



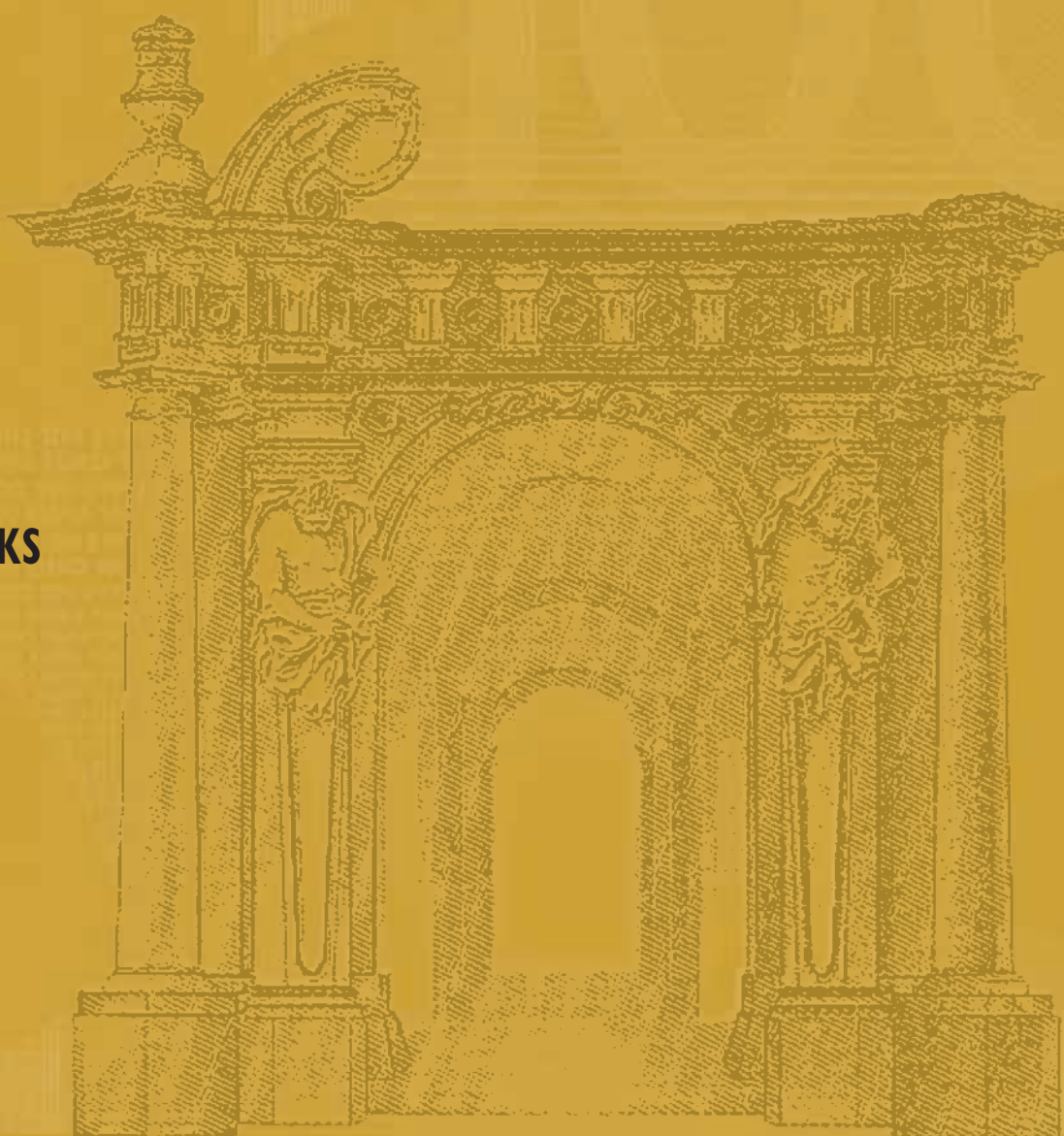
EUROPEAN CENTRAL BANK

## WORKING PAPER SERIES

NO. 352 / APRIL 2004

# FORECASTING INFLATION WITH THICK MODELS AND NEURAL NETWORKS

by Paul McNelis  
and Peter McAdam





EUROPEAN CENTRAL BANK

## WORKING PAPER SERIES

NO. 352 / APRIL 2004

# FORECASTING INFLATION WITH THICK MODELS AND NEURAL NETWORKS<sup>1</sup>

by Paul McNelis<sup>2</sup>  
and Peter McAdam<sup>3</sup>



In 2004 all publications will carry a motif taken from the €100 banknote.

This paper can be downloaded without charge from  
<http://www.ecb.int> or from the Social Science Research Network  
electronic library at [http://ssrn.com/abstract\\_id=533014](http://ssrn.com/abstract_id=533014).

<sup>1</sup> Without implicating, we thank Gonzalo Camba-Méndez, Ricardo Mestre, Jim Stock, one anonymous ECB working-paper referee and participants at the ECB Forecasting Techniques Workshop, December 2002 for helpful comments and suggestions. The opinions expressed are not necessarily those of the ECB. McAdam is also honorary lecturer in macroeconomics at the University of Kent and a CEPR and EABCN affiliate.

<sup>2</sup> Department of Economics, Georgetown University.

<sup>3</sup> DG Research, European Central Bank. Correspondence: Peter McAdam, European Central Bank, D-G Research, Econometric Modeling Unit, Kaiserstrasse 29, D-60311 Frankfurt, Germany. Tel: +49.69.13.44.6434. Fax: +49.69.13.44.6575. email: [peter.mcadam@ecb.int](mailto:peter.mcadam@ecb.int)

© European Central Bank, 2004

**Address**

Kaiserstrasse 29  
60311 Frankfurt am Main, Germany

**Postal address**

Postfach 16 03 19  
60066 Frankfurt am Main, Germany

**Telephone**

+49 69 1344 0

**Internet**

<http://www.ecb.int>

**Fax**

+49 69 1344 6000

**Telex**

411 144 ecb d

*All rights reserved.*

*Reproduction for educational and non-commercial purposes is permitted provided that the source is acknowledged.*

*The views expressed in this paper do not necessarily reflect those of the European Central Bank.*

*The statement of purpose for the ECB Working Paper Series is available from the ECB website, <http://www.ecb.int>.*

ISSN 1561-0810 (print)

ISSN 1725-2806 (online)

# CONTENTS

Abstract	4
Non-technical summary	5
1. Introduction	6
2. The Phillips curve	7
3. Non-linear inflation processes	7
3.1 Neural networks specifications	9
3.2 Neural network estimation and thick models	11
3.3 Adjustment and scaling of data	12
3.4 The benchmark model and evaluation criteria	13
3.4.1 In-sample diagnostics	14
3.4.2 Out-of-sample forecasting performance	14
4. Results	15
5. Conclusions	19
References	20
Appendix	22
European Central Bank working paper series	25

**Abstract:** This paper applies linear and neural network-based “thick” models for forecasting inflation based on Phillips–curve formulations in the USA, Japan and the euro area. Thick models represent “trimmed mean” forecasts from several neural network models. They outperform the best performing linear models for “real-time” and “bootstrap” forecasts for service indices for the euro area, and do well, sometimes better, for the more general consumer and producer price indices across a variety of countries.

**JEL:** C12, E31.

**Keywords:** Neural Networks, Thick Models, Phillips curves, real-time forecasting, bootstrap.

## Non-Technical Summary

This paper applies linear and neural network-based “thick” models (NN) for forecasting inflation based on Phillips–curve formulations in the USA, Japan and the euro area. Thick models represent “trimmed mean” forecasts from several neural network models. This “thick model” based on the NN specification is benchmarked against a well-performing linear process. The appeal of the NN is that it efficiently approximates a wide class of non-linear relations.

Our results show that non-linear Phillips curve specifications based on thick NN models can be competitive with the linear specification. We have attempted a high degree of robustness in our results by using different countries, different indices and sub-indices as well as performing different types of out-of-sample forecasts using a variety of supporting metrics. The “thick” NN models show the best “real-time” and bootstrap forecasting performance for the service-price indices for the Euro area. The performance of the neural network relative to a recursively-updated well-specified linear model should not be taken for granted. Given that the linear coefficients are changing each period, there is no reason not to expect good performance, especially in periods when there is little or no structural change taking place. We show in this paper that the linear and neural network specifications converge in their forecasts in such periods. The payoff of the neural network “thick modeling” strategy comes in periods of structural change and uncertainty, such as the early 1990’s in the USA and Germany, and after 2000 in the USA. When we examine the components of the CPI, we note that the non-linear models prove more accurate especially for forecasting inflation in the services sector. Since the service sector is, by definition, a highly labor-intensive industry and closely related to labor-market developments, this result appears to be consistent with recent research on relative labor-market rigidities and asymmetric adjustment.

## 1. Introduction

Forecasting is a key activity for policy makers. Given the possible complexity of the processes underlying policy targets, such as inflation, output gaps, or employment, and the difficulty of forecasting in real-time, recourse is often taken to simple models. A dominant feature of such models is their linearity. However, recent evidence suggests that simple, though non-linear, models may be at least as competitive as linear ones for forecasting macro variables. Marcellino (2002), for example, reported that non-linear models outperform linear and time-varying parameter models for forecasting inflation, industrial production and unemployment in the euro area. Indeed, after evaluating the performance of the Phillips curve for forecasting US inflation, Stock and Watson (1999) acknowledged that “to the extent that the relation between inflation and some of the candidate variables is non-linear”, their results may “understate the forecasting improvements that might be obtained, relative to the conventional linear Phillips curve” (p327). Moreover, Chen *et al.* (2001) examined linear and (highly non-linear) Neural Network Phillips-curve approaches for forecasting US inflation, and found that the latter models outperformed linear models for ten years of “real-time” one-period rolling forecasts.

Our paper contributes to this important debate in a number of respects. We follow Stock and Watson and concentrate on Phillips curves for forecasting inflation. However, we do so using linear and encompassing non-linear approaches. We further use a transparent comparison methodology. To avoid “model-mining”, our approach first identifies the best performing linear model and then compares that against a trimmed-mean forecast of simple non-linear models, which Granger and Jeon (2004) call a “thick model”. We examine the robustness of our inflation forecasting results by using different countries (and country aggregates), with different indices and sub-indices as well as conducting several types of out-of-sample comparisons using a variety of metrics.

