1999

Stochastic Threshold Models on Interest Rate

Huirong Li

Jian Yang

RESEARCH REPORT 9903

Stochastic Threshold Models on Interest Rate

by

Huirong Li   Jian Yang

December 1998

Department of Economics
Social Science Centre
University of Western Ontario
London, Ontario, Canada
N6A 5C2
econref@julian.uwo.ca

Stochastic Threshold Models on Interest Rate

Huirong Li and Jian Yang[1]
Department of Economics
The University of Western Ontario
London, Ontario, Canada N6A 5C2

December 1998

Abstract

Threshold models have been found useful in modelling nonlinearities in many financial time series. In this framework, the financial variable of interest evolves according to different dynamics, which is solely determined by the threshold regimes that the observed indicator variable falls into. This paper generalizes the threshold models to a class of stochastic threshold models, which allow for stochastic dependence of the current economic state on the threshold regimes. In a stochastic threshold model, different economic states are possible to occur within a certain threshold regime and each state occurs with some probability depend on the threshold regime and other recently observed information. Model identification and maximum likelihood estimation are developed. An study on short-term interest rate is conducted. We find that the short-term interest rate behaves asymmetrically in a rising versus a declining market. Declining market has significantly negative duration (in "return clock") dependence and rising market has insignificantly positive duration dependence. In the comparison of generalized autoregressive conditional heteroskedasticity models. threshold autoregressive models. generalized regime-switching models and stochastic threshold models. we find that our stochastic threshold model fits the data best in terms of alternative model selection criteria and in-sample forecasting. It also provides the best out-of-sample forecasting.

Key words: Stochastic threshold: Short-term interest rate: Conditional volatility; Maximum likelihood
JEL classification: G12: C22

1. Introduction

There has been an increasing interest in studying nonlinearities of financial time series in the literature. Many time series models have been used to capture different nonlinear features in financial data. Among them, threshold models (Tong, 1983) have drawn much attention in modelling financial dynamics. Threshold models approximate the nonlinearity of financial time series in a piece-wise linear fashion. They are often applied to the analysis of asymmetric patterns in financial data, sudden bursts at irregular time epochs and time irreversibility. The threshold autoregressive (TAR) models are originally illustrated in Tong (1983), while the popularly used ones are the self-exciting threshold autoregressive (SETAR) models proposed by Tong (1990). SETAR models apply autoregression in each piece of regime and use some lagged dependent variable as a threshold variable. While Tong's threshold model is used in modelling nonlinearities only in the conditional mean of a time series, Gourieroux and Monfort (1992) develop a class of dynamic models in which both the conditional mean and the conditional variance are endogenous stepwise functions. This is the so-called qualitative threshold ARCH (QTARCH) models. which are capable of capturing nonlinear behaviors in both the mean process and the variance of financial series. Applications of

TAR models include Kunst (1992) and Pfann, Schotman and Tschernig (1996) on interest rate, Cao and Tsay (1992) and Li and Li (1996) on daily stock returns, Lin and Yang (1998) on intraday stock returns and Siddiqui (1998) on exchange rates.

However, the existing threshold models have their limitations. The most serious problem which hinders the use of those models in financial studies is that the current state of the financial market is fully determined by a specific threshold, or equivalently speaking, those threshold models are not able to distinguish the current economic states from the threshold regimes determined by the indicator variable, usually the lagged dependent variable in TAR models. We would argue that although what we observe in past periods can provide useful information to help describe the dynamics of the current period, it can never fully reflect the state of today. In this paper, we establish a new class of threshold models, which generalize the existing threshold autoregressive models by distinguishing the threshold regimes and the economic states and by allowing for stochastic dependence of current economic states on the threshold regimes of the indicator variable. The proposed models are called the stochastic threshold autoregressive (STAR) models. In the STAR models, economic states are treated differently from the threshold regimes and different states are allowed within a certain threshold

2

regime. Each state occurs with some probability determined by the threshold regimes and other recently observed information. Meanwhile, the STAR model keeps the threshold setting to capture discrete changes and asymmetric patterns in financial series.

There are a few important features of our model. First, STAR models improve the flexibility of TAR models by describing the dynamics of financial variables as a double-probabilistic process[1]. Within each threshold regime, different states are allowed to occur with some *probability* and these economic states can shift from time to time according to a specified probability process. Second, when applied into the study of volatility behavior, our model sits between the deterministic and stochastic volatility models, which gives us a better approximation to the stochastic process of volatility than the usual deterministic models, for example, the generalized autoregressive conditional heteroskedasticity or GARCH models (Bollerslev, 1986) and allows higher complexity of model specifications than the stochastic volatility models. Lastly, the probabilities in our model are obtained through the exploration of observed information, which yields a natural way to investigate the impact of different information on the dynamics of interest rate,

---

[1]Many researchers in the literature have tried to improve the flexibility of TAR models by increasing the number of threshold regimes and/or the number of threshold varibales. However, any generalization along these directions remains assuming that observed information fully determines the current state.

stock returns and other financial variables.

It is important to distinguish STAR models from regime-switching (R-S) models (Hamilton, 1988) since both types of models allow for different states of a financial market. Obviously, first of all, like the usual TAR models, STAR models explicitly model market asymmetries, sudden changes and short-run fluctuations in financial markets. Secondly, STAR models differ from the R-S models in the nature of economic states and in the way of the determination of the states. In the R-S models, the states are assumed to be dependent over time, characterized by the state transition probabilities. The probabilities that the states occur at each period depend on the whole history of the state probabilities. However, in STAR models, the states are indeed "short-lived". At every period, current state is determined (with some probability), conditioned on all the observed information up to that period: there is no directly specified transitions among the states over time. The transitions from one state to another are indirectly carried through the observed information process from current period to next period. This major difference makes our specification of the STAR models fit into a lot of interesting situations in financial markets. in which short-run fluctuations [2] are more impor-

---

[2] By "short-run fluctuations". we mean the changes that do not have very persistent impact on the subsequent financial markets.

4

tant than the long-run structural changes. These include the situations when (1) a financial series contains too much noise that makes it impossible to identify the long-run states as defined in the R-S models (e.g. in high frequency data) and/or (2) there is no long-run changes occurring in the period of study[3].

