An evaluation of UK risky money: an artificial intelligence approach

Jane M. Binner*

Economics and Strategy Group
Aston Business School
Birmingham, B4 7ET, UK
E-mail: j.m.binner@aston.ac.uk
*Corresponding author

Alicia M. Gazely

Department of Information Management & Systems
The Nottingham Trent University
Nottingham, NG1 4BU, UK
E-mail: alicia.gazely@ntu.ac.uk

Graham Kendall

School of Computer Science
Jubilee Campus
The University of Nottingham
Nottingham, NG8 1BB, UK
E-mail: gxk@cs.nott.ac.uk

Abstract: In this paper we compare the performance of three indices in an inflation forecasting experiment. The evidence not only suggests that an evolved neural network is superior to *traditionally* trained networks in the majority of cases, but also that a risky money index performs at least as well as the Bank of England Divisia index when combined with interest rate information. Notably, the provision of long-term interest rates improves the out-of-sample forecasting performance of the Bank of England Divisia index in all cases examined.

Keywords: risky money; artificial intelligence: forecasting; neural networks: evolution strategies.

Biographical notes: Jane M. Binner applies traditional methods such as investment appraisal analysis and also state-of-the-art modelling methods such as dynamic game theory modelling. Her work centres on the application of advanced multivariate techniques to a range of data types, particularly the econometric and time series analysis of financial data. The work, which is

largely multidisciplinary in nature, involves both commercial applications and more academic approaches. She has over fifty refereed publications and three books. She has a successful track record in managing large international multidisciplinary projects. She serves regularly on programme committees for international artificial intelligence conferences.

Alicia M. Gazely graduated from the University of Manchester (UK) in 1970. She received her PhD in Ergonomics from the University of Loughborough in 1977 and qualified as a Chartered Accountant in 1979. She is currently a Senior Lecturer of Information Systems at the Nottingham Business School. Her research interests include the application of artificial intelligence techniques (specifically, neural networks) to economic and business problems.

Graham Kendall is a member of the Automated Scheduling, Optimisation and Planning Research Group at the University of Nottingham, UK. He is a member of the UK Engineering and Physical Sciences Research Council (EPSRC) Peer Review College and an Associate Editor of six international journals. He also chairs the steering committee of the *Multidisciplinary International Conference on Scheduling: Theory and Applications* (MISTA) and has chaired several other international conferences. He has been a member of the programme (or refereeing) committees of over 70 international conferences over the last few years. During his career, he has edited/authored ten books and has published over 80 refereed papers. He has been awarded externally funded grants worth over £5 million from a variety of sources, including EPSRC and commercial organisations. His research interests include adaptive learning (with an emphasis on game intelligence), heuristic development, optimisation, scheduling and artificial intelligence.

1   Introduction

The objective of current monetary policy is to identify indicators of macroeconomic conditions that will alert policy makers to impending inflationary pressures sufficiently early to allow the necessary action to be taken to control and remedy the problem. Owing to lags in the effect of monetary policy this requires a forecast of inflation up to 24 months ahead and the Bank of England provides a regular review of the prospects for inflation in the form of a quarterly 'Inflation Report', which includes quantified forecasts for inflation. The current approach to inflation forecasting utilises a range of indicators of inflationary pressures such as money supply, interest rates, exchange rates and the situation in the labour market. The role of money as a policy control tool has thus been relegated from the role of the main tool of macroeconomic control to being just one of many leading indicators of inflation selected from a range of key macroeconomic indicators. However, Nelson (2003) makes a strong case for the monetary aggregates within modern macroeconomic models. He contends that Friedman's famous proposition "inflation is always and everywhere a monetary phenomenon" remains valid in present-day models and can be given a similar interpretation irrespective of the monetary policy regime. Further, he states (p.1045) that the long-run relationship between money and inflation is:

> "of relevance to central banks in considering the implications of their policies for the steady-state rate of inflation; and this long-run relationship does deserve special attention as it has no direct counterpart in the key equations governing inflation dynamics."

The Divisia approach to monetary construction propounded by Barnett (1980; 1982) attempts to allow for the varying transactions and liquidity properties of the components of a monetary aggregate by giving them different weights. The share weight of each component depends on its size relative to the other components of the monetary aggregate and on its user cost. Consequently, currency and non-interest bearing deposits receive the highest weights because they are highly liquid assets and (correspondingly) have high user costs (in terms of interest foregone). Interest-bearing time deposits by contrast pay a relatively high rate of interest and are less liquid, and so attract a lower weight than might be expected from the size of such deposits alone. To the extent that these weights reflect the differences in transaction services provided by various monetary assets, the resulting Divisia index should be more closely related to the overall level of liquidity in the economy (and therefore to total spending in the economy) than conventional monetary aggregates. Evidence is accumulating from a wide range of empirical studies around the world to demonstrate that broad Divisia weighted monetary aggregates are superior to their simple sum monetary counterparts: see, for example Belongia (1996) for the USA, Herrman *et al*. (2000) for Germany and Ishida (1984) for Japan. In keeping with Binner *et al*. (2005) and Binner and Wattam (2003) the main stimulus for the current work arises directly from the conclusions reached in the Bank of England study, Fisher *et al*. (1993) that:

> "A Divisia measure of money appears to have some leading indicator properties for predicting both nominal output and inflation ..... a case can clearly be made for including Divisia in the range of indicators analysed by the authorities when forming their judgements on monetary conditions."

This finding was corroborated by Drake *et al*. (2000) who concluded, "the Divisia aggregate has better leading indicator properties of inflation than does simple sum M4".

