

**Moscow State University of Economy, Statistics and Information Technologies**

**Center for Interdisciplinary Scientific Research,  
Managerial Innovations and Information Technologies**

**Roman Melnikov**

**The Problem of the Yield Curve Shifts Prediction:  
A Neural Network Approach.**

**Moscow, Russian Federation, February 1999.**

## **1. The Problem of the Yield Curve Shifts Prediction within the Backpropagation Neural Network Framework.**

The procedure of the yield curve shifts prediction is one of the most important elements of the bonds portfolio active management system. The traditional method of forecasting is based on informal groundwork. The future shift of the curve in this approach is estimated via the expert appraisal of the impact of various economical, political and market factors on the yield curve position. The decision maker is guided by definite logic rules constructed on the base of recent experience and external knowledge which provide reasonable estimations in the cases of market regularity violation and significant structural changes in the macroeconomical environment.

The financial markets of highly developed countries of Europe and America function under the conditions of relatively stable environment, and their parameters fluctuate within definite ranges. Therefore the laws of their evolution may be formally studied. Accumulated experience falls for processing not only by the mind of a human being but also by mathematical techniques which provide an opportunity to make optimal estimations. A convenient tool enabling to formalise historical experience and to use it for forecasting purposes is the backpropagation neural network.

The idea of neural networks employment for analysis and prediction of the financial market conjuncture accorded wide recognition in 1990th<sup>1</sup>. Nevertheless, formalized technologies of neural network applications creation for many standard analytical problems have not been developed (or opened for usage of the general public).

Firstly, practitioners which successfully develop efficient forecasting models are not interested in revelation of their «know-how». Secondly, distributors of neurosoftware and neurocomputers are interested to position neural networks as very simple, miraculous and self-sufficient tool (the author's opinion is that this statement is nothing more than the advertisement dodge). Thirdly, neural network is a «young» method of financial market forecasting (in the comparison with technical and fundamental analysis), and universal methodology of its application has not been developed because of the short history of neurocomputing in finance. Therefore the problem of development of the analytical technology of creation of the yield curve shift prediction neural network model is the task of current interest.

First steps on a road to the solution of this problem were made by the author in the beginning of 1998. The matter of interest was the prediction of Russian GKO (short-term zero-coupon bonds) prices for several trade sessions. The general principles of this problem solution for volatile emerging markets were formulated as a result of series of experiments:

1. Tuning of the model should be carried out on the basis of the whole spectrum of bond's issues. It enables to employ all available data and apply price history of matured bonds for prediction of new bonds price trajectory.

2. Model specification should incorporate the principle of the yield curve smoothness. The violation of parity relations between prices (and interest rates) of bonds with close maturity date engenders strictly predictable correcting processes.

3. Model framework should reflect the micro-cyclical nature of emerging bonds market conditioned by infrastructure particularity.

4. Input variables vector should include the leading indicators of adjacent markets and the inflow of capital to bonds market.

Practical realization of these principles enabled to achieve satisfactory quality of short-term prediction of GKO market conjuncture in the first quarter of 1998.

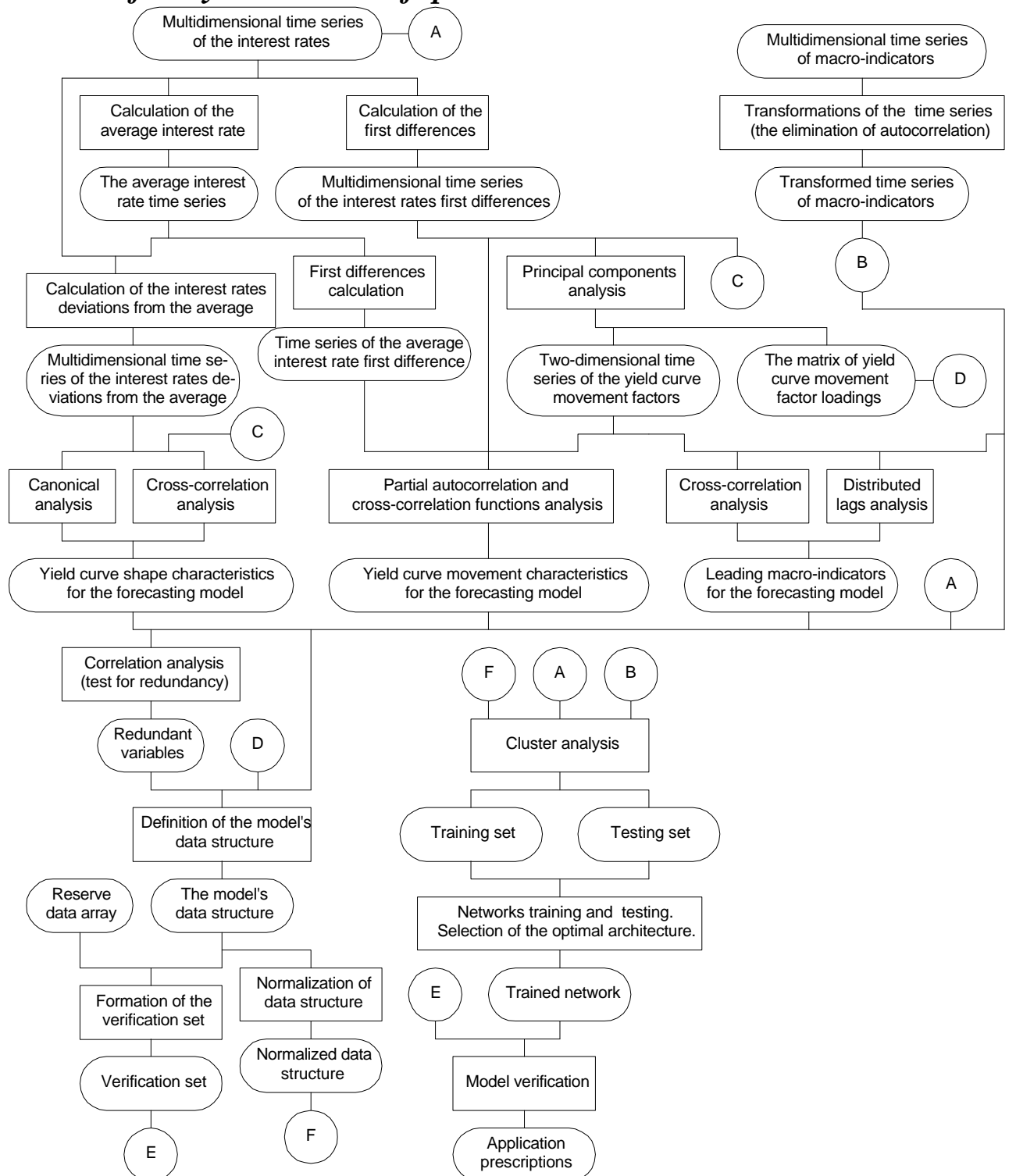
The abrupt increase of GKO market volatility in May of 1998 eliminated any possibility of its conjuncture forecasting on the basis of any formal techniques. The results of research in adjacent areas testified the inevitability of the Russian ruble devaluation and the swell of the Russian governmental debt refinancing difficulties. Both the neural network analysis and the fundamental macroeconomical analysis provided strong arguments in favor of possible default hypothesis. The events of August 1998 confirmed the validity of these conclusions.

---

<sup>1</sup> Baestaens, Van den Bergh & Wood (1994), «Neural Network Solutions for Trading in Financial Markets», London, Longman.

Russian financial crises does not permit to demonstrate applied analytical technology on a domestic example. Therefore the author decided to construct a neural network model capable to predict how the US Treasury obligations yield curve will move in a month.

**Figure 1. The analytical technology of the construction of the yield curve shift prediction neural network model.**



On the one hand, the US bonds market is one of the most stable financial markets in the world, and therefore an attempt to forecast its principal indicators with the help of a neural network model looks quite reasonable. On the other hand, this market is considered one of the most efficient. This circumstance impugns the possibility of its conjuncture forecasting on the basis of free-for-all data. Moreover, numerous attempts to predict the interest rates of the US bonds using classical econometrics techniques failed to achieve any significant result<sup>2</sup>. In this connection any attempt to

<sup>2</sup> Sharpe, Alexander & Bailey (1995), «Investments», New Jersey, Englewood Cliffs.

obtain the witness of market inefficiency within the neural networks approach is a matter of great theoretical and practical importance.

Available data<sup>3</sup> includes monthly time series of financial and macroeconomical indicators of the USA covering a period from May 31, 1990 to August 31, 1998. The yield curve is represented by the interest rates of 9 composite bonds with 3 month, 6 month, 1 year, 2 year, 3 year, 5 year, 7 year, 10 year and 30 year maturity. The macroeconomical data includes the aggregates of money supply such as M1, M2, M3 and quasi-money, the size of the governmental debt, the indexes of consumer prices (CPI), wholesale prices (WPI) and producer prices (PPI), the unemployment rate and the exchange rates of the US dollar to Deutschmark (DM) and Japanese yen.

These indicators do not cover all essential aspects of the USA macroeconomical situation which influence the Treasury obligations market. For example, it is beyond any doubt that the stock market indexes and the Gross Domestic Product (GDP), not included into the list of available indicators, are factors of great importance. However, the lack of GDP in data array does not deteriorate the quality of the forecasting model, because, according to Okun's law, there is a strong linear regression dependence between the changes of GDP and unemployment rate in the United States.

Before running any analytical procedures we have to exclude some part of data from the available array. Otherwise the opportunity of strict model verification would be lost, because the capability of the forecasting model to predict on the basis of data used to construct this model does not testify its high level of generalization and efficiency. Therefore 12 last observations were excluded from the available multidimensional data array to provide a basis for further verification.