To illustrate the properties of the STAR models, we study the dynamics of short-term interest rate within the proposed framework. The nominal short-term interest rate plays a key role in the valuation of almost all securities, which has made it one of the most frequently modeled variables in financial economics (Gray, 1996). Alternative GARCH models are often applied to capture the persistence in the conditional variance of the change of short rates. One potential source of misspecification of these GARCH models is that the structural forms of the conditional means and variances are held fixed throughout the entire sample period. In other words. these models are linear models and lack of power to capture sudden jumps, market asymmetry, and other nonlinear properties observed in the short rate. To relax the linearity inherent in these models and to examine the asymmetry in interest rates. TAR models have been applied by Kunst (1992) and Pfann. Schotman and Tschernig (1996). Meanwhile, R-S models of

---

[3]As in the S-R models. by long-run or structural changes. we only mean the changes that have a "long-lived" or persistent impact on the subsequent financial markets.

5

interest rates have been put forward by many researchers (Hamilton, 1988; Cai, 1994; Gray, 1996). The R-S models of short-term interest rates are motivated, in part, by the OPEC oil crisis (1973-1975) and the Federal Reserve experiment of 1979 to 1982. Specifically, Gray (1996) develops the so-called generalized regime-switching (GRS) model, which advances the regime-switching literature in many directions. A GRS model takes into consideration within-regime mean reversion and level effect on the conditional variance of the changes of short rates by introducing the level of short rate directly into the GARCH specification. It allows for state-dependence of all the GARCH parameters and time-varying transition probabilities (see, for example, Diebold, Lee, and Weinbach, 1994). The transition probabilities are assumed to depend on again the level of interest rates. In this paper, we will pursue the study of short rate in a STAR framework. Two states are proposed for the change of short rates characterized by both the mean parameters and the conditional variance parameters. Following Gray (1996), within-state mean reversion and level effect on conditional variance are both taken into consideration in our model. The state-occurring probabilities are allowed to be different (asymmetric) when following a rising market (positive change of short rates) verses a declining market (negative change of short rates) and are assumed to be dependent on the one-direction cumulative change of short rates. The one-direction

6

cumulative change of short rates measures the duration of a rising market or a declining market in terms of "return clock" instead of "time clock". Incorporating this variable into the state occurring probabilities will provide alternative insights on duration dependence, which has been studied by, among others, Durland and McCurdy (1994) on U.S. GDP growth, and Maheu and McCurdy (1997) on stock returns.

The rest of the paper is organized as follows. In section 2, we briefly review the existing threshold models and propose our stochastic threshold model. Section 3 discusses estimation and identification problems of the stochastic threshold model. In section 4, we demonstrate the properties of our model through a study on short-term interest rate. Section 5 concludes.

2. Model Development

2.1. The Threshold Autoregressive Models

The TAR models are first proposed by Tong (1983), while the popularly used ones are the SETAR models developed by Tong (1990). A simple SETAR process takes the form of. for a time series y_t,

$$y_t = \begin{cases} \phi_{01} + \phi_{11}y_{t-1} + u_t & \text{if } y_{t-1} \geq 0 \\ \\ \phi_{02} + \phi_{12}y_{t-1} + u_t & \text{if } y_{t-1} < 0 \end{cases}, \qquad (2.1)$$

where the second subscripts of the parameters indicate the two types of dynamics that y_t may follow according to the threshold regimes determined by the lagged value of y_t. The two types of dynamics are characterized by (ϕ_{01}, ϕ_{11}) and (ϕ_{02}, ϕ_{12}), respectively and y_{t-1} is the indicator variable or threshold variable. Tong (1990) provides a detailed description about the properties of the SETAR models. The features that SETAR models are able to capture include time irreversibility, asymmetric limit cycle and jump phenomenon. The SETAR models were mainly used, in the early stage, for the mean process and the variance is usually assumed to be constant over time. Tong (1990) combines the SETAR models with an ARCH specification and proposes the SETAR-ARCH models, in which the conditional mean takes the SETAR process and the conditional variance is allowed to have heteroskedasticity in an ARCH form.

While SETAR-ARCH models only capture certain nonlinearities in the conditional mean of a time series. Gourieroux and Monfort (1992) develop a class of threshold models in which both the conditional mean and the conditional variance are stepwise functions. Their models are the so-called *qualitative threshold ARCH*

8

(QTARCH) models, which are capable of capturing asymmetric patterns in both the conditional mean and the conditional variance. A simple QTARCH takes the following form:

$$
y_t = \begin{cases} \phi_{01} + \phi_{11}y_{t-1} + u_{t1} & \text{if } y_{t-1} \geq 0 \\ \phi_{02} + \phi_{12}y_{t-1} + u_{t2} & \text{if } y_{t-1} < 0 \end{cases}, \tag{2.2}
$$

where $u_{tr} = z_t \sqrt{h_t(\alpha_r)}$, $z_t \sim i.i.d.N(0,1)$, for $r = 1, 2$, and α_r is the vector of parameters in the conditional variance h_t under regime r; the conditional variance h_t has a GARCH specification in a threshold setting,

$$
h_t \equiv h_t(\alpha_r) = \begin{cases} \alpha_{01} + \alpha_{11}u_{t-1}^2 + \alpha_2 h_{t-1} & \text{if } y_{t-1} \geq 0 \\ \alpha_{02} + \alpha_{12}u_{t-1}^2 + \alpha_2 h_{t-1} & \text{if } y_{t-1} < 0 \end{cases}, \tag{2.3}
$$

where $\alpha_{ir} > 0$, for $i, r = 1, 2$, and $\alpha_2 > 0$. α_2 is assumed to be constant for two regimes to avoid non-invertability problem (Gourieroux and Monfort, 1992). In the QTARCH model, the mean process and the variance are each characterized by two sets of parameters, (ϕ_{01}, ϕ_{11}) and (ϕ_{02}, ϕ_{12}) for the mean, and $(\alpha_{01}, \alpha_{11}, \alpha_2)$ and $(\alpha_{02}, \alpha_{12}, \alpha_2)$ for the variance. Gourieroux and Monfort consider the statistical properties of their QTARCH models and use them to investigate the conditional variance of the daily relative change of the Paris stock index. Li and Li (1996)

9

define a double-threshold autoregressive ARCH (DTARCH) model, which shares the same idea of the QTARCH models, namely that the conditional mean takes a SETAR process and the conditional variance takes a SETAR-ARCH process. They find evidence of asymmetry in the conditional variance in Hong Kong Hang Seng index, using their DTARCH model.

