WORKING PAPER SERIES 11

2005

Marek Hlaváček, Michael Koňák and Josef Čada:
The Application of Structured Feedforward Neural Networks to the Modelling of Daily Series of Currency in Circulation

CNB CZECH NATIONAL BANK

WORKING PAPER SERIES

The Application of Structured Feedforward Neural Networks to the Modelling of Daily Series of Currency in Circulation

Marek Hlaváček
Michael Koňák
Josef Čada

11/2005

Printed and distributed by the Czech National Bank. Available at http://www.cnb.cz.

Reviewed by: Ladislav Mochán (Czech National Bank)
Petr Zemčík (CERGE-EI, Prague)
Marcel Jiřina (Institut of Informatics, ASCR)

Project Coordinator: Juraj Antal

The Application of Structured Feedforward Neural Networks to the Modelling of Daily Series of Currency in Circulation

Marek Hlaváček, Michael Koňák and Josef Čada*

Abstract

One of the most significant factors influencing the liquidity of the financial market is the amount of currency in circulation. Although the central bank is responsible for the distribution of the currency it cannot assess the demand for the currency, as that demand is influenced by the non-banking sector. Therefore, the amount of currency in circulation has to be forecasted. This paper introduces a feedforward structured neural network model and discusses its applicability to the forecasting of currency in circulation. The forecasting performance of the new neural network model is compared with an ARIMA model. The results indicate that the performance of the neural network model is better and that both models might be applied at least as supportive tools for liquidity forecasting.

JEL Codes: C45, C53.
Keywords: Neural network, seasonal time series, currency in circulation.

* Marek Hlaváček, Czech National Bank, Interventions Division (marek.hlavacek@gmail.com), Michael Koňák, Czech National Bank, Transactions Systems Division (michael.konak@cnb.cz), Josef Čada, Czech National Bank, Operations Analyses Division (josef.cada@cnb.cz).
This paper has been written within the Czech National Bank's Research Project No. C8/03. The authors would like to thank Martin Peřina from ECB, Market Operations Analysis Division (until January 2005 head of the CNB's Interventions Division) and Vladimír Pulkráb from CNB's Operations Analyses Division for valuable suggestions and comments. Thanks also go to Marian Hlaváček from CVUT, Faculty of Nuclear Sciences and Physical Engineering, for his contribution to the development of the ARIMA model as part of his bachelor thesis.

Nontechnical Summary

Central banks have recently been maintaining price stability through a set of monetary policy instruments. To pursue its objectives effectively, a central bank needs an accurate estimate of money market liquidity. However, this liquidity is influenced by several autonomous factors that are not under the full control of the central bank. One of the most important autonomous factors is currency in circulation, which is quite difficult to assess, as it is strongly influenced by many seasonal factors.

To get over the problem, central banks such as the Federal Reserve System, the European Central Bank and other central banks within the European Monetary Union already use mathematical models of currency in circulation at least as supportive tools. Most of them use linear ARIMA models, but they are simultaneously developing non-linear models to describe the seasonal influence more accurately.

This paper follows the idea of non-linear model applications and investigates the applicability of a new neural network-based model on the Czech series of currency in circulation. It introduces a special kind of neural network model which has not been used for time series forecasting before. The new model is compared with an ARIMA model with respect to out-of-sample forecasting performance. It is shown that both models outperform the currently used expert forecasts and that the neural network model is more accurate than the ARIMA model. Moreover, the comparison yields several suggestions for improving the neural network model's forecasting accuracy.

1. Introduction

Nowadays, central banks control the economic conditions mainly indirectly. Central banks pursue their statutory objective of maintaining price stability or exchange rate stability through different sets of monetary policy instruments. They usually endeavour to steer money market interest rates through proper liquidity management. Therefore, an accurate estimate of money market liquidity is essential for effective monetary policy implementation.

Although only transactions with the central bank have an impact on money market liquidity, some of them are out of the central bank's control. These factors are called autonomous factors. One of the most important autonomous factors is currency in circulation (CIC). The demand for CIC is influenced by the non-banking sector, which means it is rather volatile and depends on various seasonal factors. The influence of seasonal factors make the assessment of the demand for CIC very knotty.

For that reason, central banks employ various mathematical models to deal with this typical seasonal time series. Most of the models used are based on the principles of the Box-Jenkins methodology (Box and Jenkins, 1976) and on the subsequent improvements made to it (e.g. Hamilton (1994), Gourieroux (1997)). Recently, however, banks have also been developing new non-linear models that are supposed to approximate seasonal patterns with higher accuracy than linear models.

This paper follows the idea of non-linear model applications and investigates the applicability of a special neural network model called the structured feedforward neural network. This neural network model is derived from networks with switching units originally developed as data classifiers (Bitzan, Šmejkalová, Kučera (1995), Hakl, Hlaváček and Kalous (2003)). However, experiments with time series forecasting have shown that the model can be applied to time series forecasting as well. The comparison with the ARIMA model presented later on shows that the out-of-sample forecasting performance of the neural network is better, although the neural network model could be further developed with respect to time series forecasting.

The working paper is organised as follows: First, the paper briefly summarises the basic ideas of liquidity management and describes the CIC series in sections 2 and 3 respectively. Then, the two models are defined separately in sections 4 and 5. Finally, a comparison of the models follows in section 6

2. Liquidity Management

This and the following section sketch the main ideas of liquidity management from the Czech National Bank's point of view and summarise the necessary background to the development of the Czech CIC model.

The Czech National Bank (CNB) pursues its statutory objective of maintaining price stability through a set of monetary policy instruments. The instruments steer money market interest rates and thus control the monetary conditions for economic agents.

The most significant monetary policy instruments used by the Czech National Bank are open market operations, in particular repo tenders with a duration of 14 days. The repo tenders are the main tools for managing the liquidity of the banking sector, which means the amount of reserves. The reserves are money deposited by banks on their clearing accounts with the CNB. Reserve money is used for settling banks' transactions and for covering their cash currency operations with the central bank.

The CNB endeavours to ensure that the actual aggregate amount of the reserves corresponds to banks' needs, meaning that there is neither a deficit nor a surplus of liquidity. The banks' demand for reserve money is determined by the minimum reserve requirements imposed on them by the central bank.

The actual aggregate amount of reserves is determined by transactions between banks and the central bank, since only the central bank can issue reserve money. For example, buying foreign currency by the central bank implies issuing its equivalent in domestic currency to the clearing accounts of the banks involved in the trade, i.e. the issuance of reserve money. Open market operations have the same impact on banks' reserves, hence they can be used to control the aggregate amount of reserve money so that it is in equilibrium with the minimum reserve requirements or meets the CNB's objectives.