Specifically, using the Phillips-curve framework, this paper applies linear and “thick” neural networks (NN) to forecast monthly inflation rates in the USA, Japan and the euro area. For the latter, we examine relatively long time series for Germany, France, Italy and Spain (comprising over 80% of the aggregate) as well as for the euro-area aggregate. The appeal of the NN is that it efficiently approximates a wide class of non-linear relations. Our goal is to see how well this approach performs relative to the standard linear one, for forecasting with “real-time” and randomly-generated “split sample” or “bootstrap” methods. In the “real-time” approach, the coefficients are updated period-by-period in a rolling window, to generate a sequence of one-period-ahead predictions. Since policy makers are usually interested in predicting inflation at twelve-month horizons, we estimate competing models for this horizon, with the bootstrap and real-time forecasting approaches. It turns out that the “thick model” based on trimmed-mean forecasts of several NN models dominates in many cases the linear model for the out-of-sample forecasting with the bootstrap and the “real-time” method.

Our “thick model” approach to neural network forecasting follows on recent reviews of neural network forecasting methods by Zhang *et al.*, (1998). They acknowledge that the proper specification of the structure of a neural network is a “complicated one” and note that there is no theoretical basis for selecting one specification or another for a neural network [Zhang *et al.*, (1998) p. 44]. We acknowledge this model uncertainty and consequently make use of the “thick model”

as a sensible way to utilize alternative neural network specifications and “training methods” in a “learning” context.

The paper proceeds as follows. The next section lays out the basic model. Section 3 discusses key properties of the data and the methodological background. Section 4 presents the empirical results for the US, Japan, the euro area, and Germany, France, Italy and Spain for the in-sample analysis, as well as for the twelve-month split-sample forecasts and examines the “real-time” forecasting properties for the same set of countries. Section 5 concludes.

## 2. The Phillips Curve

We begin with the following forecasting model for inflation:

$$\pi_{t+h}^h - \pi_t = f(\Delta u_t, \dots, \Delta u_{t-k}, \Delta \pi_t, \dots, \Delta \pi_{t-m}) + e_{t+h} \quad (1)$$

$$\pi_{t+h}^h = \frac{1200}{h} \ln \left( \frac{P_t}{P_{t-h}} \right) \quad (2)$$

where  $\pi_{t+h}$  is the percentage rate of inflation for the price level  $P$ , at an annualized value, at horizon  $t+h$ ,  $u$  is the unemployment rate,  $e_{t+h}$  is a random disturbance term, while  $k$  and  $m$  represent lag lengths for unemployment and inflation. We estimate the model for  $h=12$ . Given the discussion on the appropriate measure of inflation for monetary policy (e.g., Mankiw and Reis, 2003) we forecast using both the Consumer Price Index (CPI) and the producer price index (PPI) as well as indices for food, energy and services.

The data employed are monthly and seasonally adjusted. US data comes from the Federal Reserve of St. Louis FRED data base, while the Euro Area is from the European Central Bank.<sup>1</sup> The data for the remaining countries come from the OECD Main Economic Indicators.

## 3. Non-linear Inflation Processes

Should the inflation/unemployment relation or inflation/economic activity relation be linear? Figures 1 and 2 picture the inflation unemployment relation in the euro area and the USA, respectively and Table I lists summary statistics.

---

<sup>1</sup> Euro-area data, notably, has been widely used – inter alia, Galí *et al.* (2001), Smets and Wouters (2003), McAdam and Willman (2004).



Figure 1— Euro-Area Phillips curves: 1988-2001

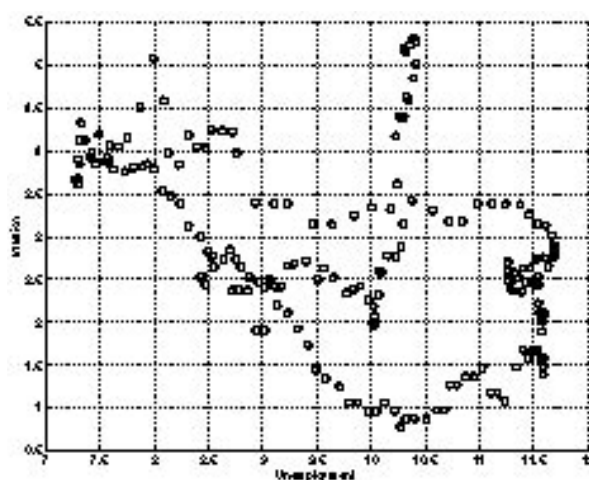


Figure 2— USA Phillips curves: 1988-2001

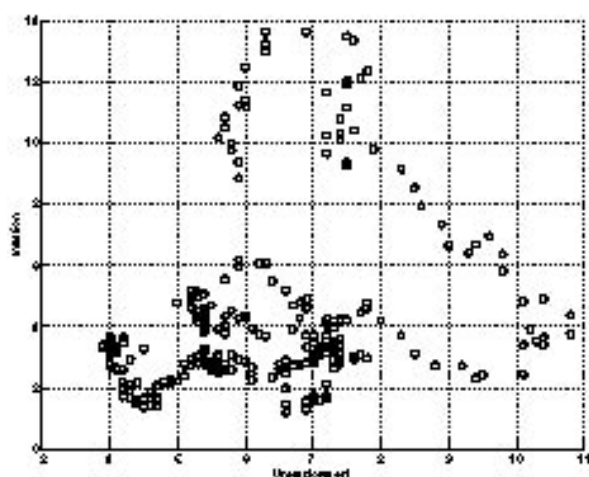


Table I—Summary Statistics

	Euro area		USA	
	<i>Inflation</i>	<i>Unemployment</i>	<i>Inflation</i>	<i>Unemployment</i>
<b>Mean</b>	2.84	9.83	3.16	5.76
<b>Std. Dev.</b>	1.07	1.39	1.07	1.07
<b>Coeff. Var.</b>	0.37	0.14	0.34	0.18

As we see, the average unemployment rate is more than four percentage points higher in the Euro Area than in the USA, and, as shown by the coefficient of variation, is less volatile. U.S. inflation, however, is only slightly higher than in the euro area, and its volatility is not appreciably different.

Need less to say, such differences in national economic performance have attracted considerable interest. In one influential analysis, for instance, Ljungqvist and Sargent (2001) point out that not only the average level but also the duration of euro-

area unemployment have exceeded the rest of the OECD during the past two decades – a feature they attribute to differences in unemployment compensation. Though, during the less turbulent 1950's and 1960's, European unemployment was lower than that of the US, with high lay-off costs, through a high tax on “job destruction”, they note that this lower unemployment may have been purchased at an “efficiency cost” by “making workers stay in jobs that had gone sour” (p. 19). When turbulence increased, and job destruction finally began to take place, older workers could be expected to choose extended periods of unemployment, after spending so many years in jobs in which both skills and adaptability significantly depreciated. This suggests that a labor market characterized by high layoff costs and generous unemployment benefits will exhibit asymmetries and “threshold behavior” in its adjustment process. Following periods of low turbulence, unemployment may be expected to remain low, even as shocks begin to increase. However, once a critical threshold is crossed, when the costs of staying employed far exceed layoff costs, unemployment will graduate to a higher level; those older workers whose skills markedly depreciated may be expected to seek long-term unemployment benefits.

The Ljungqvist and Sargent explanation of European unemployment is by no means exhaustive. Such unemployment dynamics may reflect a complex interaction among many factors, e.g., Lindbeck (1997), Blanchard and Wolfers (2000). However, notwithstanding the different emphasis of such many explanations, the general implication is that we might expect a non-linear estimation process with threshold effects, such as NNs, to outperform linear methods, for detecting underlying relations between unemployment and inflation in the euro area. At the very least, we expect (and in fact find) that non-linear approximation works better than linear models for inflation indices most closely related to changes in the labor market in the euro area – inflation in the price index for services.

### 3.1 *Neural Networks Specifications*

In this paper, we make use of a hybrid alternative formulation of the NN methodology: the basic multi-layer perceptron or feed-forward network, coupled with a linear jump connection or a linear neuron activation function. Following McAdam and Hughes-Hallett (1999), an encompassing NN can be written as:

$$n_{k,t} = \omega_0 + \sum_{i=1}^I \omega_i x_{t,i} + \sum_{j=1}^J \phi_j N_{t-1,j} \quad (3)$$

$$N_{k,t} = h(n_{k,t}) \quad (4)$$

$$y_{i,t} = \gamma_{i,0} + \sum_{k=1}^K \gamma_{i,k} N_{k,t} + \sum_{i=1}^I \beta_i x_{i,t} \quad (5)$$

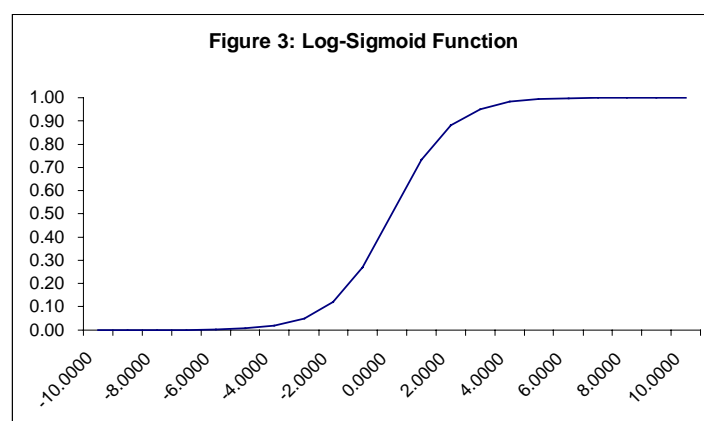
where inputs ( $x$ ) represent the current and lagged values of inflation and unemployment, and the outputs ( $y$ ) are their forecasts and where the  $I$  regressors are combined linearly to form  $K$  neurons, which are transformed or “encoded” by the

“squashing” function. The  $K$  neurons, in turn, are combined linearly to produce the “output” forecast.<sup>2</sup>

Within this system, (3)–(5), we can identify representative forms. **Simple** (or standard) **Feed-Forward**,  $\phi_j = \beta_i = 0, \forall i, j$ , namely links inputs ( $x$ ) to outputs ( $y$ ) via the hidden layer. Processing is thus *parallel* (as well as sequential); in equation (5) we have both a linear combination of the inputs and a limited-domain mapping of these through a “squashing” function,  $h$ , in equation (4). Common choices for  $h$  include the log-sigmoid form,  $N_{k,t} = h(n_{k,t}) = \frac{1}{1 + e^{-n_{k,t}}}$  (Figure 3) which transforms data to

within a unit interval:  $h: R \rightarrow [0,1]$  ,  $\left\{ \begin{array}{l} \lim_{n \rightarrow \infty} h(n) \rightarrow 1 \\ \lim_{n \rightarrow -\infty} h(n) \rightarrow 0 \end{array} \right.$  . Other, more sophisticated,

choices of the squashing function are considered in section 3.3.



