



HAL
open science

Neural Network Models for Inflation Forecasting: An Appraisal

M. Ali Choudhary

► **To cite this version:**

M. Ali Choudhary. Neural Network Models for Inflation Forecasting: An Appraisal. Applied Economics, 2011, pp.1. 10.1080/00036846.2011.566190 . hal-00704670

HAL Id: hal-00704670

<https://hal.science/hal-00704670>

Submitted on 6 Jun 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Neural Network Models for Inflation Forecasting: An Appraisal

Journal:	<i>Applied Economics</i>
Manuscript ID:	APE-2009-0317.R1
Journal Selection:	Applied Economics
Date Submitted by the Author:	01-May-2010
Complete List of Authors:	Choudhary, M. Ali; University of Surrey, Economics
JEL Code:	C51 - Model Construction and Estimation < C5 - Econometric Modeling < C - Mathematical and Quantitative Methods, C52 - Model Evaluation and Testing < C5 - Econometric Modeling < C - Mathematical and Quantitative Methods, C53 - Forecasting and Other Model Applications < C5 - Econometric Modeling < C - Mathematical and Quantitative Methods, E31 - Price Level Inflation Deflation < E3 - Prices, Business Fluctuations, and Cycles < E - Macroeconomics and Monetary Economics, E37 - Forecasting and Simulation < E3 - Prices, Business Fluctuations, and Cycles < E - Macroeconomics and Monetary Economics
Keywords:	Artificial Neural Network, Forecasting, Inflation

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



For Peer Review

Neural Network Models for Inflation Forecasting: An Appraisal*

M. Ali Choudhary[‡] and Adnan Haider[§]

[§] [‡] State Bank Pakistan, Research Department, I.I. Chundrigar Road, Karachi 74000, Pakistan

[‡] University of Surrey, Department of Economics, Guildford, GU2 7XH, U.K; email:

ali.choudhary@sbp.org.pk; adnan.haider@sbp.org.pk

Abstract

We assess the power of diverse artificial neural-network models (ANN) as forecasting tools for monthly inflation rates for 28 OECD countries. In the context of short out-of-sample forecasting horizon we find that, on average, the ANN models were a superior predictor for inflation for 45% while the AR1 model performed better for 23% of the countries. Furthermore, we develop arithmetic combinations of several ANN models and find that these may also serve as credible tools for forecasting inflation.

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

Keywords: Artificial Neural Networks; Forecasting; Inflation

*We would like to thank the Editor, the anonymous Referee, Ron Smith, Vasco Gabriel and Peter McAdam for useful suggestions. We thank Amna Saeed for a patient research assistance.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Contents

1 Introduction	3
2 Methodology	3
3 Empirical Results	5
4 Conclusion	7

For Peer Review

1 Introduction

There is growing interest in using artificial neural network (ANN henceforth) 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 important work by Nakamura (2006)', McNelis and McAdam (2005)', Binner et al. (2005) Gazely and Binner (2000) and Moshiri et al., (1999)' which highlight the potential role of ANN in the context of forecasting inflation. We contribute to this literature in two ways.

First, unlike previous literature which tended to focus on a particular type of ANN and on a specific country, we evaluate competing ANN techniques to forecast inflation for a large set of countries against the well known AR1 procedure. Second, we introduce arithmetic combinations of established ANNs, we call these quasi-ANNs, to improve the overall predictability of neural methodologies.

We forecast monthly inflation rates for 28 OECD countries using two established and two quasi-ANN techniques. The first two are commonly known as hybrid and dynamic ANN models. A third method simply averages the forecasts of the hybrid and dynamic ANN to predict inflation. A final method produces forecasts using the minimum distance criterion, whereby our algorithm generates a series by selecting points from the projections of either the hybrid or the dynamic models that are closest to the average forecasts of these two models.

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

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

2 Methodology

Neural networks are particularly useful for future predictions of 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 formation² processes across the economy and as a result has become a popular candidate variable used in the study of neural nets as a forecasting tool.³

Neural nets consist of layers of interconnected nodes which combine the data in a way to minimize the root mean squared error (RMSE) but the researcher may also employ some other minimizing criteria such as mean absolute percentage error (MAPE). One simple example of a network is a pyramid type structure⁴ where each brick represents a node. Raw information is

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

²See for example Brock and Hommes (1997).

³Other examples are Alvarez-Diaz (2008) and Lin and Yeh (2007) studies that use ANNs to forecast exchange rates and stock markets option-pricing respectively.

⁴Formally known as a 'feed-forward' mechanism.

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 subsequent layer then process the processed data from the previous layer and then pass on their weighted outputs to the nodes in the layer sitting above. This process continues until a final node at the top of the pyramid transmits the output of interest to the researcher. The final output/series 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 in 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 using a trigonometric function.