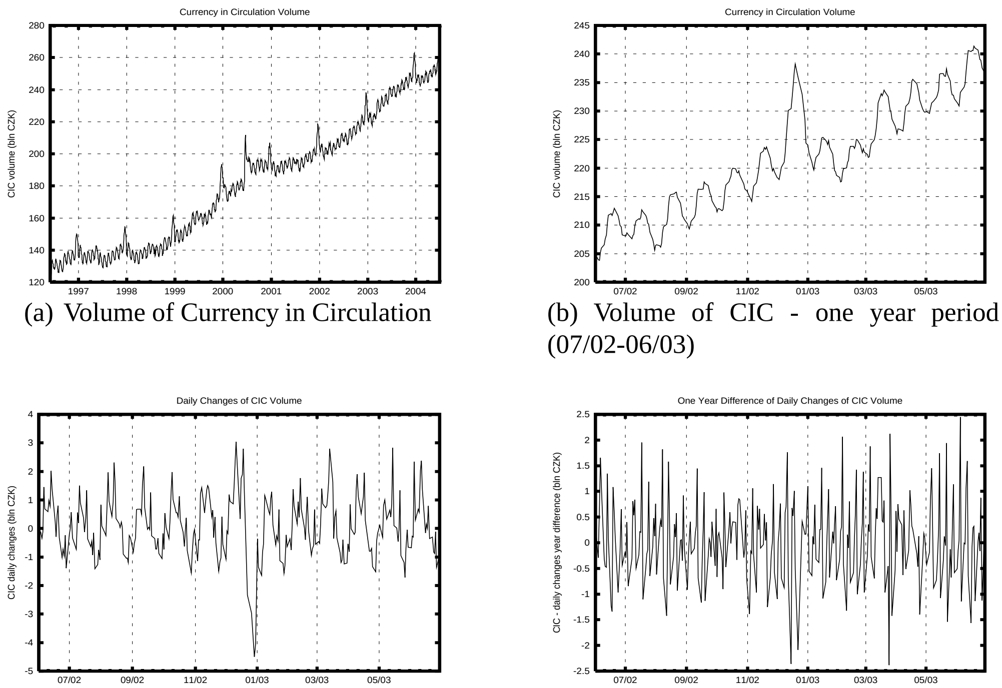
The basic mechanism is very simple. When the actual amount of the reserves equals banks' needs, there is no need for significant deviations of short-term money market rates from the desired level. In the case of a large liquidity deficit, banks would demand the necessary funds and market rates would rise significantly. And vice versa, in the case of a liquidity surplus market rates would decline. This basic idea indicates that it is crucial to have the most accurate possible estimates of the amount of reserve money the banks have at their disposal in order to determine the amount of repo tenders fitting the balance of banking sector as a whole from a liquidity point of view.

Although only transactions with the central bank have an impact on banks' reserves, some of them are out of the control of central bank management of liquidity and its counterparts. These are called autonomous factors. The most important autonomous factors are net foreign assets, government deposits and currency in circulation. While net foreign assets are quite easily predicted, based on data from the Interventions Division, and the cash management of the State Treasury Account has substantially improved the forecast of government deposits, the CIC factor remains rather volatile. Therefore, it is important to forecast the actual volume of CIC to minimise the errors in liquidity management.

The CNB conducts repo tenders daily at the beginning of the trading day, hence one-step-ahead liquidity forecasts are sufficient for setting the tender volume. The forecasts are currently estimated by experts according to historical values, the current money market situation and, in particular, the experts' experience.

Although the performance of these forecasts is quite good, it strongly depends on the experience of the particular expert, and that is undesirable. Moreover, the performance of these estimations goes down rapidly at the longer forecast horizon that is necessary for potentially

Figure 3.1: Currency in Circulation



(a) Volume of Currency in Circulation

(b) Volume of CIC - one year period (07/02-06/03)

(c) Daily changes in CIC Volume

(d) Seasonal difference in daily CIC changes

reducing the frequency of the tenders. These are the main reasons for the need for a mathematical model that should systemise the forecasting process and that might also improve the forecasting performance.

3. Currency in Circulation

The volume of currency in circulation is one of the most important factors in the liquidity forecasting process. Unfortunately, the CIC volume is out of the control of the central bank, hence it cannot be determined exactly. Therefore, it is necessary to approximate the behaviour of CIC using a mathematical model. The volume of currency in circulation is a typical seasonal time series which is significantly influenced by numerous seasonal factors (see fig. 3.1). A description of CIC behaviour and a summary of all the factors that influence the CIC follow.

3.1 Stochastic Behaviour of Currency in Circulation

For the purposes of this paper, currency in circulation is defined as banknotes and coins held outside the central bank. The distribution of banknotes and coins to the non-banking system is mainly carried out by commercial banks. They have to supply their branches and ATM networks, as clients assume that any nearby ATM is always ready and full of banknotes, and no one can even imagine it being impossible to withdraw cash from a bank branch—unless something has gone wrong.

Commercial banks are allowed to withdraw cash from their accounts with the central bank[1]. However, such a cash withdrawal (or deposit) influences the liquidity of the bank and thus of the banking sector as a whole, as it is a transaction between a commercial bank and the central bank. Therefore, changes in the total amount of currency in circulation directly influence the liquidity of the banking sector.

Moreover, it is evident that commercial banks aim to return spare cash to the central bank as soon as possible, because cash in an ATM or a safe cannot be further invested. In other words, banks flexibly follow clients' requirements and hence the demand for cash is mainly influenced by the non-banking sector, including businesses and households. This means that the changes in the volume of currency in circulation are caused by an enormous number of factors and circumstances that are obviously uncontrollable. Thus it is impossible to assess the exact volume of CIC and it has to be estimated without prior knowledge of all the factors and circumstances.

To deal with the influence of the non-banking sector it is possible to assign the superposition of all the uncontrollable factors to stochastic and seasonal behaviour of the CIC volume. This interpretation means that the CIC volume is supposed to be a random variable following a compound process with seasonal and stochastic components.

3.2 Factors with an Impact on the CIC Volume

Both the models described and compared later on in this paper express the series of the CIC volume as a function of historical values and several seasonal and shock factors. The identification of all significant seasonal patterns and shocks and the choice of the correct form for their description proved to be crucial to the models' forecasting performance. The choice of seasonal factors was predominantly motivated by a ECB working paper (Cabrero, Camba-Mendez, Hirsch and Nieto, 2002). However, the experience of CNB experts responsible for liquidity forecasting was also considered.

Exogenous factors are included in the models in two different forms, following the idea presented by Bell and Hillmer (1983). The first type of seasonal factor is a superposition of goniometric functions and the second is a lagged polynomial of the indicator function. The approximation of the seasonal influence in both models is then based on linear combination of all the factors included.