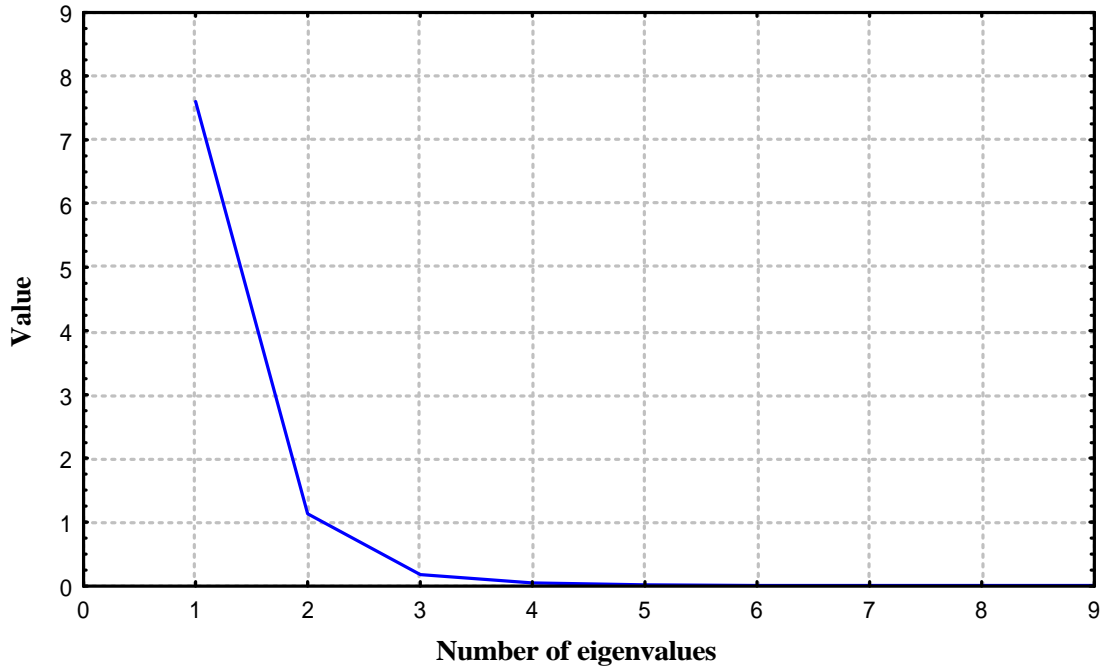
## **2. Revealing the Leading Indicators of the Yield Curve Movement Process.**

The process of leading indicators selection is the most difficult stage of the neural network application development. The matter is that only part of available data may be efficiently used for prediction purposes. Some of available indicators have no significant relation with the term structure of the interest rates, and others are the subject of the yield curve impact and therefore cannot be used within the forecasting model. We have to run special preliminary procedures to find out which variables should be included into the model, and what their lag structure should be. Various tools of applied statistics and econometrics should be used at this stage of analysis.

The consistent patterns we are looking for can be better revealed if the dimension of the yield curve movement characteristics vector (formed by the first differences of nine-dimensional time series of interest rates) is reduced. The principal components analysis was applied to find the solution of this problem.

---

<sup>3</sup> Source: Reuters (Equities 3000 & Fixed Income 3000 databases).

*Figure 2. Plot of eigenvalues*

The Kaiser criterion (only the factors with eigenvalues greater than 1 are retained) was used to determine how many principal components should be extracted. Two retained principal components accounted for 97% of total variance, and the varimax normalized rotational strategy was used to obtain a clear pattern of loadings.

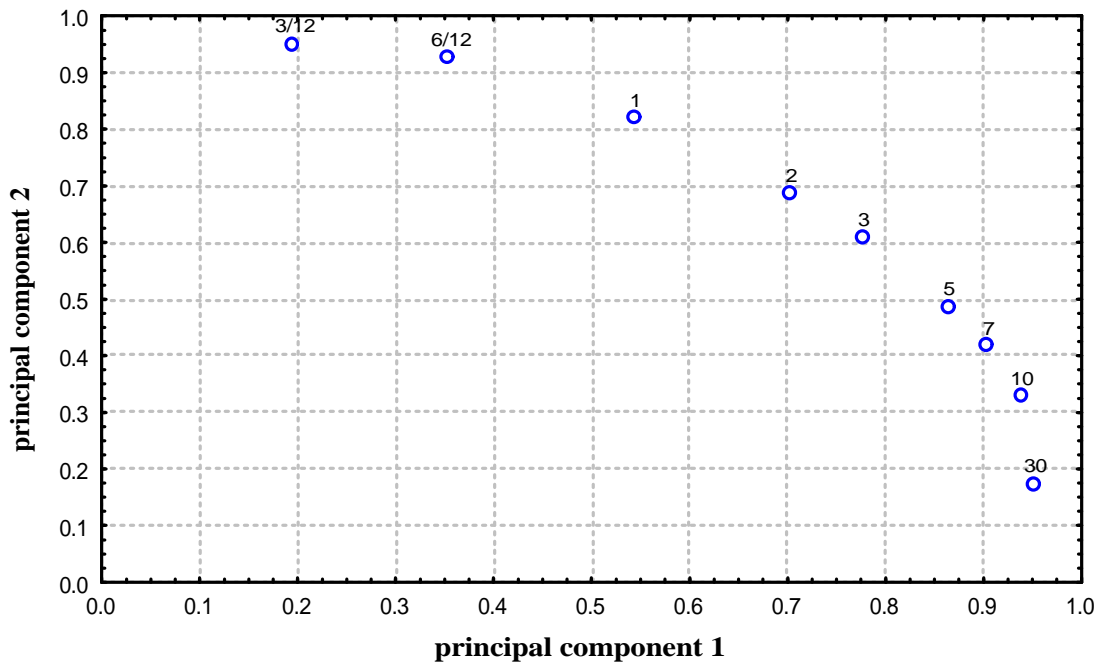
*Figure 3. Plot of rotated factor loadings*

Figure 3 demonstrates the rotated factor loadings. The changes of the short-term rates are highly correlated with the first factor, and the changes of the long-term rates – with the second factor. Therefore the first factor should be interpreted as the hidden determinant of the short-term rates movement, and the second factor – as the determinant of the long-term rates movement.

It is difficult to draw a strict demarcation line between the long-term and the short-term bonds for the purposes of the yield curve shifts analysis and forecast. Both are influenced in some degree by long-term and short-term factors. It is an important argument in favour of constructing a universal model to forecast changes in the yields to maturity for different segments of the term structure of the interest rates. At the same time it should be noted that the shifts of the long-term

rates are characterized by unique features in a lesser degree than the shifts of the short-term ones. The difference between the factor loadings for bonds with 3 and 6 month maturity is greater than between the factor loadings for bonds with 10 and 30 year maturity.

The rotated factors are useful tools for studying the relations between the macroeconomical indicators and the interest rates movements. The starting point of the research will be the relation between the aggregates of money supply and the yield curve shifts.

Various macroeconomical theories provide strong arguments in favour of the hypothesis that the increase of money supply leads to the decrease of the interest rates. The first of them is probably the indirect gear of money supply influence on prices, first formulated by H.Thornton in 1802<sup>4</sup>. It implies that the increase of money supply is accompanied by the increase of funds offered by the lending institutions to borrowers at the capital market and therefore to the decrease of the interest rate in the short-term period. However, the interest rate returns to the former level very soon. The reason is that the excess of the profit rate over the fallen interest rate leads to the increase of investments, the growth of prices for capital goods and, as a consequence, to the increase of demand for loans, because any new project requires more financial resources.

According to the monetary model of the borrowing funds, developed in 1937 by B.Olin and D.H.Robertson, the increase of money supply leads to the growth of current demand for bonds and, as a result, to the fall of the interest rates. The Keynesian theory of money demand which accorded wide recognition also implies the fall of the interest rates after the increase of money supply, except one special case - «the liquid trap», when under the conditions of the rock-bottom interest rates and clear-cut expectations of their increase the aggrandizement of money supply is not capable to generate additional demand for governmental bonds.

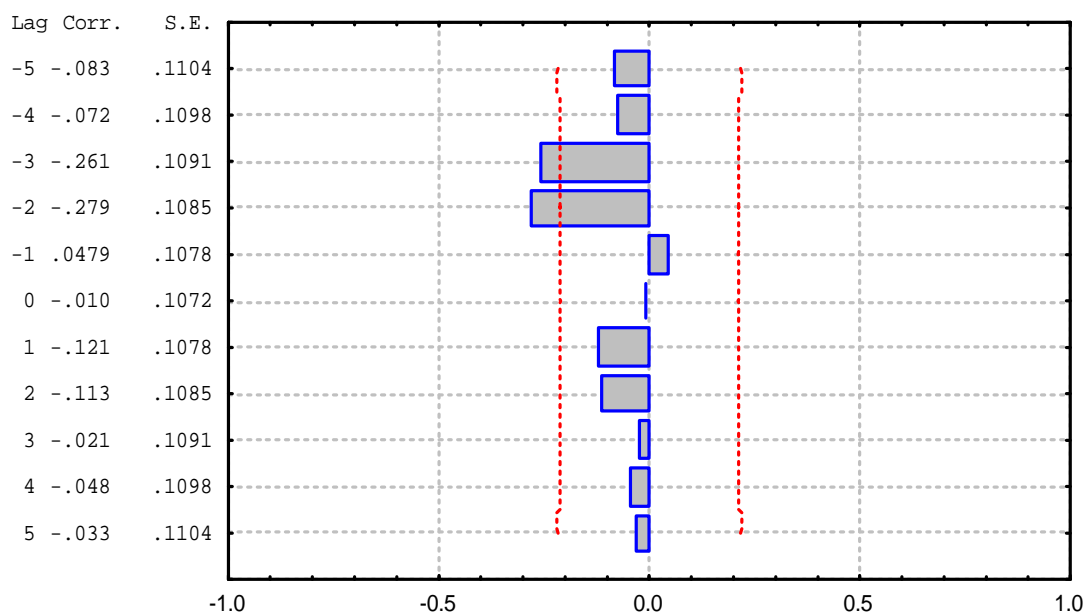
Theoretical reasons provide solid grounds to consider money supply a useful leading indicator of the yield curve shifts. But this presupposition fails to bear the empirical test.

The aggregates of M1, M2, M3 and quasi-money are characterized by a very strong serial correlation (autocorrelation coefficient for the monthly lag is equal to 0.970, 0.944, 0.950 and 0.977 correspondingly). Therefore the initial series should be transformed by calculation of their rates of increase. This transformation eliminates autocorrelation for M1 and M2 (autocorrelation coefficients become equal to 0.043 and -0.047) and significantly reduces it for M3 and quasi-money (the coefficients of the transformed series are 0.310 and 0.521 correspondingly). Now the apparatus of cross-correlation analysis may be used to study the character of interrelation between the money supply (explanatory variable) and the interest rate movement (dependent variable).