This study uses a different version of 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), Gazzely and Binner (2000) 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 help minimize the RMSE (or MAPE) in the final stages.

The functional form for the hybrid-network model is given by:

$$\hat{\pi}_{hybrid,t+j} = \sum_k \theta_{ik} \tanh(w_k x_{t-1} + b_k) \quad (1)$$

where x_{t-1} , θ_{ik} , w_k , and b_k denote the vector of lagged inflation, the weight at the k th node positioned in the i th layer, the weight of the data point assigned to k th node and biases at k th node respectively. This model produces inflation forecasts for j months ahead.

The second ANN model is the dynamic extension of hybrid neural-network model with an inbuilt 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 in this type of network is given by:

$$\hat{\pi}_{dynamic,t+j} = \psi_0 \left[v_{dc} + \sum_h v_{ho} \psi_h \left\{ v_{ch} + \sum_i v_{ih} \hat{\pi}_{dynamic,t-j} + \sum_j v_{lh} \Gamma_{j,t-1} \right\} \right] \quad (2)$$

where v_{dc} denotes the weight of the direct connection between the constant input and the output. v_{ho} denotes the weight for the connection between the constant input and nodes. The terms v_{ch} , v_{ih} and v_{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). Hence, to produce sharper forecasts, we introduce two quasi-neural network procedures. The first simply averages forecasts of (1) and (2) as:

$$\hat{\pi}_{average,t+j} = \frac{1}{2} [\hat{\pi}_{hybrid,t+j} + \hat{\pi}_{dynamic,t+j}] \quad (3)$$

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

$$\hat{\pi}_{min-dist,t+j} = \min \{ (\hat{\pi}_{hybrid,t+j} - \hat{\pi}_{average,t+j}), (\hat{\pi}_{dynamic,t+j} - \hat{\pi}_{average,t+j}) \} \quad (4)$$

The implementation of neural and quasi-neural network models (1), (2), (3) and (4), require the following steps. First, to identify the variables, which help forecast the target variable and to process the input data. Second, to layer 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 steps are training the network followed by the actual 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 Results⁵

We use monthly year-on-year inflation rates for 28 OECD countries⁶ based on IFS' consumer price index series from July-1991 to June 2008. We trained both neural and quasi-neural algorithms in MATLAB.⁷ Initially, we normalize the data to bring it within the unit circle using:

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

The normalized data is used as input of neural algorithms and hence transfers training function by using specified trigonometric function. A downside of such normalization is that it restricts forecast values to remain inside the unit circle ruling out predictions for unusual inflation movements. A noteworthy implication is that forecasters of developing countries, where inflation tends to be relatively more volatile, have therefore got to be cautious whilst interpreting ANN generated inflation-forecasts for their countries.⁸

The Matlab neural network toolkit procedure 'trainlm' is extended with our specific neural algorithms provided in (1), (2), (3) and (4) to train the data. To validate our results we compare

⁵Detailed results and codes are available upon request.

⁶The countries are: Austria, Belgium, Canada, Czech Republic, Denmark, Finland, France, Germany, Greece Hungary, Iceland, Ireland, Italy, Japan, Korea, Luxembourg, Mexico, Netherlands, Norway, Poland, Slovak Republic, Spain, Sweden, Switzerland, Turkey, United Kingdom and United States.

⁷In 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.

⁸We thank the Referee for raising this point.

the forecast from our ANN models with that of a simple AR1 regressions. Our Matlab codes report the procedure that minimizes the RMSE and MAPE for each country and forecasting horizons. In cases where the RMSE (or MAPE) is numerically similar for two or more forecasting procedures for a given country and time horizon, our codes identify and report all those techniques. The entire set of results are lengthy and tedious.⁹ Instead, in Table 1 below we provide a summary of the success, in terms of RMSE and MAPE criterion, of each procedure in forecasting one-, three- and twelve-steps ahead out-of-sample forecasts for July 07 to June 08 for 28 countries.

Table 1: Out-of-sample forecast comparisons

	1 Step		3 Steps		12 Steps	
	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE
1. AR1	43	71	43	50	29	25
2. Hybrid, Eq. (1)	43	68	57	57	18	18
3. Dynamic, Eq. (2)	25	39	25	32	14	32
4. Quasi-Avg, Eq. (3)	04	36	18	32	21	21
5. Quasi Min-Dist, Eq. (4)	43	61	50	57	04	11
6. AR1 Only	29	21	25	18	25	21
7. Neural Only	43	29	57	50	71	71

Note: The percentage of countries for which a procedure minimizes either RMSE or MAPE

Rows one through five in Table 1 show the percentage of countries for which the forecast of a given procedure is as good as any other competing technique. For example, using the RMSE criterion for 1-step forecast the AR1 performs as good as any other competing procedure for 43% of countries in the sample. Rows six and seven are more focused in that they report the percentage of countries for which either the AR1 or any other neural network provided a clear superior forecast in terms of our criterion.