A seasonal factor expressed as a superposition of goniometric functions has the form:

$$d_{t,i} = \sum_{j=1}^{p} \left(a_{i,j} \sin\left(\frac{2j\pi m_i, t}{M_{i,t}} \right) + b_{i,j} \cos\left(\frac{2j\pi m_{i,t}}{M_{i,t}} \right) \right), \tag{3.1}$$

where $d_{t,i}$ is the value of the ith factor at time t, $M_{i,t}$ is the length of the current cycle (for example the length of the month) and $m_{i,t}$ is the position in the current cycle at time t (the current day in the month). Finally, the positive number p sets the number of different frequencies

[1] The Czech National Bank uses so-called cash deposit and withdrawal accounts.

Table 3.1: Seasonal factors and shocks

seasonal factors, shocks	**order of** Γ	**power** F	**number of frequencies**
intra-monthly effect	—	—	8
day of week	3	0	—
Easter	10	5	—
fixed holiday	10	5	—
non-working days	10	5	—
Christmas	15	5	—
New Year	5	5	—
bank failure	15	5	—
Y2K	10	5	—

forming the factor. Naturally, the more frequencies are considered the better the approximation is. However, it is also necessary to keep the number of model parameters low, hence the number of frequencies p has to be chosen carefully.

The second group of seasonal factors is mostly applied to model-isolated events such as national holidays or shocks and is of the form:[2]

$$d_{t,i} = \Gamma_i \left(\boldsymbol{B} \right) \boldsymbol{B}^{-F_i} \tau_i \left(t \right), \tag{3.2}$$

where $\boldsymbol{B}$ is the backshift operator ($\boldsymbol{B} y_t = y_{t-1}$), Γ_i is a polynomial in $\boldsymbol{B}$ and τ_i is the seasonal indicator function ($\tau_i \left(t \right) = 1$ if the ith season occurs at time t and $\tau_i \left(t \right) = 0$ otherwise). Finally, F_i is a positive power of $\boldsymbol{B}$. The combination of the polynomial Γ_i and $\boldsymbol{B}^{-F_i}$ guarantees that particular seasons can influence future and past observations.

Finally, all the factors described below were included in both models in the initial set of explanatory variables. A brief overview of all the factors, together with the numbers of lags, is then summarised in table 3.1.

Intra-monthly Effect The only seasonal effect of the form (3.1) included in the model is the intra-monthly effect. The monthly cycle is particularly influenced by salary payments. This means that the demand for cash is higher around pay day and then decreases until salaries are paid again. For the intra-monthly effect, $M_{i,t}$ from formula (3.1) is the number of working days in the corresponding month, while $m_{i,t}$ from the same formula is the order of the working day in the month. To describe the intra-monthly effect, eight different frequencies were considered.

[2] The form of the lagged seasonal indicator is the only difference between the ARIMA models presented herein and in Cabrero et al. (2002), where the lag operator Γ_i is supposed to be a fraction with a first-order polynomial as the denominator too.

Day of Week The day of the week effect is similar to the intra-monthly effect. Again, the demand oscillates during the week and reaches a maximum before the weekend as the ATM network has to withstand all the shopping activity. However, the seasonal indicator function is used instead of the approximation via goniometric functions. The indicator function corresponds to Friday and three lags to Tuesday, Wednesday and Thursday. The effect of the weekday is then measured as the difference between the effect of Monday and the given day.

Floating Holiday – Easter Easter is probably the trickiest season. Although Easter Monday is always on a Monday, its date varies year by year. Hence it interferes with the intra-monthly seasonality and it seems that the influence of the Easter holiday depends on its position in the month. To catch the effect of Easter, the lagged polynomial of the Easter Monday indicator function was included in the model.

Fixed Holiday Other national holidays are, unlike Easter, fixed to a particular date, hence their positions in the month are given and do not change. On the other hand, their position in the week is different year by year. The influence of national holidays is modelled using lagged polynomials of the indicator functions.

Number of Non-working Days A national holiday falling on a Friday or a Monday might influence the demand for CIC more than one falling midweek. This is the opinion of the CNB's experts, and it is also stated in Cabrero et al. (2002). Therefore, we tried to include also the number of non-working days following the particular day as an explanatory variable. This additional factor that might model the interference between weekends and fixed holidays is also included with its lagged and future values.

Christmas and New Year The most significant season is around Christmas and New Year, when shopping activity rises dramatically. The Christmas effect is again approximated by the lagged polynomial of the indicator function, although more lags are considered to capture the influence of the New Year, too.

Shocks Apart from the seasonal effects listed above, shocks were also modelled via indicator functions. Two significant shocks were identified in the series. The first is the Y2K effect and the second is a bank failure on 16 June 2000 preceded by significant growth in withdrawals.

To conclude the definition of the CIC data, the sample range should be specified. Data from between January 1996 and June 2004 are used for the purposes of this paper. Both models were optimised on the observations up to June 2003 and the out-of-sample performance is assessed on the rest of the sample.

4. Seasonal ARIMA Model

One of the most common classes of seasonal time series models is based on the methodology proposed by Box and Jenkins (1976). Various generalisations of Box-Jenkins ARMA models are widely approved, so an ARMA based model is employed as a benchmark model here.

4.1 Box-Jenkins Methodology – Seasonal ARIMA Model

The Box-Jenkins methodology is based on Wold's theorem (e.g. Hamilton (1994)), which says that any weakly stationary process can be written as a superposition of autoregressive processes and processes of moving averages[3]. These two basic kinds of weakly stationary processes are in short called AR and MA processes, hence Box-Jenkins models are also referred to as ARMA models. Specifically, a time series $(y_t)_{t=1}^{\infty}$ is an ARMA process of orders p and q, or, in short, an **ARMA(p,q)** process, if it can be written in the following form:

$$y_t = \sum_{i=1}^{p} \alpha_i y_{t-i} + \sum_{i=1}^{q} \beta_i \epsilon_{t-i} + \epsilon_t \tag{4.3}$$

where $\alpha = (\alpha_1, \ldots, \alpha_p) \in \mathcal{R}^p$, $\beta = (\beta_1, \ldots, \beta_q) \in \mathcal{R}^q$ and $(\epsilon_t)_{t=1}^{\infty}$ is a process of independent, identically distributed random variables usually called white noise. The first sum in equation (4.3) is then the AR component, while the second one is the MA component.