The construction of Divisia money in the UK has been the topic of recent Bank of England research; see Hancock (2005), and remains a subject of ongoing debate. In keeping with the recent work by Drake and Mills (2005) for the USA, we are interested in incorporating financial innovations and changes to wealth holder preferences into the monetary aggregate. The acceleration of financial innovation has lead economists to question whether the boundary between financial assets such as bonds, shares or mutual funds and monetary assets has been eroded to the point where it is now non-existent. It has been suggested that it might be appropriate to include such risky assets into monetary aggregates (see for example Barnett and Zhou (1995) and Drake *et al*. (1999)), although the substitutability between risky and capital certain assets decreases as the level of risk aversion increases. Interest rate risk arises in the construction of a weighted monetary index such as Divisia due to the fact that the interest rates used as weights are not known until the end of each period. Under risk neutrality, this would not constitute a problem. Consumers are, however, expected to be risk averse. Thus interest rate risk in combination with risk aversion complicates the construction of Divisia aggregates. However, Barnett and Liu (1995), Barnett *et al*. (1997) and Barnett and Liu (2000) have developed a method to account for risk in combination with risk aversion within the Divisia index construction utilising results from the consumption capital asset pricing

model in finance. The equations for constructing the risk adjustment needed for the returns on each asset are provided in Barnett *et al*. (1997) along with a demonstration that the new risk adjusted Divisia exactly tracks the underlying aggregator function under risk aversion. Since that extension is based upon the Consumption Capital-Asset-Pricing Model (CCAPM), the extension is subject to the 'equity premium puzzle' problem of smaller than necessary adjustment for risk. More recently, Barnett and Wu (2005) have extended Barnett *et al*.'s result to the case of risk aversion with intertemporally nonseparable tastes.

Drake *et al*. (1998) studied the leading indicator properties of various UK simple sum, Divisia and risky Divisia monetary aggregates over the time period 1979Q1 to 1994Q2. They found that using risky aggregates offers an improvement over both the simple sum and the standard Divisia monetary aggregates. Elger and Binner (2004) is devoted solely to the construction of the new risky Divisia aggregate and they evaluate the econometric performance using a money demand system that includes the UK household sector holdings of risky money over the period 1980Q1 to 1999Q1. They found that the risky aggregate performed in accordance with economic theory, producing well specified, remarkably stable long run demand for money functions.

The novelty of this paper lies in the use of evolved neural networks utilising evolution strategies to examine the UK's recent experience of inflation. To our knowledge this is the first time that this methodology has been investigated in this domain. Our preference for evolving neural networks, utilising evolution strategies as the mechanism to adjust the weights, is due to their success in other domains. For example, the work of Chellapilla and Fogel in evolving a world class checkers player, see Chellapilla and Fogel (2001) and Fogel (2002). The use of the approach is further motivated by the fact that we can evolve a network that searches for the correct set of inputs and a good set of network weights rather than having to hand tune the parameters, which is the case when using an approach such as back propagation. Whilst a single evolution run can take significantly longer than a single back propagation run, this is balanced by the fact that we only carry out a single run of the evolution method but may require significant tuning when using other approaches. In addition, the evolutionary method may work for a variety of problem instances whereas other approaches may need tailoring for each individual problem.

The artificial neural network technology is an unusual tool in this context, although the application of neural networks in the field of economics is growing in popularity, as indicated by the diverse range of applications surveyed in Zhang *et al*. (1998) and more recently in Binner *et al*. (2004). Neural networks allow approximation of highly non-linear functions and so offer more promise in the context of econometric modelling than standard linear models, especially since there is no requirement to specify regression parameters and assumptions about data distribution are less rigorous. Where time series such as those in the present study are concerned, neural networks are limited by the shortage of data points on which to train the network. However, promising results in earlier neural network studies have encouraged us to believe that the technique holds great potential and that exploratory studies such as this one are worthwhile; see for example, Binner *et al*. (2002).

This approach is a unique tool in the context of predicting inflation and its use is highly exploratory although results presented here give us confidence to believe that significant advances in macroeconomic forecasting and policy making are possible using techniques such as this. This is the first attempt to evaluate the empirical performance of a risk-adjusted weighted index number measure of money using such an approach.

This application of artificial intelligence techniques to explore the money – inflation link is highly experimental in nature and the overriding feature of this research is very much one of simplicity. It is virtually certain in this context that more accurate inflation forecasting models could be achieved with the inclusion of additional explanatory variables, particularly those currently used by monetary authorities around the world as leading indicator components of inflation.

2   Data and forecasting model

In this study the performance of a UK risky monetary index is compared with that of the two capital certain indices, simple sum and Divisia. The effect of including information on interest rates is also studied, as the Bank seeks to meet the Governments inflation target through the decisions on interest rates taken by the Monetary Policy Committee.

The household sector holdings of the M4 components (£ millions) included in the official Bank of England index and their returns were obtained from the Bank of England Monetary and Financial Statistics division website. Following Drake *et al*. (1999), the standard Bank of England Divisia components are assumed to be capital certain.

These assets are; Notes and Coins (N/C), Non-Interest Bearing Deposits (NIBD), Interest Bearing Bank Sight Deposits (IBSD), Interest Bearing Bank Time Deposits (IBTD) and Building Society Deposits (BSD). This data set is described in detail in Fisher *et al*. (1993). Recent studies have extended the list of capital certain assets to include National Savings (NS) and certificates of deposit, see for example, Chrystal and Drake (1997). In the construction of the capital certain indices used in this study, only NS deposits are added to the assets used by the Bank of England. The reason is that a consistent NS series is available for the entire sample, while certificates of deposits are only available from 1986. Personal sector holdings of NS have been downloaded from DataStream and the Office for National Savings has contributed by supplying the deposit rates. The benchmark rate for the capital certain index was constructed by adding 2.8 percentage points to the three-month local authority rate. This figure is 0.8 percentage points above that used by the Bank of England in order to avoid negative user costs when national savings are included.

