ACTA UNIVERSITATIS LODZIENSIS
FOLIA OECONOMICA 194, 2005

Aneta Wlodarczyk, Tomasz Szmigiel**

THEORETICAL ASPECTS OF USING MARKOV MODELS IN RESEARCH OF EXCHANGE RATE VOLATILITY

Abstract

During modeling of short-run exchange rate fluctuations, there is usually a need for taking into consideration some random-type conditions, i.e. it is necessary to abandon the fundamental exchange rate theories in favor of probabilistic modeling. Among stochastic models, of special interest are Markov models. The main advantages of Markov models include a relative simplicity of construction, easy inferences, well-known estimation methods and especially consistence of properties of these models with the observed properties of many real phenomena. Application of switching models is based on a general assumption that the examined time series can be presented as sequences of random variables of a known type of conditional distribution in all regimes. Known from literature propositions concerning the modeling of exchange rate with the use of switching models did not provide sufficiently good forecasts of the future exchange rate levels because of, among others, low frequency of data used for the construction of the model (quarterly or monthly data). The authors are going to continue the examination of the PLN exchange rate fluctuation with the use of Markov models that was started in this paper. The next stage of their work will be connected with conducting empirical research concerning the occurrence of calendar anomalies in the Polish currency market. For this purpose, a new method based on the Markov chains theory will be applied, which offers a new perspective to this problem. Testing of the calendar time hypothesis has been considered so far mostly in the aspect of comparison of daily expected values and variances of exchange rate return rates. Then, on the basis of the data concerning exchange rates for high measurement frequency, a Markov switching model will be constructed and used for description of the PLN depreciation and appreciation period.

Key words: Markov models, time series, exchange rate volatility, calendar anomalies.

* Ph.D., student, Econometrics and Statistics Department, Technical University in Częstochowa.

I. INTRODUCTION

Recently, some significant changes have taken place in financial markets. According to Jajuga (2000), the most important of these changes include: technological development, globalization of financial markets, resignation from mediation in financial markets, dynamic development of financial innovations, rise and development of risk transfer market and integration of various segments of the financial market. These events resulted in an increase in the fluctuation of exchange rates. Similar transformations could also be observed in the Polish foreign currency market, especially after the year 1995, when the process of a gradual elasticizing of the rate mechanism started. The new regulation of the law of the foreign exchange, a new currency basket based only on EURO and dollar and a PLN floating exchange rate favored bigger and more frequent currency fluctuations. In spite of the fact that, on 1st January, 2002, 12 national currencies have been replaced by EUR, the information about the exchange rate fluctuations will remain one of the most important pieces of information.

The setting of the exchange rate is explained by, among others, such economic theories as the Purchasing Power Parity theory, the law of one price, the International Fisher Effect theory, and the Interest Rate Parity theory. Using the criterion of time as a basis, it is possible to divide the exchange rate theories into those that explain changes occurring over a long period of time (PPP), those that refer to a medium period (connected with the NBP – Polish National Bank policy) and those that approach the rate from the point of view of short-term changes (the overshoot theory) (discussed in Chrabonszczewska et al., 1996). According to this distinction, the level of exchange rate during a long period is determined by many economic factors, among others by the GNP increase rate, inflation rate, the central bank interventions, situation of the balance of payments or the economy development level and structure. In the short run, the exchange rate level may be also influenced by psychological and political factors that can suddenly increase the demand for a given currency, therefore strengthening its exchange rate. It is often believed that the exchange rate changes are caused by rational expectations of the investors. It means that occurrence in the currency market of information about a factor that should cause an increase in the exchange rate in the future results in a short-term appreciation of the exchange rate. The rational expectations theory and its application to finance was described by M. Osińska (2000). The exchange rate may also be influenced by unpredictable factors, such as disturbances in international exchange markets that can have a passing, but important influence on the exchange rate and to a large extent are connected with the phenomenon of globalization of financial markets.

For example, losses suffered by international investment banks in one market are compensated by taking appropriate investment actions in other markets. One should also bear in mind that short-term investors, in making their decisions, seldom rely upon an assessment of the economic situation and balance of payments of a given country. They take investment actions on the basis of information, sometimes unreliable, that appear in the exchange market, using at most some technical analysis instruments.

Therefore, for some time we have been able to observe a significant increase in the scope of a conducted research concerning the use of a rich apparatus of stochastic processes for the forecasting of exchange rate behavior in time. The authors of this paper present a proposal concerning the use of Markov models for detecting and describing regularities that govern the process of exchange rate fluctuation. In the first part of the paper, the Markov chains theory will be used for examination of calendar anomalies connected with the weekend effect that occur in the exchange market. This article also describes a method based on the Markov chains theory that can be used for examination of mutual relationships between business volume volatility and price volatility for futures. In the second part of the paper, presented will be some problems connected with construction and estimation of switching parameters of Markov models that can be used for the forecasting of exchange rate fluctuations.

II. PROPERTIES OF THE RATE OF RETURN OF EXCHANGE RATE DISTRIBUTION

Over a longer period of time, exchange return rates are characterized by the following properties (discussed in Jajuga, 2000):

1. Occurrence of the volatility clustering phenomenon – both large and small changes of exchange rates occur in series. After a large volatility period, there occurs a period of smaller volatility. This dependence is also known under the name of data concentration effect. In the case of long time series and data of high observation frequency it usually results in an increased random factor variance in appropriate classic models.