The attractive feature of such functions is that they represent threshold behavior of the type previously discussed. For instance, they model representative non-linearities (e.g. a Keynesian liquidity trap where “low” interest rates fail to stimulate the economy or “labor-hoarding” where economic downturns have a less than proportional effect on layoffs etc). Further, they exemplify agent learning – at extremes of non-linearity, movements of economic variables (e.g., interest rates, asset prices) will generate a less than proportionate response to other variables. However if this movement continues, agents learn about their environment and start reacting more proportionately to such changes.

We might also have **Jump Connections**,  $\phi_j \neq 0, \forall j, \beta_i = 0, \forall i$ : direct links from the inputs,  $x$ , to the outputs. An appealing advantage of such a network is that it nests the pure linear model as well as the feed-forward NN. If the underlying relationship between the inputs and the output is a pure linear one, then only the direct jump connectors, given by  $\{\beta_i\}$ ,  $i = 1, \dots, I$ , should be significant. However, if the true

<sup>2</sup> Stock (1999) points out that the LSTAR (logistic smooth transition autoregressive) method is a special case of NN estimation. In this case,  $y_{t+h} = \alpha(L)y_t + d_t\beta(L)y_t + u_{t+h}$ , the switching variable  $d_t$  is a log-sigmoid function of past data, and determines the “threshold” at which the series switches.

relationship is a complex non-linear one, then one would expect  $\{\omega\}$  and  $\{\gamma\}$  to be highly significant, while the coefficient set  $\{\beta\}$  to be relatively insignificant. Finally, if the underlying relationship between the inputs variables  $\{x\}$  and the output variable  $\{y\}$  can be decomposed into linear and non-linear components, then we would expect all three sets of coefficients,  $\{\beta, \omega, \gamma\}$  to be significant. A practical use of the jump connection network is that it is a useful test for neglected non-linearity in a relationship between the input variables  $x$  and the output variable  $y$ .<sup>3</sup>

In this study, we examine this network with varying specifications for the number of neurons in the hidden layers, jump connections. The lag lengths for inflation and unemployment changes are selected on the basis of in-sample information criteria.

### 3.2 Neural Network Estimation and Thick Models

The parameter vectors of the network,  $\{\omega\}, \{\gamma\}, \{\beta\}$  may be estimated with non-linear least squares. However, given its possible convergence to local minima or saddle points (e.g., see the discussion in Stock, 1999), we follow the hybrid approach of Quagliarella and Vicini (1998): we use the genetic algorithm for a reasonably large number of generations, one hundred, then use the final weight vector  $[\{\hat{\omega}\}, \{\hat{\gamma}\}, \{\hat{\beta}\}]$  as the initialization vector for the gradient-descent minimization based on the quasi-Newton method. In particular, we use the algorithm advocated by Sims (2003).

The genetic algorithm proceeds in the following steps: (1) create an initial population of coefficient vectors as candidate solutions for the model; (2) have a selection process in which two different candidates are selected by a fitness criterion (minimum sum of squared errors) from the initial population; (3) have a cross-over of the two selected candidates from step (3) in which they create two offspring; (4) mutate the offspring; (5) have a "tournament", in which the parents and offspring compete to pass to the next generation, on the basis of the fitness criterion. This process is repeated until the population of the next generation is equal to the population of the first. The process stops after "convergence" takes place with the passing of one hundred generations or more. A description of this algorithm appears in the appendix.<sup>4</sup>

Quagliarella and Vicini (1998) point out that hybridization may lead to better solutions than those obtainable using the two methods individually. They argue that it is not necessary to carry out the gradient descent optimization until convergence, if one is going to repeat the process several times. The utility of the gradient-descent algorithm is its ability to improve the individuals it treats, so its beneficial effects can be obtained just performing a few iterations each time.

Notably, following Granger and Jeon (2004), we make use of a "thick modeling" strategy: combining forecasts of several NNs, based on different numbers of neurons in the hidden layer, and different network architectures (feedforward and jump connections) to compete against that of the linear model. The combination forecast is the "trimmed mean" forecast at each period, coming from an ensemble of networks, usually the same network estimated several times with different starting values for the parameter sets in the genetic algorithm, or slightly different networks.

<sup>3</sup> For completeness, a final case in this encompassing framework is Recurrent networks, (Elman, 1988),  $\phi_j = 0 \forall j, \beta_i \neq 0 \forall i$ , with current and lagged values of the inputs into system (memory). Although, this less popular network, is not used in this exercise. For an overview of NNs, see White (1992).

<sup>4</sup> See Duffy and McNelis (2001) for an example of the genetic algorithm with real, as opposed to binary, encoding.



We numerically rank the predictions of the forecasting model then remove the 100\* $\alpha$ % largest and smallest cases, leaving the remaining 100\*(2- $\alpha$ )% to be averaged. In our case, we set  $\alpha$  at 5%. Such an approach is similar to forecast combinations. The trimmed mean, however, is fundamentally more practical since it bypasses the complication of finding the optimal combination (weights) of the various forecasts.

### 3.3 Adjustment and Scaling of Data

For estimation, the inflation and unemployment “inputs” are stationary transformations of the underlying series. As in equation (1), the relevant forecast variables are the one-period-ahead first differences of inflation.<sup>5</sup>

Besides stationary transformation, and seasonal adjustment, scaling is also important for non-linear NN estimation. When input variables  $\{x_t\}$  and stationary output variables  $\{y_t\}$  are used in a NN, “scaling” facilitates the non-linear estimation process. The reason why scaling is helpful is that the use of very high or small numbers, or series with a few very high or very low outliers, can cause underflow or overflow problems, with the computer stopping, or even worse, or as Judd (1998, p.99) points out, the computer continuing by assigning a value of zero to the values being minimized.

There are two main ranges used in linear scaling functions: as before, in the unit interval,  $[0, 1]$ , and  $[-1, 1]$ . Linear scaling functions make use of the maximum and minimum values of series. The linear scaling function for the  $[0, 1]$  case transforms a variable  $x_k$  into  $x_k^*$  in the following way:<sup>6</sup>

$$x_{k,t}^* = \frac{x_{k,t} - \min(x_k)}{\max(x_k) - \min(x_k)} \quad (6)$$

A non-linear scaling method proposed by Helge Petersohn (University of Leipzig), transforming a variable  $x_k$  to  $z_k$  allows one to specify the range  $0 < z_k < 1$ , or  $(0,1)$ , given by  $[\max(z_k), \min(z_k)] = [\bar{z}_k, \underline{z}_k]$ :

$$z_{k,t} = \left( 1 + \exp \left[ \left( \frac{\ln(\bar{z}_k^{-1} - 1) - \ln(\underline{z}_k^{-1} - 1)}{\max(x_k) - \min(x_k)} \right) [x_{k,t} - \min(x_k)] + \ln(\underline{z}_k^{-1} - 1) \right] \right)^{-1} \quad (7)$$

Finally, Dayhoff and De Leo (2001) suggest scaling the data in a two step procedure: first, standardizing the series  $x$ , to obtain  $z$ , then taking the log-sigmoid transformation of  $z$ :

$$z = \frac{x - \bar{x}}{\sigma_x} \quad (8)$$

<sup>5</sup> As in Stock and Watson (1999), we find that there are little noticeable differences in results using seasonally adjusted or unadjusted data. Consequently, we report results for the seasonally adjusted data.

<sup>6</sup> The linear scaling function for  $[-1,1]$ , transforming  $x_k$  into  $x_k^{**}$ , has the form,  $x_{k,t}^{**} = 2 \frac{x_{k,t} - \min(x_k)}{\max(x_k) - \min(x_k)} - 1$ .

$$x^* = \frac{1}{1 + \exp(-z)} \quad (9)$$

Since there is no a priori way to decide which scaling function works best, the choice depends critically on the data. The best strategy is to estimate the model with different types of scaling functions to find out which one gives the best performance. When we repeatedly estimate various networks for the “ensemble” or trimmed mean forecast, we use identical networks employing different scaling function.

In our “thick model” approach, we use all three scaling functions for transforming the input variables. For the hidden layer neurons, we use the logsigmoid functions for the neural network forecasts. The networks are simple, with one, two or three neurons in one hidden-layer, with randomly-generated starting values<sup>7</sup>, using the feedforward and jump connection network types. We thus make use of twenty different neural network “architectures” in our thick model approach. These are twenty different randomly-generated integer values for the number of neurons in the hidden layer, combined with different randomly generated indicators for the network types and indicators for the scaling functions. Obviously, our thick model approach can be extended to a wider variety of specifications but we show, even with this smaller set, the power of this approach.<sup>8</sup>

In nonlinear neural network estimation, there is no closed-form solution for obtaining the parameter values of the network. The final values of the parameter estimates, and thus the predicted values of inflation, even with convergence, may be slightly different, depending on the choice of the scaling function and the starting values of the estimates, for a given neural network structure. Since we are also varying the network structure, of course, we will have a spectrum of predicted values. From this set we derived the trimmed mean forecast. This “thick model” approach is similar to “bagging predictors” in the machine learning and artificial intelligence literature (see Breiman, 1996).