The theory of ARMA processes has become very popular during recent decades and has been further developed in many respects and applied in various branches. This progress has allowed the Box-Jenkins methodology to be applied not only to weakly stationary processes, but also to non-stationary processes that besides an ARMA process include various trends and seasonal and other deterministic or stochastic components. Bell and Hillmer (1983) suggested using the model (4.4) for series with calendar variations, which is a linear regression model with errors following an ARIMA process:

$$y_t = D_{t,i} + \frac{\Theta(\boldsymbol{B})}{\Phi(\boldsymbol{B})\,\Delta(\boldsymbol{B})}\,\epsilon_t. \tag{4.4}$$

Here y_t is the modelled series, $D_{t,i}$ is the regression part, $\boldsymbol{B}$ is the backshift operator and Θ, Φ, Δ are polynomials in $\boldsymbol{B}$. The polynomials Θ and Φ are moving-average and autoregressive operators, respectively. The polynomial Δ is a difference operator that can also include a seasonal difference operator. The regression part $D_{t,i} = \sum_{i=1}^{s} d_{t,i}$ is the superposition of all seasonal factors $d_{t,i}$ included in the model, as described in section 3.2.

The formula (4.4) defines a quite general model that might still be optimised by a least square estimator as shown by Pierce (1971). However, to identify the model it is necessary to identify the orders of the ARIMA process as well as the appropriate seasonal factors and their lags. This cannot be done in one step, hence a more general methodology should be applied.

4.2 Seasonal ARIMA Model of Czech Currency in Circulation

The model described in this section was identified using the two-step approach proposed by Koreisha and Pukkila (1998) repetitively. This means that the appropriate form of the difference operator Δ was found first. Once the difference operator was known, the regression part of the model was identified and after that the form of the autoregressive and moving averages

[3] For any chosen order of AR process there exists an MA process of infinite order, such that the superposition of these two processes is equal to the given weakly stationary process.

operators was estimated. The last two steps were repeated until all the parameters included in the model were not classified as insignificant.

According to the procedure outlined above, the sample autocorrelation and partial autocorrelation functions (SACF, SPACF) were computed and investigated to specify the difference operator. First, the correlogram of the series of the CIC volume indicated that the daily changes in CIC have to be examined instead of the CIC volume. Peaks with one year frequency and linearly decreasing amplitude were then detected in the correlogram of the differenced series, so the year difference was also applied. Altogether, the values of the SACF and SPACF indicated Δ in the form:

$$\Delta = (\boldsymbol{I} - \boldsymbol{B})\left(\boldsymbol{I} - \boldsymbol{B}^{252}\right)$$

where $\left(\boldsymbol{I} - \boldsymbol{B}^{252}\right)$ is a first-order seasonal difference operator that corresponds to the one-year difference.

Apart from the seasonal differencing, seasonal factors were included in the regression part of the model to deal with the calendar variations. The factors included were selected from the set of variables summarised in section 3.2. However, only factors that were not refused by a significance test remained in the final model. An overview of the seasonal factors and their lags that pass the tests is provided in table 4.1.

The number of subsequent non-working days was ultimately not included in the model. The out-of-sample forecasting performance of the model with this dummy variable was the same as without it. Therefore, we decided not to use this variable in the benchmark ARIMA model, as it makes the interpretation of the influence of the number of non-working days more difficult.

The ARMA structure of the stochastic component was investigated simultaneously with the identification of the appropriate set of seasonal factors. The lags of the MA and AR processes were chosen with respect to the SACF and SPACF. Finally the autoregressive operator:

$$\Phi = \left(\boldsymbol{I} - \boldsymbol{B} - \boldsymbol{B}^{7} - \boldsymbol{B}^{9}\right)\left(\boldsymbol{I} - \boldsymbol{B}^{20}\right)\left(\boldsymbol{I} - \boldsymbol{B}^{42}\right)\left(\boldsymbol{I} - \boldsymbol{B}^{65}\right)$$

and the moving averages operator:

$$\Theta = \left(\boldsymbol{I} - \boldsymbol{B}^{10} - \boldsymbol{B}^{15}\right)\left(\boldsymbol{I} - \boldsymbol{B}^{252}\right)$$

were identified.

The final ARIMA model is described by 71 parameters. The correlogram in figure 6.4 shows that a tiny correlation still remained in the residuals. Also the Ljung-Box test (e.g. Hamilton (1994)) indicates a higher correlation particularly around lags 60 and 252, although the addition of appropriate lags does not improve the forecasting performance of the model. Finally, the Jarque-Bera test (e.g. Hamilton (1994)) disproved the normality of standardised residuals and verified the difference of $N(0,1)$ density and the residuals histogram (see figure 6.1).

Table 4.1: Seasonal factors and shocks included in the ARIMA model

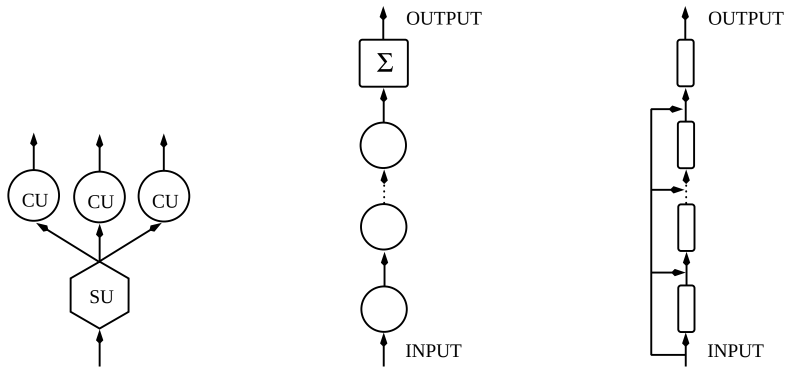
seasonal factors, shocks	$\Gamma(\boldsymbol{B})\,\boldsymbol{B}^{-F}$	**sines / cosines frequencies**	**number of parameters**
intra-monthly effect	—	$\frac{1}{4\pi}, \frac{1}{6\pi}, \frac{1}{8\pi}, \frac{1}{10\pi}, \frac{1}{12\pi}$ $\frac{1}{2\pi}, \frac{1}{4\pi}$	7
day of week	$\omega_0\boldsymbol{I} + \omega_1\boldsymbol{B} + \omega_2\boldsymbol{B}_2 + \omega_3\boldsymbol{B}_3$	—	4
Easter	$(\omega_0\boldsymbol{I} + \omega_2\boldsymbol{B}_2 + \omega_3\boldsymbol{B}_3 + \omega_4\boldsymbol{B}_4$ $+\omega_5\boldsymbol{B}_5 + \omega_6\boldsymbol{B}_6 + \omega_7\boldsymbol{B}_7 + \omega_8\boldsymbol{B}_8)\,\boldsymbol{B}^{-4}$	—	8
fixed holiday	$(\omega_0\boldsymbol{I} + \omega_2\boldsymbol{B}_2 + \omega_3\boldsymbol{B}_3 + \omega_4\boldsymbol{B}_4$ $+\omega_5\boldsymbol{B}_5 + \omega_6\boldsymbol{B}_6 + \omega_7\boldsymbol{B}_7$ $+\omega_8\boldsymbol{B}_8 + \omega_9\boldsymbol{B}_9)\,\boldsymbol{B}^{-4}$	—	9
non-working days	—	—	—
Christmas	$(\omega_0\boldsymbol{I} + \omega_2\boldsymbol{B}_2 + \omega_3\boldsymbol{B}_3 + \omega_4\boldsymbol{B}_4$ $+\omega_5\boldsymbol{B}_5 + \omega_6\boldsymbol{B}_6 + \omega_7\boldsymbol{B}_7 + \omega_8\boldsymbol{B}_8$ $+\omega_9\boldsymbol{B}_9 + \omega_{10}\boldsymbol{B}_{10} + \omega_{11}\boldsymbol{B}_{11} + \omega_{12}\boldsymbol{B}_{12}$ $+\omega_{13}\boldsymbol{B}_{13})\,\boldsymbol{B}^{-3}$	—	14
New Year	$(\omega_0\boldsymbol{B}_0 + \omega_1\boldsymbol{B}_1 + \omega_2\boldsymbol{B}_2)\,\boldsymbol{B}^{-3}$	—	3
bank failure	$(\omega_0\boldsymbol{I} + \omega_1\boldsymbol{B}_1 + \omega_2\boldsymbol{B}_2 + \omega_3\boldsymbol{B}_3$ $+\omega_4\boldsymbol{B}_4 + \omega_5\boldsymbol{B}_5 + \omega_6\boldsymbol{B}_6 + \omega_7\boldsymbol{B}_7$ $+\omega_8\boldsymbol{B}_8 + \omega_9\boldsymbol{B}_9 + \omega_{10}\boldsymbol{B}_{10} + \omega_{11}\boldsymbol{B}_{11}$ $+\omega_{12}\boldsymbol{B}_{12})\,\boldsymbol{B}^{-4}$	—	13
Y2K	$(\omega_0\boldsymbol{B}_0 + \omega_1\boldsymbol{B}_1 + \omega_4\boldsymbol{B}_4 + \omega_5\boldsymbol{B}_5)\,\boldsymbol{B}^{-5}$	—	4
total	—	—	62

