Discussion Papers in Economics

NEURAL NETWORK MODELS FOR INFLATION FORECASTING: AN APPRAISAL

By

Ali Choudhary
(University of Surrey and State Bank of Pakistan)
&
Adnan Haider
(State Bank of Pakistan)

DP 08/08

Department of Economics
University of Surrey
Guildford
Surrey GU2 7XH, UK
Telephone +44 (0)1483 689380
Facsimile +44 (0)1483 689548
Web www.econ.surrey.ac.uk
ISSN: 1749-5075

Neural Network Models for Inflation Forecasting: An Appraisal

Ali Choudhary*‡ and Adnan Haider†

*University of Surrey and ‡†State Bank of Pakistan

ABSTRACT

We assess the power of artificial neural network models as forecasting tools for monthly inflation rates for 28 OECD countries. For short out-of-sample forecasting horizons, we find that, on average, for 45% of the countries the ANN models were a superior predictor while the AR1 model performed better for 21%. Furthermore, arithmetic combinations of several ANN models can also serve as a credible tool for forecasting inflation.

JEL Classification: C51, C52, C53, E31, E37

Keywords: Artificial Neural Networks; Forecasting; Inflation

† Corresponding Author: adnan.haider@sbp.org.pk.

1. Introduction

There is growing interest in using artificial neural network (ANN) as a complimentary approach to forecast macroeconomic series[1]. The reason for this rising popularity is that ANN pays particular attention to nonlinearities and learning processes both of which can help improve predictions for complex variables.

In this paper we extend Nakamura (2006)'s, McNelis and McAdam (2005)'s and Moshiri et al., (1999)'s efforts on highlighting the potential role of ANN for forecasting inflation in two ways. First, we evaluate these techniques to forecast inflation for large set of countries. Second, we introduce arithmetic combinations of several established ANNs to improve predictability.

We forecast monthly inflation rates for 28 OECD countries using two ANN and two quasi-ANN techniques. The first two are the commonly known as hybrid and dynamic ANN models. A third method averages the forecasts of the hybrid and dynamic ANN to predict inflation. A final method produces forecasts using the minimum distance criterion in that our algorithm selects those points from either the hybrid or the dynamic models that are closest to their mean forecasts.

Two results standout: (i) that the neural nets considerably outcompete the AR(1) tool to predict short horizons up to 3 months (ii) that simple hybrid learning rule and the minimum distance quasi-ANN rules dominate other forms of neural nets.

In the next section we present the methodology which is followed by the results. We end with concluding remarks.

2. Methodology

Neural networks are particularly useful for future predictions for variables for which the data generating process is not well known and may also be subject to nonlinearities. For example, inflation is an amalgamation of complex expectation formation[2] processes across the economy and as a result has become a popular candidate in the study of neural nets as a forecasting tool.

Neural nets consist of layers of interconnected nodes which combine the data in a way to minimize root mean squared error (RMSE) but the researcher may also use some other minimizing criteria such mean absolute percentage error (MAPE). One simple example of a network is a pyramid type structure[3] where each brick represents a node. Raw information is fed at the bottom of the pyramid where each node independently processes information and then transmits output, weighted by the importance of the node in question, to *all* the nodes sitting in the layer above. The nodes in this new layer then process the already processed-

[1] See for example Fernandez-Rodriguez et al. (2000) and Redenes and White (1998), Nakamura (2006) and Moshiri and Cameron (2000)), see Chen, Racin and Swanson (2001), Swanson and White (1997) and Stock and Watson (1998)

[2] See for example Brock and Hommes (1997).

[3]Formally known as a 'feed-forward' mechanism.

data and then pass on their weighted outputs to nodes on the layer above. This process continues until the node at the top of the pyramid finally transmits output of interest to the researcher. The final output is then checked against a RMSE criterion and if the criterion is not met, learning happens by taking into consideration the size of the error and a rule which allows adjusting initial weights assigned to each node and each layer in the pyramid. One key point that deserves mentioning is that each node is equipped with a combination function which combines various data points into a single value using weights. These single values are then transformed into the unit circle normally using a trigonometric function.

This study extends the pyramid type structure to forecast inflation rates using two neural and two quasi-neural architectures. The first is known as a hybrid-network (see Nakamura (2006) for example) whereby the properties of the pyramid like structure are retained with the advantage that the nodes sitting in between and the top and bottom layers can communicate with one another and pass on combined values that can ease the way for minimizing RMSE in the final stages.