2. Exchange return rate distributions are leptokurtic. The probability of occurrence of untypical (very large or very small) exchange rate fluctuations is higher than in the case of normal distribution. In the literature, another term for this phenomenon is also used, i.e. occurrence of "thick tails" in the rates of return of exchange rates distribution.

3. These distributions are in many cases skew, which means that the rate of return distribution is not symmetrical around the average.

4. Exchange rate fluctuations are negatively correlated with the changeability of their variance. Process variance depends on the previous rates of return, so when the exchange rate drops, there is a tendency towards an increase in the rates of return. This dependence is known as the lever effect.

5. There is a relationship between the variance of the exchange return rate and autocorrelation. The autocorrelation usually accompanies a small exchange rate variance and large volatility results in the lack of autocorrelation.

6. Long-run data dependence, which means that after significant increases there are further increases, after which sudden decreases occur, followed by further decreases.

Such setting of exchange rates is a result of the nature of the very processes taking place in the currency markets. Among the causes that result in, among others, occurrence of the phenomenon of exchange rates volatility clustering, it is necessary to mention the specificity of information inflow to the market and their subjective interpretation by individual market participants who also set the exchange rate (psychological factor). The inflow of information usually occurs in an irregular way, in special cases in series which means that they can be correlated. That is why their meaning and power of influence on the change of market exchange rates are diversified. As a result, it is possible to observe sub-periods of decreased and increased exchange rates volatility. Exchange rates volatility depends on expected new information, among others on the decisions of Rada Polityki Pieniężnej (the Financial Policy Council), messages about the macroeconomic situation (e.g. concerning the inflation level, interest rates, gross national product, etc.). Their announcement dates are usually known in advance, and the uncertainty that concerns them results in an increase of the rates of return variance. This situation has a direct influence on the demand and supply of foreign exchange and currency and results in an increased exchange rate volatility. Such behavior of exchange rates requires the use of special methods that would take into account the dependencies discussed above. The authors of this paper suggest that Markov chains and switching regression be used for examination of volatility of exchange rate.

III. MARKOV PROCESSES

Price models (or return rates) of financial instruments are usually considered in the category of stochastic processes. In the financial practice, we usually deal with stochastic processes with discrete time, as the values of this process are observed at specific moments, e.g. once a day or once a minute.

Markov processes belong to a class of stochastic processes, where the past and the future – with a fixed present, are independent, or the future depends on the past only through the present. According to Iosifescu (1988), the following Markov chain definition is given:

Let the stochastic process $X(t)$ be a process with discrete time, i.e. a sequence of random variables $X_0, X_1 ...$ It is assumed that the value of random variable X_k $(k = 0, 1, ...)$ defines the state of the process at moment k.

The sequence of random variables $X_0, X_1, X_2 ...$ constitutes a Markov chain if for any i, j, n and for any real numbers $x_1, x_2, ...$, the following relation is true:

$$P(X_n = j/X_{n-1} = i, X_{n-2} = x_{n-2}, ..., X_0 = x_0) = P(X_n = j/X_{n-1} = i) \quad (1)$$

i.e. conditional distribution $X(t_n)$ with given values of $X(t_1), ..., X(t_{n-1})$ depends only on $X(t_{n-1})$ and does not depend on the previous process course. Construction of Markov chains is based on three basic qualities:

a) finite set S, the elements of which, called states, are assumed to be numbered in a certain specified way;

b) probability distribution $p = [p_i(0) = P(X_0 = i)]_{i \in S}$ on S, called initial probability;

c) stochastic matrix $P = [P_{i,j}(n, s)]_{i,j \in S}$, the elements of which, called transition probability, are defined in the following way:

$$P_{i,j}(n, s) = P\{X_n = j/X_s = i\} \quad \text{for} \quad n > s. \quad (2)$$

The above relationship defines conditional probability of the fact that the process is in state j at moment n on condition that, at moment s, it was in state i.

A homogeneous Markov chain is such a Markov chain $\{X_n\}_{n \in N}$ that for n, $k \in N$, the following equality is true:

$$P\{X_n = j/X_{n-1} = i\} = P\{X_{n+k} = j/X_{n+k-1} = i\} = p_{ij}. \quad (3)$$

Therefore, for a homogeneous Markov chain, the transition probabilities between states in one step are constant in time and independent of the process duration time.

The properties of Markov chains:

$$p_i(n) = P\{X(n, \omega) = i\}, \quad i \in S \quad (4)$$

$$\sum_{i \in S} p_i(n) = 1 \tag{5}$$

$$P_{i,j}(n, s) \geqslant 0, \quad i, j \in S \tag{6}$$

$$\sum_{j \in S} p_{i,j}(n, s) = 1 \quad \text{for} \quad i \in S \tag{7}$$

$$p_j(n) = \sum_{i \in S} p_i(s) \cdot P_{i,j}(n, s) \quad \text{for} \quad j \in S \tag{8}$$

The formula (8) defines an unconditional probability distribution for reaching state j by a homogeneous Markov chain at moment n: $p_j(n) = P(X_n = j)$. This dependence was derived from the formula for total probability. I-verse of the matrix **P**

$$p_i = [p_{i1}\, p_{i2} \cdots p_{ir}]$$

is a conditional probability distribution of variable X_n on condition that $X_{n-1} = i$.