A few other alternative threshold models have been proposed to increase the flexibility of the original threshold models. Among them, *smooth transition threshold* autoregressive models are proposed by Terasvirta (1990), in which the variant parameters in the autoregression are specified as "smooth" functions of the indicator variable rather than step functions of the indicator variable. However, threshold models are motivated to capture abrupt changes and asymmetric patterns in time series. The use of the "smooth" function instead of the step function will reduce the ability of threshold models in capturing *discrete sudden changes* and asymmetries in financial time series.

2.2. The Stochastic Threshold Autoregressive Models

Threshold models deal with asymmetric patterns explicitly and are able to capture sudden changes in financial time series. However. in the previous threshold models, the current economic states are treated equivalently as the observed threshold

regimes. In other words, the current economic states are assumed to depend deterministically on the observed threshold regimes of the indicator variable. We have argued that such treatment is not appropriate in many cases. We propose in the following a stochastic threshold model in which the economic states are treated differently from the threshold regimes determined by the threshold variable. It allows for stochastic dependence of current state on the observed threshold regimes of the indicator variable and provides much more flexibility in terms of model specification. Nevertheless, it retains the power of analyzing the market asymmetry and provides a natural way to examine how observed information might affect the dynamics of financial variables by allowing the information to influence the determination of state-occurring probabilities.

Without loss of generality, we consider a simple STAR model with one lagged dependent variable and two threshold regimes. For a time series y_t, the model takes the following form

$$\text{if } y_{t-1} \geq 0 \tag{2.4}$$

$$y_t = \begin{cases} \dot{\phi}_{01} + \phi_{11} y_{t-1} + u_{t1} & \text{with probability } p(x_{t-1}) \\ \dot{\phi}_{02} + \dot{\phi}_{12} y_{t-1} + u_{t2} & \text{with probability } 1 - p(x_{t-1}), \end{cases}$$

11

$$\text{if } y_{t-1} < 0 \tag{2.5}$$

$$y_t = \begin{cases} \phi_{01} + \phi_{11} y_{t-1} + u_{t1} & \text{with probability } 1 - q(x_{t-1}) \\ \phi_{02} + \phi_{12} y_{t-1} + u_{t2} & \text{with probability } q(x_{t-1}) \end{cases}$$

where the second subscripts refer to two economic states under each threshold regime, state 1 and state 2 ($i = 1, 2$), which are characterized by two sets of parameters, (ϕ_{01}, ϕ_{11}) and (ϕ_{02}, ϕ_{12}). The indicator (threshold) variable in this model is y_{t-1}. There are two threshold regimes in this model, $y_{t-1} \geq 0$ and $y_{t-1} < 0$. It is clear that in this setting, the current state of the market is distinguished from the threshold regime that is determined by last observation of the indicator variable. The occurrences of state 1 and state 2 are both possible given the threshold regime. The probabilities of the occurrences of state 1 and state 2 are governed by two probability processes, p under regime $y_{t-1} \geq 0$, and q under regime $y_{t-1} < 0$. We note that the probabilities are not trivially constant. instead they are time-variant and depend on all the relevant information. x_{t-1}. Introducing x_{t-1}, a vector of exogenous or predetermined variables. into the processes of probabilities p and q allows us to examine how these variables. x_{t-1} will affect the realizations of the state of the world. or the probabilities of being in any state under a certain threshold regime. Therefore. the market asymmetry is represented by different

12

structural parameters determining p and q. The market innovations, u_{t1} and u_{t2}, are allowed to follow different processes in different states and might be correlated across time, but are required to be independent across the states.

The STAR model generalizes Tong's threshold models in a natural way by distinguishing the threshold regimes from the economic states and allowing for the stochastic dependence of the current state on the observed threshold regimes. The economic states describe the important features of a financial market at each time period. However, threshold regimes are only segments of a pre-determined threshold variable. While keeping the properties of Tong's threshold model, the STAR model will be able to produce richer implications by allowing the states to occur in a non-pre-determined way. As the traditional threshold models, the STAR model can be easily adapted to have more than two number of regimes, states, and more autoregressive orders.

2.3. A Simple Stochastic Threshold Autoregressive "GARCH" Model

In this subsection, we consider in detail a simple stochastic threshold model which incorporates GARCH effects in the conditional variance (STGARCH). Some problems associated with the usual GARCH specification in a stochastic threshold model will be avoided through the adoption of some recent results in the litera-

ture.

(1) Specification for the conditional mean

To consider the possible asymmetric patterns in the mean process, we specify the conditional mean in the form of a stochastic threshold autoregressive model, or a STAR model. It takes the general form as in the equation (2.4) and (2.5), a simple STAR(1,2) model. We denote the parameters in the conditional mean equations as $\phi_i = (\phi_{0i} \ \varphi_{1i})'$, $i = 1, 2$.