The risky index used in the study is constructed by extending the list of assets in the capital certain index to include bonds, shares and unit trusts. Data on the household sector holdings of risky assets, equity (UK quoted shares), bonds (UK government bonds) and unit trusts at market values (£ millions), were obtained from DataStream. Quarterly returns on equities and bonds were calculated from the Financial Times all-share index and the Financial Times Actuaries government bond price index respectively. The return on unit trusts was calculated as an average of quarterly returns on all unit trusts. The returns on the risky assets are quarterly returns, whilst the returns on the capital certain assets are annualised yearly returns reported on a quarterly basis. The index construction is not dependent upon whether quarterly or yearly interest rates are used, although all returns need to be measured on the same basis. Thus, the yearly rates of return on the capital certain assets have been transformed to quarterly returns.

Studies have shown that returns on risky assets are autocorrelated. Hence, it may be possible to forecast these returns using an AR-model see, for example, Drake *et al.* (1999), Lo and MacKinlay (1988) and Poterba and Summers (1986). In the current work, the returns on all three assets are modelled as AR(2)-processes. One step ahead forecasts are then used to construct the actual interest rate forecasts used in the index calculation.

It is well known that empirical estimates often yield unreasonably high estimates of the degree of relative risk aversion. In the famous 'equity premium puzzle' research undertaken by Mehra and Prescott (1985), the estimated degree of relative risk aversion was as high as 25. Since it is difficult to find estimates of the degree of relative risk aversion that make sense economically, it has become common to choose a value directly. Drake *et al.* (1999) assume values in the range 0 to 7. In the absence of better estimates, the mid-value 3.5 from this study was chosen. It should, however, be noted that the importance of the chosen value for the coefficient of relative risk aversion is marginal depending on the very low correlations between the real rates of the risky assets and the growth rate of real consumption. In this study, the estimated covariances are 0.00012272, 0.00016532 and 0.00018979 for bonds, shares and unit trusts respectively. The chosen measure of consumption is real Private Final Consumption Expenditure (PFCE). In some cases, the estimated covariance between the real return on bonds and growth in real consumption was found to be negative. This implies that households can diversify consumption risk by holding risky assets, which is counterintuitive Barnett and Liu (2000). For the risky index, the benchmark rate used in calculating the user costs was constructed using an envelope approach, $R_t^* = \max (E_t r_{it}^* - \phi_{it}, i = 1 \ldots l)$. The benchmark rate, R_t^* is calculated from the maximum of the risk adjusted expected own rates. Following Anderson *et al.* (1997) the monetary asset quantities have been seasonally adjusted using the default values in the Bureau of Census X11 programme. A break adjustment was also made to levels in the risky index, to deal with the extraordinary volatility in stock market returns that occurred in 1987 ('Black Monday') without changing the actual model used to forecast the interest rates.

Monetary data was available for the period 1980Q1 to 1999Q4. The starting point is restricted by the fact that the personal sector data for the risky assets is only available from 1980 onwards. Of the total quarterly data points available, after loss of data points due to the time lags implicit in our model of up to four quarters, 77 quarters remained, meaning that the period covered in terms of the inflation rate being predicted is 1981Q1 to 2000Q1, or just over 19 years.

The RPI price series was obtained from DataStream. Inflation was constructed for each quarter as year-on-year growth rates of prices.

The choice of interest rates is a complex matter and the rates used are often based upon arbitrary assumptions. Considering the fact that most recent simple sum money demand studies contain one short and one long interest rate, one might ask what really is the own-rate of the monetary assets and what is the opportunity cost? For the short rate, it has become common to use an arbitrary average of deposit rates or a short-term interest rate, and here, following Elger and Binner (2004), the 30-day interest rate was chosen as a proxy for the own-rate since it appeared to be sufficiently short to be strongly correlated with the returns on all capital certain assets. It is also common to use a five or ten year interest rate as a proxy for the long rate and hence we chose the ten year rate, though since the yield curve tends to flatten out over longer maturities, the exact choice of maturity for the long rate is less important. Seasonally unadjusted quarterly interest rate series were obtained from DataStream.

2.1 The model

A simple model of the relationship between money and inflation was employed as the basis for our work (see Equation 1 below) that takes inflation, Π_t, in the current quarter to be a function of money measures in the four preceding quarters, where M_t represents money. The model also includes an autoregressive term, representing inflation for the preceding period. This model is the preferred specification in our earlier neural network studies; see for example Binner *et al.* (2002), since it consistently outperforms even simpler model constructs, although of course in the same vein, we recognise that a more complex model design will almost certainly yield superior results. A final variable was included to represent time (here taking a value between 1 and 77) that allows for the possibility that external factors, not catered for by the autoregressive variable, might affect the inflation rate.

$$\Pi_t = F(M_{t-1}, M_{t-2}, M_{t-3}, M_{t-4}, \Pi_{t-1}). \tag{1}$$

The simple model above was tested in four variations; monetary index alone; short interest rate added; long interest rate added; both long and short interest rates added. Interest rates were also lagged for four quarters to be consistent with the monetary indices. Given three monetary indices, the result was 12 variations to be tested using both the trained and the evolved neural network models employed in this study.

3 Trained neural network methodology

Neural networks are constructions made up of many relatively simple, interconnected processors. Establishing the architecture for the neural network is analogous to curve fitting, Refenes and Azema-Barac (1994), in that choosing the number of hidden layers and hidden neurons is like choosing the order of a polynomial. Choosing a lower order of polynomial than required leads to a poor fit with the data (the network fails to converge for the training data and prediction for new data is poor). Choosing a higher order than required leads to a good or perfect fit to the data (overfitting to the training data) but poor prediction for new data (poor generalisation).

The advantage of a neural network methodology for this investigation, however, is that neural networks are inductive. Thus, even when there is no exact knowledge of the rules determining the features of a given phenomenon, knowledge of empirical regularities can still allow the phenomenon to be modelled and this is the strength of the neural network. The process of training allows the network, in effect, to ignore excess input variables. The technique seems ideal for economic phenomena where the central task is to model a system of immense complexity without losing predictive power. It seems, however, that the potential of non-traditional techniques such as the neural network has not been exploited.