5. Neural Network Model

Since the ARIMA model is linear in seasonal effects, it might not sufficiently cover the calendar variation. Therefore, the application of non-linear models has been increasingly investigated recently. Unfortunately, the functions realised by typical neural networks such as MLP or RBF networks are higher-order superpositions of non-linear functions with a large number of parameters. Consequently, the application of neural networks raises many doubts, as they are too complicated and dealing with them is like dealing with black boxes.

To avoid such doubts, the concept of neural networks with switching units was chosen as an alternative approach. Switching units allow the use of linear transfer functions and the model as a whole is then a combination of two common stochastic methodslinear regression and cluster analysis. Such a model is relatively simple and can be further analysed using common stochastic tools.

Figure 5.1: Neural Network with Switching Units



(a) Neuron with a switching unit (SU) and three computational units (CU)

(b) Basic string topology

(c) Hyper-topology of the feedforward structured network with basic strings as nodes

The feedforward structured neural network described in this section is derived from the neural network with switching units introduced in Bitzan, Šmejkalová, Kučera (1995). For the purposes of this paper, this original model is referred to as the *Basic String*. The generalisation is based on the connection of more basic strings into a hyper structure. In the following subsections the final neural network model is described step by step.

5.1 Neuron with Switching Unit

The main idea of a neuron with a switching unit is to control the flow of the input. Any neuron with a switching unit consists of one switching unit and several computational units as shown in figure 5.2(a). The switching units only choose which computational unit will process the given input, while the computational units apply a transfer function to the inputs.

The switching unit splits the input space into several disjoint clusters. The number of clusters is the same as the number of computational units, and each cluster is associated with a different computational unit. Inputs from a cluster are processed only by the associated computational unit. The clusters are found during the training process using cluster analysis methods. The character of the clusters depends on the metric or pseudometric of the input space. This metric or pseudometric, together with number of computational units, defines the switching unit.

Computational units are neurons in the common sense. They could be of any type, such as a perceptron or RBF unit, but such a general case might lead again to a black-box model. Therefore, only computational units with linear transfer functions are considered. The linear transfer function is defined by the formula:

$$(y_0, y_1, \ldots, y_n) = (\alpha_0, \alpha_1 x_1, \ldots, \alpha_n x_n) \tag{5.5}$$

where $y = (y_0, \ldots, y_n)$ is an output, $x\,(x_1, \ldots, x_n)$ is an input, and parameter $\alpha = (\alpha_0, \ldots, \alpha_n)$ is estimated from the training set using the linear regression equation:

$$Y = X\alpha + \epsilon. \tag{5.6}$$

where Y is a column vector of the modelled quantity, X is a matrix of appropriate inputs, the column vector α is the estimated parameter, and ϵ is the error or residuum. This implies that the sum of the components of output y from (5.5) is the approximation or estimation of the correct value of the modelled quantity. Nevertheless, the row vector y, instead of the sum of its components, is the output of the neuron, because summing the components leads to the loss of information contained in the vector.

5.2 Basic String

The original model of a neural network with switching units proposed in Bitzan, Šmejkalová, Kučera (1995) is quite a simple one. It is referred to here as a basic string because of its string topology. The topology is defined by a graph that is a path exactly. This means that every neuron except the input and output ones has one parent and one child. The input neuron is a neuron with a switching unit with a Euclidean norm and linear transfer function. The output neuron only sums its inputs, and the rest are neurons with switching units with clusters defined by the sum of the input components and with a linear transfer function. A sketch of the string architecture is provided in figure 5.2(b). The Euclidean norm is used in the input neuron because it reflects the original structure of the network input space. The rest of the neurons split the inputs according to the expected output, as the sum of the input components is the approximation of the modelled quantity.

Even though these models are simple, they are quite effective. From the theoretical point of view, strings are capable, for example, of approximating any smooth or measurable function (Hlaváček (2002)), and they can realise any AR process as well. The model has also been applied to several problems with encouraging results (Hakl et al. (2003)). However, more complicated problems need more complicated models. Therefore, more general models called feedforward structured networks have been defined as models with a topology described by a hyper-structure with basic strings as its nodes (see 5.2(c)).

The advantage of the hyper-structure is that the use of basic strings as nodes helps to define and control the topology. This is particularly important in the context of genetic optimisation (see Kalous (2004)), but also when the topology is defined according to the results of models derived by an expert.