First, comparing rows six and seven we find that at the short-end of forecasts, i.e. 1- and 3-steps ahead, on average, for 45%¹⁰ of the countries neural networks are clearly a superior predictor using both the RMSE and the MAPE criterion. The AR1 process performs superior for only 23% of countries. Importantly, for the remaining 32% of countries either the AR1 or any other neural network technique perform equally well. Though these results lean towards a superior overall performance for neural networks, there exists a sizeable proportion of countries for which both types of procedure perform equally well. Consequently, it is not straight forward to rule in favor of one method over another. Second, continuing with the short-end of forecasts, the hybrid (1) and our newly developed quasi-minimum distance technique (4) dominate all other forms of forecasting but the AR1 technique is not far behind either. Finally, the last two columns show that there is no single technique that dominates long term forecasting; a result consistent with the literature.

⁹The authors will provide results by country upon request.

¹⁰This figure is obtained by simply averaging out row seven excluding the 12-step columns.

4 Conclusion

We show that, overall, neural network models and their combinations dominate the simple AR1 process for forecasting inflation rates in OECD countries over the short-term horizon. However, for a good number of countries the AR1 technique provided sound results as well. It may therefore always be advisable to continuously compare econometric procedures with that of neural networks when choosing a forecasting tool.

For Peer Review

References

- [1] Alvarez-Diaz, M. (2008) Exchange rates forecasting: local or global methods? *Applied Economics*, **40**, 1969 – 1984.
- [2] Balkin, S. D. (1997) Using recurrent neural network for Time series forecasting, Conference Paper, *International Symposium on forecasting*, Barbados.
- [3] Binner, J. M., Bissoondeal R. K. , Elger, T., Gazely, A. M., and Mullineux, A. W. (2005) A comparison of linear forecasting models and neural networks: an application to Euro inflation and Euro Divisia, *Applied Economics*, **37**, 665 – 680.
- [4] Binner, J. M. and Gazely A. M. (2000) The application of neural networks to the divisia index debate: evidence from three countries, *Applied Economics*, **32**, 1607 – 1615.
- [5] Brock, W. A. and Hommes, C. H. (1997) A rational route to randomness, *Econometrica*, **65**, 1059–1095.
- [6] Cameron, N. and Moshiri, S. (2000), Econometrics Versus ANN Models in Forecasting Inflation, *Journal of Forecasting*, , **19**, 201 – 217.
- [7] Cameron, N., Moshiri, S. and Scuse, D. (1999) Static, dynamic and hybrid neural networks in forecasting inflation, *Computational Economics*, **14**, 219 – 235.
- [8] Elman, J. L. (1990) Finding structure in time, *Cognitive Science*, **14**, 179-211.
- [9] Fernandez-Rodriguez, F., Gonzalez-Martel, C. and Sosvilla-Rivero, S. (2000) On the profitability of technical trading rules based on artificial neural networks: evidence from the madrid stock market, *Economics Letters*, **69**, 89–94.
- [10] Kuan, C. M. and Liu, T. (1995) Forecasting exchange rates using feedforward and recurrent neural networks, *Journal of Applied Econometrics*, **10**, 347 – 364.
- [11] Lin, C-T. and Yeh, H-Y. (2009) Empirical of the Taiwan stock index option price forecasting model - applied artificial neural network, *Applied Economics*, **41**, 285 – 293,.
- [12] 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.
- [13] McNelis, P. and McAdam, P. (2005) Forecasting inflation with thick models and neural networks, *Economic Modeling*, **22**, 548 – 567.
- [14] Nakamura, E. (2006) Inflation forecasting using a neural network, *Economics Letter*, **86**, 373-378.
- [15] Redenes, A. P. and White, H. (1998) Neural networks and financial economics, *International Journal of Forecasting*, **6**, 541—551.
- [16] Stock, J. H. and Watson, M. W. (1998) A comparison of linear and nonlinear univariate models for forecasting macroeconomic time Series, *NBER Working Paper 6607*.

- 1
2
3
4
5
6
7 [17] Swanson, N. R. and White H. (1997) A model selection approach to real-time macroeconomic forecasting using linear models and artificial neural networks, *The Review of Economics and Statistics*, **79**, 540–550.
8
9
10
11 [18] Xiaohong, C., Racin, J. and Swanson, N. (2001) Semiparametric ARX neural network models with an application to forecasting inflation, *IEEE Transactions on Neural Networks*,
12 **12**, 674–683.
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60