A neural architecture of four hidden units in a standard back-propagation network was employed throughout. Preliminary experiments with 4, 5 and 6 hidden units confirmed findings from our earlier work, that a larger network gives no advantage and can lead to over-specialisation. These experiments also led to the surprising conclusion that four hidden units is the optimum size of the network for this data for all model variations, even though some contain many more input variables than others. The

smallest network had six inputs (four lags of monetary index plus the autoregressive term and the time variable) and the largest had 14 inputs (as above plus four lags for each of the two interest variables).

A policy of deliberate extended training was used Masters (1993) rather than risk the unreliability of premature stopping techniques since this policy had proved successful in earlier studies. All networks were trained for 100 000 epochs. Seventy-seven quarters of data were available, of which the first sixty-seven were used for training and the last ten for testing (forecasting). This proportion of training to testing is higher than that conventionally used for neural networks, but at this exploratory stage we are primarily interested in the ability of the network to model the data as a precursor to predictive ability, rather than focusing exclusively on predictive accuracy *per se*.

4   Evolution strategies

Evolution Strategies (ES) are closely related to genetic algorithms, with (possibly) some of the same inherent difficulties Szipro (2002) and Horn and Goldberg (1994). Originally ES's used only mutation and a population of a single individual and were used to optimise real valued variables. More recently, ES's have used a population size greater then one, they have used crossover and have also been applied to discrete variables, see Back *et al*. (1991). However, their main use is still in finding values for real variables by a process of mutation, rather than crossover.

An individual in an ES is represented as a pair of real vectors, $v = (x, \sigma)$. The first vector, x, represents a point in the search space and consists of a number of real valued variables. The second vector, σ, represents a vector of standard deviations.

Mutation is performed by replacing x by:

$$x^{t+1} = x^t + N(0, \sigma) \tag{2}$$

where $N(0, \sigma)$ is a random Gaussian number with a mean of zero and a standard deviation of σ. This mimics the evolutionary process that small changes occur more often than larger ones.

In evolutionary computation there are two variations with regard to how the new generation is formed. The first, termed $(\mu + \lambda)$, uses μ parents and creates λ offspring. Therefore, after mutation, there will be $\mu + \lambda$ members in the population. All these solutions compete for survival, with the μ best selected as parents for the next generation. An alternative scheme, termed (μ, λ), works by the μ parents producing λ offspring (where $\lambda > \mu$). Only the λ compete for survival. Thus, the parents are completely replaced at each new generation. Or, to put it another way, a single solution only has a life span of a single generation. In this work, we use a 10 + 10 strategy.

5   Evolved neural network methodology

The artificial neural networks we utilise are feed forward networks that have their weights adapted using evolution strategies, rather than a supervised learning technique such as back propagation. The networks comprise an input layer, with the neurons within that layer using an identify function. A hidden layer with a varying number of neurons is used, with the number of nodes being a matter of experimentation. The network has a

single output neuron, which is the predicted inflation rate. The hidden and output neurons use a non-linear activation function, which, in these experiments is either sigmoid (Equation 3) or tanh (Equation 4). In Equations (3) and (4), *x* is the input value to the given neuron:

$$f(x) = 1/(1 + exp(-x)) \tag{3}$$

$$f(x) = 1/(1 + exp(x)) - 1. \tag{4}$$

In the initial experiments a population of 20 neural networks are created. The networks have a fixed structure and the population is homogeneous. The value for σ, for each neural network, is initially set to 0.05.

Each neural network is presented with the entire set of training data (67 samples) and the output for each sample is compared against the actual inflation rate for that quarter. The fitness for a particular neural network is then given by:

$$\sum_{i=1 \ to \ 67}(t_i - o_i)^2 \tag{5}$$

where t = the training value presented to the network and o = the actual output from the network.

The aim is to minimise Equation (5) and the network that produces the minimum value is the one that is used on the validation set (10 samples).

Once all 20 neural networks have had the training data presented, they are sorted based on their fitness (Equation 5) and the best ten are selected to progress to the next generation (the worse performing networks die off). These are copied (in order to maintain a population size of 20) and the copied networks have their weights mutated in order to try and improve their predictive ability. This is done using Equation 2. In addition, the value of σ is also mutated using Equation 6:

$$\sigma' = \sigma * \exp(\tau * N(0, 1)) \tag{6}$$

where $\tau = (2(N_W)^{0.5})^{-0.5}$ and N_W = the number of weights in the neural network.

This process continues for 1 000 000 iterations. This figure was chosen as a good trade off between time taken (about 1 h to run on a 1.5 GHz PC) and to give the possibility of a suitably good neural network evolving. There is a danger that we could 'overfit' the data, Lawrence *et al.* (1997). That is, the validation set can produce a larger error than the training data. We did carry out some experiments, in order to find out if this was the case, and we do not believe that this is happening.

As we were not, at this stage, evolving the structure of the network (this is ongoing work); just mutating the weights, we needed to conduct a series of experiments that tested various permutations of parameters. The experiments involved interchanging Risky Divisia M4, Bank of England Divisia M4 and Simple Sum M4 with no interest rate, the short interest rate, long interest rate and both short and long interest rates. We also experimented with different activation functions and a different number of hidden neurons. Each experiment was run twice. The various experimental parameters are summarised below:

- Measures: {Risky M4, BoE Divisia M4, Simple Sum M4} = |3|

- Population Size of Networks: 20

- Iterations: 1 000 000

- Networks Retained and Mutated: 10

- Input Neurons: {5,9,9,13} = |4|

- Hidden Neurons: {3,4,5} = |3|

- Activation fn (hidden): {sigmoid, tanh} = |2|

- Activation fn (input/output): identity

- Results averaged over 2 runs = |2|.