5.3 Feedforward Structured Neural Network for CIC Forecasting

The model used for CIC modelling is a simple feedforward structured network with the hyper-structure described by a graph that is a path extended by additional connections between the network input and each basic string as shown in figure 5.2(c). The main role of the additional connections is to keep the original structure in the analysed data, because the input space structure can be heavily damaged by non-invertible transfer functions. Although this model

Table 5.1: Summary of final neural network model topology

string order	string length	number of clusters	string order	string length	number of clusters
1	1	4	6	3	3,3,1
2	4	2,2,2,2	7	2	4,3
3	1	4	8	3	2,2,2
4	4	2,3,2,2	9	3	3,2,2
5	1	3	10	2	2,2

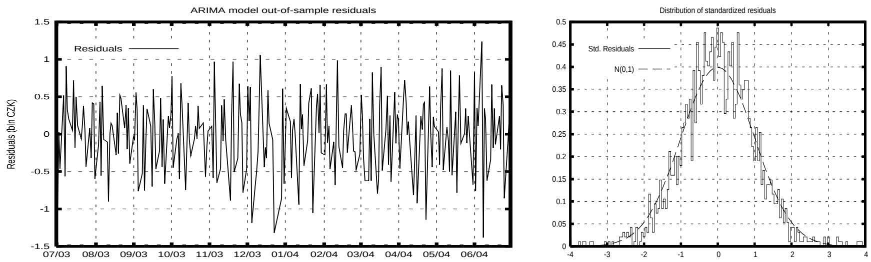
still cannot realise an MA process, the additional connections allow the use of more neurons effectively, hence better analysis of the input data is possible.

The number of strings in the network itself, their length, and the numbers of clusters in the neurons can be chosen arbitrarily, and unfortunately no guidelines on how to choose the best combination are available. The neural network model described here is a product of an iterative process based on analysis and comparison of the derived models. First, a few randomly generated networks were analysed and, according to their performances, some restrictions were applied. The length of the strings was limited to five neurons and the number of clusters per neuron was restricted to between two and five. In line with these restrictions, ten networks were randomly generated and then particular basic strings from the best four networks were combined in different orders until the final topology described in table 5.1 was chosen as the best one.

All the inputs summarised in table 3.1, together with the lagged values of CIC, comprise the input to the neural network model. The lags included in the model were selected according to the SACF and SPACF functions (see fig.6.5) using the same methodology as the one used for the Box-Jenkins ARIMA model construction (Box and Jenkins (1976),Hamilton (1994)). However, the inability of the neural network model to realise the MA process was also considered, and hence, for example, five lags around one year instead of a single one are used to deal with the strong one-year correlation. Finally, the following lags were used: 1, 2, 5, 10, 15, 20, 21, 22, 250, 251, 252, 253 and 254.

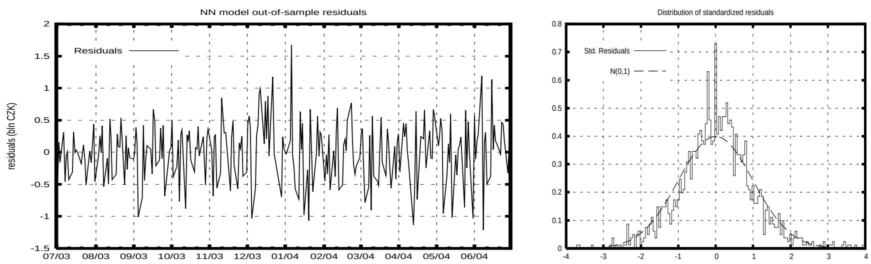
The model was then applied to the series of daily changes and to its one-year seasonal difference. The daily changes were forecasted with significantly higher accuracy then the seasonally differenced series. The probable explanation is that the seasonally differenced series contains a prominent MA(252) component, which can be hard to approximate using the neural network. On the other hand, the non-linear neural network model might stabilise the non-stationary series of CIC daily changes, as it approximates the seasonal character of the series with higher accuracy than the linear regression model.

Figure 6.1: ARIMA out-of-sample residuals



(a) Residuals (b) Histogram

Figure 6.2: Neural network out-of-sample residuals



(a) Residuals (b) Histogram

Anyway, the correlogram of the residuals (see fig. 6.5) does not suggest a strong correlation of the residuals. As in the case of the ARIMA model, additional lags do not significantly improve the model performance. Histogram in fig. 6.2) show the residuals are hardly normal which was also confirmed by normality tests. Normality is also rejected even if the peaks around zero are removed, as they are caused by clusters with only a few observations that might be classified as outliers.

6. Comparison

Two models—an ARIMA one and a neural network one—were described in the previous sections 4 and 5. In addition to the model definitions and a description of their functionality, the results of the in-sample residual analysis are presented there. Moving the analysis along, we discuss and compare the out-of-sample forecasting performance and the general applicability of the two models in this section. First, the forecasting performance is compared, and then a discussion of the applicability follows.

The sample used for forecasting performance qualification is the one-year period from July 2003 to June 2004 and is the same for both models. The comparison is focused on one-step-ahead forecasts, because experiments showed that the forecasting horizon does not affect the relative performance of the models. The starting point for the analysis is the one-step-ahead residual plots (see figures 6.1 and 6.2 respectively).

Figure 6.3: Differences between the ARIMA and the NN model squared residuals

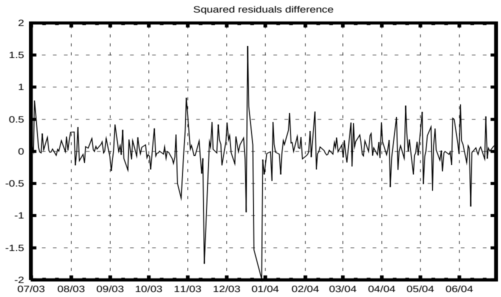


Table 6.1: RMSE and Diebold-Mariano test results

horizon	ARIMA RMSE	NN RMSE	D-M p-value
1	0.491	0.454	0.975
5	0.476	0.442	—
10	0.484	0.448	—

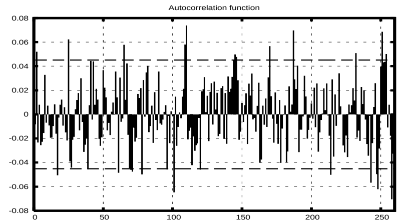
Figures 6.2 and 6.3 show that the neural network model badly miscalculates the forecast in a few cases, particularly around Christmas. Focusing on these unfitted events, it was found they all fell into a cluster with only a few observations in it. This means that the neural network cannot be reliably applied to such sparse events, as it is strongly overlearned with regard to them. However the table 6.1 shows that the neural network is more accurate on average for all the horizons considered, although the Diebold-Mariano test (e.g. Hamilton (1994)) does not classify the difference as significant (the test p-value is also reported in table 6.1).

Comparison with the CNB expert on the same data is not possible, as the ARIMA model was already used as a supportive tool for the prediction. However, the RMSE for the period from July 2002 to June 2003 was 0.66, while the forecasting performance of the neural network and the ARIMA model was almost the same as presented in table 6.1.