The Markov chain is definite if the transition probability matrix $\mathbf{P} = (p_{i,j})$ is given and initial probabilities $p_i = P\{X_0 = i\}$, $i = 0, 1, 2, \ldots$, are given, which is dealt with by the following theorem (see Sobczyk, 1973):

Theorem 1. Let us have a countable set S and sequence $\{p_i\}$ and matrix $\mathbf{P} = (p_{i,j})$ that meet the following conditions:

$$p_i \geqslant 0, \quad \sum_i p_i = 1, \quad p_{ij} \geqslant 0, \quad \sum_j = p_{ij} = 1$$

Then there exists probabilistic space (Γ, F, P) and a Markov chain $\{X_n\}_{n \geqslant 0}$ defined in this space. The chain has state space S, initial probabilities p_i and transitional probability matrix $\mathbf{P} = (p_{i,j})$.
Depending on the characteristic values of matrices **P**, a sequence of powers $\mathbf{P}^n$ has different properties at $n \to \infty$ that are the basis for inference about the chain behavior in a long period of time.
In Podgórska et al. (2000), the following characteristics for matrix **P** are given:

1. Stochastic matrix $\mathbf{P} \in M(n, n)$ is called an irreductible matrix, if the characteristic value $\lambda_1 = 1$ is a single root of the characteristic equation of this matrix.

2. Stochastic matrix $\mathbf{P} \in M(n, n)$ is called an aperiodic matrix if the characteristic value $\lambda_1 = 1$ is the only root of the characteristic equation with the module 1.

3. The irreducible and aperiodic matrix is called a regular matrix. Under the Fréchet theorem (this theorem can be found in, among others, Josifescu, 1988), there is a limit

$$\lim_{n \to \infty} \mathbf{P}^n = \mathbf{\Pi} \tag{9}$$

where $\mathbf{\Pi}$ denotes a stochastic ergodic matrix (of identical rows).
Because of the above properties of regular matrix, the following theorem (Podgórska et al., 2000) is true.

Theorem 2. If transition probability matrix **P** of a finite homogeneous Markov chain is regular, then the chain is ergodic, which means that there is a limit

$$\lim_{n \to \infty} \mathrm{p}_{ij}^{n} = \pi_j \quad \text{for} \quad i, j \in S$$

where

$$\sum_{j \in S} \pi_j = 1 \quad \text{and} \quad 0 \leqslant \pi_j \leqslant 1.$$

In this way, a row vector of probabilities of reaching balance states $\boldsymbol{\pi} = (\pi_1, \pi_2, \pi_3, \ldots, \pi_r)$ was defined. It meets the following relation:

$$\boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{P}. \tag{10}$$

On the basis of the definition of the long-term probabilities vector, it is possible to infer that, regardless of the state of process at moment 0, the chances of reaching a given state stabilize with time and equal individual elements of vector $\boldsymbol{\pi}$.

The notion of a Markov chain can be extended to the case of multiple Markov dependence.

We can say that the sequence of random variables $(X_n)_{n \geqslant 0}$ with the values in (finite or countable) set S is a multiple second-order Markov chain (double Markov chain) then and only then, if:

$$\mathrm{P}(X_{n+1} = i_{n+1}/X_n = i_n, \ldots, X_0 = i_0) = \mathrm{P}(X_{n+1} = i_{n+1}/X_n = i_n, X_{n-1} = i_{n-1}) \tag{11}$$

for all $i_0, i_1, ..., i_{n+1} \in S$, whereas the last conditional probability depends on i_n and i_{n-1}. It is easy to notice that if $(X_n)_{n \geq 0}$ is a double Markov chain, then the sequence $(X_i, X_{i+1})_{i \geq 0}$ is a simple Markov chain of state space $S^2 = \{(i, j): i, j \in S\}$.

In this way, with any multiple Markov chain, a certain single chain can be connected. Therefore, in some cases, investigation of a multiple Markov chain can be reduced to investigation of simple Markov chains.

The process of estimation of a transition probability matrix **P** that is usually unknown can be based on two types of data: microdata or macrodata. For estimation of these probabilities, a ordinary or generalized restricted least squares method with may be used. The maximum likelihood method may also be used.

A detailed description of procedures used for estimation of elements of a transition probability matrix can be found in, among others, Lee et al. (1997) and Podgórska et al. (2000).

IV. USE OF MARKOV CHAINS FOR EXAMINATION OF CALENDAR ANOMALIES

Research conducted on effective markets has revealed some anomalies present in these markets. These anomalies include the day-of-the-week-effect, the January effect, the scale effect and other effects.

During analyzing of exchange rate quotations in the interbank market, one can notice a stabilizing of the exchange rate level in the non-trading periods on the Warsaw Stock Exchange, especially in the secondary market, i.e. during weekends and holidays. At the same time it is known that the information hidden during the period when the market is closed affects the value of exchange rates immediately after the market has been opened (see in Dunis et al., 2001). If the information accumulated linearly, the variance within the period from the closure of the market on Friday to its closure on Monday should be three times higher than, for example, variance within the period from the closure of the market on Monday to its closure on Tuesday. However, the research results show that the information flows in more slowly during periods when the market is closed (the weekend effect or the Monday effect) (see in Fama, 1965). Therefore, the following hypotheses can be verified:

a) fluctuations of exchange rates occur in accordance with the flow of calendar time,

b) fluctuations of exchange rates occur in accordance with the flow of session time (passing over non-transaction days).

The hypothesis involving calendar time should be confirmed by the occurrence of different distributions of "one-day" return rates in situations when consecutive transaction days are separated by non-transaction days (weekends, holidays) in comparison with sessions taking place on consecutive calendar days. However, if the price changes occur, in accordance with the session time hypothesis, only during the market activity period, the distributions of the "one-day" return rates should be similar, regardless of whether there were non-transaction days between the transaction days or not.