Therefore, the total number of experiments, which took about an hour each, was 144 (3 * 4 * 3 * 2 * 2). In fact, we carried out some initial tests, using a larger number of hidden neurons, but these gave similar (or worse) results than using 3, 4 or 5. We also found that the activation function did not really affect the result and in future work we have just used the sigmoid function.

6   Results

Tables 1A and 1B shows the results from our two alternative experiments (trained network vs. evolved network). The results are divided between in-sample (the training set) and out-of-sample (the validation set). Within these two categories, three standard forecasting evaluation measures were used to compare the predicted inflation rate with the actual inflation rate, namely, Root Mean Squared Error (RMSE), Mean Absolute Difference (MAD) and Mean Absolute Percent Error (MAPE). The experiments comprise the four lagged measures of money and the previous quarter's inflation rate.

In the case of the trained neural networks, Table 1A, looking first at MAPE, within sample results are reasonable, and there is not a great deal to choose between the indices. The best 'fit' is Simple Sum with both types of interest variable included at 4.2%, but this condition also has the worst out-of-sample result.

Out-of-sample results for ten quarters (2.5 years) are also reasonable. The best in terms of MAPE is BoE Divisia with long-term interest rates (9.2%) (see Figure 1); the next best is risky, also with long-term rates (10.0%). Simple Sum does poorly out-of-sample across the board. Looking at RMSE and MAD instead, the two results above are reversed so that risky does slightly better. These two results are also clearly better than for any other condition.

Comparing the three monetary indices, simple sum performs poorly in all conditions but particularly badly where information about both types of interest rate is provided. The training error is small for this network and so the poor out-of-sample performance may be an artefact of over fitting, and BoE Divisia is also affected in this way. Since the networks in this condition are really small compared to the number of variables (4 hidden units for 14 variables) this is a strange result. Comparing the performance of the risky index with that of the BoE index, in both cases there is one condition where the index performs poorly (risky in the 'none' condition and BoE in the 'both' condition) but both indices do well when information on long term interest rates (alone) is included. This may be a result of the high statistical correlation between the series, –0.83 and –0.82 for the risky money and traditional Bank of England Divisia respectively. Also, there is the simple explanation that the neural network favours the stability of the long term interest rates in preference to the more volatile short term rates.

Table 1A Trained neural network results

	Risky interest				Bank of England Divisia interest				Simple sum interest			
In-sample	*Both*	*Long*	*Short*	*None*	*Both*	*Long*	*Short*	*None*	*Both*	*Long*	*Short*	*None*
RMSE	0.00325	0.00433	0.00506	0.00555	0.00198	0.00421	0.00495	0.00848	0.00183	0.00433	0.00528	0.00580
MAD	0.0026	0.0035	0.0038	0.0043	0.0015	0.0033	0.0039	0.0070	0.0015	0.0034	0.0041	0.0048
MAPE	7.1%	8.9%	12.5%	10.1%	4.5%	8.6%	10.3%	17.1%	4.2%	9.6%	12.1%	13.0%
Out-of-sample												
RMSE	0.00649	0.00315	0.01220	0.01820	0.01887	0.00360	0.00841	0.00749	0.02187	0.00982	0.00593	0.00925
MAD	0.0059	0.0025	0.0079	0.0138	0.0160	0.0026	0.0074	0.0063	0.0173	0.0074	0.0053	0.0076
MAPE	22.6%	10.0%	29.1%	53.6%	61.1%	9.2%	28.7%	24.1%	72.8%	29.0%	20.3%	27.8%

Table 1B Evolved neural network results (averages, 2 runs)

|	Risky interest				Bank of England Divisia interest				Simple sum interest			
In-sample	Both	Long	Short	None	Both	Long	Short	None	Both	Long	Short	None
RMSE	0.00680	0.00706	0.00723	0.00764	0.00687	0.00690	0.00706	0.00739	0.00703	0.00715	0.00731	0.00748
MAD	0.0053	0.0057	0.0057	0.0062	0.0054	0.0055	0.0057	0.0059	0.0054	0.0056	0.0058	0.0060
MAPE	14.8%	15.5%	16.6%	17.9%	15.2%	15.1%	16.5%	17.4%	15.4%	15.5%	16.7%	17.2%
Out-of-sample												
RMSE	0.01330	0.00693	0.00765	0.00828	0.00838	0.00689	0.00813	0.00737	0.00966	0.00769	0.00698	0.00760
MAD	0.0122	0.0063	0.0066	0.0070	0.0073	0.0061	0.0069	0.0063	0.0089	0.0068	0.0060	0.00682
MAPE	45.7%	22.9%	24.5%	25.8%	28.2%	22.7%	26.5%	24.5%	33.9%	25.0%	23.3%	26.2%