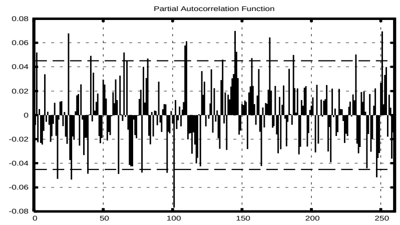
The neural network model outperforms the ARIMA model particularly at the beginning of the testing sample, where the neural network RMSE is really low. This can be viewed in figure 6.6, which shows the RMSE for the particular months. The figure 6.6 also indicates that the average error changes during the forecasted period in both models. The growth of the error at the end of the period is probably caused by obsolescence of the models, while the changes in the middle of the period correspond to the Christmas season, whose effect might not be well approximated.

Another interesting fact is that the forecast error does not increase with the forecast horizon. On the contrary, the RMSE of both models for the five-day horizon is lower than that for the one-day horizon. This apparent paradox means that the models cannot approximate intra-

Figure 6.4: Correlogram of ARIMA in sample residuals

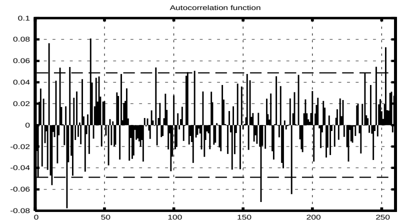


(a) Autocorrelation Function

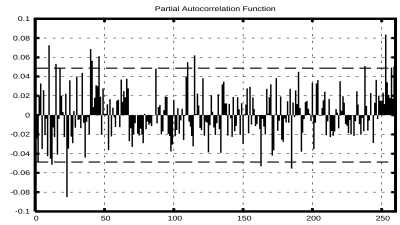


(b) Partial Autocorrelation Function

Figure 6.5: Correlogram of NN in sample residuals

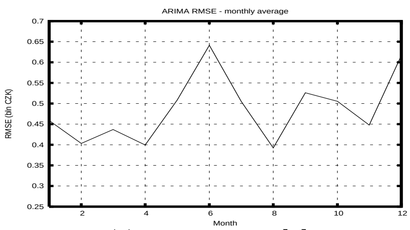


(a) Autocorrelation Function

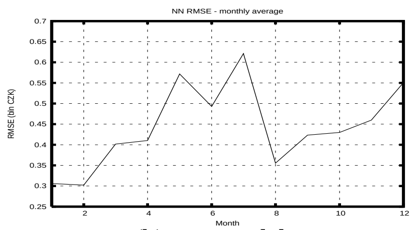


(b) Partial Autocorrelation Function

Figure 6.6: RMSE in particular months



(a) ARIMA model



(b) NN model

weekly effects with sufficient accuracy. The reason is probably that the intra-weekly effect changes a lot during the period.

Moving to the general point of view, the interpretability of the results is the next important attribute of any model. The regression component of the ARIMA model (4.4) allows the approximation of the seasonal influence to be analysed effectively. This cannot be done easily for the neural network model, as the influence of the given season might be modelled miscellaneously by different computational units. On the other hand, proper analysis of the neural network model is also possible and might lead to more specific findings.

Conversely, the advantage of the neural network application is that it can easily be reoptimised even if the set of exogenous factors changes. In the case of the ARIMA model, any change in the set of inputs considered means that the model has to be rebuilt completely. However, the topology of a neural network model can be preserved and it is only necessary to learn the network again using the new set of inputs. The learning process of a neural network is then a fully automated compact algorithm that can be run without any expert.

7. Conclusion

The paper introduces a new kind of neural network model for currency in circulation forecasting. A feedforward structured neural network model that might be suitable for the analysis of arbitrary seasonal time series is compared with the more conventional Box-Jenkins ARIMA model. The characteristic properties of both models were discussed, with an emphasis on out-of-sample forecasting performance.

First it was found that both models are more stable and accurate than the forecasts of CNB experts, although the in-sample residuals do not meet the standard conditions of normal distribution and a weak correlation still remains in the residuals. The analysis of out-of-sample residuals then showed that the neural network model outperforms the ARIMA model on average and particularly in the first free month of the testing sample. However the Diebold-Mariano test does not classify the difference as significant. On the other hand, the neural network model badly fits a few observations in the testing sample, probably because the model is overlearned in sparse observations. Nonetheless, the neural network model is a competitive alternative to the Box-Jenkins ARIMA models and is worth improving.

Regarding the properties of the structured neural network model, a few improvements that might improve the model are obvious. First, feedback would be included to deal with MA processes. Second, the selection of relevant inputs could be improved through the application of stochastic tests. Next, a more general network architecture might be considered simultaneously with the use of genetic algorithms for architecture optimisation. Finally, a tool for analysing the network, data flows and other model properties might make the model more transparent.

Although all these extensions would improve the model performance, applying the neural network model in its current stage of development, or, better still, combining it with the ARIMA model, is relevant at least as a supportive tool for liquidity forecasting.

References

ANDERSON, R. L. (1954): "The problems of Autocorrelation in Regression Analysis." *Journal of the American Statistical Association,* 49:113–129.

BALKIN, S. D. (1997): "Using Recurrent Networks for Time Series Forecasting." Working paper 11, Pennsylvania State University

BELL, W. R. AND S. C. HILLMER (1984): "Modeling Time Series With Calendar Variation." *Journal of the American Statistical Association,* 78:526–534.

BOX, G. E. P. AND G. M. JENKINS (1976): *Time Series Analysis: Forecasting and Control.* San Francisco: Holden Day.

BITZAN, P., ŠMEJKALOVÁ J., AND M. KUČERA (1995): "Neural Networks with Switching Units." *Neural Network World,* 4:515–523.

CABRERO, A., CAMBA-MENDEZ, G., HIRSCH, A., AND F. NIETO (2002): "Modelling the Daily Series of Banknotes in Circulation in the Context of the Liquidity Management of the European Central Bank." Working paper 142, European Central Bank

DIEBOLD, F. X AND R. S. MARIANO (1995): "Comparing Predictive Accuracy." *Journal of Bussines and Economic Statistics,* 13:253–263.

DORFFNER, G. (1996): "Neural Networks for Time Series Processing." *Neural Network World,* 4:447–468.

GOURIEROUX, CH. (1997): *Time Series and Dynamic Models*Series Temporelles at Modeles Dynamiques by Economica. *Cambridge University Press.*

HAMILTON, J. D. (1994): *Time Series Analysis.* Princeton University Press.

HAKL, F., HLAVÁČEK, M., AND R. KALOUS (2003): "Application of Neural Networks to Higgs Boson Search." *Nuclear Instruments nad Methods in Physics Research,* 489–491.

HALLAS, M. AND G. DORFFNER (1998): "A Comparative Study on Feedforward and Recurrent Neural Networks in Time Series Prediction Using Gradient Descent Learning." In *Proceedings of 14th European Meeting on Cybernetics and Systems Research,* Vienna.