The original data set consists of quotations of exchange rates in the interbank market. 24-hour volatility was defined as absolute return rates from quotations of the exchange rates in accordance with the following formula:

$$v_t = |\ln(P_t/P_{t-1})| \tag{12}$$

where:

P_t is opening (closing) price on the day t.

P_{t-1} is closing (opening) price on the day $t-1$.

The data set obtained in this way was grouped into three volatility states: H for a high volatility state; N for a normal volatility state and L for a low volatility state. Grouping was carried out in such a way as to obtain a more or less equal number of volatility states in each of the groups. The above criterion was used for determination of the limits of three intervals: high, normal and low volatility, that is, for a correct selection of the values of parameters k_1 and k_2 that are shown in Table 1.

Table 1. Limits of intervals of the high, normal and low volatility

State	Meaning	Volatility limits for exchange rate
H	High volatility	$v_t > \bar{v}_t + k_1 d$
N	Medium volatility	$v_t \leqslant \bar{v}_t + k_1 d$ and $v_t > \bar{v}_t - k_2 d$
L	Low volatility	$v_t \leqslant \bar{v}_t - k_2 d$

Where:

v_t – mean absolute rate of the exchange rate change

d – mean absolute deviation of v_t.

The subject of the analysis will be a daily structure of the exchange rate fluctuation, to which the Markov chain theory was applied. The authors of this paper made an attempt to show a nonclassical method of verification of the hypothesis according to which fluctuations of a given exchange rate occur in accordance with the flow of calendar time. The calendar time hypothesis assumes that the volatility distributions for days following the weekend are different.

According to this pattern, the market is more active and volatilities are higher on Monday than on other weekdays. That is why high volatility states should occur more often on Mondays after high volatility states than after high volatility states on other working days.

Therefore, we should assess with what probability the transition between pairs of adjacent states occurs, e.g. *HH* or *HN*, where, for instance, *HN* denotes a state of low volatility occurring after a state of high volatility.

In this paper, a model of a long-run probability distribution was used for investigation of daily structure of transition between states, obtained on the basis of the methodology of second-order Markov chains. A second-order Markov chain was constructed for volatility of exchange rates. In the model presented, the second-order Markov chain was used for obtaining a long-run, finite probability distribution of occurrence of pairs of adjacent states of the *HN* type. As a result of this action, we obtain the following **P** matrix: the matrix of second-order transition probability.

Table 2. Diagram of the matrix of second-order transition probability

Past states	Present states								
	HH	*HN*	*HL*	*NH*	*NN*	*NL*	*LH*	*LN*	*LL*
HH	$P_{HH,HH}$	$P_{HN,HH}$	$P_{HL,HH}$	0	0	0	0	0	0
HN	0	0	0	$P_{NH,HN}$	$P_{NN,HN}$	$P_{NL,HN}$	0	0	0
HL	0	0	0	0	0	0	$P_{LH,HL}$	$P_{LN,HL}$	$P_{LL,HL}$
NH	$P_{HH,NH}$	$P_{HN,NH}$	$P_{HL,NH}$	0	0	0	0	0	0
NN	0	0	0	$P_{NH,NN}$	$P_{NN,NN}$	$P_{NL,NN}$	0	0	0
NL	0	0	0	0	0	0	$P_{LH,NL}$	$P_{LN,NL}$	$P_{LL,NL}$
LH	$P_{HH,LH}$	$P_{HN,LH}$	$P_{HL,LH}$	0	0	0	0	0	0
LN	0	0	0	$P_{NH,LN}$	$P_{NN,LN}$	$P_{NL,LN}$	0	0	0
LL	0	0	0	0	0	0	$P_{LH,LL}$	$P_{LN,LL}$	$P_{LL,LL}$

In a given column, state *XY* refers to *X* as first state and *Y* as second state, whereas in a given row, *YZ* refers to *Y* as second state and *Z* as third state. An element in matrix **P**, $\mathbf{P}_{YZ,XY}$ indicates the probability of transition from state *Y* to *Z*, on condition that in the previous step, a transition from *X* to *Y* had occurred. The first and second state provides information about the prior state in the Markov chain (row); the second and third state provides information about the posterior state in the Markov chain (column). Let vector $\boldsymbol{\pi}$ denotes a long-run probability distribution:

$$\boldsymbol{\pi} = (\pi_{HH}\,\pi_{HN}\,\pi_{HL}\,\pi_{NH}\,\pi_{NN}\,\pi_{NL}\,\pi_{LH}\,\pi_{LN}\,\pi_{LL})^T,$$

or, with additional symbols introduced: $HH \equiv 1$, $HN \equiv 2$, $HL \equiv 3$, $NH \equiv 4$, $NN \equiv 5$, $NL \equiv 6$, $LH \equiv 7$, $LN \equiv 8$, $LL \equiv 9$, $\{\pi_j\}$ for $j = 1, ..., 9$. According to the relationship (10):

$$\boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{P}$$

and stochastic properties of the vector $\boldsymbol{\pi}$:

$$\sum_{j \in S} \pi_j = 1 \quad \text{and} \quad 0 \leqslant \pi_j \leqslant 1,$$

it is possible to determine a long-run distribution of states in the Markov chain.