Figure 1   Out-of sample forecast of best fitting trained neural network model (see online version for colours)
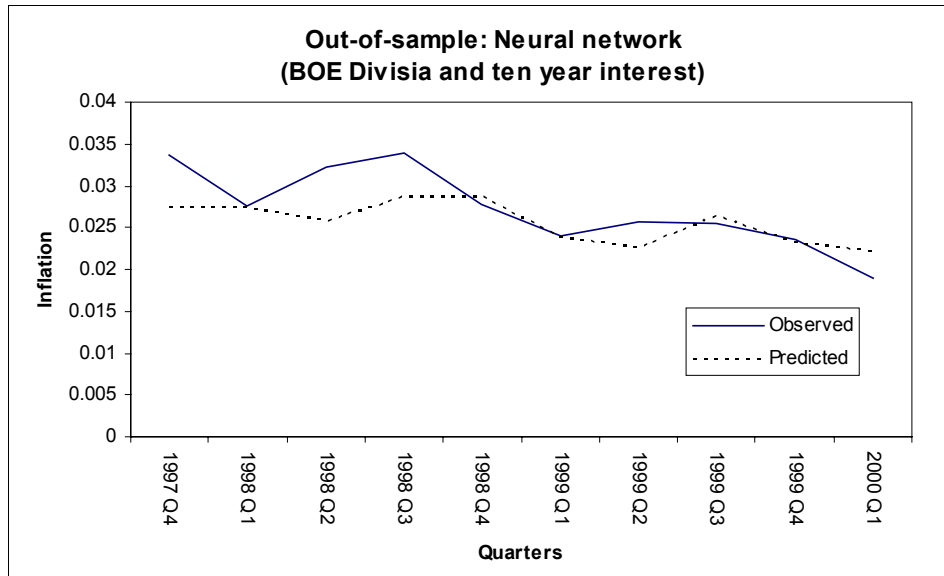
Results confirm that long term (ten year) interest rates appear to be helpful in predicting inflation in conjunction with the risky and Bank of England indices, but not with Simple Sum. In fact, the best neural network result achieved does indeed combine the risky Divisia with long term interest rates using the RMSE criterion. Short term rates do not seem to be helpful, despite the fact that they are used by government to try to control inflation. A possible, simplistic, explanation for this finding is that, while short term interest rates are mainly determined by central bank's policy actions, long term rates reflect market beliefs of how inflation will develop.

It is also interesting to note that a similar pattern emerges when the evolved neural networks are considered. Table 1B shows the results obtained when averages are taken across the set of four experiments, *i.e.*, those having no interest rates inputted, those with the short interest as input, those with the long interest rate as input and those using both long and short rates. It is again interesting to note that, the best out-of-sample results were achieved, on average, when long interest rates were combined with the Bank of England Divisia and with the risky measure of money, see Figure 2.

A comparison of the trained neural network results (Table 1A) with the evolved results (Table 1B) reveals that, while the neural network consistently provides a better within-sample fit, the process of evolving neural networks using evolution strategies yields more accurate out-of-sample results in 8 out of the 12 cases considered. On average, the evolving strategy RMSE results are 21% lower than those produced by trained neural networks.

Figure 2   Out-of sample forecast of best fitting evolved neural network model (see online version for colours)
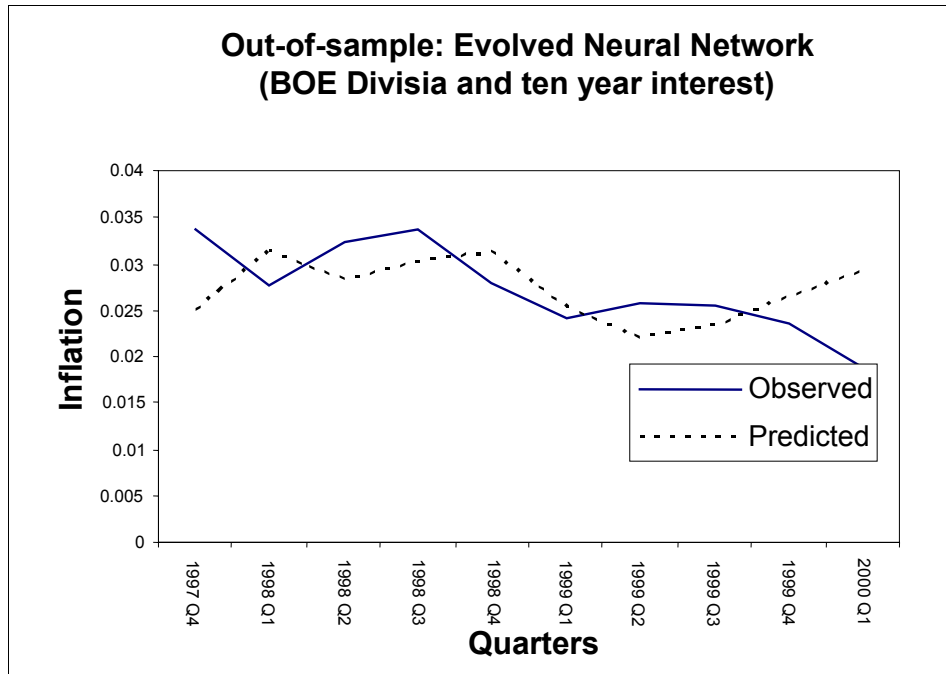
7   Concluding remarks

Evolving neural networks using evolution strategies have been found to compete very favourably with trained neural networks and have the potential to beat neural networks in terms of superior predictive performance. The evolution strategies presented here have the added advantage that we only need to carry out a single run whilst other approaches require significant tuning. In addition, the evolutionary method may work for a variety of problem instances whereas other approaches may need tailoring for each individual problem Artificial Intelligence techniques in general and evolution in particular are highly effective tools for predicting future movements in inflation. There is tremendous scope for further research into the development of these methods as new macroeconomic forecasting models.

This study compares the empirical performance of three different monetary services indices within a simple macroeconomic framework using the novel approach in the context of artificial neural networks. A capital certain monetary services index is constructed using the same components included in the Bank of England's monetary services index plus national savings. A risky index is constructed by adding bonds, shares and unit trusts to the list of assets included in the capital certain index. Including such assets is motivated both by financial innovations and recent theoretical developments in the field of monetary aggregation theory. With the acceleration of financial innovations, it is legitimate to ask if the frontier between financial assets such as bonds, shares or mutual funds and capital certain assets has become non-existent. Theoretical

developments have addressed the problem that the returns on many assets, in particular risky assets, are not known until the end of each period with certainty. In particular, a method has been developed to explicitly adjust the user costs used in the calculation of the index based upon results from the Consumption CAPM (CCAPM) framework.