(2) Specification for the conditional variance

To accommodate discrete changes, volatility clustering and asymmetric patterns in the conditional variance, we employ the GARCH specification in a STAR model. If we can observe the state of last period, then we have the usual way to employ the GARCH specification. In this ideal case, we are able to compute the *true* residual and the *true* variance of last period. denoted as u_{t-1}^* and h_{t-1}^*[4], respectively. A stochastic threshold model with a GARCH specification is then,

$$
\text{if } y_{t-1} \geq 0 \tag{2.6}
$$

$$
h_t^* = \begin{cases} \alpha_{01} + \alpha_{11} u_{t-1}^{*}{}^2 + \alpha_2 h_{t-1}^* & \text{with probability } p(x_{t-1}) \\ \alpha_{02} + \alpha_{12} u_{t-1}^{*}{}^2 + \alpha_2 h_{t-1}^* & \text{with probability } 1 - p(x_{t-1}), \end{cases}
$$

---

[4]From now on. we will use h_t^* to denote the true but unobserved variance and h_t to denote the expected value as defined by Gray in all STGARCH models. u_t^* and u_t are defined similarly.

if $y_{t-1} < 0$ (2.7)

$$h_t^* = \begin{cases} \alpha_{01} + \alpha_{11}{u_{t-1}^*}^2 + \alpha_2 h_{t-1}^* & \text{with probability } 1 - q(x_{t-1}) \\ \\ \alpha_{02} + \alpha_{12}{u_{t-1}^*}^2 + \alpha_2 h_{t-1}^* & \text{with probability } q(x_{t-1}) \end{cases},$$

where $u_{ti}^* = z_t\sqrt{h_t^*(\alpha_i)}$, $z_t \sim i.i.d.$ $N(0,1)$ and $h_t^* \equiv h_t^*(\alpha_i) = \alpha_{0i} + \alpha_{1i}{u_{t-1}^*}^2 + \alpha_2 h_{t-1}^*$, $\alpha_{0i} > 0$, $\alpha_{1i} \geq 0$, $\alpha_2 \geq 0$, and $\alpha_{1i} + \alpha_2 < 1$, for $i = 1, 2$. However, in the case of a stochastic threshold model for the conditional variance, the state at the beginning of each time period is not observed and the usual GARCH specification will cause the path-dependence problem in the following way. The conditional variance at time t in the above specification, h_t^*, depends on the state at time t and the conditional variance at time $t-1$, h_{t-1}^*, which depends on the state at time $t-1$ and the conditional variance at time $t-2$, h_{t-2}^*, and so on. Consequently, the conditional variance at time t depends on the entire sequence of states up to time t. The likelihood function is constructed by integrating over all the possible paths. We therefore face the similar path-dependence problem as in a switching-regime GARCH model. Hamilton and Susmel (1994) point out that switching-regime GARCH models are essentially intractable and impossible to estimate due to the dependence of the conditional variance on the entire past history of the data in a GARCH model.

To deal with the path-dependence problem, we apply the *generalized* GARCH specification proposed by Gray (1996), in which the conditional variance at current period is assumed to be determined by the *expected* conditional residual of last period and the *expected* conditional variance of last period, u_{t-1} and h_{t-1}, rather than the *unobserved true* conditional residual and the *unobserved true* conditional variance, u_{t-1}^* and h_{t-1}^*. Our specification for the conditional variance with a STAR process is thus,

if $y_{t-1} \geq 0$

$$h_t^* = \begin{cases} \alpha_{01} + \alpha_{11}u_{t-1}^2 + \alpha_2 h_{t-1} & \text{with probability } p(x_{t-1}) \\ \alpha_{02} + \alpha_{12}u_{t-1}^2 + \alpha_2 h_{t-1} & \text{with probability } 1 - p(x_{t-1}), \end{cases} \tag{2.8}$$

if $y_{t-1} < 0$

$$h_t^* = \begin{cases} \alpha_{01} + \alpha_{11}u_{t-1}^2 + \alpha_2 h_{t-1} & \text{with probability } 1 - q(x_{t-1}) \\ \alpha_{02} + \alpha_{12}u_{t-1}^2 + \alpha_2 h_{t-1} & \text{with probability } q(x_{t-1}), \end{cases} \tag{2.9}$$

where h_t^* is the true but unobserved conditional variance of return and h_{t-1} is defined as the following. As shown by Gray (1996). conditional on the normality within each state. i.e. assuming the conditional normality of u_{ti} (for all t) under

16

each state i, the conditional variance at time $t-1$, h_{t-1}, is given by

if $y_{t-2} \geq 0$ $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ (2.10)

$$
\begin{aligned}
h_{t-1} &\equiv E[y_{t-1}^2|\Omega_{t-2}] - [E[y_{t-1}|\Omega_{t-2}]]^2 \\
&= p_{t-1}((\mu_{t-1,1})^2 + h_{t-1,1}) + (1 - p_{t-1})((\mu_{t-1,2})^2 + h_{t-1}^{(2)}) \\
&\quad -(p_{t-1}\mu_{t-1,1} + (1 - p_{t-1})\mu_{t-1,2})^2,
\end{aligned}
$$

if $y_{t-2} < 0$ $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ (2.11)

$$
\begin{aligned}
h_{t-1} &\equiv E[y_{t-1}^2|\Omega_{t-2}] - [E[y_{t-1}|\Omega_{t-2}]]^2 \\
&= (1 - q_{t-1})((\mu_{t-1,1})^2 + h_{t-1,1}) + q_{t-1}((\mu_{t-1,2})^2 + h_{t-1}^{(2)}) \\
&\quad -((1 - q_{t-1})\mu_{t-1,1} + q_{t-1}\mu_{t-1,2})^2,
\end{aligned}
$$

where Ω_{t-2} is the information set at period $t-1$, $p_{t-1} \equiv p(x_{t-2})$, $q_{t-1} \equiv q(x_{t-2})$, and $\mu_{t-1,i} = \omega_{0i} + \omega_{1i}y_{t-2}$, for $i = 1, 2$. Now h_{t-1} is not path-dependent and can be used to construct h_{t1} and h_{t2} as the following:

$$
h_{ti} = \alpha_{0i} + \alpha_{1i}u_{t-1}^2 + \alpha_2 h_{t-1}, \quad i = 1, 2.
$$

17

where u_{t-1} is similarly defined as follows:

if $y_{t-2} \geq 0$ $\hspace{8cm}$ (2.12)

$$u_{t-1} = y_{t-1} - E[y_{t-1}|\Omega_{t-2}]$$

$$= y_{t-1} - (p_{t-1}(\phi_{01} + \phi_{11}y_{t-2}) + (1 - p_{t-1})(\phi_{02} + \phi_{12}y_{t-2}))$$

if $y_{t-2} < 0$

$$u_{t-1} = y_{t-1} - E[y_{t-1}|\Omega_{t-2}]$$

$$= y_{t-1} - ((1 - q_{t-1})(\phi_{01} + \phi_{11}y_{t-2}) + q_{t-1}(\phi_{02} + \phi_{12}y_{t-2})).$$

After knowing h_{t1} and h_{t2}, equations (2.10, 2.11) are used to construct h_t, which will be used to replace the unobserved h_t^* in the model. In this "GARCH" specification, we overcome the path-dependence by justifying the usual GARCH specification with the values of the conditional variance and the conditional mean under the assumption of conditional normality in each state, while still capture the nature of persistence in the conditional variance.