In order to test if $\pi_{HH} > \pi_j$, where $j \neq HH$, it is necessary to test the null hypothesis that implies that the long-run probabilities for all transitions between two states are identical. A test statistic is calculated:

$$\chi^2 = N \sum_{j=1}^{9} \frac{(\pi_j - 1/9)^2}{1/9} \tag{13}$$

where N is the total number of observations. Under the null hypothesis, the χ^2 statistic follows an asymptotic χ^2 distribution with $(T-1)$ degrees of freedom, where T is the number of states, in this case, $T = 9$. If there is no basis to reject the null hypothesis, it means that there is no significant difference among the long-run probabilities of reaching different states. Thus, it is not possible to draw the conclusion that $\pi_{HH} > \pi_j$, where $j \neq HH$.

$$\text{Let } \boldsymbol{\pi}_h = (\pi_{h1}\, \pi_{h2}\, \pi_{h3}\, \pi_{h4}\, \pi_{h5}\, \pi_{h6}\, \pi_{h7}\, \pi_{h8}\, \pi_{h9})^{\mathrm{T}}$$

where $\boldsymbol{\pi}_h$ is the 9×1 limiting probability vector on day h, N_h is the sample size and $\boldsymbol{\pi}_{hj}$ denotes the limiting probability of transition j ($j \equiv HH$ or HN ... etc) on day h.

It is necessary to show that in the long run, there occur different volatility distributions for days after the weekend in relation to working days. The null hypothesis H_0: $\pi_1 = \pi_2 = ... = \pi_5$ was set, where, for example, vector π_1 contains limiting probabilities for Monday. The null hypothesis informs us then that

$$\pi_{1f}^{e} = \dots = \pi_{hj}^{e} = \dots = \pi_{5j}^{e} = \frac{\sum_{h=1}^{5} \pi_{hj} N_h}{\sum_{h=1}^{5} N_h} \quad \text{for every } j, \tag{14}$$

where π_{hj}^{e} denotes the limiting probability of transition j on day h. The test statistic was defined:

$$\chi^2 = \sum_{h=1}^{5} N_h \times \sum_{j=1}^{9} \frac{(\pi_{hj} - \pi_{hj}^{e})^2}{\pi_{hj}^{e}}. \tag{15}$$

Under the null hypothesis, the χ^2 statistics follows an asymptotic χ^2 distribution with $(H-1)$ $(T-1)$ degrees of freedom, where T is the number of states and H denotes the total number of days of the week. Even if the null hypothesis should be rejected, it means that vector π for Monday should be different from that in other periods, which means that in the finance market, the weekend effect occurs. This method was used for examination of calendar anomalies on the Warsaw Stock Exchange, where the subject of the examination was the WIG 20 (see in Skrodzka et al., 2002) index; it was also used for examination of the hourly price behavior structure volatility of cash prices and futures prices for the Nikkei index (see in Sheyun et al., 1999).

V. ANALYZING THE RELATIONSHIPS BETWEEN VOLATILITY AND TRADING VOLUME FOR FUTURES

The exchange rate level has a significant influence on the value of derivative instruments applied to the exchange rate of a given foreign exchange. The derivatives are increasingly more often used by Polish companies and financial institutions for managing the exchange rate risk. That is why the authors of this paper present the possibility of application of the Markov chains for examination of the futures price volatility on the stock exchange in Warsaw.

One of the most important cause-and-effect relationships occurring in the financial markets is the relationship between stock price (indexes) and trading volume. The opening and closing effects, also connected with calendar anomalies, can be examined from the point of view of the relationship between the trading volume and the stock price volatility. The calendar anomalies hypothesis will be confirmed by proving that both the trading volume and price volatility are higher during the periods of stock market opening and closure that follow "non-session" days (weekend, holidays). Taking into consideration the calendar time theory, one can draw the conclusion that a high trading volume should accompany high volatility at the beginning and the end of the week.

For examination of such a relationship, a Markov chain-based method can be used, in which the information about the trading volume change is combined with the information about price volatility in order to construct a composite state of the first-order Markov chain. Composite states are defined in the following way:

Table 3. Construction of composite states of the Markov chain for the price-volume relationship

	HH	*HN*	*HL*	*NH*	*NN*	*NL*	*LH*	*LN*	*LL*
Δv_t ΔV_t	high high	high normal	high low	normal high	normal normal	normal low	low high	low normal	low low

where: $\Delta V_t = V_t - V_{t-1}$ is related to absolute daily volume of business changes, whereas $\Delta v_t = v_t - v_{t-1}$ is related to absolute daily price volatility increase[1].

In this way, it is possible to construct a composite state first-order Markov chain. Next, a probability matrix of transition among states is constructed, in a similar way as in the case of exchange rates. Discussing the information models theory, one has pointed out that a high trading volume should accompany high price volatility. In the information model that they created, the opening and closing effects result in a concentration of turnover in these periods, the aftermath of which is a decrease of trade liquidity and an increase in the securities volatility rate. Because of this, on the days on which the price volatility of the futures increases, their trading volume should also increase. That is why in the vector of balanced states distribution for the Markov chain, the long-run probabilities of states *HH* and *LL* should be higher than those of other states. To check this, we will test the null hypothesis assuming that the time proportions in all balanced states do not differ significantly. For this purpose, the χ^2 test that was discussed earlier will be used. This method was used for examination of relationships between exchange rates and turnover for futures issued for the Nikkei index in the years 1993–1994 (see in Shiyun et al., 1999).