Risk-adjusted Divisia's are not estimation free. Under a set of CCAPM special case assumptions, the need for estimation reduces to making forecasts of returns and estimating an Arrow-Pratt measure of relative risk aversion. In empirical work, this may not constitute a major obstacle, at least from a practical point of view. The researcher can simply choose some method to obtain forecasts of returns along with some reasonable risk-aversion coefficient. Our own experimentation using alternative models to forecast returns revealed no dramatic changes in the aggregate risky Divisia depending upon the model used to obtain the forecasts. It should be noted that the risk adjustments obtained in this current work are much smaller than one would expect *a priori*.

Since the inclusion of risky assets into monetary services indices is a new research field, there is tremendous scope for future research in this area. This paper does not examine the ability of such an index to forecast changes in output or the price level from a macroeconomic policy viewpoint; although recent evidence has been provided by Schunk (2001) using out-of-sample forecasting techniques on US data. Similarly, causal relationships between changes in money and changes in income and the price level are not investigated here. An empirical application of causality tests in a multivariate cointegration framework based on monetary services indices is presented in Funding la Cour (2005). She also develops the idea of encompassing in this multivariate cointegration framework. It may well be the case that the information content of the risky index is greater than the information content in the capital certain index, though the present results do not provide strong evidence one way or the other. This conclusion is reinforced by the difficulties faced in the construction of a risky index in order to account for interest rate risk in combination with risk aversion. Specifically, two problems are identified in the construction of the risky monetary services index. Firstly, the equity premium puzzle is strongly present in UK data, making the risk adjustment minute. Much work is needed to incorporate suggested solutions from the finance literature into the monetary aggregation literature. Secondly, the Black Monday in October 1987 caused a large shift in the risky index. Even though this might appear to be a big problem, the size of the break is largely determined by the model used to estimate the forecasts of real interest rates.

Evidence presented here suggests the new risky money index performs at least as well as the Bank of England Divisia index when used in combination with long interest rates. It seems from the present results that such information is indeed useful, even though monetary services aggregates already internalise interest rates in their construction. There is clear evidence of an interaction effect between the monetary indices and interest rate information. The improvement in prediction with the benefit of any interest rate information is particularly marked in the case of the risky index, since this index performs particularly poorly on its own. The results may demonstrate that forecasting in general can be improved by the inclusion of additional variables, even within a simple macroeconomic modelling framework. In any case, the question merits further investigation. Given the increased liquidity of many financial assets and the increased monetary services provided by broader assets not previously regarded as money, increased substitution between monetary and non-monetary assets is taking place. Central

banks run the risk that conventional broad monetary assets will become increasingly unstable unless such substitutions between official and unofficial monetary assets are internalised into the construction of money. The problem of destabilisation by substitutions between assets within the official aggregate and risky assets, such as unit trusts, was faced by the US authorities in the so-called 'missing M2' episode in the early 1990s. As Barnett (1995) has pointed out, however, such destabilising substitutions could be completely internalised by constructing a suitably wide aggregate using a procedure such as the risk-adjusted Divisia approach.

References

Anderson, R., Jones, B. and Nesmith, T. (1997) 'Building new monetary services indexes: concepts, data and methods', *Federal Reserve Bank of St. Louis Review*, January–February, pp.53–82.

Back, T., Hoffmeister, F. and Schwefel, H-P. (1991) 'A survey of evolution strategies', in R. Belew and L. Booker (Eds.) *Proceedings of the Fourth Conference on Genetic Algorithms*, San Diego, CA: Morgan Kaufman Publishers, pp.2–9.

Barnett, W.A. (1980) 'Economic monetary aggregates: an application of index number theory and aggregation theory', *Journal of Econometrics*, Vol. 14, pp.11–48, reprinted in W.A. Barnett and A. Serletis (Eds.) (2000) *The Theory of Monetary Aggregation*, Chap. 2, North Holland: Elsevier, pp.11–47.

Barnett, W.A. (1982) 'The optimal level of monetary aggregation', *Journal of Money, Credit and Banking*, Vol. 14, pp.687–710, reprinted in W.A. Barnett and A. Serletis (Eds.) (2000) *The Theory of Monetary Aggregation*, Chap. 7, North Holland: Elsevier, pp.125–148.

Barnett, W.A. (1995) 'Exact aggregation under risk', in W. Barnett, H. Moulin, M. Salles and N. Schofield (Eds.) *Social Choice, Welfare and Ethics*, Cambridge: Cambridge University Press, pp.353–374, reprinted in W.A. Barnett and A. Serletis (Eds.) (2000) *The Theory of Monetary Aggregation*, Chap. 10, Amsterdam, North Holland, pp.195–216.

Barnett, W.A. and Liu, Y. (1995) 'The CAPM extended Divisia monetary aggregate with exact tracking ability under risk', Working Paper No. 195, Washington University, St. Louis, Missouri, USA.

Barnett, W.A. and Liu, Y. (2000) 'Beyond the risk neutrality function', Chap. 1 in M.T. Belongia and J.M. Binner (Eds.) *Divisia Monetary Aggregates: Theory and Practice*, Basingstoke, UK: Palgrave, pp.11–27.

Barnett, W.A., Liu, Y. and Jensen, M. (1997) 'The CAPM risk adjustment for exact aggregation over financial assets', *Macroeconomic Dynamics*, Vol. 1, pp.485–512, reprinted in W. Barnett and A. Serletis (Eds.) (2000) *The Theory of Monetary Aggregation*, Chap. 12, Amsterdam, North Holland, pp.245–273.

Barnett, W.A. and Wu, S. (2005) 'On user costs of risky monetary assets', *Annals of Finance*, Vol. 1, No. 1, pp.35–50.

Barnett, W.A. and Zhou, G. (1995) 'Mutual funds and monetary aggregates: commentary', *Federal Reserve Bank of St. Louis Review*, Vol. 76, November–December, pp.53–62, reprinted in W. Barnett and A. Serletis (Eds.) (2000) *The Theory of Monetary Aggregation*, Chap. 15, Amsterdam, North Holland, pp.307–322.