(3) Specification for the state probabilities

We use a logistic function for the probability specification (as in Diebold, Lee

and Weinbach (1994)), that is,

$$p_t \equiv p(x_{t-1}) = \frac{\exp(\beta_p x_{t-1})}{1 + \exp(\beta_p x_{t-1})}$$

$$q_t \equiv q(x_{t-1}) = \frac{\exp(\beta_q x_{t-1})}{1 + \exp(\beta_q x_{t-1})}$$

(2.13)

where x_{t-1} is a vector of exogenous or pre-determined variables. p_t is the probability of being in state 1 at time t, if $y_{t-1} \geq 0$; q_t is the probability of being in state 2 at time t, if $y_{t-1} < 0$. Specifically, the probabilities of the occurrence of state 1 and state 2 can be different when $y_{t-1} \geq 0$ and when $y_{t-1} < 0$, which are identified by the two sets of parameter vectors in the probabilities, β_p and β_q. In other words, the impact of x_{t-1} on the probabilities of being in any state could be different under the regime of $y_{t-1} \geq 0$ from under the regime of $y_{t-1} < 0$. Again, the specifications of p_t and q_t distinguish our model from the switching-regime model in the sense that our probabilities are not state-dependent. instead they are determined by some observed variables. In this way, our model considers the importance of short-run information instead of long-run information which is carried through the whole history in the switching-regime models.

3. Estimation and Identification

We assume conditional Gaussianity within each state and perform maximum likelihood estimation. The conditional likelihood function consists of two parts,

$$L_{p,n}(\theta, \beta) \tag{3.1}$$

$$= -\frac{1}{2n} \sum_{t=s+1}^{n} \{p_t(\ln h_{t1} + (u_{t1}^2/h_{t1})) + (1 - p_t)(\ln h_{t2} + (u_{t2}^2/h_{t2}))\},$$

$$L_{q,n}(\theta, \beta)$$

$$= -\frac{1}{2n} \sum_{t=s+1}^{n} \{(1-q_t)(\ln h_{t1} + (u_{t1}^2/h_{t1})) + q_t(\ln h_{t2} + (u_{t2}^2/h_{t2}))\},$$

and the complete conditional likelihood function is

$$L_n(\theta, \beta) = L_{p,n}(\theta, \beta) \cdot I(y_{t-1} \geq 0) + L_{q,n}(\theta, \beta) \cdot I(y_{t-1} < 0),$$

where θ is the vector of parameters to be estimated, $I(\cdot)$ is an indicator function, h_{ti} and $u_{ti}(i = 1, 2)$ are obtained as the way defined in equation (2.10, 2.11) and (2.12), and p_t and q_t are defined as in equation (2.13).

The usual assumptions for a threshold autoregressive GARCH process are summarized as the following:

(1) The time series $\{y_t\}$ is stationary and ergodic.

(2) $E[y_t^2] < \infty$.

(3) All the parameters in the conditional variance are positive or non-negative, $\alpha_2 > 0$ and $\alpha_{ki} \geq 0$ for $i = 1, 2$, and $k = 0, 1$.

(4) Let $\theta_i = \begin{pmatrix} \phi_{0i} & \phi_{1i} & \alpha_{0i} & \alpha_{1i} & \alpha_2 \end{pmatrix}^T$, for $i = 1, 2$, then $\theta_1 \neq \theta_2$.

Besides the above assumptions, we require one additional assumption in order to make the STAR model fully identified.

(5) $\theta_1 > \theta_2$, where the inequality is defined as $\theta_1 > \theta_2$ if $\theta_{d1} > \theta_{d2}$, where $(\theta_{d1}, \theta_{d2})$ is the first pair of the element-by-element-different parameters in θ_1 and θ_2. The last assumption is needed to identify the probability parameters, β, over states. In our specification of the logistic probability function, we find that

$$\frac{\exp(-\gamma z)}{1 + \exp(-\gamma z)} = \frac{1}{1 + \exp(\gamma z)} = 1 - \frac{\exp(\gamma z)}{1 + \exp(\gamma z)},$$

for any γ and z. Therefore. we cannot identify the model without the assumption (5) since the log-likelihood value at (θ_1, θ_2) associated with β will be exactly the same as the value at (θ_2, θ_1) associated with $-\beta$.

4. Modelling Short-Term Interest Rates

4.1. Model Specification

A stylized fact of short-term interest rate is mean-reverting, which is commonly modelled by letting next period's change in the short rate depend linearly on the current short rate level. A commonly used model of short rates with GARCH variance is as follows.

$$dr_t = \phi_{01} + \phi_{11} r_{t-1} + u_t, \tag{4.1}$$

$$h_t = \alpha_{01} + \alpha_{11} u_{t-1}^2 + \alpha_{21} h_{t-1}, \tag{4.2}$$

where dr_t is the change of short rates at time t, $dr_t = r_t - r_{t-1}$, r_{t-1} is the short rate at time $t - 1$, and u_t follows a normal distribution with mean zero and conditional variance h_t. To take into account the level effect on the conditional variance, which is studied by Cox, Ingersoll, and Ross (1985) in a continuous time setup, Gray (1996) suggests to add the level of short rate directly into the GARCH specification. Then the conditional variance becomes

$$h_t = \alpha_{01} + \alpha_{11} u_{t-1}^2 + \alpha_{21} h_{t-1} + \alpha_{31}^2 r_{t-1}. \tag{4.3}$$

In experimenting this model and the following models, α_{01} term always converges to the boundary ($\alpha_{01} = 0$), which is consistent with the finding reported by Gray (1996). Therefore, in the following study, we will assume the GARCH process is simply defined as

$$h_t = \alpha_{11} {u_{t-1}}^2 + \alpha_{21} h_{t-1} + \alpha_{31}^2 r_{t-1}. \tag{4.4}$$

