hidden neurons (from 1 to 20), and the best structure was selected. It was found that the most accurate results involved use of the Feed Forward Back Propagation with five neurons in hidden layer and architecture of configuration: 5-5-1.

In multiple hidden-layer ANN prediction optimal architecture determined by varying the number of hidden neurons, and the best structure was selected. It was found that the most accurate results involved use of the Feed Forward Back Propagation with two hidden layer and architecture of configuration: 5-1-2-1.

To evaluate the performance of the ANN prediction, observed role of public expenditure in growth rate (taken at frontier of best practice) are plotted against the predicted ones for single and multiple hidden layer. Figure 2 and 3 illustrate the results with the performance indices between predicted and observed data for the training and testing data sets, respectively.

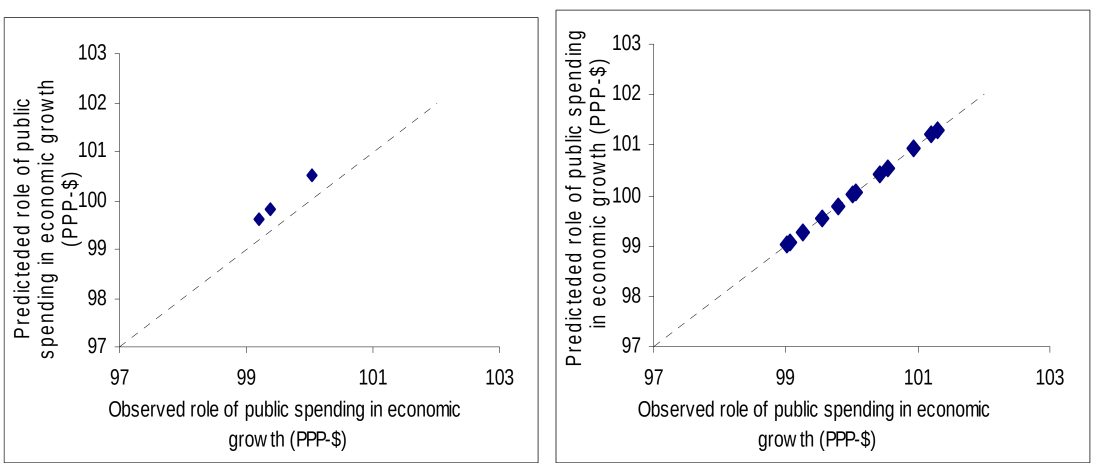

Figure 2: performance of ANN training and testing of single hidden layer model.

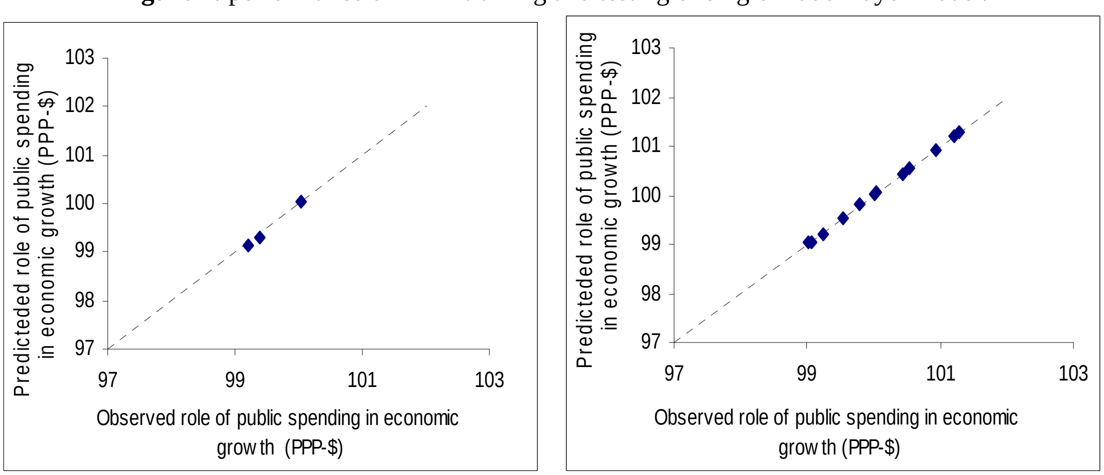

Figure 3: performance of ANN training and testing of multiple hidden layer models.

Table 2. The gist of results of ANN model.

		Single	Multiple	OLS
Validation (out of sample)	MAE	0,004317	0,000450	0,569691
	RMSE	0,001675	0,000208	0,698202
	RSQ	0,998973	0,998826	0,495291
Training	MAE	2,25E-05	0,000122	0,001414
	RMSE	1,1E-05	6,2E-05	0,001789
	RSQ	0,999991	0,999554	0,944152

The figures exhibits that best structure of ANN with multiple hidden layers have lower validation error compared with best structure of single hidden layer. The results of ordinary least squares are lower that ANNs for prediction, both results from out of sample and training come to improve the deficit degree of freedom. With basing the Granger–Newbold test, results from ANNs is more powerful than OLS, but the difference between results from best structure of single layer instead of multiple layers is not significant.