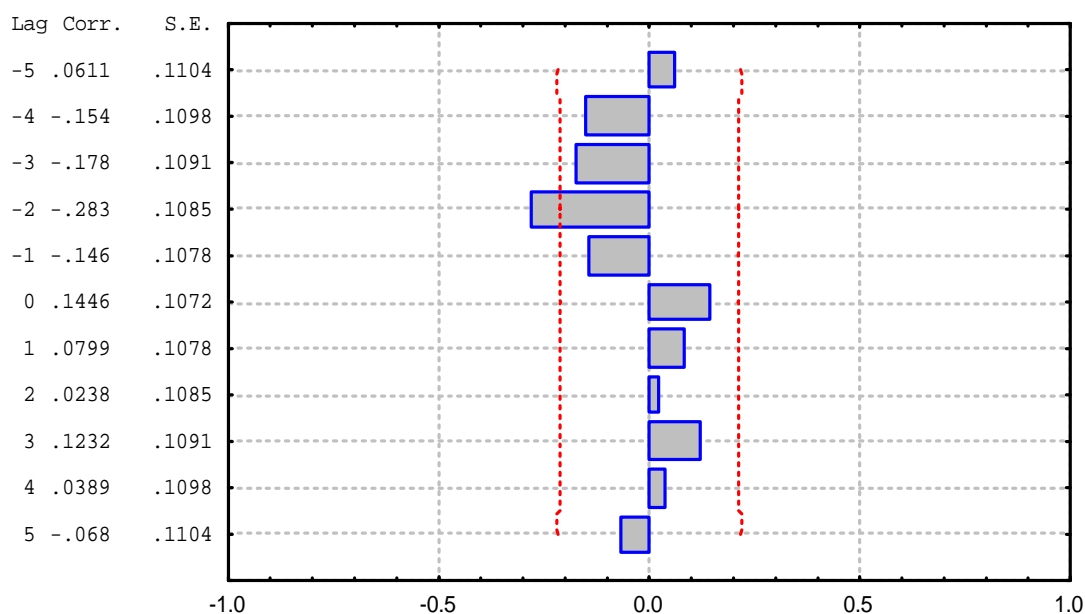
---

<sup>4</sup> Blaug (1985), «Economic theory in retrospect», Cambridge University Press.

**Figure 4. Cross-correlation between the rate of increase of M1 and the short-term rates movement factor**



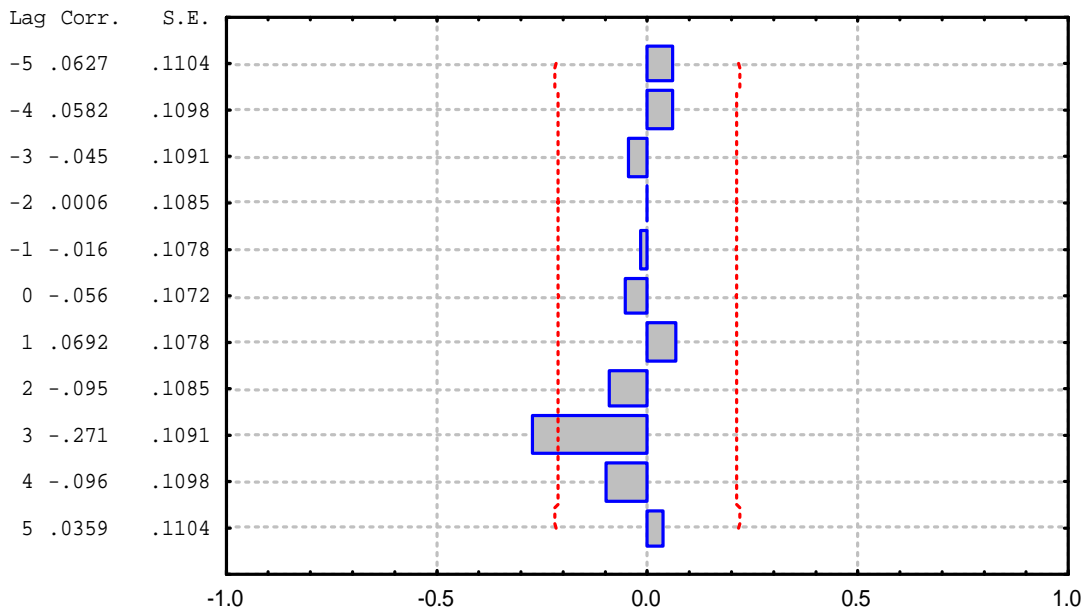
**Figure 5. Cross-correlation between the rate of increase of M3 and the long-term rates movement factor**



Correlation coefficients with the largest absolute values are observed in the negative lags area. Therefore the money supply is not a leading, but a lagging indicator for the yield curve movements. This conclusion contradicts theoretical models and makes information about money supply dynamics absolutely useless for prediction of the interest rates changes. The author's opinion is that the remarked phenomenon may be explained by the divergence between the assumption of money supply inelasticity used in the above-mentioned theories and the tactical purposes of the Federal Reserve System determining the priority of the interest rates over the money supply.

The analysis of structural shifts in the money supply provides no useful leading indicator for the interest rates changes. Therefore another group of macroeconomical indicators – the indexes of producer price, wholesale price and consumer price – should be tested. As the price indexes are strongly expressed trending indicators, their series should be transformed previously by calculation of the rates of increase.

**Figure 6. Cross-correlation between the rate of increase of PPI and the short-term rates movement factor**



Cross-correlation analysis testifies that the relation between the quarterly lagged transformed producer price index and the short-term interest rates movement factor has the highest level of statistical significance. The acceleration of producer prices rise leads to the fall of the short-term interest rates in a quarter. A similar relation is observed between the 4-month lagged transformed consumer price index and the long-term interest rates movement factor.

The opinion of the author is that this dependence is stipulated by the cyclic character of macroeconomical processes and the features of the inflationary expectations formation. Peak of the prices rise having stayed in the past enables to expect low inflation rate in the future; recent stabilization of prices generates expectations of their growth and acts as a factor of the interest rates increase.

The next macroeconomical indicator rendering important effect to the yield curve is the unemployment rate. The influence of this variable is very specific, and we should run more powerful technique – the distributed lags analysis – to study it carefully. The purpose of this technique is the search of leading indicators for the interest rates changes, and therefore the zero lag should be excluded from the consideration.

Table 1.

The levels of statistical significance for betas of the lagged explanatory variables for relation between the change of unemployment rate and the long-term rates movement factor

Lag order number	The total number of considered lags			
	2	3	4	5
1	<b>0.7970</b>	<b>0.9378</b>	0.7281	0.6449
2	0.0472	0.0603	0.0455	0.0260
3	-	0.0980	0.1026	0.0719
4	-	-	0.2458	0.2137
5	-	-	-	0.2843



Table 2.

The levels of statistical significance for betas of the lagged explanatory variables for relation between the change of unemployment rate and the short-term rates movement factor

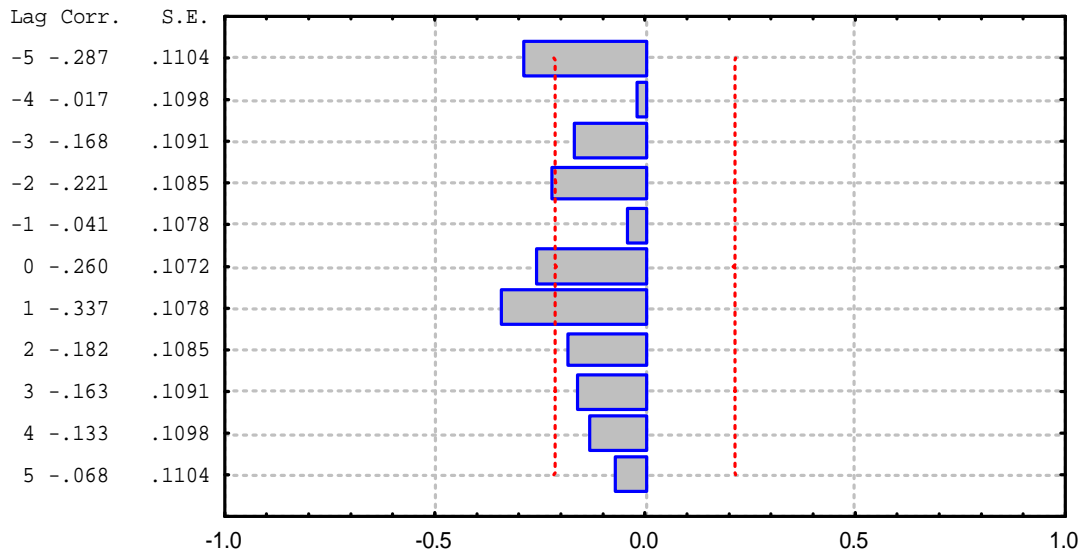
Lag order number	The total number of considered lags			
	2	3	4	5
1	0.0463	0.0238	0.0307	0.0320
2	0.6439	0.7554	0.7560	0.0680
3	-	0.0434	0.0479	0.0476
4	-	-	0.8645	0.8658
5	-	-	-	0.6285

Tables 1 and 2 demonstrate the levels of statistical significance for betas of the lagged explanatory variables. Tested model specification includes the factor of the yield curve shift as the endogenous variable and the change of the unemployment rate as the exogenous variable. The levels of statistical significance for variables with positive betas are given in black and for variables with negative betas are given in red.

Tables 1 and 2 testify that long-term and short-term bonds react to the change of unemployment rate quite differently. The long-term interest rates increase and the short-term rates decrease after the rise of unemployment rate during the last quarter. The fall of unemployment rate leads to the decrease of the long-term rates and to the growth of the short-term rates. Therefore the slope of the yield curve is significantly effected by recent changes in unemployment rate.

The revealed consistent pattern falls for logical explanation. The decrease of employment compels Federal Reserve System to put expansionary policy into effect: to buy Treasury bills at the open market and to pull down the discount rate. These measures find reflection in the decrease of the short-term interest rates. But market valuation of the adequate long-term rates is based on the expectations of economic growth, a great demand for borrowing funds and stabilizational policy of expensive money in the future, because the peak of unemployment has stayed in the past. On the contrary, the recent increase of employment is traditionally accompanied by restrictionary monetary policy aimed to prevent the overheat of economy and by the expectations of cyclic recession with low level of the short-term interest rates in the future. This explanation is confirmed by p-values of the lagged variables. The short-term rates movement factor is predicted on the basis of monthly (most recent changes) and quarterly lagged explanatory variable, while the long-term rates movement factor is influenced with lags of 2 month and especially 3 month.

**Figure 7. Cross-correlation between  
the rate of increase of the governmental debt  
and the short-term rates movement factor**



The phenomenon of influence of the governmental debt growth to the yield curve shifts sets a very interesting task for a financial analyst. The rates of the governmental debt increase have negative correlation with the short-term rates movement factor and no correlation with the long-term rates movement factor. The peak of correlation coefficient between the rate of increase of governmental debt and the shift factor of the short-term interest rates is achieved at monthly lag.