An analysis of the long-run probabilities of reaching composite states π_j reveals that states of high and low volatility, i.e. *HH* and *LL* occurred much more frequently than the other states. It is also confirmed by a test

[1] The limits of volatility intervals of individual states in the case of volume of business increase and increase in volatility of the prices of futures are determined according to the same principle as in the case of exchange rates volatility.

of distribution χ^2, in which the null hypothesis of an identical long-run state distribution was rejected. On this basis, the conclusion was drawn that in the case of futures for the Nikkei index, this method confirms the existence of a relationship between the trading volume and the contract price. Therefore, the days on which high changes of trading volume occur are also the days of high price volatility changes and vice versa.

VI. MARKOV SWITCHING MODEL

Switching models belong to the class of tools that are appropriate for description of dynamics of processes, the characteristics of which are subject to discrete changes with time, so during their modeling, one must take into consideration appropriate discrete parameter changes. In other words, in an observation time series of such a process, it is possible to observe periods, during which the process values are generated by various regimes. In the case of switching models it is assumed that both the mechanism that controls changes within the limits of individual regimes and the regime change mechanism is random. As a consequence of such assumptions, it is impossible to decide on the current regime exclusively on the basis of the knowledge of the process state at a given moment. Applications of switching models are based on a general assumption that the investigated time series can be modeled with use of stochastic processes defined as sequences of random variables of a known conditional distribution in each of the regimes.

In numerous applications of switching models, the stochastic process that controls the regime changes is a homogeneous Markov chain. This category of switching models is called in the literature "Markov switching models".

Switching models appeared in econometric literature about 30 years ago. The first mentions of a switching model can be found in a work by Goldfeld and Quandt (1973), in which the authors considered a linear regression model with coefficients that changed along with the change of the process-controlling regime. Applications of Markov switching models examining exchange rate fluctuations, are presented in works on, among others, Engel et al., 1990, and Engel, 1992.

The Markov switching model proposed by Hamilton (1989) is closely connected with the TAR (threshold autoregression) model. The basic difference between those two models is that in the case of the Markov switching model, the regime changes are not determined by the process level (so-called threshold value), but through a non-observable state variable that is usually modeled as a Markov chain. For instance:

$$x_t = \begin{cases} \alpha_1 + \beta_1 x_{t-1} + \varepsilon_{1t} & \text{for} \quad s_t = 1 \\ \alpha_2 + \beta_2 x_{t-1} + \varepsilon_{2t} & \text{for} \quad s_t = 0 \end{cases} \tag{16}$$

where s_t is a non-observable, two-state Markov chain, with a certain transition probability matrix **P**. For both regimes, x_t is a first-order autoregression process AR(1), but the parameters of this process (containing a variance of a random component) differ in individual regimes and the regime change is random and subject to autocorrelation.

Regime changes are cause by factors other than those occurring in series that are currently being modeled (s_t determined the regime, not x_t), it is seldom known in which regime the process is (s_t is unobservable), but *post factum* it is often possible to determine in which regime the process was, at a certain trust level s_t can be estimated by means of Hamilton filtering process (Hamilton, 1989). Moreover, regime changes can be identified through interactions between the data and Markov chain, not through examination of the *a priori* data. Difficulties in measurement of the factor that has a significant influence on the examined phenomenon justify treating this variable as an unobservable one and modeling it as a Markov chain. The Markov switching model proposed by Hamilton (1989, 1990) and Engel et al. (1990) has the following form:

$$y_t = \mu_0(1 - s_t) + \mu_1{}^1 s_t + \varepsilon_t, \qquad \varepsilon_t \sim N(\mathbf{0}, \mathbf{\Omega}_{s_t}). \tag{17}$$

The examined phenomenon was presented as a vector stochastic process $(y_1, y_2, ..., y_T)$. The unobservable variable s_t describes the state of the variable explained at moment t. Variable s_t is a Markov chain of two states and it takes on scalar values $\{0, 1\}$. The first state (regime) describes a situation in which an event that interests us occurred, the second regime means an occurrence of an opposite event. When variable $s_t = 0$, the explained variable is subject to normal distribution of parameters μ_0, Ω_0; when $s_t = 1$, then the explained variable is subject to normal distribution of parameters μ_1, Ω_1. In other words, the stochastic process that generates such a phenomenon is a mixture of two normal distributions having different averages.

The transition probabilities do not change with time and amount to:

$$\begin{aligned} P(s_t = 0/s_{t-1} = 0) &= p_{00} \\ P(s_t = 1/s_{t-1} = 0) &= p_{01} = 1 - p_{00} \\ P(s_t = 0/s_{t-1} = 1) &= p_{10} = 1 - p_{11} \\ P(s_t = 1/s_{t-1} = 1) &= p_{11} \end{aligned} \tag{18}$$

The Markov switching model has a few interesting properties from the point of view of the purpose of the investigation. As a Markov chain describes state changes, the process may switch suddenly. It imitates sudden changes in the level of the explaining variable, resulting, for example, from the change of investors' expectations.

The Markov switching model that is described in (17) models the explained variable as a mixture of two normal distributions. As Titterington et al. (1985) demonstrated, a nondegenerate finite mixture of two normal distributions does not have normal distribution. It is consistent with the investigations carried out so far, that revealed that financial price distribution differs from normal distribution. An additional advantage of the selected model is that it belongs to the non-linear model class.