To model the asymmetry and short-run fluctuations in short rate, a modified STGARCH model will be adopted, which is as follows.

if $dr_{t-1} \geq 0$ $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ (4.5)

$$dr_t = \begin{cases} \Phi_{01} + \phi_{11} r_{t-1} + u_{t1} & \text{with probability } p(x_{t-1}) \\ \Phi_{02} + \Phi_{12} r_{t-1} + u_{t2} & \text{with probability } 1 - p(x_{t-1}), \end{cases}$$

if $dr_{t-1} < 0$ $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ (4.6)

$$dr_t = \begin{cases} \phi_{01} + \Phi_{11} r_{t-1} + u_{t1} & \text{with probability } 1 - q(x_{t-1}) \\ \phi_{02} + \Phi_{12} r_{t-1} + u_{t2} & \text{with probability } q(x_{t-1}), \end{cases}$$

23

if $dr_{t-1} \geq 0$

$$h_t^* = \begin{cases} \alpha_{11}{u_{t-1}}^2 + \alpha_2 h_{t-1} + \alpha_{31}^2 r_{t-1} & \text{with probability } p(x_{t-1}) \\ \\ \alpha_{12}{u_{t-1}}^2 + \alpha_2 h_{t-1} + \alpha_{32}^2 r_{t-1} & \text{with probability } 1 - p(x_{t-1}), \end{cases} \quad (4.7)$$

if $dr_{t-1} < 0$

$$h_t^* = \begin{cases} \alpha_{11}{u_{t-1}}^2 + \alpha_2 h_{t-1} + \alpha_{31}^2 r_{t-1} & \text{with probability } 1 - q(x_{t-1}) \\ \\ \alpha_{12}{u_{t-1}}^2 + \alpha_2 h_{t-1} + \alpha_{32}^2 r_{t-1} & \text{with probability } q(x_{t-1}), \end{cases} \quad (4.8)$$

where h_t^*, u_{t-1} and h_{t-1} are defined as in section 2.3. The state-occurring probability specification is that

$$p_t \equiv p(cdr_{t-1}) = \frac{\exp(\beta_{0p} + \beta_{1p}cdr_{t-1})}{1 + \exp(\beta_{0p} + \beta_{1p}cdr_{t-1})} \quad (4.9)$$

$$q_t \equiv q(cdr_{t-1}) = \frac{\exp(\beta_{0q} + \beta_{1q}cdr_{t-1})}{1 + \exp(\beta_{0q} + \beta_{1q}cdr_{t-1})},$$

where cdr_{t-1} is the one-direction cumulative change of short rates at period $t-1$ defined as follows.

$$cdr_t = cdr_{t-1} + dr_t, \quad \text{if } dr_t \cdot cdr_{t-1} \geq 0; \quad (4.10)$$

$$cdr_t = dr_t, \quad \text{if } dr_t \cdot cdr_{t-1} < 0.$$

24

cdr_t measures the one-direction cumulative positive changes or negative changes of short rates and can be interpreted as the duration of a rising or a declining market using a "return clock" rather than a "time clock". Allowing cdr_{t-1} to enter the state-occurring probability process will enable us to study the duration dependence in short rate. In the above model, under each threshold regime, there are two states in the market, state 1 and state 2, which are characterized by two sets of parameters $\theta_1 = (\phi_{01}, \phi_{11}, \alpha_{11}, \alpha_2, \alpha_{31})$ and $\theta_2 = (\phi_{02}, \phi_{12}, \alpha_{12}, \alpha_2, \alpha_{32})$. At each time period t, state 1 occurs with probability p_t when following a positive change of short rate and it occurs with probability $1 - q_t$ when following a negative change of short rate; state 2 occurs with probabilities $1 - p_t$ and q_t respectively under different threshold regimes.

4.2. Data Description

The data used in this study is the weekly annualized percentage yields on one-month U.S. Treasure bill rates recorded on every Friday from May 1986 through May 1996, which is extracted from the DataStreams provided by Ivey Business School at the University of Western Ontario. The data set includes 525 observations. The first two observations will not be included in the estimation since we need to construct the first order difference of the rates and the lagged difference of

25

the rate is selected to be the threshold variable. Therefore, the sample size in our study is 523. Summary statistics are presented in Table 1. Figure 1 and Figure 2 plot the short rates and the changes of short rates respectively.

4.3. Empirical Results

A number of specifications for the short rate are estimated and compared to our STGARCH model. These include a constant-variance model, a GARCH model defined by (4.1) and (4.2), a double threshold model (DTGARCH) as a special case of STGARCH model with $p(cdr_{t-1}) = 1$ and $q(cdr_{t-1}) = 1$, and a GRS model of Gray (1996).

In the GRS model estimated, there are two possible states of the financial market. With u_{t-1}, h_{t-1}, and h_t^* defined in section 2.3, a GRS model takes the following form.

$$dr_t = \begin{cases} \phi_{01} + \phi_{11} r_{t-1} + u_{t1}, & \text{if } S_t = 1. \\ \phi_{02} + \phi_{12} r_{t-1} + u_{t2}, & \text{if } S_t = 2: \end{cases} \tag{4.11}$$

$$h_t^* = \begin{cases} \alpha_{11} u_{t-1}^2 + \alpha_2 h_{t-1} + \alpha_{31}^2 r_{t-1}. & \text{if } S_t = 1. \\ \alpha_{12} u_{t-1}^2 + \alpha_2 h_{t-1} + \alpha_{32}^2 r_{t-1}. & \text{if } S_t = 2. \end{cases} \tag{4.12}$$

26

where the coefficient of h_{t-1} is assumed to be constant over two states since it is found difficult to identify the model with different coefficients of h_{t-1} over two states and the benefit (according to alternative model selection criteria) from this more general specification is marginal for the data we study. The transition probabilities are time-varying and defined as

$$
\begin{aligned}
prob(S_t &= 1|S_{t-1} = 1) \equiv p_t(cdr_{t-1}) = \frac{\exp(\beta_{0p} + \beta_{1p}cdr_{t-1})}{1 + \exp(\beta_{0p} + \beta_{1p}cdr_{t-1})}, \quad (4.13)\\
prob(S_t &= 2|S_{t-1} = 2) \equiv q_t(cdr_{t-1}) = \frac{\exp(\beta_{0q} + \beta_{1q}cdr_{t-1})}{1 + \exp(\beta_{0q} + \beta_{1q}cdr_{t-1})}.
\end{aligned}
$$