### 3.4 The Benchmark Model and Evaluation Criteria

We examine the performance of the NN method relative to the benchmark linear model. In order to have a fair “race” between the linear and NN approaches, we first estimate the linear auto-regressive model, with varying lag structures for both inflation and unemployment. The optimal lag length for each variable, for each data set, is chosen based on the Hannan-Quinn criterion. We then evaluate the in-sample diagnostics of the best linear model to show that it is relatively free of specification error. For most of the data sets, we found that the best lag length for inflation, with the monthly data, was ten or eleven months, while one lag was needed for unemployment.

After selecting the best linear model and examining its in-sample properties, we then apply NN estimation and forecasting with the “thick model” approach discussed above, for the same lag length of the variables, with alternative NN

<sup>7</sup> We use different starting values as well as different scaling functions in order to increase the likelihood of finding the global, rather than a local, minimum.

<sup>8</sup> We use the same lag structure for both the neural network and linear models. Admittedly, we do this as simplifying computational short cut. Our goal is thus to find the “value added” of the neural network specification, given the benchmark best linear specification. This does not mean that alternative lag structures may work even better for neural network forecasting, relative to the benchmark best linear specification of the lag structure.



structures of two, three, or four neurons, with different scaling functions, and with feedforward, jump connection and We estimate this network alternative for thirty different iterations, and take the “trimmed mean” forecasts of this “thick model” or network ensemble, and compare the forecasting properties with those of the linear model.

### 3.4.1 *In-sample diagnostics*

We apply the following in-sample criteria to the linear auto-regressive and NN approaches:

- $R^2$  goodness-of-fit measure - denoted  $R^2$ ;
- Ljung-Box (1978) and McLeod-Li (1983) tests for autocorrelation and heteroskedasticity - **LB** and **ML**, respectively;
- Engle-Ng (1993) LM test for symmetry of residuals - **EN**;
- Jarque-Bera test for Normality of regression residuals - **JB**;
- Lee-White-Granger (1992) test for neglected non-linearity - **LWG**;
- Brock-Dechert-Scheinkman (1987) test for independence, based on the “correlation dimension” - **BDS**;

### 3.4.2 *Out-of-sample forecasting performance*

The following statistics examine the out-of-sample performance of the competing models:

- The root mean squared error estimate - **RMSQ**;
- The Diebold-Mariano (1995) test of forecasting performance of competing models - **DM**;
- The Persaran-Timmerman (1992) test of directional accuracy of the signs of the out-of-sample forecasts, as well as the corresponding success ratios, for the signs of forecasts - **SR**;
- The bootstrap test for “in-sample” bias.

For the first three criteria, we estimate the models recursively and obtain “real-time” forecasts. For the US data, we estimate the model from 1970.01 through 1990.01 and continuously update the sample, one month at a time, until 2003.01. For the euro-area data, we begin at 1980.01 and start the recursive real-time forecasts at 1995.01.

The bootstrap method is different. This is based on the original bootstrapping due to Efron (1983), but serves another purpose: out-of-sample forecast evaluation. The reason for doing out-of-sample tests, of course, is to see how well a model

generalizes beyond the original training or estimation set or historical sample, for a reasonable number of observations. As mentioned, the recursive methodology allows only one out-of-sample error for each training set. The point of any out-of-sample test is to estimate the “in-sample bias” of the estimates, with a sufficiently ample set of data. LeBaron (1997) proposes a variant of the original bootstrap test, the “0.632 bootstrap” (described in Table II).<sup>9</sup> The procedure is to estimate the original in-sample bias by repeatedly drawing new samples from the original sample, with replacement, and using the new samples as estimation sets, with the remaining data from the original sample, not appearing in the new estimation sets, as clean test or out-of-sample data sets. However, the bootstrap test does not have a well-defined distribution, so there are no “confidence intervals” that we can use to assess if one method of estimation dominates another in terms of this test of “bias”.

**Table II—“0.632” Bootstrap Test for In-Sample Bias**

Obtain mean square error from estimation set	$SSE(n) = \frac{1}{n} \sum_{i=1}^n [y_i - \hat{y}_i]^2$
Draw B samples of length n from estimation set	$z_1, z_2, \dots, z_B$
Estimate coefficients of model for each set	$\Omega^1, \Omega^2, \dots, \Omega^B$
Obtain “out of sample” matrix for each sample	$\tilde{z}_1, \tilde{z}_2, \dots, \tilde{z}_B$
Calculate average mean square error for “out of sample”	$SSE(n_b) = \frac{1}{n_b} \sum_{i=1}^{n_b} [\tilde{z}_b - \hat{z}_b(\Omega^b)]^2$
Calculate average mean square error for B bootstraps	$SSE(B) = \frac{1}{B} \sum_{b=1}^B SSE(n_b)$
Calculate “bias adjustment”	$\tilde{\omega}^{(0.632)} = 0.632[SSE(n) - SSE(B)]$
Calculate “adjusted error estimate”	$SSE^{(0.632)} = (1 - 0.632)SSE(n) + 0.632SSE(B)$

## 4 Results<sup>10</sup>

Table III contains the empirical results for the broad inflation indices for the USA, the euro area (as well as Germany, France, Spain and Italy) and Japan. The data set for the USA begins in 1970 while the European and Japanese series start in 1980. We “break” the USA sample to start “real-time forecasts” at 1990.01 while the other countries break at 1995.01.

<sup>9</sup> LeBaron (1997) notes that the weighting 0.632 comes from the probability that a given point is actually in a given bootstrap draw,  $1 - \left[1 - \frac{1}{n}\right]^n \approx 0.632$ .

<sup>10</sup> The (Matlab) code and the data set used in this paper is available on request.



Table III—Diagnostic / Forecasting Results

	USA CPI	PPI	Euro Area CPI	PPI	Germany CPI	PPI	France CPI	PPI	Spain CPI	PPI	Italy CPI	PPI	Japan CPI	WPI
LAGS-Inf	10	10	11	11	10	10	10	10	11	10	10	10	11	11
LAGS-Un	1	1	1	1	1	1	1	1	1	1	1	1	1	1
RSQ-LS	0.992	0.992	0.998	0.997	0.993	0.993	0.993	0.994	0.995	0.994	0.994	0.995	0.996	0.992
L-B *	0.948	0.851	0.414	0.094	0.956	0.892	0.992	0.910	0.937	0.799	0.828	0.667	0.885	0.985
McL-L *	0.829	<b>0.000</b>	0.003	0.867	0.880	0.835	0.592	0.318	0.452	0.818	0.258	0.491	0.976	0.854
E-N *	0.628	<b>0.000</b>	0.019	0.640	0.984	0.832	0.758	0.031	0.516	0.713	0.669	0.216	0.273	0.769
J-B *	<b>0.001</b>	<b>0.000</b>	<b>0.016</b>	<b>0.005</b>	<b>0.234</b>	<b>0.000</b>	<b>0.020</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>0.989</b>	<b>0.000</b>	0.284	<b>0.000</b>
LWG	0	1	7	1	0	0	0	0	1	1	1	1	2	1
BDS *	0.083	<b>0.000</b>	0.117	0.360	0.819	0.637	0.215	0.416	0.128	0.091	0.531	0.346	0.993	0.528
RSQ-NET	0.992	0.992	0.998	0.997	0.994	0.993	0.993	0.994	0.995	0.994	0.994	0.995	0.996	0.992
RMSQ-LS	0.214	0.386	0.167	0.358	0.308	0.303	0.225	0.368	0.178	0.368	0.207	0.305	0.340	0.340
RMSQ-NET	0.213	0.385	0.167	0.343	0.307	0.302	0.224	0.371	0.180	0.371	0.206	0.304	0.339	0.333
SR-LS	0.986	0.971	0.973	0.973	0.978	0.940	0.963	0.989	1.000	0.989	0.988	0.989	0.986	0.943
SR-NET	0.986	0.971	0.973	0.973	0.978	0.940	0.976	0.989	1.000	0.989	0.988	0.989	0.986	0.943
DM-1 *	<b>0.036</b>	<b>0.088</b>	0.568	<b>0.000</b>	<b>0.092</b>	0.218	0.344	0.768	0.600	0.768	0.267	<b>0.067</b>	0.117	<b>0.014</b>
DM-2 *	<b>0.043</b>	0.104	0.565	<b>0.002</b>	<b>0.073</b>	0.221	0.335	0.807	0.591	0.807	0.235	<b>0.091</b>	<b>0.098</b>	<b>0.048</b>
DM-3 *	<b>0.029</b>	<b>0.087</b>	0.571	<b>0.002</b>	0.108	0.230	0.358	0.796	0.599	0.796	0.228	<b>0.099</b>	<b>0.074</b>	<b>0.060</b>
DM-4 *	<b>0.033</b>	0.118	0.571	<b>0.000</b>	<b>0.086</b>	0.261	0.356	0.773	0.601	0.773	<b>0.042</b>	<b>0.099</b>	<b>0.076</b>	<b>0.087</b>
DM-5 *	<b>0.019</b>	0.108	0.584	<b>0.000</b>	<b>0.076</b>	0.243	0.345	0.778	0.611	0.778	0.220	<b>0.096</b>	<b>0.080</b>	<b>0.100</b>
Bootstrap SSE-LS	0.079	0.182	0.031	0.116	0.078	0.101	0.043	0.068	0.117	0.091	0.041	0.106	0.136	0.100
Bootstrap SSE-NET	0.079	0.181	0.030	0.116	0.078	0.101	0.043	0.068	0.117	0.091	0.039	0.106	0.136	0.100
Ratio	<b>0.997</b>	<b>0.993</b>	<b>0.990</b>	<b>0.996</b>	1.000	<b>0.999</b>	<b>0.998</b>	<b>0.999</b>	1.003	<b>0.993</b>	<b>0.954</b>	1.002	<b>1.000</b>	1.002

\*: represents probability values

**Note:** Bold indicates those series which show superior performance of the network, either in terms of Diebold-Mariano or bootstrap ratios.

*DM\_1*, ... *DM\_5* etc allow for the out of sample forecast errors to be corrected for autocorrelations at lags one through 5.