VII. ESTIMATION OF PARAMETERS OF THE MARKOV SWITCHING MODEL

The discussed model can be estimated with the use of the version of EM (Expectations Maximization) algorithm proposed by Hamilton (1990). As a result of the estimation, a parameter vector Θ is obtained that contains elements of the following matrices:

$\mu_j - (k \times 1)$ vector of mean values of explained variable in state j, where $j \in \{0, 1\}$, k – number of components of vector y_t,

$\mathbf{\Omega} - (k \times k)$ variance-covariance matrix,

p_{ij} – probability of transition from the state i to j,

p – probability of the fact that $\mathbf{y}_t$ was in state $s_t = 0$ at time $t = 1$.

Additionally, as a result of the estimation, we obtain the probability of the fact that the process was in state s_t at moment t:

$$\mathrm{P}(s_t/\mathbf{y}_1, ..., \mathbf{y}_t;\ \mathbf{\Theta}) \tag{19a}$$

If information goes beyond moment t (e.g. the whole test $t = 1, 2, ..., T$) used for the probability estimation, then it would be possible to obtain smoothened probability:

$$\mathrm{P}(s_t/\mathbf{y}_1, ..., \mathbf{y}_T;\ \mathbf{\Theta}) \tag{19b}$$

One of the ways to maximize the likelihood function construed for the requirements of this model is an application of an appropriate version of the Expectations Maximization algorithm (discussed, among others, in Dempster et al., 1977; Hamilton, 1990). The application of the EM algorithm to the switching models class requires deriving a dependence among the probabilities that were assessed on the basis of the available information that observation $\mathbf{y}_t$ was generated by regime j and the conditions imposed on the parameters by a system of equations resulting from the need for maximization of the likelihood functions:

$$\frac{\partial L}{\partial p_{kj}}(\mathbf{y}_1, ..., \mathbf{y}_T, \theta) = 0 \quad \text{for} \quad k = 1, 2, ..., N, \quad j = 1, 2, ..., r, \tag{20}$$

$$\sum_{k=1}^{N} p_{kj} = 1 \quad for \quad j = 1, 2, ..., r.$$

This dependence allows us to find in the consecutive iterations of the algorithm an increasingly better (as regards the criterion connected with the likelihood function) estimation of the parameter vector of the model. Hamilton (1990) proves that the estimator of the maximum likelihood method of vector $\boldsymbol{\Theta}$ is a solution of the following system of equations:

$$(\text{for } \boldsymbol{\Omega}_0 = \boldsymbol{\Omega}_1 = \boldsymbol{\Omega})$$

$$\mu_j = \frac{\sum_{t=1}^{T} \mathbf{y}_t \, \mathrm{p}(s_t = j/\mathbf{y}_1, ..., \mathbf{y}_T; \theta)}{\sum_{t=1}^{T} \mathrm{p}(s_t = j/\mathbf{y}_1, ..., \mathbf{y}_T; \theta)} \quad j = 0, 1 \tag{21}$$

$$\boldsymbol{\Omega} = \frac{\sum_{j=0}^{K} \sum_{t=1}^{T} (\mathbf{y}_t - \mu_j)(\mathbf{y}_t - \mu_j)^T \, \mathrm{p}(s_t = j/\mathbf{y}_1, ..., \mathbf{y}_T; \theta)}{T} \quad K = 1 \tag{22}$$

$$\mathrm{p}_{ij} = \frac{\sum_{t=2}^{T} \mathrm{p}(s_t = j, s_{t-1} = i/\mathbf{y}_1, ..., \mathbf{y}_T; \theta)}{\sum_{t=2}^{T} \mathrm{p}(s_{t-1} = i/\mathbf{y}_1, ..., \mathbf{y}_T; \theta)} \tag{23}$$

$$\mathrm{p} = \mathrm{p}(s_1 = 0/\mathbf{y}_1, ..., \mathbf{y}_T; \theta) \tag{24}$$

The procedure starts from the giving of any initial values of vector $\boldsymbol{\Theta}$, that will be used for estimation of probability smoothening (19b) (see Hamilton, 1990).

The estimated probability is then used for estimation of the parameters of vector $\boldsymbol{\Theta}$ described by equations (21)–(24). When the initial values have been determined for all model parameters, in each iteration of the EM algorithm, two steps are made. The first step (expectations) consists in determination of probabilities $P(S_t = j/\mathbf{Y}_t = \mathbf{y}_t)$ according to formula (17). The second step (maximization) leads to determination, with the help of a system of equations (18), of a vector that maximizes the likelihood function. Estimation stops when a given convergence criterion is met. In order to prevent the situation when the estimation is dominated by a small number of very high observations, the values that exceed the threshold of three average absolute deviations from the sample average were replaced with the height of this threshold. It is possible to show (Hamilton, 1994), that a sequence of estimations obtained in this way is convergent towards the local maximum of the likelihood function.

VIII. AN ATTEMPT OF APPLICATION OF MARKOV SWITCHING MODEL TO FORECASTING OF EXCHANGE RATE CHANGE

In analyzing exchange rates behavior, one can notice the occurrence of so-called long swings, i.e. periods of keeping up appreciative or depreciative rate tendencies. This phenomenon can be explained by an inflow of information to the currency market and their subjective interpretation by individual participants in the market, who also set exchange rates. The inflow of information occurs in an irregular way, in special cases in series (which means that the information can be correlated) and their meaning and influence on the change of the exchange rates are diverse. As a result, there are sub-periods of decreased and increased exchange rate change periods. Changeability of exchange rates depends on new, expected information (e.g. messages concerning the macroeconomic situation in the country). Dates of their announcement are usually known in advance, and uncertainty that concerns them results in an increase of exchange rate fluctuation.