It is worth noting the following. First of all, p_t and q_t in the STGARCH model are state occurring probabilities, which are solely determined by recently observed information including the one-direction cumulative change of short rates, cdr_{t-1}, and the sign of the change of short rates. dr_{t-1}. However. in the GRS model. p_t and q_t are state transition probabilities. which illustrate the probabilistic evolution path of the unobserved states. The state occurring probabilities are determined by the whole probabilistic history of the economic states. Therefore. the R-S model might not perform well if the market fluctuations are dominated by short-run fluctuations. in other words. there are very few changes which have long-lived or highly persistent impact on the subsequent markets. Second. the above GRS

27

model is different from the one in Gray (1996) in the way that Gray (1996) specifies the time-varying transition probability depending on the level of the short rate r_{t-1}. We find his specification hard to be estimated for this data since the impact of the level of the short rate on the volatility of the change of short rates has been already captured by specifying a mean reverting and allowing the level of short rate in the conditional variance equation.

The estimation results are presented in Table 2 for these four models and for our STGARCH model. The first column of Table 2 reports maximum likelihood estimates of the single-state constant-variance model. The conditional mean terms are significantly different from zero and ϕ_{11} is negative, showing reversion to the mean. The implied long-run mean $(-\phi_{01}/\phi_{11})$ is 5.1968%. The second column of Table 3 reports the results for the GARCH model with the level of short rate in the conditional variance (α_{31}). Although the estimates in the mean terms are quite different from that in the single-state constant-variance model, they are still significantly different from zero and with mean reverting. The persistence of the volatility implied by the GARCH model is 0.8312 $(\alpha_{11} + \alpha_2)$, which is relatively lower than reported in the literature of short rates. For example. Kees, Nissen, Schotman. and Wolff (1994) report $\alpha_{11} + \alpha_2 = 1.10$ for one-month T-bills, and Hong (1988) reports $\alpha_{11}+\alpha_2 = 1.073$ for excess returns on three-month T-bills over

28

one-month T-bills. The reason is that our data does not cover the periods of OPEC oil crisis and the Federal Reserve Experiment. As Lamoureux and Lastrapes (1990) argue, the high persistence implied by the GARCH model might be mis-specification of sudden changes and/or long-run shifts. Excluding the periods with long-run shifts which have very persistent impact on the future market makes us fortunately avoid the possible mis-specifications implied by GARCH models. In this GARCH specification, the parameter α_{31} represents the impact of the level of the short rate on the volatility, which is found statistically significant from the estimation results.

The DTGARCH model explicitly considers the market asymmetry by assuming alternative dynamics following a rising market (positive change of short rates) versus a declining market (negative change of short rates). The third column of Table 2 presents the estimation results of the DTGARCH model. Clearly, market asymmetry appears in both the mean and the conditional variance. Following a rising market. the implied long-run mean $-\omega_{01}/\varphi_{11} = 0.4192\%$ of short rate is lower than that following a declining market with $-\omega_{02}/\phi_{12} = 6.4180\%$. The mean reversion parameters in both markets are not statistically significant. which is due to that the mean reversion effect can be migrated partially by assuming asymmetric dynamics for different markets. The high change of short rates is associated

with high persistence in volatility ($0.9253 = \alpha_{12} + \alpha_2 > \alpha_{11} + \alpha_2 = 0.6041$). And we also observe that the estimate of α_{32} is much greater than that of α_{31}, this observation together with the high persistence following a declining market simply indicates high volatility following a declining market. The great differences between the parameter estimates following a rising market and the parameter estimates following a declining market suggest that market asymmetry is an important issue when studying short rates. However. the first two models assume linearity in the mean and variance, ignoring possible short-run jumps and long-run shifts. The DTGARCH model deals with the nonlinearities by assuming nonstochastic dependence of the current state of the market on the last period observed information (threshold regimes of the indicator variable), which certainly limits its power of describing market fluctuations appearing as transitions from one state to another state. A better and more flexible model would allow for stochastic dependence of the current state of the market on the previous observed information. which motivates both the regime-switching model and the stochastic threshold model in this paper. The most important difference between these two models is that the stochastic threshold model uses the recently observed information to determine the state occurring probabilities. while the regime-switching model takes into account the information back to the very past by assuming state-dependence.

Neverthless, given the similarities between these two models, the estimation results at the forth column for GRS and the fifth column for STGARCH have similar patterns. State 1 is characterized by higher rate, higher mean reversion, and higher volatility. Within each state, the GARCH process is stationary and the implied persistence is much reduced with $\alpha_{i2} + \alpha_2 < 0.5$ in the GRS model and $\alpha_{i2} + \alpha_2 < 0.4$ in the STGARCH model for state $i = 1, 2$. For the state-occurring probabilities in the STGARCH model and the transition probabilities in the GRS model, the estimates of β_{1p} are positive in both models, which implies insignificantly positive duration (in return clock) dependence, i.e. the longer in a high volatile state, the more likely the market will stay in this state. However, the estimates of β_{1q} are negative in both models and very significant in the STGARCH model. which implies significantly negative duration (in return clock) dependence, i.e. the longer in a less volatile state. the more likely the market will switch to the highly volatile state.