What is clear across a variety of countries is that the lag lengths for both inflation and unemployment are practically identical. With such a lag length, not surprisingly, the overall in-sample explanatory power of all of the linear models is quite high, over 0.99. The marginal significance levels of the Ljung-Box indicate that we cannot reject serial independence in the residuals.<sup>11</sup> The McLeod-Li tests for autocorrelation in the squared residuals are insignificant except for the US producer price index and the aggregate euro-area CPI. For most countries, we can reject normality in the regression residuals of the linear model (except for Germany, Italian and Japanese CPI). Furthermore, the Lee-White-Granger and Brock-Decker-Scheinkman tests do not indicate “neglected non-linearity”, suggesting that the linear auto-regressive model, with lag length appropriately chosen, is not subject to obvious specification error. This model, then, is a “fit” competitor for the neural network “thick model” for out-of-sample forecasting performance.

The forecasting statistics based on the root mean squared error and success ratios are quite close for the linear and network thick model. What matters, of course, is the significance: are the real-time forecast errors statistically “smaller” for the network model, in comparison with the linear model? The answer is not always. At the ten percent level, the forecast errors, for given autocorrelation corrections with the Diebold-Mariano statistics, are significantly better with the neural network approach for the US CPI and PPI, the euro area PPI, the German CPI, the Italian PPI and the Japanese CPI and WPI.

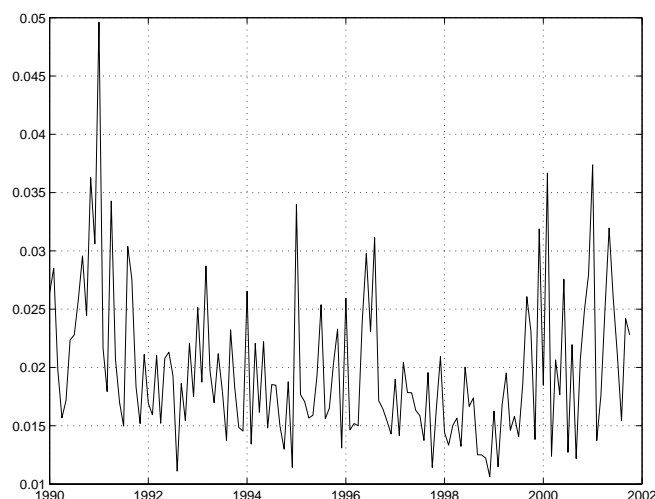
To be sure, the reduction in the root mean squared error statistic from moving to network methods is not dramatic, but the “forecasting improvement” is significant for the USA, Germany, Italy, and Japan. The bootstrapping sum of squared errors shows a small gain (in terms of percentage improvement) from moving to network

<sup>11</sup> Since our dependent variable is a 12-month-ahead forecast of inflation, the model by construction has a moving average error process of order 12, one current disturbance and 11 lagged disturbances. We approximate the MA representation with an AR (12) process, which effectively removes the serial dependence.

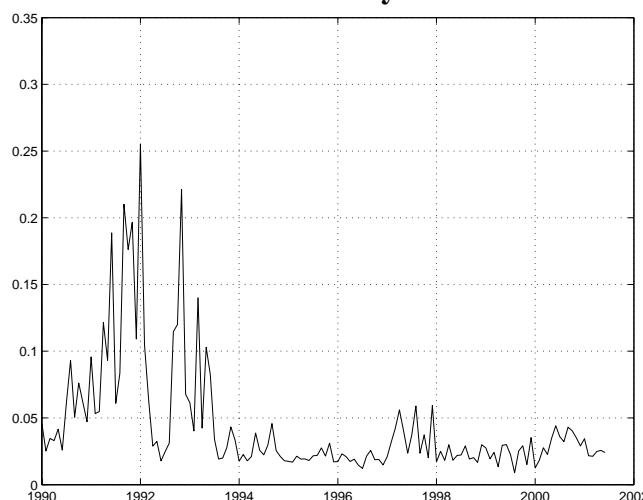
methods for the USA CPI and PPI, the euro area CPI and PPI, France CPI and PPI, Spain PPI and Italian CPI and PPI. For Italy, the percentage improvement in the forecasting is greatest for the CPI, with a gain or percentage reduction of almost five percent. For the other countries, the network error-reduction gain is less than one percent.

The usefulness of this “think modeling” strategy for forecasting is evident from an examination of Figures 4 and 5. In these figures, we plot the standard deviations of the set of forecasts for each out-of-sample period of all of the models. This comprises at each period twenty-two different forecasts, one linear, one based on the trimmed mean, and the remaining twenty neural network forecasts.

**Figure 4: Thick Model Forecast Uncertainty:  
USA**



**Figure 5: Thick Model Forecast Uncertainty:  
Germany**



We see in these two figures that the thick model forecast uncertainty is highest in the early 1990's in the USA and Germany, and after 2000 in the USA. In Germany,

this highlights the period of German unification. In the USA, the earlier period of uncertainty is likely due to the first Gulf War oil price shocks. The uncertainty after 2000 in the USA is likely due to the collapse of the US share market.

What is most interesting about these two figures is that models diverge in their forecasts in times of abrupt structural change. It is, of course, in these times that the thick model approach is especially useful. When there is little or no structural change, models converge to similar forecasts, and one approach does about as equally well as any other.

What about sub-indices? In Table IV, we examine the performance of the two estimation and forecasting approaches for food, energy and service components for the CPI for the USA and euro area.

**Table IV—Food, Energy and Services Indices, Diagnostics and Forecasting**

	USA			Euro Area		
	Food	Energy	Services	Food	Energy	Services
LAGS-INFLATION	10	11	10	10	10	10
LAGS-UNEMPLOY	1	6	1	1	1	1
RSQ-LS	0.992	0.993	0.993	0.994	0.993	0.996
L-B *	0.728	0.971	0.465	0.565	0.217	0.696
McL-L *	0.000	0.043	0.001	0.498	0.583	0.619
E-N *	0.000	0.075	0.000	0.442	0.374	0.883
J-B *	0.000	0.000	0.000	0.386	0.005	0.742
LWG	5	0	15	1	1	0
BDS *	0.000	0.000	0.000	0.092	0.938	0.342
RSQ-NET	0.991	0.994	0.993	0.996	0.993	0.997
RMSQ-LS	0.322	2.123	0.129	0.333	0.770	0.246
RMSQ-NET	0.32	2.144	0.129	0.334	0.775	0.230
SR-LS	0.949	0.974	0.961	0.961	0.941	0.941
SR-NET	0.955	0.974	0.955	0.961	0.941	0.941
DM-1 *	0.511	0.882	0.354	0.900	0.846	<b>0.000</b>
DM-2 *	0.512	0.854	0.313	0.876	0.801	<b>0.000</b>
DM-3 *	0.513	0.848	0.339	0.891	0.800	<b>0.000</b>
DM-4 *	0.513	0.839	0.324	0.934	0.793	<b>0.001</b>
DM-5 *	0.514	0.812	0.348	0.936	0.829	<b>0.002</b>
Bootstrap SSE-LS	0.402	3.001	0.049	0.067	0.428	0.086
Bootstrap SSE-NET	0.41	2.992	0.048	0.067	0.426	0.080
Ratio	<b>0.998</b>	<b>0.994</b>	<b>0.981</b>	<b>0.997</b>	<b>0.995</b>	<b>0.934</b>

\*: represents probability values

**Note:** See notes to Table III

The lag structures are about the same for these models as the overall CPI indices, except for the USA energy index, which has a lag length of unemployment of six. The results only show a market “real-time forecasting” improvement for the service component of the euro area. However the bootstrap method shows a reduction in the forecasting error “bias” for all of the indices, with the greatest reductions in forecasting error, of almost seven percent, for the services component of the euro area.

## 5 Conclusions

Forecasting inflation other industrialized countries is a challenging task. Notwithstanding the costs of developing tractable forecasting models, accurate forecasting is a key component of successful monetary policy. All our chosen countries have undergone major structural and economic-policy regime changes over the past two to three decades, some more dramatically than others. Any model, however complex, cannot capture all of the major structural characteristics affecting the underlying inflationary process. Economic forecasting is a learning process, in which we search for better subsets of approximating models for the true underlying process. Here, we examined one set of approximating alternative, a “thick model” based on the NN specification, benchmarked against a well-performing linear process. We do not suggest that the network approximation is the only alternative or the best among a variety of alternatives<sup>12</sup>. However, the appeal of the NN is that it efficiently approximates a wide class of non-linear relations.

Our results show that non-linear Phillips curve specifications based on thick NN models can be competitive with the linear specification. We have attempted a high degree of robustness in our results by using different countries, different indices and sub-indices as well as performing different types of out-of-sample forecasts using a variety of supporting metrics. The “thick” NN models show the best “real-time” and bootstrap forecasting performance for the service-price indices for the Euro area, consistent with, for instance, the analysis of Ljungqvist and Sargent (2001). However, these approaches also do well, sometimes better, for the more general consumer and producer price indices for the US, Japan and European countries.

The performance of the neural network relative to a recursively-updated well-specified linear model should not be taken for granted. Given that the linear coefficients are changing each period, there is no reason not to expect good performance, especially in periods when there is little or no structural change taking place. We show in this paper that the linear and neural network specifications converge in their forecasts in such periods. The payoff of the neural network “thick modeling” strategy comes in periods of structural change and uncertainty, such as the early 1990’s in the USA and Germany, and after 2000 in the USA.

When we examine the components of the CPI, we note that the nonlinear models work especially for forecasting inflation in the services sector. Since the service sector is, by definition, a highly labor-intensive industry and closely related to labor-market developments, this result appears to be consistent with recent research on relative labor-market rigidities and asymmetric adjustment.

---

<sup>12</sup> One interesting competing approximating model is the auto-regressive model with drifting coefficients and stochastic volatilities, e.g., Cogley and Sargent (2002).