This interrelation may be partly explained by the actions of the Federal Reserve System pressing for the counteraction to investment displacement effect through putting down the Treasury bills yields in the conditions of expansionary fiscal policy connected with the increase of governmental debt. Besides, the negative correlation between the short-term rates movement factor and no correlation with the long-term rates movement factor testify the growth of the yield curve slope in the conditions of governmental debt growth. According to the expectations theory of the yield curve formation it should be interpreted as the result of market expectation of the short-term rates increase in the future.

Time series of Deutschmark and Japanese yen exchange rates to the US dollar do not affect the yield curve shifts. On the contrary, they are influenced by the changes in the American interest rates and play the role of lagging indicators. Thus, the whole spectrum of available macroeconomical indicators has been tested for presence of statistically significant relations with the rotated principal components of the yield curve shifts. The result of these tests are 4 leading indicators for the short-term interest rates movement factor such as the rate of increase of producer price index (lag 3), the changes of unemployment rate (lags 1 and 3), the rate of increase of governmental debt (lag 1) and 3 indicators useful for the long-term interest rates forecasting such as the rate of increase of consumer price index (lag 4) and the changes of unemployment rate (lags 2 and 3).

As all available macroeconomical parameters have been studied, additional leading indicators may be revealed only through processing of the yield curve parameters. First of all, partial autocorrelation functions of the rotated principal components should be studied.

**Figure 8. Partial autocorrelation function  
of the long-term rates movement factor**

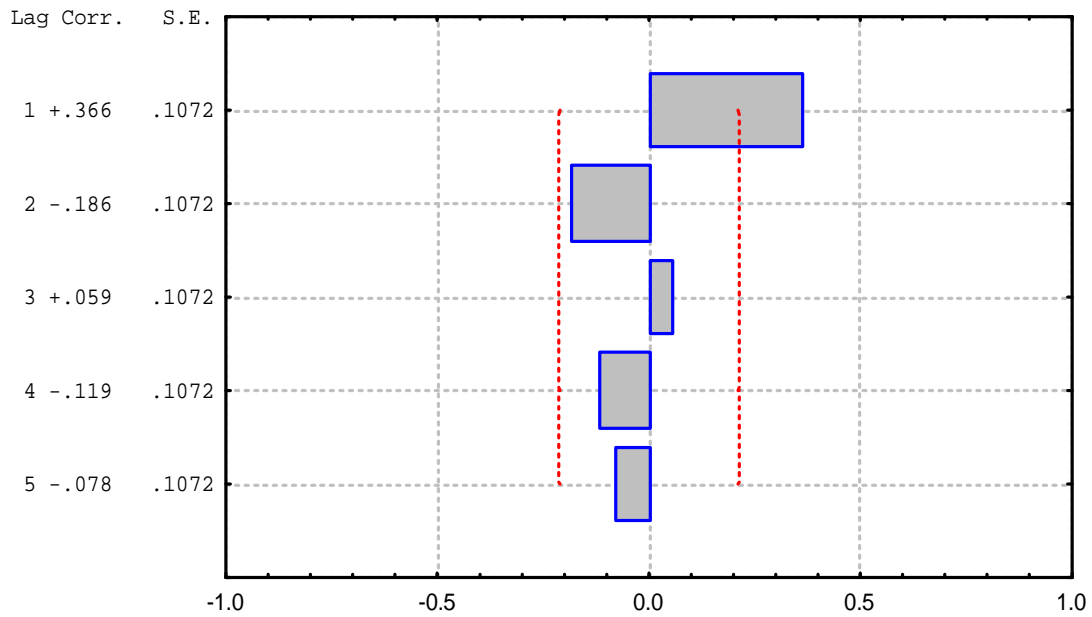


Figure 8 shows that autocorrelation coefficient of the long-term interest rates movement factor is statistically significant at 5% p-level only for monthly lag. Partial autocorrelation functions of the other rotated principal component and the interest rates first differences are of the same character, but the significance of autocorrelation for monthly lag increases with the decrease of bond's maturity.

The changes of yields of both the short-term and the long-term bonds are determined by the common laws of movement of every point of the yield curve. Therefore the lagged values of the average interest rate change (calculated on the basis of all representatives of the yield curve) may serve as very useful indicators for forecasting purposes.

Table 3.

The levels of statistical significance for betas of the lagged explanatory variables for relation between the change of average interest rate and the long-term rates movement factor

Lag order number	The total number of considered lags			
	2	3	4	5
1	0.0029	0.0042	0.0034	0.0046
2	0.0079	0.0825	0.0581	0.0757
3	-	0.3186	0.8550	0.7806
4	-	-	0.1937	0.4006
5	-	-	-	0.4736

Table 4.

The levels of statistical significance for betas of the lagged explanatory variables for relation between the change of average interest rate and the short-term rates movement factor

Lag order number	The total number of considered lags			
	2	3	4	5
1	0.0004	0.0001	0.0001	0.0001
2	0.0976	0.7782	0.7756	0.8252
3	-	0.0281	0.0650	0.0592
4	-	-	0.9503	0.7941
5	-	-	-	0.6065

The distributed lags analysis proves the validity of this hypothesis. Lags of 1 and 2 months are statistically significant for the long-term interest rates movement factor, and lags of 1 and 3 months – for the short-term interest rates movement factor at 8% p-level.

According to the expectation theory of the yield curve formation gear, the future changes of the interest rates are dependent of the current term structure. The segmentation theory also assumes that significant contravention of parities between the yields of different segments of the term structure stipulates the deformation of the curve at the next steps of market functioning through the negative feedback gear caused by partial restructuring of institutional investors' portfolios. The task of revealing the most useful parameters of the current yield curve for short-term forecasting purposes may be solved through calculation of the correlation matrix describing the relations between the deviations of interest rates from the average and the changes of these rates in the next month.

Table 5.

The correlation coefficients between the deviations of interest rates from the average and the changes of these rates in the next month.

	$\Delta r(3m)$	$\Delta r(6m)$	$\Delta r(1y)$	$\Delta r(2y)$	$\Delta r(3y)$	$\Delta r(5y)$	$\Delta r(7y)$	$\Delta r(10y)$	$\Delta r(30y)$
d(3m)	-0.2723	-0.2662	-0.2184	-0.1436	-0.1110	-0.0694	-0.0338	-0.0185	0.0000
d(6m)	-0.1445	-0.1748	-0.1603	-0.1085	-0.0874	-0.0582	-0.0281	-0.0178	-0.0045
d(1y)	0.0392	-0.0183	-0.0639	-0.0537	-0.0469	-0.0321	-0.0110	-0.0044	0.0035
d(2y)	0.1493	0.0718	-0.0106	-0.0446	-0.0451	-0.0366	-0.0243	-0.0145	0.0001
d(3y)	0.3047	0.2154	0.0882	0.0008	-0.0227	-0.0323	-0.0331	-0.0241	-0.0068
d(5y)	0.1566	0.1627	0.1295	0.0618	0.0388	0.0099	-0.0139	-0.0168	-0.0131
d(7y)	0.0526	0.0966	0.1104	0.0821	0.0681	0.0484	0.0221	0.0178	0.0112
d(10y)	0.0338	0.0829	0.1102	0.0909	0.0776	0.0551	0.0290	0.0172	0.0052
d(30y)	0.0390	0.0895	0.1254	0.1131	0.0995	0.0721	0.0447	0.0264	-0.0025

Table 5 and the following ones apply the notation:  $T=3m...30y$  - bond's maturity,  $d(T)$  - the deviation of yield of the bond with maturity  $T$  from the average,  $\Delta r(T)$  - the change of yield of the bond with maturity  $T$  in the next month; statistically significant at 5% p-level correlation coefficients are given in red.

Table 5 testifies that the current yield curve has a greater influence to the change of the short-term rates than to the movement of yields of the long-term bonds. This conclusion corresponds with the soft variant of the expectation hypothesis of the yield curve formation. The leading role among the parameters of the yield curve as prediction indicators is played by the deviations of 3 year and 3 month interest rates from the average.

The canonical analysis, maximizing the correlation coefficients between the linear combinations of 8 interest rates deviations from the average (the 9<sup>th</sup> one is redundant, because the sum of all deviations is equal to zero) and the changes of 9 interest rates in the next month, confirms the significance of 3 year and 3 month interest rates as leading indicators.

Table 6.

Canonical loadings of the endogenous variables for relation between the interest rates deviations from the average and the changes of these rates in the next month.

Rate change	Canonical root's order number							
	1	2	3	4	5	6	7	8
$\Delta r(3m)$	0.76	0.35	-0.10	-0.13	-0.17	-0.27	-0.13	0.00
$\Delta r(6m)$	0.63	0.47	0.11	-0.09	-0.20	-0.24	-0.13	-0.10
$\Delta r(1y)$	0.38	0.44	0.07	-0.09	-0.30	-0.30	-0.16	-0.18
$\Delta r(2y)$	0.20	0.47	0.11	-0.06	-0.28	-0.42	-0.09	-0.06
$\Delta r(3y)$	0.13	0.43	0.15	-0.13	-0.26	-0.41	-0.06	0.00
$\Delta r(5y)$	0.02	0.46	0.15	-0.13	-0.23	-0.34	-0.03	0.11
$\Delta r(7y)$	0.01	0.38	0.15	-0.08	-0.17	-0.29	-0.02	0.17
$\Delta r(10y)$	-0.05	0.39	0.12	-0.11	-0.05	-0.22	0.04	0.17
$\Delta r(30y)$	-0.09	0.27	0.09	-0.06	-0.02	-0.09	0.20	0.17

Table 7.

Canonical loadings of the exogenous variables for relation between the interest rates deviations from the average and the changes of these rates in the next month.

Deviation	Canonical root's order number							
	1	2	3	4	5	6	7	8
d(3m)	-0.34	-0.23	0.08	0.14	0.26	-0.12	-0.22	0.82
d(6m)	-0.16	-0.20	-0.06	0.09	0.24	-0.16	-0.15	0.90
d(1y)	0.11	-0.09	-0.11	0.02	0.20	-0.03	0.01	0.96
d(2y)	0.29	-0.02	-0.20	-0.14	0.11	0.27	0.15	0.87
d(3y)	0.53	0.13	-0.29	-0.13	0.03	0.51	0.32	0.49
d(5y)	0.23	0.15	-0.03	-0.11	-0.17	0.28	0.25	-0.86
d(7y)	0.02	0.18	0.07	-0.06	-0.19	0.10	0.17	-0.94
d(10y)	-0.02	0.13	0.10	-0.02	-0.25	-0.01	0.09	-0.95

According to the values of canonical loadings of the interest rates changes, the first canonical root characterizes relation between the yield curve shape and the movement of the short-term rates in the next month, the second canonical root describes dependence between the current term structure and its general shift in the next month, and the other canonical roots have no meaningful economical interpretation. The deviations of 3 month and 3 year interest rates have the largest by absolute value and contrary by sign correlation coefficients with the first canonical variable. Moreover, 3 month interest rate has the maximal by absolute value correlation coefficient with the second canonical variable, contrary by sign to the corresponding parameter of 3 year interest rate. Because the deviations of 3 year and 3 months interest rates from the average play the role of useful leading indicators for the yield curve shifts, the spread (or difference) between them should be included into neural network forecasting model as the input variable.