Conclusion

This study performs for determined ability of the DEA-ANN chain to forecast the role of public expenditure in economic growth. The results show that this approach is a powerful method. First, DEA method allows us to develop a neutral evaluation, unbiased a priori by any type of criteria, of the proportions in which the goal of productive spending is pursued, for any expenditure. This evaluation is coherent with the standard way of analyzing production functions: identification of best practices, as a prerequisite for the identification of a production frontier. Second, we are then able to re-calculate the role of public spending by using ANN on these data taken at frontier. The prediction-power test appears positive, although of debatable reliability since it utilizes a small sample. Last, best structure of multiple hidden layer indicate more ability to forecast according to best structure of single hidden layer, but the difference between those is not much (maybe because of our small sample…).

References

Afriat, S.N. (1972) Efficiency estimation of production function. *International Economic Review* 13: 568–598.

Aschauer, D.A. (1988) Governement spending and the falling rate of profit. *Economic Perspectives* (12).

Aschauer, D.A. (1989) Is public expenditure productive? *Journal of Monetary Economics* 23: 177–200.

Athanassopoulos, A.D. & Curram, S.P. (1996) A comparison of data envelopment analysis and artificial neural networks as tool for assessing the efficiency of decision making units. *Journal of the Operational Research Society* 47(8): 1000–1016.

Barro, R. J. (1990) Government spending in a simple model of endogenous growth. *Journal of Political Economy* 98: 103–125.

Banker, R.D., Charnes, A. & Cooper, W. (1984) Some model for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science* 30: 1078–1092.

Bleaney, M., Gemmel, N., & Kneller, R. (2001) Testing endogenous growth model: public expenditure, taxation and growth over the long run. *Canadian Journal of Economics* 34(1).

Charnes, A. & Cooper, W. (1984) The non-Archimedean CCR ratio for efficiency analysis: a rejoinder to Boyd and Fare, *European Journal of operational Research* 15: 333–334.

Charnes, A., Cooper, W. & Rhodes S.E. (1978) Measuring efficiency of decision-making units. *European Journal of operational Research* 2(6).

Charnes, A., Cooper, W. & Rhodes S.E. (1979) Measuring efficiency of decision-making units. *European Journal of operational Research* 3: 339.

Devarajan, S., Swaroop, V. & Zou, H.F. (1996) The composition of public expenditure and economic growth. *Journal of Monetary Economics* 37: 313–334.

Costa, A. & Markellos, R.N. (1997) Evaluating public transport efficiency with neural network models. *Transportation Research* C, 5(5): 301–312.

Farrell, M.J. (1957) The measurement of productive efficiency. *Journal of Royal Statistical Society*, A CXX, Part 3: 253–290.

Fleissig, A.R., Kastens, T. & Terrell, D. (2000) Evaluating the seminonparametric fourier, aim, and neural networks cost functions. *Economics Letters* 68(3): 235–244.

Hamiltona, C. & Turton, H. (2002) Determinants of emissions growth in OECD countries. *Energy Policy* 30: 63–71.

Hecht Nielsen, R. (1990) Neural computing Reading. MA: Addison Wesley, p. 124–133

Kneller, R., Bleaney, M.F. & Gemmell, N. (1999) Fiscal policy and growth: evidence from OECD countries. *Journal of Public Economics* 74: 171–190.

Liang, L. & Wu, D. (2005) An application of pattern recognition on scoring Chinese corporations financial conditions based on backpropagation neural network. *Computers and Operations Research* 32(5).

Lynde, C. (1992) Private profit and public capital. *Journal of Macroeconomics* 14(1).

Pendharkar, P.C. & Rodger, J.A. (2003) Technical efficiency-based selection of learning cases to improve forecasting accuracy of neural networks under monotonicity assumption. *Decision Support Systems* 36(1): 117–136.

Poli, I. & Jones, R.D. (1994) A neural net model for prediction. *Journal of the American Statistical Association* 89: 117–121.

Shavlik, J.W., Mooney, R.J. & Towell, G.G. (1991) Symbolic and neural learning algorithms: An experimental comparison. *Machine Learning* 6: 111–143.

Ventelou, B. & Bry, X. (2006) The role of public spending in economic growth: Envelopment methods. *Journal of Policy Modeling* 28: 403–413.

Wu, D., Yang, Z. & Liang, L. (2006) Using DEA-neural network approach to evaluate branch efficiency of a large Canadian bank. *Expert Systems with Applications* 31: 108–115.