## References

- Blanchard, O. J. and Wolfers, J. (2000) "The role of shocks and institutions in the rise of European unemployment", *Economic Journal*, 110, 462, C1-C33.
- Breiman, L. (1996) "Bagging Predictors", *Machine Learning* 24, 123-140.
- Brock, W., W. Dechert, and J. Scheinkman (1987) "A Test for Independence Based on the Correlation Dimension", Working Paper, Economics Department, University of Wisconsin at Madison.
- Chen, X., J. Racine, and N. R. Swanson (2001) "Semiparametric ARX Neural Network Models with an Application to Forecasting Inflation", Working Paper, Economics Department, Rutgers University.
- Cogley, T. and T. J. Sargent (2002) "Drifts and Volatilities: Monetary Policies and Outcomes in Post-WWII US", Available at: [www.stanford.edu/~sargent](http://www.stanford.edu/~sargent).
- Dayhoff, Judith E. and James M. De Leo (2001) "Artificial Neural Networks: Opening the Black Box". *Cancer*, 91, 8, 1615-1635.
- Diebold, F. X. and R. Mariano (1995) "Comparing Predictive Accuracy", *Journal of Business and Economic Statistics*, 3, 253-263.
- Duffy, J. and P. D. McNelis (2001) "Approximating and Simulating the Stochastic Growth Model: Parameterized Expectations, Neural Networks and the Genetic Algorithm", *Journal of Economic Dynamics and Control*, 25, 1273-1303.
- Efron, B. (1983) "Estimating the Error Rate of a Prediction Rule: Improvement on Cross Validation", *Journal of the American Statistical Association* 78(382), 316-331.
- Elman J. (1988) "Finding Structure in time", University Of California, mimeo.
- Engle, R. and V. Ng (1993) "Measuring the Impact of News on Volatility", *Journal of Finance*, 48, 1749-1778.
- Fogel, D. and Z. Michalewicz (2000) *How to Solve It: Modern Heuristics*, New York: Springer.
- Galí, J., M. Gertler, and J. D. López-Salido (2001) "European Inflation Dynamics", *European Economic Review*, 45, 7, 1237-1270.
- Granger, C. W. J. and Y. Jeon (2004) "Thick Modeling", *Economic Modeling*, 21, 2, 323-343.
- Granger, C. W. J., M. L. King, and H. L. White (1995) "Comments on Testing Economic Theories and the Use of Model Selection Criteria", *Journal of Econometrics*, 67, 173-188.
- Judd, K. L. (1998) *Numerical Methods in Economics*, MIT Press.
- LeBaron, B. (1997) "An Evolutionary Bootstrap Approach to Neural Network Pruning and Generalization", Working Paper, Economics Department, Brandeis University.
- Lee, T. H., H. White, and C. W. J. Granger (1992) "Testing for Neglected Nonlinearity in Times Series Models: A Comparison of Neural Network Models and Standard Tests", *Journal of Econometrics*, 56, 269-290.
- Lindbeck, A. (1997) "The European Unemployment Problem". Stockholm: Institute for International Economic Studies, Working Paper 616.
- Ljunqvist, L. and T. J. Sargent (2001) "European Unemployment: From a Worker's Perspective", Working Paper, Economics Department, Stanford University.
- Mankiw, N. Gregory and R. Reis (2003) "What measure of inflation should a central bank target", *Journal of European Economic Association*, 1, 5, 1058-1086.
- Marcellino, M. (2002) "Instability and Non-Linearity in the EMU", Working Paper 211, Bocconi University, IGIER.

- Marcellino, M., J. H. Stock, and M. W. Watson (2003) "Macroeconomic Forecasting in the Euro Area: Country Specific versus Area-Wide Information", *European Economic Review*, 47, 1-18.
- McAdam, P. and A. J. Hughes Hallett (1999) "Non Linearity, Computational Complexity and Macro Economic modeling", *Journal of Economic Surveys*, 13, 5, 577-618.
- McAdam, P. and A. Willman (2004) "Supply, Factor Shares And Inflation Persistence: Re-Examining Euro-Area New-Keynesian Phillips Curves", *Oxford Bulletin of Economics and Statistics*, forthcoming.
- McLeod, A. I. and W. K. Li (1983) "Diagnostic Checking ARMA Time Series Models Using Squared-Residual Autocorrelations", *Journal of Time Series Analysis*, 4, 269-273.
- Michalewicz, Z (1996) *Genetic Algorithms + Data Structures=Evolution Programs*. Third Edition. Berlin: Springer.
- Pesaran, M. H. and A. Timmermann (1992) "A Simple Nonparametric Test of Predictive Performance", *Journal of Business and Economic Statistics*, 10, 461-65.
- Quagliarella, D. and A. Vicini (1998) "Coupling Genetic Algorithms and Gradient Based Optimization Techniques" in Quagliarella, D. J. *et al.* (Eds.) *Genetic Algorithms and Evolution Strategy in Engineering and Computer Science*, John Wiles and Sons.
- Sargent, T. J. (2002) "Reaction to the Berkeley Story". Web Page: [www.stanford.edu/~sargent](http://www.stanford.edu/~sargent).
- Sims, C. S. (2003) "Optimization Software: CSMINWEL". Webpage: <http://eco-072399b.princeton.edu/yftp/optimize>.
- Smets, F. and Wouters, R. (2003) "An estimated stochastic dynamic general equilibrium model of the euro area", *Journal of European Economic Association*, 1, 5, 1123-1175.
- Stock, J. H. (1999) "Forecasting Economic Time Series", in Badi Baltagi (Ed.), *Companion in Theoretical Econometrics*, Basil Blackwell.
- Stock, J. H. and M. W. Watson (1998) "A Comparison of Linear and Non-linear Univariate Models for Forecasting Macroeconomic Time Series", NBER WP 6607.
- Stock, J. H. and M. W. Watson (1999) "Forecasting Inflation", *Journal of Monetary Economics*, 44, 293-335.
- Stock, J. H. and M. W. Watson (2001) "Forecasting Output and Inflation", NBER WP 8180.
- White, H. L. (1992) *Artificial Neural Networks*, Basil Blackwell.
- Zhang, G. B. Eddy Patuwo and M. Y. Hu (1998) "Forecasting with artificial neural networks: The state of the art", *International Journal of Forecasting*, 14, 1, 1, 35-62.

## Appendix:

### Evolutionary Stochastic Search: The Genetic Algorithm

Both the Newton-based optimization (including back propagation) and Simulated Annealing (SA) start with a random initialization vector  $\Omega_0$ . It should be clear that the usefulness of both of these approaches to optimization crucially depend on how “good” this initial parameter guess really is. The genetic algorithm (GA) helps us come up with a better “guess” for using either of these search processes. In addition, the GA avoids the problems of landing in a local minimum, or having to approximate the Hessians. Like Simulated Annealing, it is a statistical search process, but it goes beyond SA, since it is an *evolutionary search process*. The GA proceeds in the following steps.

#### Population creation

This method starts not with one random coefficient vector  $\Omega$ , but with a population  $N^*$  (an even number) of random vectors. Letting  $p$  be the size of each vector, representing the total number of coefficients to be estimated in the NN, one creates a population  $N^*$  of  $p$  by 1 random vectors:

$$\begin{pmatrix} \Omega_1 \\ \Omega_2 \\ \vdots \\ \Omega_p \end{pmatrix}_1 \begin{pmatrix} \Omega_1 \\ \Omega_2 \\ \vdots \\ \Omega_p \end{pmatrix}_2 \begin{pmatrix} \Omega_1 \\ \Omega_2 \\ \vdots \\ \Omega_p \end{pmatrix}_i \dots \begin{pmatrix} \Omega_1 \\ \Omega_2 \\ \vdots \\ \Omega_p \end{pmatrix}_{N^*} \quad (11)$$

#### Selection

The next step is to select two pairs of coefficients from the population at random, with replacement. Evaluate the “fitness” of these four coefficient vectors according to the sum of squared error function given above. Coefficient vectors which come closer to minimizing the sum of squared errors receive “better” fitness values.

One conducts a simple fitness “tournament” between the two pairs of vectors: the winner of each tournament is the vector with the best “fitness”. These two winning vectors (i, j) are retained for “breeding” purposes:



$$\begin{pmatrix} \Omega_1 \\ \Omega_2 \\ \vdots \\ \Omega_p \end{pmatrix}_i \begin{pmatrix} \Omega_1 \\ \Omega_2 \\ \vdots \\ \Omega_p \end{pmatrix}_j \quad (12)$$

### Crossover

The next step is crossover, in which the two parents “breed” two children. The algorithm allows “crossover” to be performed on each pair of coefficient vectors  $i$  and  $j$ , with a fixed probability  $p > 0$ . If crossover is to be performed, the algorithm uses one of three difference crossover operations, with each method having an equal (1/3) probability of being chosen:

- *Shuffle crossover.* For each pair of vectors,  $k$  random draws are made from a binomial distribution. If the  $k^{th}$  draw is equal to 1, the coefficients  $\Omega_{i,p}$  and  $\Omega_{j,p}$  are swapped; otherwise, no change is made.
- *Arithmetic crossover.* For each pair of vectors, a random number is chosen,  $\omega \in (0,1)$ . This number is used to create two new parameter vectors that are linear combinations of the two parent factors,  $\omega\Omega_{i,p} + (1-\omega)\Omega_{j,p}$ ,  $(1-\omega)\Omega_{i,p} + \omega\Omega_{j,p}$ .
- *Single-point crossover.* For each pair of vectors, an integer  $I$  is randomly chosen from the set  $[1, k-1]$ . The two vectors are then cut at integer  $I$  and the coefficients to the right of this cut point,  $\Omega_{i,I+1}, \Omega_{j,I+1}$  are swapped.

In binary-encoded genetic algorithms, single-point crossover is the standard method. There is no consensus in the genetic algorithm literature on which method is best for real-valued encoding.

Following the operation of the crossover operation, each pair of “parent” vectors is associated with two “children” coefficient vectors, which are denoted C1(i) and C2(j). If crossover has been applied to the pair of parents, the children vectors will generally differ from the parent vectors.

### Mutation

The fifth step is mutation of the children. With some small probability  $\tilde{p}\tilde{r}$ , which decreases over time, each element or coefficient of the two children's vectors is subjected to a mutation. The probability of each element is subject to mutation in generation  $G = 1, 2, \dots, G^*$ , given by the probability  $\tilde{p}\tilde{r} = 0.15 + 0.33/G$ .