Belongia, M.T. (1996) 'Measurement matters: recent results from monetary economics re-examined', *Journal of Political Economy*, Vol. 104, pp.1065–1083.

Binner, J.M., Bissoondeeal, R. and Mullineux, A.W. (2005) *A Composite Leading Indicator of the Inflation Cycle for the Euro Area Applied Economics*, Vol. 37, pp.1257–1266.

Binner, J.M., Gazely, A.M. and Chen, S.H. (2002) 'Financial innovation and Divisia monetary indices in Taiwan: a neural network approach', *The European Journal of Finance*, Vol. 8, pp.238–247.

Binner, J.M., Kendall, G. and Chen, S.H. (2004) 'Applications of artificial intelligence in finance and economics', *Advances in Econometrics*, Vol. 19.

Binner, J.M. and Wattam, S.I. (2003) 'A new composite leading indicator of inflation for the UK', *Global Business and Economics Review*, Vol. 5, No. 2, pp.242–264.

Chellapilla, K. and Fogel, D.B. (2001) 'Evolving an expert checkers playing program without using human expertise', *IEEE Transactions on Evolutionary Computation*, Vol. 5, No. 4, pp.422–428.

Chrystal, K.A. and Drake, L. (1997) 'Personal sector money demand in the UK', *Oxford Economic Studies*, Vol. 59, pp.188–206.

Drake, L., Chrystal, K.A. and Binner, J.M. (2000) 'Weighted monetary aggregates for the UK', Chap. 3 in M.T. Belongia and J.M. Binner (Eds.) *Divisia Monetary Aggregates; Theory and Practice*, Basingstoke, UK: Palgrave, pp.47–78.

Drake, L., Fleissig, A.R. and Mullineux, A.W. (1999) 'Are risky assets substitutes for monetary assets?', *Economic Inquiry*, Vol. 37, No. 3, pp.510–526.

Drake, L. and Mills, T.C. (2005) 'A new empirically weighted monetary aggregate for the United States', *Economic Inquiry*, Vol. 43, No. 1, pp.138–157.

Drake, L., Mullineux, A.W. and Agung, J. (1998) 'Incorporating risky assets in Divisia monetary aggregates', Economics Research Paper, No. 98/15, Department of Economics, Loughborough University, UK.

Elger, T. and Binner, J.M. (2004) 'The UK household sector demand for risky money', *Topics in Macroeconomics*, UC Berkeley Electronic Press, http://www.bepress.com/bejm/topics/vol4/iss1/art3.

Fisher, P., Hudson, S. and Pradhan, M. (1993) 'Divisia indices for money: an appraisal of theory and practice', Bank of England Working Paper No. 9, April.

Fogel, D.B. (2002) *Blondie24: Playing at the Edge of AI*, San Francisco, CA: Morgan Kaufmann.

Funding la Cour, L. (2005) 'The problem of measuring "money": results from an analysis of Divisia monetary aggregates for Denmark', Chap. 8 in M.T. Belongia and J.M. Binner (Eds.) *Money, Measurement and Computation*, Basingstoke, UK: Palgrave.

Hancock, M. (2005) 'Divisia money', *Bank of England Quarterly Bulletin*, Spring 2005, pp.39–46.

Herrmann, H., Reimers, H.E. and Toedter, K.H. (2000) 'Weighted monetary aggregates for Germany', Chap. 4 in M.T. Belongia and J.M. Binner (Eds.) *Divisia Monetary Aggregates; Theory and Practice*, Basingstoke, UK: Palgrave, pp.79–101.

Horn, J. and Goldberg, D. (1994) 'Genetic algorithm difficulty and the modality of fitness landscapes', *Proceedings of Foundations of Genetic Algorithms (FOGA)*, Workshop held 30 July – 2 August 1994, Colarado, USA, Vol. 3.

Ishida, K. (1984) 'Divisia monetary aggregates and demand for money: a Japanese case', *Bank of Japan Monetary and Economic Studies*, June, pp.49–80.

Lawrence, S., Giles, C.L. and Tsoi, A.C. (1997) 'Lessons in neural network training: overfitting may be harder than expected', *Proceedings of the Fourteenth National Conference on Artificial Intelligence (AAAI-97)*, Menlo Park, CA: AAAI Press, pp.540–545.

Lo, A. and MacKinlay, A. (1988) 'Stock markets do not follow random walks: evidence from a simple specification test', *Review of Financial Studies*, Vol. 1, pp.41–66.

Masters, T. (1993) *Practical Neural Network Recipes in C++*, London, UK: Academic Press.

Mehra, R. and Prescott, E. (1985) 'The equity premium: a puzzle', *Journal of Monetary Economics*, Vol. 15, pp.145–161.

Nelson, E. (2003) 'The future of monetary aggregates in monetary policy analysis', *Journal of Monetary Economics*, Vol. 50, pp.1029–1059.

Poterba, J. and Summers, L. (1986) 'The persistence of volatility and stock market fluctuations', *American Economic Review*, Vol. 76, pp.1142–1151.

Refenes, A.N. and Azema-Barac, M. (1994) 'Neural network applications in financial asset management', *Neural Computing and Applications*, Vol. 2, pp.13–39.

Schunk, D. (2001) 'The relative forecasting performance of the Divisia and simple sum monetary aggregates', *Journal of Money, Credit and Banking*, Vol. 33, pp.272–283.

Szipro, G.G. (2002) 'Tinkering with genetic algorithms: forecasting and data mining in finance and economics', in S-H. Chen (Ed.) *Evolutionary Computation in Economics and Finance*, Heidelberg: Physica Verlag.

Zhang, G., Patuwo, B.E. and Hu, M.Y. (1998) 'Forecasting with artificial neural networks: the state of the art', *International Journal of Forecasting*, Vol. 14, pp.35–62.