The functional form of hybrid-network model is given as:

$$\hat{\pi}_{hybrid,t+i} = \sum_{k} \Theta_{ik} \tanh(w_k x_{t-1} + b_k) \tag{1}$$

where x_{t-1} is the vector of lagged inflation variables, Θ_{ik}, w_k, b_k denote weight at the k^{th} node's weight positioned in the i^{th} layer and the weight of the data point assigned to k^{th} node and biases at k^{th} node respectively. This model produces inflation forecasts for *i*-months ahead.

The second ANN model is the dynamic extension of hybrid neural network model with recursive behavior. Studies such as Elman (1990), Kuan and Liu (1995), Balkin (1997) and Moshiri et al., (1999) use this architecture to predict the economic variables. This model includes the lags of the dependent variable as an explanatory variable in the hybrid network to capture richer dynamics. The functional form for each node is:

$$\hat{\pi}_{dynamic,t+j} = \Psi_0 \left(\upsilon_{dc} + \sum_{h} \upsilon_{ho} \Psi_h (\upsilon_{ch} + \sum_{i} \upsilon_{ih} \hat{\pi}_{dynamic,t-j_i} + \sum_{j} \upsilon_{lh} \Gamma_{j,t-1}) \right) \tag{2}$$

where υ_{dc} denotes the weight of the direct connection between the constant input and the output. υ_{ho} denote the weight for the connections between the constant input and nodes. The terms υ_{ch}, υ_{ih} and υ_{lh} are weights of other connections. The functions Ψ_0 and Ψ_h are activation functions and $\Gamma_{j,t-1}$ represents the value of network output from the previous time unit of a dynamic network. The analytical algorithmic description of this model is extensively explained in Kuan and Liu (1995) and Balkin (1997).

One caveat in both networks above is that they may produce wide forecasts, especially if data is volatile or contains a number of structural breaks, see, Medeiros et al., (2002). In order to produce sharper forecasts we introduce two quasi-neural network procedures. The first averages forecasts of (1) and (2) as:

$$\hat{\pi}_{average,i,t+j} = \frac{1}{2}\left[\hat{\pi}_{hybrid,t+j} + \hat{\pi}_{dynamic,t+j}\right] \tag{3}$$

The second is an algorithm that selects values on the basis of minimum distance of hybrid (1) and dynamic (2) network forecasts from the average forecast (3). This can be written as:

$$\hat{\pi}_{\min\_dis,t+j} = \min\left[(\hat{\pi}_{hybrid,t+j} - \hat{\pi}_{average,t+j}), (\hat{\pi}_{dynamic,t+j} - \hat{\pi}_{average,t+j})\right] \tag{4}$$

Implementing neural and quasi-neural network models (1), (2), (3) and (4), require the following steps: First, identifying the variables, which help forecast the target variable, and processing the input data. Second, layering network architecture where a minimum of three layers are required and the decision on the maximum number of layers needs experimentation. Since this study considers monthly data on inflation rate, the number of layers is twelve. We could use more layers but that would make the training time costly. Furthermore, for a dynamic network the literature recommends at most fifteen layers. At the network specification stage we can adjust a number of default parameters or values that influence the behavior of the training process. These deal with the learning, error tolerance rates of the network, the maximum number of runs, stop value for terminating training and randomizing weights with some specified dispersion.

The final step is training the network and forecasting. We train our specified ANN models using the Levenberg-Marquardt (LM) algorithm, a standard training algorithm used in the relevant literature. The algorithm is terminated according to an early stopping procedure that avoids over fitting (see Nakamura (2006)). The forecast evaluation use root mean of squared errors (RMSE) and mean absolute percentage error (MAPE) criteria. The training algorithm is run on the training set until the RMSE / MAPE starts to decrease on the validation set.

3. Empirical Results[4]

We use monthly inflation rates for 28 OECD countries based on IFS' consumer price index data from July-1991 to June 2008 from. We trained both neural and quasi-neural algorithms in MATLAB.[5] Initially, we normalize the data to bring it within the unit circle using:

$$\pi_t^n = 2*(\pi_t - \pi_t^{\max})/(\pi_t^{\min} - \pi_t^{\max}) \tag{5}$$

The normalized data is used as input of neural algorithms and hence transfers training function by using specified trigonometric function. The MATLAB neural network toolkit

[4] Detailed results and codes are available upon request.

[5] In order to simulate our algorithms we used MATLAB neural network toolkit. The default parameter values are assigned as: hidden layers = 12; max lag = 12; training set = 80; forecast period = 12; learning rate = 0.25; learning increment = 1.05; learning decrement = 0.07; training parameter epochs = 1000 and target RMSE = 0.00005.

procedure 'trainlm' is extended with our specific neural algorithms to train the data. To validate we also compare the forecast performance of our ANN models with AR (1). We estimate one step, three steps and twelve steps ahead out-of-sample forecasts for July 07 to June 08.