HARVEY A., KOOPMAN, S. J., AND M. RIANI (1997): "The Modeling and Seasonal Adjustment of Weekly Observations." *Journal of Bussines & Economic Statistics,* 15(3): 354–368.

HLAVÁČEK, M. (2002): "Design and Analysis of Neural Network with Switching Units Suitable for the Study of Elementary Particlas Decay with the Application of Genetic Algorithms." Master thesis, Czech Technical University, Prague.

HORNIK, K., STINCHCOMBE, M., AND H. WHITE (1989): "Multilayer feedforward networks are universal approximators." *Neural Networks,* 2:359–366.

HILLMER, S. C. AND G. C. TIAO (1982): "An ARIMA-Model-Based Approach to Seasonal Adjustment." *Journal of the American Statistical Association,* 77(377):63–70.

KALOUS, R. (2004): "Evolutionary Operators on Neural Networks Architecture." In *Proceedings of PhD. Conference 2004*, Prague. Institute of Computer Science, MatfyzPress.

KOREISHA, S. G. AND T. PUKKILA (1998): "A Two-Step Approach for Identifying Seasonal Autoregressive Time Series Forecasting Models." *International Journal of Forecasting*, 14:483–496.

PIERCE, D. A. (1971): "Least Squares Estimation in the Regression Model with Autoregressive-Moving Average Errors." *Biometrika*, 58(2):299–312.

TRAPLETTI, A., LEISCH, F., AND K. HORNIK (1999): "On the Ergodicity and Stationarity of the ARMA(1,1) Recurrent Neural Network Process." Working paper 37, SFB, Vienna.

CNB WORKING PAPER SERIES

11/2005	Marek Hlaváček: Michael Koňák Josef Čada	*The application of structured feedforward neural networks to the modelling of daily series of currency in circulation*
10/2005	Ondřej Kameník:	*Solving SDGE models: A new algorithm for the sylvester equation*
9/2005	Roman Šustek:	*Plant-level nonconvexities and the monetary transmission mechanism*
8/2005	Roman Horváth:	*Exchange rate variability, pressures and optimum currency area criteria: Implications for the central and eastern european countries*
7/2005	Balázs Égert: Luboš Komárek	*Foreign exchange interventions and interest rate policy in the Czech Republic: Hand in glove?*
6/2005	Anca Podpiera: Jiří Podpiera	*Deteriorating cost efficiency in commercial banks signals an increasing risk of failure*
5/2005	Luboš Komárek: Martin Melecký	*The behavioural equilibrium exchange rate of the Czech koruna*
4/2005	Kateřina Arnoštová: Jaromír Hurník	*The monetary transmission mechanism in the Czech Republic (evidence from VAR analysis)*
3/2005	Vladimír Benáček: Jiří Podpiera Ladislav Prokop	*Determining factors of Czech foreign trade: A cross-section time series perspective*
2/2005	Kamil Galuščák: Daniel Münich	*Structural and cyclical unemployment: What can we derive from the matching function?*
1/2005	Ivan Babouček: Martin Jančar	*Effects of macroeconomic shocks to the quality of the aggregate loan portfolio*

10/2004	Aleš Bulíř: Kateřina Šmídková	*Exchange rates in the new EU accession countries: What have we learned from the forerunners*
9/2004	Martin Cincibuch: Jiří Podpiera	*Beyond Balassa-Samuelson: Real appreciation in tradables in transition countries*
8/2004	Jaromír Beneš: David Vávra	*Eigenvalue decomposition of time series with application to the Czech business cycle*
7/2004	Vladislav Flek, ed.:	*Anatomy of the Czech labour market: From over-employment to under-employment in ten years?*
6/2004	Narcisa Kadlčáková: Joerg Keplinger	*Credit risk and bank lending in the Czech Republic*
5/2004	Petr Král:	*Identification and measurement of relationships concerning inflow of FDI: The case of the Czech Republic*
4/2004	Jiří Podpiera:	*Consumers, consumer prices and the Czech business cycle identification*
3/2004	Anca Pruteanu:	*The role of banks in the Czech monetary policy transmission mechanism*
2/2004	Ian Babetskii:	*EU enlargement and endogeneity of some OCA criteria: Evidence from the CEECs*

1/2004	Alexis Derviz: Jiří Podpiera	*Predicting bank CAMELS and S&P ratings: The case of the Czech Republic*
12/2003	Tibor Hlédik:	*Modelling the second-round effects of supply-side shocks on inflation*
11/2003	Luboš Komárek: Zdeněk Čech Roman Horváth	*ERM II membership – the view of the accession countries*
10/2003	Luboš Komárek: Zdeněk Čech Roman Horváth	*Optimum currency area indices – how close is the Czech Republic to the eurozone?*
9/2003	Alexis Derviz: Narcisa Kadlčáková Lucie Kobzová	*Credit risk, systemic uncertainties and economic capital requirements for an artificial bank loan portfolio*
8/2003	Tomáš Holub: Martin Čihák	*Price convergence: What can the Balassa–Samuelson model tell us?*
7/2003	Vladimír Bezděk: Kamil Dybczak Aleš Krejdl	*Czech fiscal policy: Introductory analysis*
6/2003	Alexis Derviz:	*FOREX microstructure, invisible price determinants, and the central bank's understanding of exchange rate formation*
5/2003	Aleš Bulíř:	*Some exchange rates are more stable than others: Short-run evidence from transition countries*
4/2003	Alexis Derviz:	*Components of the Czech koruna risk premium in a multiple-dealer FX market*
3/2003	Vladimír Benáček: Ladislav Prokop Jan Á. Víšek	*Determining factors of the Czech foreign trade balance: Structural Issues in Trade Creation*
2/2003	Martin Čihák: Tomáš Holub	*Price convergence to the EU: What do the 1999 ICP data tell us?*
1/2003	Kamil Galuščák: Daniel Münich	*Microfoundations of the wage inflation in the Czech Republic*
4/2002	Vladislav Flek: Lenka Marková Jiří Podpiera	*Sectoral productivity and real exchange rate appreciation: Much ado about nothing?*
3/2002	Kateřina Šmídková: Ray Barrell Dawn Holland	*Estimates of fundamental real exchange rates for the five EU pre-accession countries*
2/2002	Martin Hlušek:	*Estimating market probabilities of future interest rate changes*
1/2002	Viktor Kotlán:	*Monetary policy and the term spread in a macro model of a small open economy*

CNB Research and Policy Notes

CNB Economic Research Bulletin

Czech National Bank
Economic Research Department
Na Příkopě 28, 115 03 Praha 1
Czech Republic
phone: +420 2 244 12 321
fax: +420 2 244 14 278
http://www.cnb.cz
e-mail: research@cnb.cz