If mutation is to be performed on a vector element, one uses the following non-uniform mutation operation, due to Michalewicz (1996). Begin by randomly drawing two real numbers  $r_1$  and  $r_2$  from the  $[0,1]$  interval and one random number  $s$ ,



from a standard normal distribution. The mutated coefficient  $\tilde{\Omega}_{i,p}$  is given by the following formula:

$$\tilde{\Omega}_{i,p} = \begin{cases} \Omega_{i,p} + s \left[ 1 - r_2^{1-G/G^*b} \right] & \text{if } r_1 > 0.5 \\ \Omega_{i,p} - s \left[ 1 - r_2^{1-G/G^*b} \right] & \text{if } r_1 \leq 0.5 \end{cases} \quad (13)$$

where  $G$  is the generation number,  $G^*$  is the maximum number of generations, and  $b$  is a parameter which governs the degree to which the mutation operation is non-uniform. Usually one sets  $b = 2$  and  $G^* = 150$ . Note that the probability of creating a new coefficient via mutation, which is far from the current coefficient value, diminishes as  $G \rightarrow G^*$ . This mutation operation is non-uniform since, over time, the algorithm is sampling increasingly more intensively in a neighborhood of the existing coefficient values. This more localized search allows for some fine-tuning of the coefficient vector in the later stages of the search, when the vectors should be approaching close to a global optimum.

### Election tournament

The last step is the election tournament. Following the mutation operation, the four members of the “family” (P1, P2, C1, C2) engage in a fitness tournament. The children are evaluated by the same fitness criterion used to evaluate the parents. The two vectors with the best fitness, whether parents or children, survive and pass to the next generation, while the two with the worst fitness value are extinguished.

One repeats the above process, with parents  $i$  and  $j$  returning to the population pool for possible selection again, until the next generation is populated by  $N^*$  vectors.

### Elitism

Once the next generation is populated, introduce elitism. Evaluate all the members of the new generation and the past generation according to the fitness criterion. If the “best” member of the older generation dominated the best member of the new generation, then this member displaces the worst member of the new generation and is thus eligible for selection in the coming generation.

### Convergence

One continues this process for  $G^*$  generations, usually  $G^*=150$ . One evaluates convergence by the fitness value of the best member of each generation.

## European Central Bank working paper series

For a complete list of Working Papers published by the ECB, please visit the ECB's website (<http://www.ecb.int>).

- 202 "Aggregate loans to the euro area private sector" by A. Calza, M. Manrique and J. Sousa, January 2003.
- 203 "Myopic loss aversion, disappointment aversion and the equity premium puzzle" by D. Fielding and L. Stracca, January 2003.
- 204 "Asymmetric dynamics in the correlations of global equity and bond returns" by L. Cappiello, R.F. Engle and K. Sheppard, January 2003.
- 205 "Real exchange rate in an inter-temporal n-country-model with incomplete markets" by B. Mercereau, January 2003.
- 206 "Empirical estimates of reaction functions for the euro area" by D. Gerdesmeier and B. Roffia, January 2003.
- 207 "A comprehensive model on the euro overnight rate" by F. R. Würtz, January 2003.
- 208 "Do demographic changes affect risk premiums? Evidence from international data" by A. Ang and A. Maddaloni, January 2003.
- 209 "A framework for collateral risk control determination" by D. Cossin, Z. Huang, D. Aunon-Nerin and F. González, January 2003.
- 210 "Anticipated Ramsey reforms and the uniform taxation principle: the role of international financial markets" by S. Schmitt-Grohé and M. Uribe, January 2003.
- 211 "Self-control and savings" by P. Michel and J.P. Vidal, January 2003.
- 212 "Modelling the implied probability of stock market movements" by E. Glatzer and M. Scheicher, January 2003.
- 213 "Aggregation and euro area Phillips curves" by S. Fabiani and J. Morgan, February 2003.
- 214 "On the selection of forecasting models" by A. Inoue and L. Kilian, February 2003.
- 215 "Budget institutions and fiscal performance in Central and Eastern European countries" by H. Gleich, February 2003.
- 216 "The admission of accession countries to an enlarged monetary union: a tentative assessment" by M. Ca'Zorzi and R. A. De Santis, February 2003.
- 217 "The role of product market regulations in the process of structural change" by J. Messina, March 2003.

- 218 “The zero-interest-rate bound and the role of the exchange rate for monetary policy in Japan” by G. Coenen and V. Wieland, March 2003.
- 219 “Extra-euro area manufacturing import prices and exchange rate pass-through” by B. Anderton, March 2003.
- 220 “The allocation of competencies in an international union: a positive analysis” by M. Ruta, April 2003.
- 221 “Estimating risk premia in money market rates” by A. Durré, S. Evjen and R. Pilegaard, April 2003.
- 222 “Inflation dynamics and subjective expectations in the United States” by K. Adam and M. Padula, April 2003.
- 223 “Optimal monetary policy with imperfect common knowledge” by K. Adam, April 2003.
- 224 “The rise of the yen vis-à-vis the (“synthetic”) euro: is it supported by economic fundamentals?” by C. Osbat, R. Ruffer and B. Schnatz, April 2003.
- 225 “Productivity and the (“synthetic”) euro-dollar exchange rate” by C. Osbat, F. Visselaar and B. Schnatz, April 2003.
- 226 “The central banker as a risk manager: quantifying and forecasting inflation risks” by L. Kilian and S. Manganelli, April 2003.
- 227 “Monetary policy in a low pass-through environment” by T. Monacelli, April 2003.
- 228 “Monetary policy shocks – a nonfundamental look at the data” by M. Klaeffer, May 2003.
- 229 “How does the ECB target inflation?” by P. Surico, May 2003.
- 230 “The euro area financial system: structure, integration and policy initiatives” by P. Hartmann, A. Maddaloni and S. Manganelli, May 2003.
- 231 “Price stability and monetary policy effectiveness when nominal interest rates are bounded at zero” by G. Coenen, A. Orphanides and V. Wieland, May 2003.
- 232 “Describing the Fed’s conduct with Taylor rules: is interest rate smoothing important?” by E. Castelnuovo, May 2003.
- 233 “The natural real rate of interest in the euro area” by N. Giammarioli and N. Valla, May 2003.
- 234 “Unemployment, hysteresis and transition” by M. León-Ledesma and P. McAdam, May 2003.
- 235 “Volatility of interest rates in the euro area: evidence from high frequency data” by N. Cassola and C. Morana, June 2003.

- 236 “Swiss monetary targeting 1974-1996: the role of internal policy analysis” by G. Rich, June 2003.
- 237 “Growth expectations, capital flows and international risk sharing” by O. Castrén, M. Miller and R. Stiegert, June 2003.
- 238 “The impact of monetary union on trade prices” by R. Anderton, R. E. Baldwin and D. Taglioni, June 2003.
- 239 “Temporary shocks and unavoidable transitions to a high-unemployment regime” by W. J. Denhaan, June 2003.
- 240 “Monetary policy transmission in the euro area: any changes after EMU?” by I. Angeloni and M. Ehrmann, July 2003.
- 241 “Maintaining price stability under free-floating: a fearless way out of the corner?” by C. Detken and V. Gaspar, July 2003.
- 242 “Public sector efficiency: an international comparison” by A. Afonso, L. Schuknecht and V. Tanzi, July 2003.
- 243 “Pass-through of external shocks to euro area inflation” by E. Hahn, July 2003.
- 244 “How does the ECB allot liquidity in its weekly main refinancing operations? A look at the empirical evidence” by S. Ejerskov, C. Martin Moss and L. Stracca, July 2003.
- 245 “Money and payments: a modern perspective” by C. Holthausen and C. Monnet, July 2003.
- 246 “Public finances and long-term growth in Europe – evidence from a panel data analysis” by D. R. de Ávila Torrijos and R. Strauch, July 2003.
- 247 “Forecasting euro area inflation: does aggregating forecasts by HICP component improve forecast accuracy?” by K. Hubrich, August 2003.
- 248 “Exchange rates and fundamentals” by C. Engel and K. D. West, August 2003.
- 249 “Trade advantages and specialisation dynamics in acceding countries” by A. Zaghini, August 2003.
- 250 “Persistence, the transmission mechanism and robust monetary policy” by I. Angeloni, G. Coenen and F. Smets, August 2003.
- 251 “Consumption, habit persistence, imperfect information and the lifetime budget constraint” by A. Willman, August 2003.
- 252 “Interpolation and backdating with a large information set” by E. Angelini, J. Henry and M. Marcellino, August 2003.
- 253 “Bond market inflation expectations and longer-term trends in broad monetary growth and inflation in industrial countries, 1880-2001” by W. G. Dewald, September 2003.

- 254 "Forecasting real GDP: what role for narrow money?" by C. Brand, H.-E. Reimers and F. Seitz, September 2003.
- 255 "Is the demand for euro area M3 stable?" by A. Bruggeman, P. Donati and A. Warne, September 2003.
- 256 "Information acquisition and decision making in committees: a survey" by K. Gerling, H. P. Grüner, A. Kiel and E. Schulte, September 2003.
- 257 "Macroeconomic modelling of monetary policy" by M. Klaeßling, September 2003.
- 258 "Interest rate reaction functions and the Taylor rule in the euro area" by P. Gerlach-Kristen, September 2003.
- 259 "Implicit tax co-ordination under repeated policy interactions" by M. Catenaro and J.-P. Vidal, September 2003.
- 260 "Aggregation-theoretic monetary aggregation over the euro area, when countries are heterogeneous" by W. A. Barnett, September 2003.
- 261 "Why has broad money demand been more stable in the euro area than in other economies? A literature review" by A. Calza and J. Sousa, September 2003.
- 262 "Indeterminacy of rational expectations equilibria in sequential financial markets" by P. Donati, September 2003.
- 263 "Measuring contagion with a Bayesian, time-varying coefficient model" by M. Ciccarelli and A. Rebucci, September 2003.
- 264 "A monthly monetary model with banking intermediation for the euro area" by A. Bruggeman and M. Donnay, September 2003.
- 265 "New Keynesian Phillips Curves: a reassessment using euro area data" by P. McAdam and A. Willman, September 2003.
- 266 "Finance and growth in the EU: new evidence from the liberalisation and harmonisation of the banking industry" by D. Romero de Ávila, September 2003.
- 267 "Comparing economic dynamics in the EU and CEE accession countries" by R. Süppel, September 2003.
- 268 "The output composition puzzle: a difference in the monetary transmission mechanism in the euro area and the US" by I. Angeloni, A. K. Kashyap, B. Mojon and D. Terlizzese, September 2003.
- 269 "Zero lower bound: is it a problem with the euro area?" by G. Coenen, September 2003.
- 270 "Downward nominal wage rigidity and the long-run Phillips curve: simulation-based evidence for the euro area" by G. Coenen, September 2003.
- 271 "Indeterminacy and search theory" by N. Giammarioli, September 2003.