Now the following system of the leading indicators for the yield curve shifts may be set (variable identifiers assume the above-mentioned transformations of the original time series):

- the rate of increase of producer price index with lag 3 PPI(3);
- the rate of increase of consumer price index with lag 4 CPI(4);
- the changes of unemployment rate with lags 1, 2 and 3 UR(1), UR(2) and UR(3);
- the rate of increase of governmental debt with lag 1 Debt(1);
- the changes of average yield with lags 1, 2 and 3 CM(1), CM(2) and CM(3);
- the spread between 3 year and 3 months interest rates with lag 1 TS(1).

Table 8.

Correlation matrix of the leading indicators of the yield curve shifts.

	PPI(3)	CPI(4)	UR(1)	UR(2)	UR(3)	Debt(1)	CM(1)	CM(2)	CM(3)	TS(1)
PPI(3)	1.0000	0.2648	0.0187	0.0388	0.0684	0.0525	-0.1547	-0.0150	0.0222	-0.1576
CPI(4)	0.2648	1.0000	0.2480	-0.1151	-0.3715	0.2427	-0.1707	0.0732	-0.0400	-0.0847
UR(1)	0.0187	0.2480	1.0000	0.0222	-0.1640	0.1101	-0.1577	-0.2451	0.0121	-0.1085
UR(2)	0.0388	-0.1151	0.0222	1.0000	0.0322	0.1528	-0.1548	-0.1631	-0.2545	-0.1265
UR(3)	0.0684	-0.3715	-0.1640	0.0322	1.0000	0.0661	0.1260	-0.1485	-0.1586	-0.0329
Debt(1)	0.0525	0.2427	0.1101	0.1528	0.0661	1.0000	-0.1538	-0.1311	-0.0890	0.0944
CM(1)	-0.1547	-0.1707	-0.1577	-0.1548	0.1260	-0.1538	1.0000	0.4801	0.1379	0.3947
CM(2)	-0.0150	0.0732	-0.2451	-0.1631	-0.1485	-0.1311	0.4801	1.0000	0.4773	0.4510
CM(3)	0.0222	-0.0400	0.0121	-0.2545	-0.1586	-0.0890	0.1379	0.4773	1.0000	0.3824
TS(1)	-0.1576	-0.0847	-0.1085	-0.1265	-0.0329	0.0944	0.3947	0.4510	0.3824	1.0000

Correlation coefficients between the most of the revealed explanatory variables do not exceed 0.4 by their absolute value. The group of closely connected variables including the indicators of movement and shape of the yield curve such as CM(1), CM(2), CM(3) and TS(1) can be easily reviewed against this background. The special role within this group is played by CM(2) indicator which is characterized by the maximal correlation coefficients of 0.4801, 0.4773 and 0.4510 with 3 other variables of the group. Therefore this parameter should be considered as redundant and excluded from the list of explanatory variables.

The leading indicators for the yield curve shifts are revealed. The next step of the applied analytical technology is the specification of patterns representation for neural network model.

### 3. Neural Network Forecasting Models Specification.

There are three possible alternatives of model specification, each of which has its advantages and disadvantages. The first conceptual approach – «the market reaction model» – defines the output of the neural network as the vector of changes of 9 available interest rate indicators. The second one – «the yields reaction model» – assumes the inclusion of bonds identifier into the list of input variables and reduction of output vector to only one parameter characterizing the change of bond's yield. The third variant implies the construction of 9 various neural networks to model different reactions of different bonds to the changes of leading indicators.

In the opinion of the author, the third conceptual approach seems to be the least preferable. Firstly, the reactions for environmental changes are similar for different bonds, and the goal of the forecasting system construction may be achieved within a single model. Secondly, the interior parts of the yield curve may be modeled by other points of the term structure in senses of yield and duration, and therefore their parameters, including the structures of forecasting models, can be interpolated. Thirdly, the application of this approach reduces to minimum the number of observations and degrees of freedom of the model (the difference between the number of observations and the number of estimated parameters). Lastly, the construction of many different models requires an enormous expenditure of machine time and analyst's efforts (this argument seems to be crucial). Therefore a decision to reject this conceptual approach was made.

The first conceptual approach («the market reaction model») is based on the idea of complex relationship between the factors of market environment, inertia and structure and the multidimensional pattern of the yield curve. The identity of the interest rates reaction to environmental signals in this class of architectures finds reflection in the unity of weights of hidden neurons, and the specific character of each interest rate reaction – in the uniqueness of weights of output neurons.

The balance between the unique features and the common laws in the second approach («the yields reaction model») is achieved by displacement of emphasis from market parameters to intercross characteristics of bonds playing the role of representatives of different segments of the yield curve. It enables to significantly increase the degrees of freedom of the model and to include bond's unique characteristics into the model specification. However, this approach of one-dimensional output is connected with the danger of unification of imitated reaction to environmental

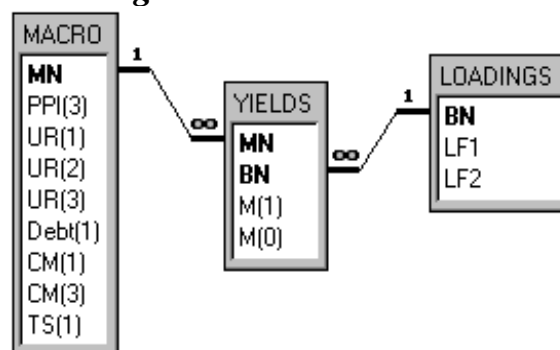
signals for different segments of the yield curve, because the larger part of parameters is identical for different input patterns fixed at the same date.

It is expedient to use loadings on the rotated principal components of the yield curve shift as bond's identifiers within the framework of the yields reaction model. The loading on the first rotated principal component may be considered as an indicator of bond's amenability to influence of the long-term rates determinants, and the second rotated principal component – as an indicator of bond's liability to the influence of the short-term rates determinants. Another useful input variable for the yields reaction model is the change of particular interest rate (lag 1), because the autocorrelation coefficient for this lag is statistically significant for all available time series of the interest rates first differences.

Now specifications of input and output vectors for the first and second conceptual approaches may be formulated. The market reaction model includes such input variables as the rate of increase of producer price index with lag 3 PPI(3); the rate of increase of consumer price index with lag 4 CPI(4); the changes of unemployment rate with lags 1, 2 and 3 UR(1), UR(2), UR(3); the rate of increase of governmental debt with lag 1 Debt(1); the average yield change with lags 1 and 3 CM(1), CM(3); the spread between 3 year and 3 month interest rates with lag 1 TS(1). The output pattern in this model specification is a nine-dimensional vector of the interest rates changes. The yields reaction model adds loadings on the rotated principal components of the yield curve shift LF1 and LF2 and the change of interest rate with lag 1 M1 – bond's unique parameters – to the list of input variables and reduces output to the change of the particular bond's interest rate with lag 0 M(0).

Then the structure of database required for estimation of neural network parameters of the wider yields reaction model should be analyzed. Its normalization enables to separate data into three tables connected by the secondary attributes. The fields of the first table MACRO are the month number MN (the unique identifier), PPI(3), CPI(4), UR(1), UR(2), UR(3), Debt(1),  $\tilde{N}_i(1)$ ,  $\tilde{N}_i(3)$ , TS(1). The second table YIELDS is composed of the month number MN and the bond number BN, forming the primary key, and also M(1) and M(0). The third table LOADINGS includes bond unique identifier BN and fields LF1 and LF2. The tables MACRO and YIELDS are connected by one-to-many relation on field MN, and the tables LOADINGS and YIELDS – by relation of the same type on field BN. The table MACRO consists of 84 records, the table LOADINGS consists of 9 records, and the table YIELDS – the process of interaction between objects MACRO and LOADINGS – includes  $84 \times 9 = 756$  records.

**Figure 9. Data structure.**



Random partitioning of available data into training and testing sets does not provide a reliable basis for construction of the yields reaction neural network forecasting model with good quality of generalization because of interior relations in the database. Almost all indicators would be presented both to training set and to testing set if random partitioning approach is run. Therefore lose of the testing benchmark performance would occur.

The special role played in data structure by table MACRO containing the maximal number of fields requires to partition the whole database into training and testing sets on the basis of this table. Besides, it is necessary to ensure proper reflection of all possible environmental conditions in the training set and especially in the testing set (the smaller of the two).

To satisfy these requirements, partitioning of the database was carried out by k-means clustering of MACRO records (without excluded identifiers) into 8 clusters. Then 8 observations selected from different clusters and representing different years and months were used to form a testing set together with the related records of tables YIELDS and LOADINGS. Other data formed a training set for the neural network forecasting application. The group of representatives of record set MACRO selected as a result of cluster analysis adequately characterizes various possible conditions of environment and different periods of market functioning, so it can be used for partitioning of data array into training and testing sets within the frameworks of both the market reaction model and the yields reaction model.

The models were constructed on the basis of classical backpropagation neural network architecture with sigmoid transfer functions and one hidden layer. The generalized Widrow-Hoff delta-rule with variable learning rate and fixed smoothing factor (equal to 0.9, as in the most neural network applications) was used as a learning law<sup>5</sup>. The learning rate tuning was based upon the ratio between «good» and «bad» output pattern estimations throughout the training set. The estimations of training examples outputs lying within the tolerance level interval (10% of variance range) were considered as good. The learning rate was tuned after each epoch of training by the rule

$$r = 1 - 0.9 * G / (B + G),$$

where  $r$  is a learning rate,  $B$  is a number of bad training patterns estimations and  $G$  is a number of good training patterns estimations. This learning strategy enables to descent quickly down the error surface at the beginning of training process and to adjust thoroughly weights of the network at the final stages of training. The correction of weights was carried out in the interactive mode after each run of the training pattern.

The choice of optimal architecture was made on the basis of series of experiments in training and testing of neural networks with different number of hidden neurons (from 2 to 12). The criterion used for the optimal architectural variant selection was the testing set standard error. Training continued for 5000 epochs, and test was run after each training epoch.