The top 5 rows of Table 1 show the percentage of countries for which the forecast for a given procedure is as good as or better than the forecast of competing techniques. The last two rows show the percentage of countries for which either the AR1 or any neural network produce superior forecasts. First, comparing rows 6 and 7 we find that using simple criterion-step average, the neural networks are superior for 45% of the countries while the simple AR1 process performs better for 21% of countries. For the remaining 34%, AR1 and any other network perform equally well. The results are therefore not conclusive. Second, the hybrid and (our newly developed) quasi-minimum distance techniques dominate all other forms of forecasting but AR1 is not far behind either. Finally, the last two columns show that there is no single technique that dominates long term forecasting; a result also found in the literature.

4. Conclusion

We show that overall neural network models and their certain combinations dominate the simple AR1 process for forecasting inflation rates for OECD countries; especially for short to medium term forecasts. However, there are some countries for which the AR1 technique provided sound results. It therefore may always be preferable to continuously compare econometric procedures with that of neural networks to make the choice for a forecasting tool.

REFERENCES

[1]. Balkin, S. D. (1997), Using recurrent neural network for Time series forecasting, Conference Paper, International Symposium on forecasting, Barbados, 1997.

[2]. Brock, William A. and Cars H. Hommes, (1997), A Rational Route to Randomness, *Econometrica*, 65 (5), pp 1059–1095.

[3]. Chen, Xiaohong, J. Racin, and N. Swanson, (2001), Semiparametric ARX Neural Network Models with an Application to Forecasting Inflation, *IEEE Transactions on Neural Networks*, 12, pp 674–683.

[4]. Elman, J. L. (1990), Finding structure in Time, *Cognitive Science*, 14, pp 179-211.

[5]. Fernandez-Rodriguez, Fernando, Christian Gonzalez-Martel, and Simon Sosvilla-Rivero, (2000), On the profitability of technical trading rules based on artificial neural networks:: Evidence from the Madrid stock market, *Economics Letters*, 69 (1), pp 89–94.

[6]. Kuan, Chung-Ming and Liu, T. (1995), Forecasting exchange rates using feedforward and Recurrent Neural Networks, Journal of Applied Econometrics, 10, pp 347 – 364.

[7]. Medeiros, M. C., Terasvirta, T., and Rech, G. (2002), Building neural network models for time series: A statistical approach, Stockholm School of Economics Working Paper No. 508.

[8]. McNelis, Paul and McAdam, Peter (2005), Forecasting inflation with Thick models and Neural Networks, *Economic Modeling*, 22 (5), 548 – 567.

[9]. Moshiri, Saeed and Norman Cameron, (2000), Econometrics Versus ANN Models in Forecasting Inflation, *Journal of Forecasting*, February Issue, 19, pp 201 – 217.

[10]. Moshiri, S., Cameron, N., and Scuse, D. (1999), Static, Dynamic and Hybrid Neural Networks in Forecasting inflation, Computational Economics, 14, pp 219 – 235.

[11]. Nakamura, Emi, (2006), Inflation forecasting using a neural network, Economics Letter, 86 (3), pp 373-378.

[12]. Redenes, A. P. and H. White, (1998), Neural Networks and Financial Economics, *International Journal of Forecasting*, 6 (17), pp 541—551.

[13]. Stock, James H. and Mark W. Watson, (1998), A Comparison of Linear and Nonlinear Univariate Models for Forecasting Macroeconomic Time Series, *NBER Working Paper* 6607, 7.

[14]. Swanson, Norman R. and Halbert White, (1997), A Model Selection Approach to Real-Time Macroeconomic Forecasting Using Linear Models and Artificial Neural Networks, *The Review of Economics and Statistics*, November Issue, 79 (4), pp 540–550.

APPENDIX

Table 1: The Percentage of Countries for Which a Procedure Minimizes RMSE or MAPE

		1 Step		3 Steps		12 Steps	
		RMSE	MAPE	RMSE	MAPE	RMSE	MAPE
1.	AR1	43%	71%	43%	50%	29%	25%
2.	**Hybrid**	**43%**	**68%**	**57%**	**57%**	18%	18%
3.	Dynamic	25%	39%	25%	32%	14%	32%
4.	Quasi-Avg	4%	36%	18%	32%	21%	21%
5.	**Quasi Min-Dist**	**43%**	**61%**	**50%**	**57%**	4%	11%
6.	AR1 Only	**29%**	**21%**	**25%**	**18%**	**25%**	**21%**
7.	Neural Only	**43%**	**29%**	**57%**	**50%**	**71%**	**71%**