- 272 “Inflation targets and the liquidity trap” by M. Klaefferling and V. López Pérez, September 2003.
- 273 “Definition of price stability, range and point inflation targets: the anchoring of long-term inflation expectations” by E. Castelnuovo, S. Nicoletti-Altimari and D. Rodríguez-Palenzuela, September 2003.
- 274 “Interpreting implied risk neutral densities: the role of risk premia” by P. Hördahl and D. Vestin, September 2003.
- 275 “Identifying the monetary transmission mechanism using structural breaks” by A. Beyer and R. Farmer, September 2003.
- 276 “Short-term estimates of euro area real GDP by means of monthly data” by G. Rünstler and F. Sédillot, September 2003.
- 277 “On the indeterminacy of determinacy and indeterminacy” by A. Beyer and R. Farmer, September 2003.
- 278 “Relevant economic issues concerning the optimal rate of inflation” by D. R. Palenzuela, G. Camba-Méndez and J. Á. García, September 2003.
- 279 “Designing targeting rules for international monetary policy cooperation” by G. Benigno and P. Benigno, October 2003.
- 280 “Inflation, factor substitution and growth” by R. Klump, October 2003.
- 281 “Identifying fiscal shocks and policy regimes in OECD countries” by G. de Arcangelis and S. Lamartina, October 2003.
- 282 “Optimal dynamic risk sharing when enforcement is a decision variable” by T. V. Koepl, October 2003.
- 283 “US, Japan and the euro area: comparing business-cycle features” by P. McAdam, November 2003.
- 284 “The credibility of the monetary policy ‘free lunch’” by J. Yetman, November 2003.
- 285 “Government deficits, wealth effects and the price level in an optimizing model” by B. Annicchiarico, November 2003.
- 286 “Country and sector-specific spillover effects in the euro area, the United States and Japan” by B. Kaltenhaeuser, November 2003.
- 287 “Consumer inflation expectations in Poland” by T. Łyziak, November 2003.
- 288 “Implementing optimal control cointegrated I(1) structural VAR models” by F. V. Monti, November 2003.
- 289 “Monetary and fiscal interactions in open economies” by G. Lombardo and A. Sutherland, November 2003.

- 290 "Inflation persistence and robust monetary policy design" by G. Coenen, November 2003.
- 291 "Measuring the time-inconsistency of US monetary policy" by P. Surico, November 2003.
- 292 "Bank mergers, competition and liquidity" by E. Carletti, P. Hartmann and G. Spagnolo, November 2003.
- 293 "Committees and special interests" by M. Felgenhauer and H. P. Grüner, November 2003.
- 294 "Does the yield spread predict recessions in the euro area?" by F. Moneta, December 2003.
- 295 "Optimal allotment policy in the eurosystem's main refinancing operations?" by C. Ewerhart, N. Cassola, S. Ejerskov and N. Valla, December 2003.
- 296 "Monetary policy analysis in a small open economy using bayesian cointegrated structural VARs?" by M. Villani and A. Warne, December 2003.
- 297 "Measurement of contagion in banks' equity prices" by R. Gropp and G. Moerman, December 2003.
- 298 "The lender of last resort: a 21st century approach" by X. Freixas, B. M. Parigi and J.-C. Rochet, December 2003.
- 299 "Import prices and pricing-to-market effects in the euro area" by T. Warmedinger, January 2004.
- 300 "Developing statistical indicators of the integration of the euro area banking system" by M. Manna, January 2004.
- 301 "Inflation and relative price asymmetry" by A. Rátfai, January 2004.
- 302 "Deposit insurance, moral hazard and market monitoring" by R. Gropp and J. Vesala, February 2004.
- 303 "Fiscal policy events and interest rate swap spreads: evidence from the EU" by A. Afonso and R. Strauch, February 2004.
- 304 "Equilibrium unemployment, job flows and inflation dynamics" by A. Trigari, February 2004.
- 305 "A structural common factor approach to core inflation estimation and forecasting" by C. Morana, February 2004.
- 306 "A markup model of inflation for the euro area" by C. Bowdler and E. S. Jansen, February 2004.
- 307 "Budgetary forecasts in Europe - the track record of stability and convergence programmes" by R. Strauch, M. Hallerberg and J. von Hagen, February 2004.
- 308 "International risk-sharing and the transmission of productivity shocks" by G. Corsetti, L. Dedola and S. Leduc, February 2004.
- 309 "Monetary policy shocks in the euro area and global liquidity spillovers" by J. Sousa and A. Zaghini, February 2004.
- 310 "International equity flows and returns: A quantitative equilibrium approach" by R. Albuquerque, G. H. Bauer and M. Schneider, February 2004.
- 311 "Current account dynamics in OECD and EU acceding countries – an intertemporal approach" by M. Bussière, M. Fratzscher and G. Müller, February 2004.

- 312 “Similarities and convergence in G-7 cycles” by F. Canova, M. Ciccarelli and E. Ortega, February 2004.
- 313 “The high-yield segment of the corporate bond market: a diffusion modelling approach for the United States, the United Kingdom and the euro area” by G. de Bondt and D. Marqués, February 2004.
- 314 “Exchange rate risks and asset prices in a small open economy” by A. Derviz, March 2004.
- 315 “Option-implied asymmetries in bond market expectations around monetary policy actions of the ECB” by S. Vähämaa, March 2004.
- 316 “Cooperation in international banking supervision” by C. Holthausen and T. Rønde, March 2004.
- 317 “Fiscal policy and inflation volatility” by P. C. Rother, March 2004.
- 318 “Gross job flows and institutions in Europe” by R. Gómez-Salvador, J. Messina and G. Vallanti, March 2004.
- 319 “Risk sharing through financial markets with endogenous enforcement of trades” by T. V. Köppl, March 2004.
- 320 “Institutions and service employment: a panel study for OECD countries” by J. Messina, March 2004.
- 321 “Frequency domain principal components estimation of fractionally cointegrated processes” by C. Morana, March 2004.
- 322 “Modelling inflation in the euro area” by E. S. Jansen, March 2004.
- 323 “On the indeterminacy of New-Keynesian economics” by A. Beyer and R. E. A. Farmer, March 2004.
- 324 “Fundamentals and joint currency crises” by P. Hartmann, S. Straetmans and C. G. de Vries, March 2004.
- 325 “What are the spill-overs from fiscal shocks in Europe? An empirical analysis” by M. Giuliodori and R. Beetsma, March 2004.
- 326 “The great depression and the Friedman-Schwartz hypothesis” by L. Christiano, R. Motto and M. Rostagno, March 2004.
- 327 “Diversification in euro area stock markets: country versus industry” by G. A. Moerman, April 2004.
- 328 “Non-fundamental exchange rate volatility and welfare” by R. Straub and I. Tchakarov, April 2004.
- 329 “On the determinants of euro area FDI to the United States: the knowledge-capital-Tobin's Q framework, by R. A. De Santis, R. Anderton and A. Hijzen, April 2004.
- 330 “The demand for euro area currencies: past, present and future” by B. Fischer, P. Köhler and F. Seitz, April 2004.
- 331 “How frequently do prices change? evidence based on the micro data underlying the Belgian CPI” by L. Aucremanne and E. Dhyne, April 2004.
- 332 “Stylised features of price setting behaviour in Portugal: 1992-2001” by M. Dias, D. Dias and P. D. Neves, April 2004.



- 333 “The pricing behaviour of Italian firms: New survey evidence on price stickiness” by S. Fabiani, A. Gattulli and R. Sabbatini, April 2004.
- 334 “Is inflation persistence intrinsic in industrial economies?” by A. T. Levin and J. M. Piger, April 2004.
- 335 “Has eura-area inflation persistence changed over time?” by G. O'Reilly and K. Whelan, April 2004.
- 336 “The great inflation of the 1970s” by F. Collard and H. Dellas, April 2004.
- 337 “The decline of activist stabilization policy: Natural rate misperceptions, learning and expectations” by A. Orphanides and J. C. Williams, April 2004.
- 338 “The optimal degree of discretion in monetary policy” by S. Athey, A. Atkeson and P. J. Kehoe, April 2004.
- 339 “Understanding the effects of government spending on consumption” by J. Galí, J. D. López-Salido and J. Vallés, April 2004.
- 340 “Indeterminacy with inflation-forecast-based rules in a two-bloc model” by N. Batini, P. Levine and J. Pearlman, April 2004.
- 341 “Benefits and spillovers of greater competition in Europe: A macroeconomic assessment” by T. Bayoumi, D. Laxton and P. Pesenti, April 2004.
- 342 “Equal size, equal role? Interest rate interdependence between the euro area and the United States” by M. Ehrmann and M. Fratzscher, April 2004.
- 343 “Monetary discretion, pricing complementarity and dynamic multiple equilibria” by R. G. King and A. L. Wolman, April 2004.
- 344 “Ramsey monetary policy and international relative prices” by E. Faia and T. Monacelli, April 2004.
- 345 “Optimal monetary and fiscal policy: A linear-quadratic approach” by P. Benigno and M. Woodford, April 2004.
- 346 “Perpetual youth and endogenous labour supply: a problem and a possible solution” by G. Ascari and N. Rankin, April 2004.
- 347 “Firms’ investment decisions in response to demand and price uncertainty” by C. Fuss and P. Vermeulen, April 2004.
- 348 “Financial openness and growth: Short-run gain, long-run pain?” by M. Fratzscher and M. Bussiere, April 2004.
- 349 “Estimating the rank of the spectral density matrix” by G. Camba-Mendez and G. Kapetanios, April 2004.
- 350 “Exchange-rate policy and the zero bound on nominal interest rates” by G. Camba-Mendez and G. Kapetanios, April 2004.

- 351 “Interest rate determination in the interbank market” by V. Gaspar, G. P. Quirós and H. R. Mendizábal, April 2004.
- 352 “Forecasting inflation with thick models and neural networks” by P. McNelis and P. McAdam, April 2004.