The results of experiments testify that the best quality of generalization is achieved by the neural networks with the architectures of 9–7–9 (the notation used is input neurons – hidden neurons – output neurons) for the market reaction model and 12–3–1 for the yields reaction model. These architectures enable to minimize the testing set standard error, which is decreased to the level of one half of the standard deviation of output variable values. The networks with less than 7 (for the market reaction model) and 3 (for the yields reaction model) neurons in the hidden layer are not able to significantly reduce the standard error even for the training set. The networks that have more hidden neurons are characterized by the redundant number of parameters. Therefore they achieved overtrained condition after the next to nothing number of epochs and lost the ability to generalize presented training patterns.

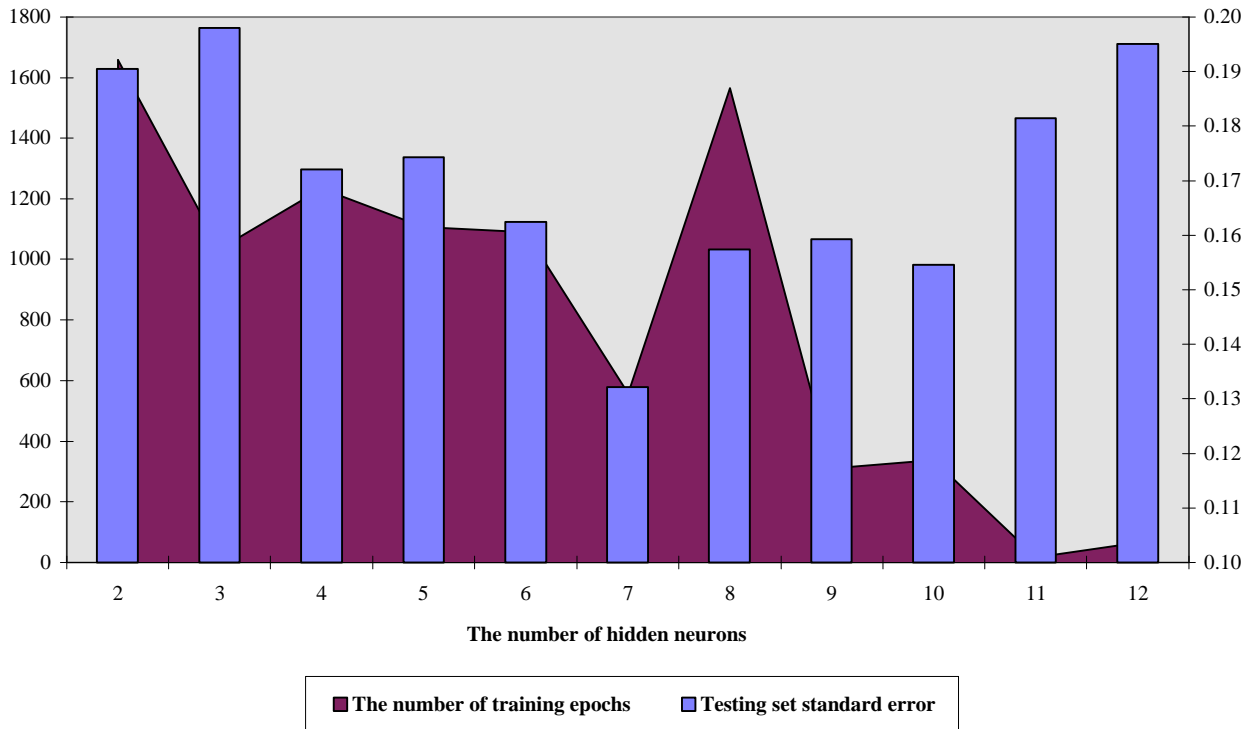
Neural networks with the multidimensional output require more neurons in the hidden layer than their analogues with the one-dimensional output to adequately reflect the relation between input and output variables. Networks with the multidimensional output achieve the overtrained condition earlier (after less epochs in the comparison with the one-dimensional output networks), because their training set has relatively small size and consists of disjoint observations.

---

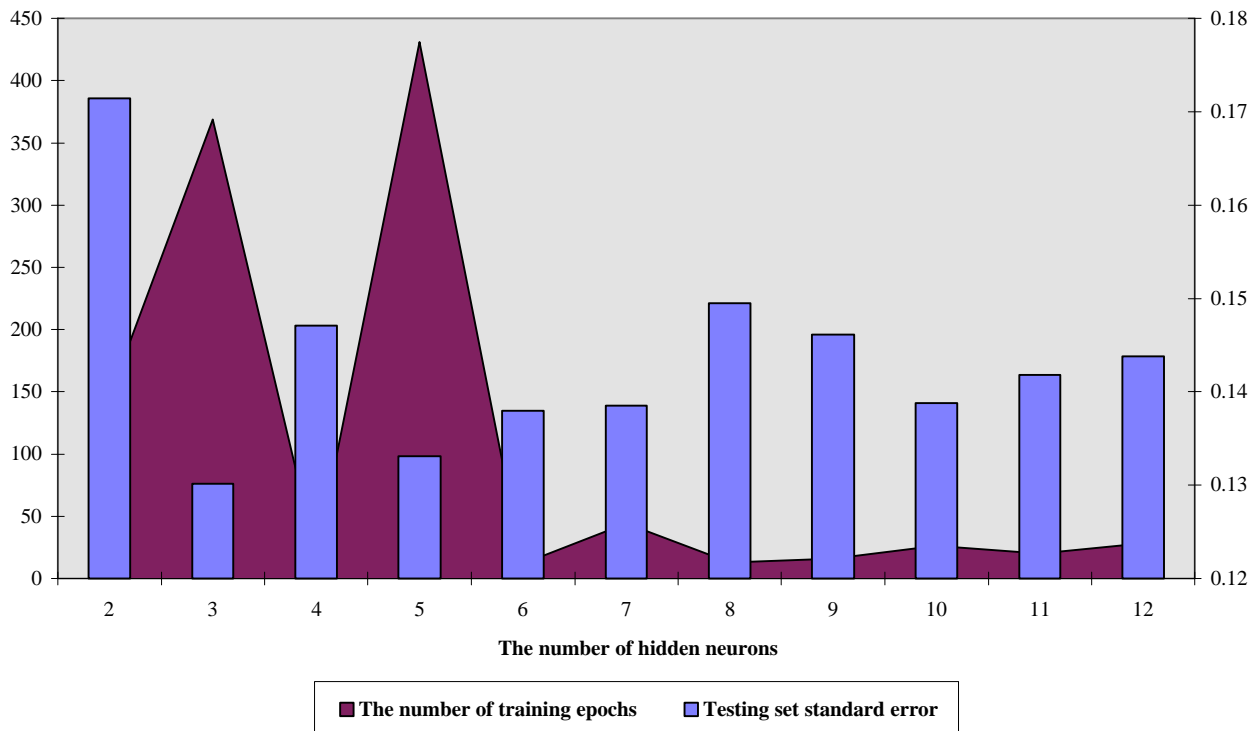
<sup>5</sup> Hecht-Nielsen (1991), «Neurocomputing», San-Diego, Addison-Wesley.



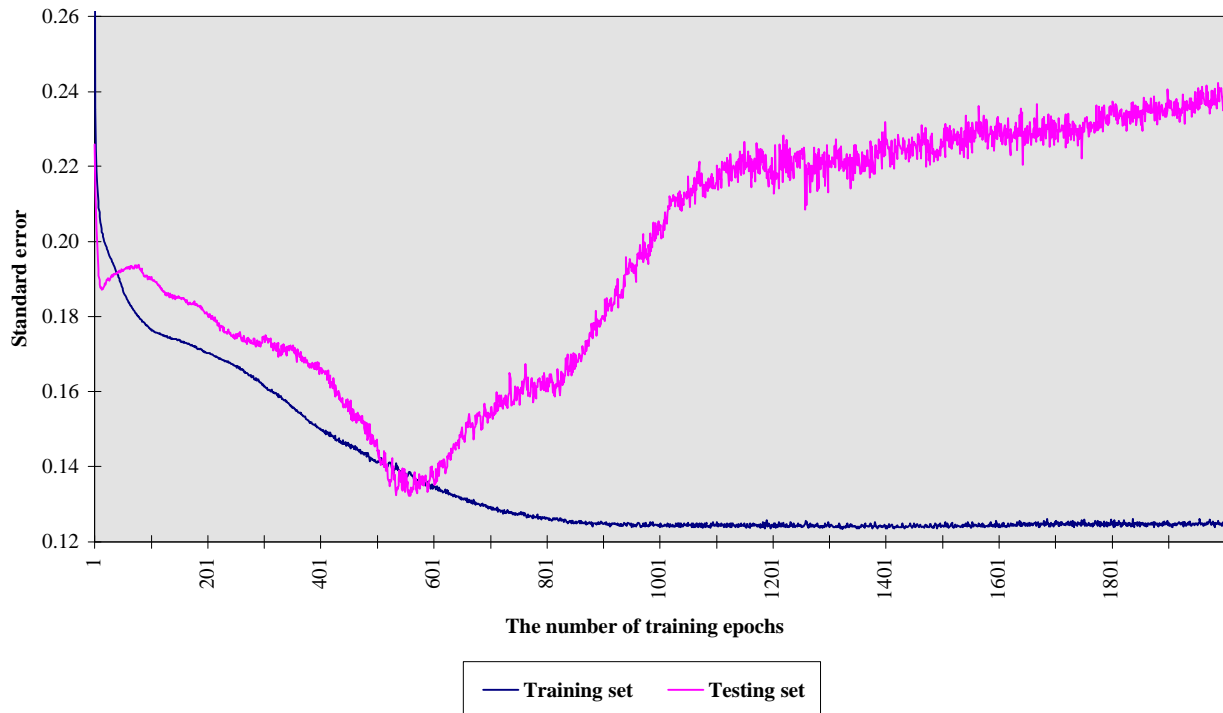
*Figure 10. Minimization of testing set standard error for the market reaction model.*



*Figure 11. Minimization of testing set standard error for the yields reaction model.*



The problem of overtraining significantly affects the technological process of neural network forecasting system construction. The training curve for the market reaction neural network model with the optimal (9-7-9) configuration is presented to illustrate this phenomenon.

*Figure 12. Training curve.*

#### 4. Neural Network Forecasting Models Verification.

The results of neural network forecasting models verification on the basis of market data that was not included into testing and training sets (for the period from 30.09.97 to 31.08.98) inspire definite, but moderate optimism. Correlation coefficient between factual and predicted interest rates throughout the whole verification set is equal to 0.4191 for the market reaction model and to 0.4798 for the yields reaction model, and these values may be considered as a proof of good efficiency of forecasting systems. However, estimations of the future yield curve shifts given by the networks are biased into the side of the interest rates overstatement.