According to the assumptions of the Hamilton model, the empirical series of logarithmic return rates for quotations of a given exchange rate should be decomposed into a phase of increase and a phase of decrease of the rate. In this way, two regimes will be distinguished: a regime corresponding to depreciation of the rate and a regime corresponding to appreciation of the rate. For each regime, its basic characteristics should be determined: the expected value of the rate change (μ_0 and μ_1) and the rate change diversity (σ_0 and σ_1) that will approximate the rate change dynamics within a given state. Transition probabilities for individual states

(p_{00} and p_{11}) should also be assessed. Transition probability values are the starting point for determination of an average length of the exchange rate appreciation and depreciation period:

$$a = \frac{1}{1 - p_{00}} \tag{25}$$

$$d = \frac{1}{1 - p_{11}} \tag{26}$$

where:

a – mean length of the appreciation period,
d – mean length of the depreciation period.

REFERENCES

Campbell J.Y., Lo A.W., MacKinlay A.C. (1997), *The Econometrics of Financial Markets*, Princeton University Press, Princeton, New Jersey.

Chrabonszczewska E., Kalicki K. (1996), *Teoria i polityka kursu walutowego*, Oficyna Wyd. SGH, Warszawa.

Dempster A.P., Laird N.M., Rubin D.B. (1977), Maximum likelihood for incomplete date via the EM algorithm, *Journal of the Royal Statistical Society*, **39**.

Dunis Ch. (ed.) (2001), *Prognozowanie rynków finansowych*, Oficyna Ekonomiczna, Kraków.

Engel C., Hamilton J.D. (1990), Long swings in the dollar: are they in data and do markets know it?, *American Economic Review*, **80**.

Engel Ch. (1992), Can the Markov switching model forecast exchange rates?, *NBER Working Paper*, no 4210.

Fama E. (1965), The behavior of stock market prices, *Journal of Business*, **38**.

Goldfeld S.M., Quandt R.E. (1973), A Markov model for switching regressions, *Journal of Econometrics*, **1**.

Hamilton J.D. (1989), A new approach to the economic analysis of nonstationary time series and the business cycle, *Econometrica*, **57**.

Hamilton J.D. (1990), Analysis of time series subject to changes in regime, *Journal of Econometrics*, **45**.

Hamilton J.D. (1994), *Time Series Analysis*, Princeton University Press, Princeton, New Jersey.

Iosifescu M. (1988), *Skończone procesy Markowa i ich zastosowania*, PWN, Warszawa.

Jajuga K. (ed.) (2000), *Metody ekonometryczne i statystyczne w analizie rynku kapitałowego*, Wyd. Akademii Ekonomicznej im. Oskara Langego, Wrocław.

Lee T.C., Judge G.G., Zellner A. (1997), *Estimating the Parameters of the Markov Probability Model from Aggregate Time Series Data*, North Holland, Amsterdam.

Osińska M. (2000), *Ekonometryczne modelowanie oczekiwań gospodarczych*, Wyd. UMK, Toruń.

Podgórska M., Śliwka P., Topolewski M., Wrzosek M. (2000), *Łańcuchy Markowa w teorii i w zastosowaniach*, Oficyna Wyd. SGH, Warszawa.

Shiyun W., Guan L.K., Chang C. (1999), A new methodology for studying intraday dynamics of Nikkei index futures using Markov chains, *Journal of International Financial Markets, Institutions and Money*, **9**.

Skrodzka W., Włodarczyk A. (2002), *Wykorzystanie łańcuchów Markowa do badania zmienności indeksu WIG 20*, II konferencja: „Rynek kapitałowy. Skuteczne inwestowanie", Wyd. Uniwersytetu Szczecińskiego, Szczecin.

Sobczyk K. (1973), *Metody dynamiki statystycznej*, PWN, Warszawa.

Titterington D.M., Smith A.F.M., Makov U.E. (1985), *Statistical Analysis of Finite Mixture Distributions*, Wiley, Chichester.

Aneta Włodarczyk, Tomasz Szmigiel

TEORETYCZNE ASPEKTY WYKORZYSTANIA MODELI MARKOWA DO BADANIA ZMIENNOŚCI KURSU WALUTOWEGO

Streszczenie

Prawidłowe oszacowanie kierunku zmian kursu wymiany może zmniejszyć ryzyko inwestycji w walutę lub może pozwolić na osiągnięcie większych dochodów z tej inwestycji. W opracowaniu tym autorzy przedstawiają propozycję zastosowania modeli Markowa do wykrycia i opisania prawidłowości rządzących procesem zmienności kursu walutowego. W pierwszej części została wykorzystana teoria łańcuchów Markowa do badania anomalii kalendarzowych występujących na rynku walutowym związanych z efektem weekendowym lub efektem stycznia. W artykule przedstawiona została również metoda oparta na teorii łańcuchów Markowa, która może posłużyć do zbadania wzajemnych powiązań pomiędzy zmiennością wolumenu obrotu oraz zmiennością cen dla terminowych kontraktów walutowych. W drugiej części zostaną przedstawione zagadnienia związane z budową i estymacją parametrów przełącznikowych modeli Markowa. W oparciu o modele przełącznikowe można prognozować zmiany kursu walutowego. Praca ma charakter teoretyczny. Badania empiryczne zostaną przeprowadzone w późniejszym terminie.