The average monthly change of the factual interest rates from 30.09.97 to 31.08.98 is equal to -0.0560. Corresponding statistics for predictions of the market reaction model is equal to -0.0172, and for predictions of the yields reaction model – to 0.0172. The one-tailed Student's t-criterion enables to reject the hypothesis of displacement absence at 1% p-level for both neural network forecasting models.

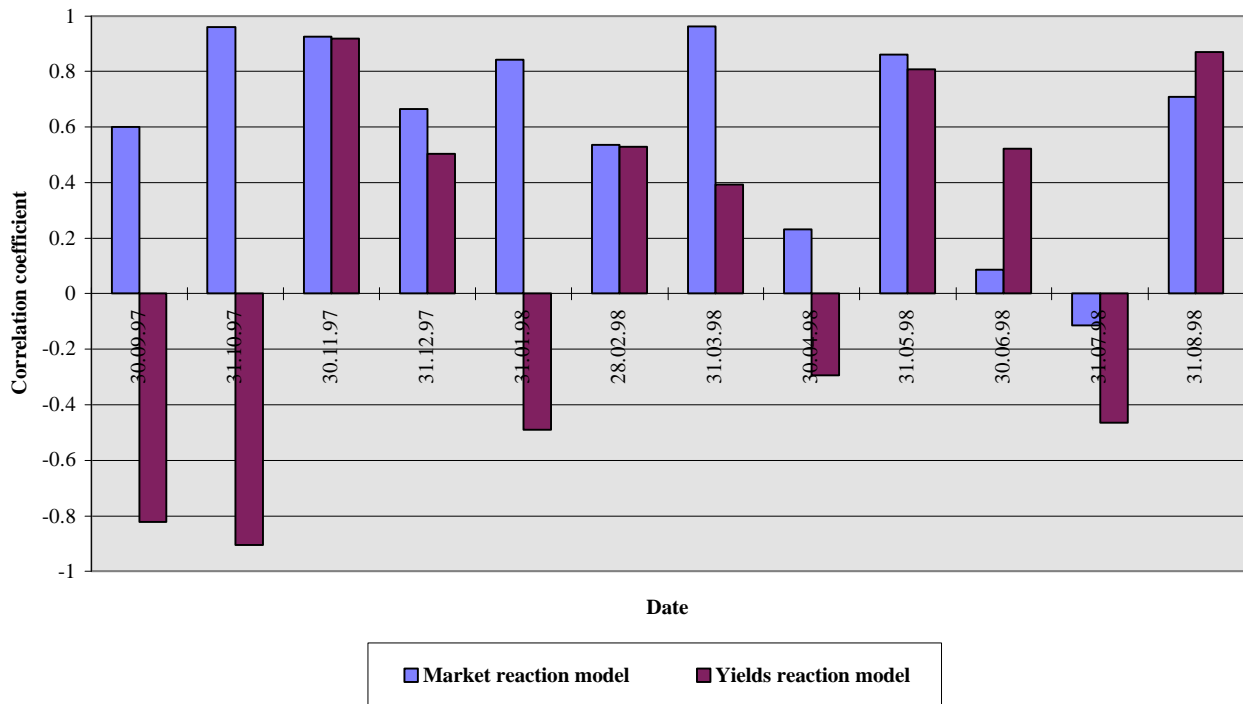
Predictions of the yields reaction model are characterized by a stronger bias than predictions of the market reaction model, but they give better estimations of future market volatility. Fisher's F-criterion enables to reject the hypothesis of predicted and factual interest rates changes variances equality for the yields reaction model, but does not enable to reject this hypothesis for the market reaction model at 5% level of statistical significance.

The opinion of the author is that the explanation of bias phenomenon is connected with the global financial crisis. The collapse of the emerging markets caused the metamorphose of risk tolerance and to portfolios restructuring in favor of governmental bonds characterized by the lowest level of risk. It resulted in switching of money flows into the US Treasury bonds market and to decrease of the interest rates that cannot be explained with the aid of information about macro-indicators of the American economy. But a qualified financial analyst could have made proper logical corrections of neural networks outputs based on general knowledge of the global situation in the world economy.

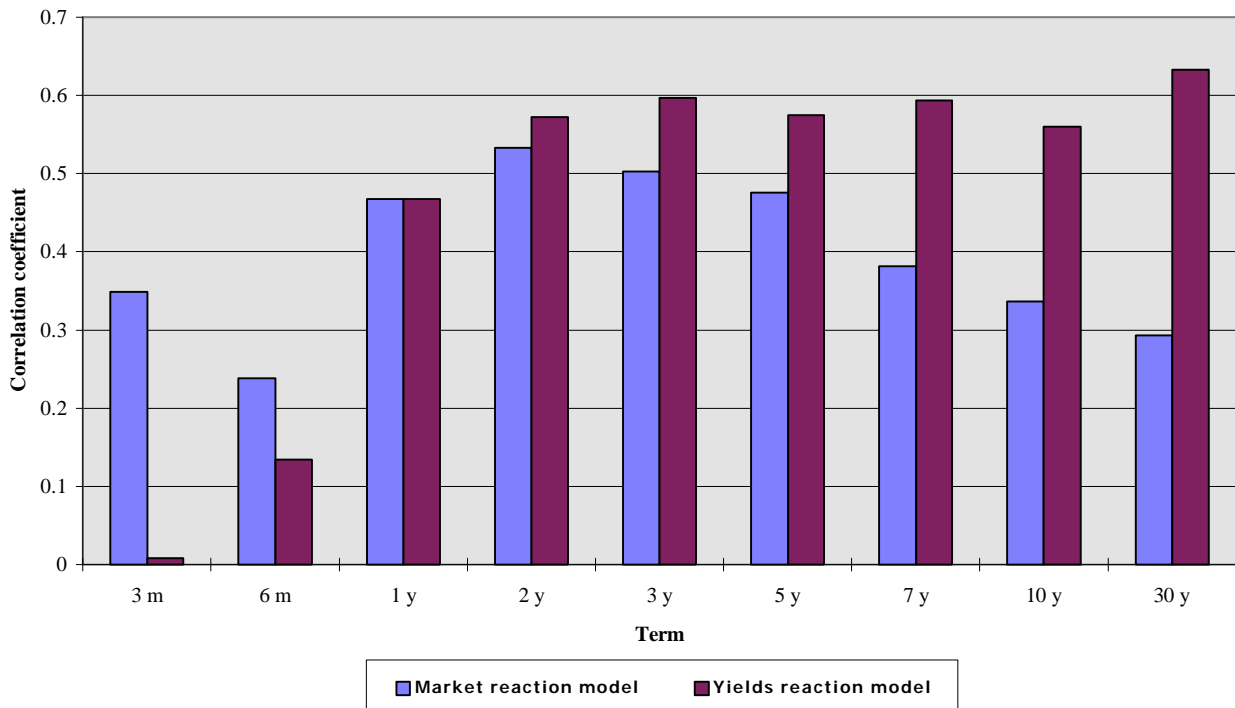
To evaluate the models performance, it is expedient to bring factual observations and outputs of the neural networks into matrixes of  $12 \times 9$  dimension in which the rows of which are formed by the changes of yields of bonds with different maturities for the same period and the columns are formed by the changes of yields of bonds with the same maturity for different periods. The

correlation coefficients between the rows of factual observations and predicted values matrixes can be interpreted as the indicators of quality of the yield curve shape movement prediction, and the correlation coefficients between the columns of these matrixes – as the indicators of quality of different (by their term) interest rates pathways prediction.

*Figure 13. Verification of capability to predict the shape of the yield curve shift.*



*Figure 14. Verification of capability to predict the yield change trajectory.*



According to figure 13, the market reaction model enables to estimate very precisely the future shift of the yield curve shape, or the relative potential of growth for short-term, medium-term and long-term bonds. The correlation coefficient between factual and predicted yields changes exceeded 0.5 in 9 periods of 12, and no gross blunder, finding reflection in large by its absolute value negative correlation coefficient, was registered for the market reaction model. The yields reaction model, on the contrary, demonstrates a very weak ability to predict the shape of the yield curve shift

and frequently makes critical mistakes. The correct estimation of comparative potential of growth for short-term and long-term interest rates was given by this model only in 7 cases of 12.

However, the comparison of quality of the interest rates pathways prediction testifies that the yields reaction model demonstrates better results. Its forecasts of the medium-term and the long-term interest rates are characterized by high accuracy, while the estimations of the short-term interest rates movements are not better than the predictions given by unsophisticated models such as «there are no changes» or «the changes will be the same». At the same time the schedule of the correlation coefficients between factual and predicted changes of yields of bonds with different maturities for the market reaction model reminds a dome of Gaussian distribution with the maximum in the area of the medium-term interest rates.

Therefore both models are characterized by their values and defects. The market reaction model gives less biased estimations and predicts the shape of the yield curve shift and the pathways of the short-term interest rates better. The yields reaction model gives better estimations of future market volatility and pathways of the medium-term and the long-term interest rates.

Neural network models as a whole should be considered as a useful tool of short-term forecasting of the governmental bonds market conjuncture. At the same time their predictions can not be used within a mechanical trading strategies and requires corrections of a qualified analyst.

### 5. Multi-objective Portfolio Optimization within the Fuzzy Decisions Theory Framework.

Neural network forecasting models provide rather adequate predictions of direction and shape of the yield curve shift. However, these predictions are very far from being perfect. Moreover, in some cases they are absolutely incorrect.

Financial markets are an area of high uncertainty, and even the most efficient forecasting models generate gross blunders from time to time. Therefore decisions of the portfolio manager should be based on the consideration of various scenarios of the yield curve shift. Otherwise the interest rate risk becomes inadmissible.

The plurality of admissible estimations is a usual practice of portfolio management. The manager is guided by several tactical objectives – to provide adequate return for different market scenarios, some of which are more likely, and others less likely in the manager's view. Therefore the fixed-income portfolio selection problem can be transformed into a problem of decision-making in a fuzzy environment. The pioneering investigation of Srichander Ramaswamy provides a framework for the solution of this problem<sup>6</sup>.

Ramaswamy offers the following linear programming model for the fixed-income portfolio selection using fuzzy decisions theory:

$$\begin{aligned}
 & I \rightarrow \max \left\{ \right. \\
 & \left. \frac{\sum_{j=1}^J r_{ij} x_j - g_i^{\min}}{g_i^{\max} - g_i^{\min}} \geq I, \quad i = \overline{1, I} \right\} \\
 & \left\{ \begin{aligned} & \sum_{j=1}^J x_j = 1 \\ & x_j \geq 0, \quad j = \overline{1, J} \end{aligned} \right.
 \end{aligned}$$

This model implies the following notation:  $I$  – the fuzzy logic certainty factor of portfolio selection,  $i$  – the number of the yield curve scenario,  $I$  – the total number of the yield curve scenarios under consideration,  $j$  – the number of the fixed-income financial instrument,  $J$  – the total number of available fixed-income financial instruments,  $g_i^{\min}$  – inferior limit of the portfolio's return under scenario  $i$  (the tactical objective),  $g_i^{\max}$  – superior limit of the portfolio's return under scenario  $i$  (the tactical objective),  $r_{ij}$  – return of the fixed-income financial instrument  $j$  under scenario  $i$ ,  $x_j$  – the

<sup>6</sup> Ramaswamy (November 1998), «Portfolio Selection Using Fuzzy Decision Theory», Basle, Bank for International Settlements, Monetary and Economic Department.

share of the fixed-income financial instrument  $j$  in the portfolio (the control vector).

The neural network outputs of the market reaction model and the yields reaction model may be used as scenarios (together with another predictions) within this model specification. But this improvement is insufficient for model transformation into efficient decision support system in the real work of transaction costs. Therefore it should be modified to provide a basis for consideration of the transaction costs effect for the portfolio return at the stage of decision making.

Let the position of a size  $Q_j^H \geq 0$  be opened at the moment of decision making on the fixed-income instrument  $j$  with the market price  $P_j^M$  (including accrued coupon). The manager's decision implies portfolio restructuring by means of buying  $Q_j^B \geq 0$  bonds at the price of  $P_j^B$  and selling  $Q_j^S \geq 0$  bonds at the price of  $Q_j^S \geq 0$ . Transaction costs procure inequality  $P_j^S < P_j^M < P_j^B$ .

Let's define the current market value of the bond's portfolio as  $CPV = \sum_{j=1}^J P_j^M \times Q_j^M$ , the starting trading account balance as  $M$  and money inflow as  $MF$  (in the case of money outflow  $MF$  becomes a negative number). A portfolio manager is obliged to support a positive trading account balance, therefore his decisions must satisfy the inequality  $M + MF + \sum_{j=1}^J (P_j^S \times Q_j^S - P_j^B \times Q_j^B) \geq 0$ ,

where  $\sum_{j=1}^J P_j^S \times Q_j^S$  is the proceeds from sell transactions and  $\sum_{j=1}^J P_j^B \times Q_j^B$  is the expenditure for buy transactions. Inadmissibility of the short positions determines the necessity of  $J$  inequalities  $Q_j^H + Q_j^B - Q_j^S \geq 0, j = \overline{1, J}$ .

The return of total portfolio consisting of bonds and trading account balance is calculated on the basis of formula

$$R_i = \frac{\sum_{j=1}^J (1 + r_{ij}) \times P_j^M \times (Q_j^H + Q_j^B - Q_j^S) + M + MF + \sum_{j=1}^J (P_j^S \times Q_j^S - P_j^B \times Q_j^B)}{CPV + M + MF} - 1$$

Then the model of transactions optimization may be written in the following form:

$$I \rightarrow \max$$

$$\frac{R_i - g_i^{\min}}{g_i^{\max} - g_i^{\min}} \geq I, i = \overline{1, I}$$

$$R_i = \frac{\sum_{j=1}^J (1 + r_{ij}) \times P_j^M \times (Q_j^H + Q_j^B - Q_j^S) + M + MF + \sum_{j=1}^J (P_j^S \times Q_j^S - P_j^B \times Q_j^B)}{CPV + M + MF} - 1$$

$$M + MF + \sum_{j=1}^J (P_j^S \times Q_j^S - P_j^B \times Q_j^B) \geq 0$$

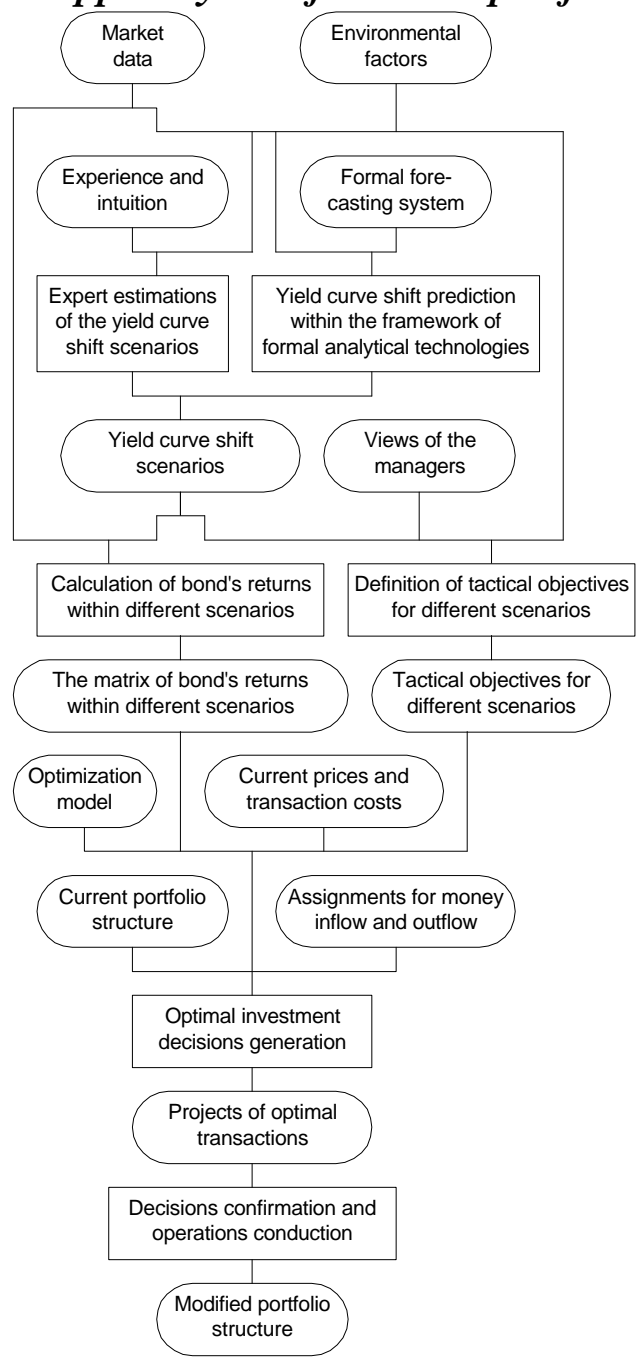
$$Q_j^H + Q_j^B - Q_j^S \geq 0, j = \overline{1, J}$$

$$Q_j^B \geq 0, j = \overline{1, J}$$

$$Q_j^S \geq 0, j = \overline{1, J}$$

The solution of this linear programming optimization model are the vectors of the best transactions for buy and sell. Figure 15 demonstrates the whole decision support system.

**Figure 15. Decision support system for bonds portfolio management.**



### Bibliography.

1. Aivazian & Mkhitarian (1998), «Applied Statistics and Essentials of Econometrics», Moscow, UNITY (in Russian).
2. Baestaens, Van den Bergh & Wood (1994), «Neural Network Solutions for Trading in Financial Markets», London, Longman.
3. Beltratti, Consiglio & Zenios (March 1998), «Scenario Modeling for the Management of International Bond Portfolios», The Wharton Financial Institutions Center.
4. Blaug (1985), «Economic Theory in Retrospect», Cambridge University Press.
5. Brown & Kritzman (1990), «Quantitative Methods for Financial Analysis», Homewood, Dow Jones-Irwin.
6. Buser, Karolyi & Sanders (April 1996), «Adjusted Forward Rates as Predictors of Future Spot Rates», Dice Center of Financial Economics.
7. Campbell, Cambell & Dolan (1988), «Money, Banking, and Monetary Policy», The Dryden Press.
8. Dougherty (1992), «Introduction to Econometrics», Oxford University Press.
9. Douglas (1990), «Bond Risk Analysis», New Jersey, Prentice-Hall.
10. Dubrov, Lagosha & Khrustalev (1999), «Risk Situations Modeling in Economy and Business», Moscow, Finance and Statistics (in Russian).
11. Eddows & Stansfield (1991), «Decision Making Techniques», London, Longman.
12. Granberg (1990), «Statistical Modeling and Forecasting», Moscow, Finance & Statistics (in Russian).
13. Hecht-Nielsen (1991), «Neurocomputing», San-Diego, Addison-Wesley.
14. Lapsina, Ambarian & Melnikov (October 1998), «Investment Decisions for the Portfolio's Risks Optimization in Russian Financial Market», Parnu, 5<sup>th</sup> Nordic-Baltic Conference in Regional Science.
15. Mankiw (1992), «Macroeconomics», New-York, Worth Publishers.
16. Marshall & Bansal (1992), «Financial Engineering: A Complete Guide to Financial Innovation», New York, Allyn & Bacon.
17. Nijkamp & Wang (August 1998), «Winners and Losers in the European Monetary Union: A Neural Network Analysis of Spatial Industrial Shifts», Vienna, 38<sup>th</sup> Congress of the European Regional Science Association.
18. O'Brian & Shrivastava (1995), «Financial Analysis and Securities Trading», Financial Trading System.
19. Payeras & Pou (August 1998), «The EMU and the Spanish Term Structure of Interest Rates», Vienna, 38<sup>th</sup> Congress of the European Regional Science Association.
20. Pervozvansky & Pervozvanskaya (1992), «Financial Market: Analysis and Risk», Moscow, INFRA-M (in Russian).
21. Ramaswamy (November 1998), «Portfolio Selection Using Fuzzy Decision Theory», Basle, Bank for International Settlements, Monetary and Economic Department.
22. Sharpe, Alexander & Bailey (1995), «Investments», New Jersey, Prentice-Hall.
23. Sidorkin & Kostyukhin (1995), «Forecasting on the Basis of Neural Networks Apparatus», Odessa, OSPU (in Russian).
24. Smets & Tsatsaronis (September 1997), «Why Does the Yield Curve Predict Economic Activity?», Basle, Bank for International Settlements, Monetary and Economic Department.
25. Thor, Lapsina & Melnikov (August 1998), «Identifying, Measuring and Management Risks in Russian Secondary Stock Markets», Vienna, 38<sup>th</sup> Congress of the European Regional Science Association.
26. Trakhtengerts (1998), «Computer Decision Support Systems», Moscow, SINTEG (in Russian).
27. Wagner (1969), «Principles of Operations Research (With Applications to Managerial Decisions)», New Jersey, Prentice-Hall.
28. Watsham & Parramore (1996), «Quantitative Methods in Finance», International Thomson Business Press.
29. Zmitrovitch (1997), «Artificial Intelligence Systems», Minsk, TetraSystems (in Russian).