4.4. Model Selection

To make comparisons of the alternative models. we report in Table 3 various model selection statistics including the maximal value of the log-likelihood function. the Akaike Information Criterion (AIC) and the Schwarz criterion. Our model shows

better fit than the alternative models according to all the criteria. We rank the five models not surprisingly in the order as in Table 3 from left to right. To have a close look at the comparison results between the GRS model and the STGARCH model, we plot out the ex ante and smoothed probabilities that the short rate process is in state 1 (the high-volatility state) at time t. The ex ante probability is based on information available at time t ($Pr[S_t = 1|\Phi_{t-1}]$) and the smoothed probability is based on the entire sample ($Pr[S_t = 1|\Phi_T]$). Certainly, the ex ante probability is more important to measure the performance of the GRS model in terms of forecasting. The state occurring probability of state 1 (ex ante probability as specified) implied by the STGARCH model is also plotted. Compared with the plot of absolute value of the change of short rates, we conclude that STGARCH approach describes much better the dynamics of the changes of short rates in terms of the probabilities of falling into different states.

We also make a comparison of the in-sample forecasting of the volatility of the change of short rates. Since the conditional variance is an expectation of squared innovations to the interest rate process, we measure the forecasting error by mean squared error (MSE) and mean absolute error (MAE) between actual volatility ($av_t = u_t^2$ where $u_t = dr_t - E_{t-1}[dr_t]$) and forecast volatility ($fv_t = E_{t-1}[u_t^2]$). It is found that our STGARCH model provides better in-sample forecasting per-

formance by both MSE and MAE measures. In general, we conclude that our model yields the best fit in terms of various model selection criteria and in-sample forecasting performance.

To avert fears of overparametrization of the STGARCH model, and to establish the economic significance of determining the state of the market by recently observed information, we conduct a comparison of the out-of-sample forecasting of alternative models, particularly the GRS model and the STGARCH model. We have used the samples prior to January 1994 to implement the estimation of alternative models and the samples after January 1994 are left for the study of out-of-sample forecasting of the volatility. We measure the forecasting error by OMSE, OMAE, TMSE, TMAE, and R^2. OMSE and OMAE are the mean squared errors and mean absolute errors, respectively, over the out-of-sample period; TMSE and TMAE are the mean squared errors and mean absolute errors over in-sample and out-of-sample period (the entire period); and R^2 is defined (see Gray, 1996) as

$$R^2 = 1 - \frac{\sum_{t=1}^{T}(av_t - fv_t)^2}{\sum_{t=1}^{T} av_t^2}. \tag{4.14}$$

which is calculated over the whole sample period and provides a direct measure of the goodness of fit. The STGARCH model performs well over five measures. How-

ever, the GRS model performs even worse than the DTGARCH model. We could conclude that short-run information matters more in determining the dynamics of short rate. This is not surprising typically when the data we study covers the period without shifts that have highly persistent impact on the subsequent markets.

5. Conclusion

This paper generalizes the threshold models into a class of stochastic threshold models, in which economic states are distinguished from statistical threshold regimes. This generalization addresses the major criticism of the existing threshold models of not-knowing the current state and tries to uncover the state of the world up to some probability.

We pursue the study of short term interest rates within the proposed framework. The proposed STGARCH model outperforms the traditional threshold model and the generalized regime-switching model in terms of both in-sample fitting and out-of-sample forecasting. Asymmetric patterns appear in the volatility of the changes of short rates in a way that highly volatile market follows a rising market while less volatile market follows a declining market. The highly volatile

market has a positive duration dependence, however, the less volatile market has a negative duration dependence. Our results point towards some promising directions for future research within the general framework proposed in the paper.

REFERENCES

Bollerslev, T. (1986), "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31, 307-327.

Cai, J. (1994), "A Markov Model of Switching-Regime ARCH," *Journal of Business & Economics Statistics*, 12, 309-316.

Cao, C. Q. and Tsay, R. S. (1992), "Nonlinear Time-series Analysis of Stock Volatilities," *Journal of Applied Econometrics*, 7, 165-185.

Dacco, R. and Satchell, S. (1995), "Why Do Regime Switching Models Forecast so Badly?" Discussion Paper August-1995, Birkbech College, Department of Economics.

Diebold, F. X., Lee, J. and Weinbach, G. C. (1994), "Regime Switching with Time-Varying Transition Probabilities," in *Nonstationary time series analysis and cointegration. Advanced Texts in Econometrics*. Hargreaves. Colin-P.. ed. Oxford and New York: Oxford University Press, 283-302.

Durland, J. M. and McCurdy,T. H. (1994), "Duration-Dependent Transitions in a Markov Model of U.S. GNP Growth," *Journal of Business & Economic Statistics*. 12, 279-88.

Gourieroux. C. and Monfort. A. (1992), "Qualitative Threshold ARCH Models." *Journal of Econometrics*. 52, 159-199.

Gray, S. F. (1996), "Modeling the Conditional Distribution of Interest Rates as A Regime-Switching Process," *Journal of Financial Economics, 42*, 27-62.

Hamilton, J. D. (1988), "Rational Expectations Econometric Analysis of Changes in Regime: An Investigation of the Term Structure of Interest Rate," *Journal of Economic Dynamics and Control, 12*, 385-423.

—— (1989), "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," *Econometrica, 57*, 357-384.

—— (1990), "Analysis of Time Series Subject to Changes in Regime," *Journal of Econometrics, 45*, 39-70.

—— and Susmel, R. (1994), "Autoregressive Conditional Heteroskedasticity and Changes in Regime." *Journal of Econometrics, 64*, 307-334.

Kunst, R. M. (1992), "Threshold Cointegration in Interest Rates." Discussion Paper 92-26. University of California. San Diego, Department of Economics.

Lamoureux. C. and Lastrapes, W. D. (1990), "Persistence in Variance, Structural Change, and the GARCH Model." *Journal of Business & Economic Statistics. 8,* 225-234.

Li, C. W. and Li. W. K. (1996), "On A Double-Threshold Autoregressive Heteroskedastic Time Series Model." *Journal of Applied Econometrics. 11.* 253-274.

Lin, S. and Yang, J. (1998), "Equity Returns Dynamics and Exogenous Buy/Sell

Trading Information," manuscript, the University of Western Ontario, Department of Economics.

Maheu, J. M. and McCurdy, T. H. (1997), "Identifying Bull and Bear Markets in Stock Returns," manuscript, Queen's University, Department of Economics.

Pfann, G. A., Schotman, P. C. and Tschernig, R. (1996), "Nonlinear Interest Rate Dynamics and Implications for the Term Structure," *Journal of Econometrics*, 74, 149-76.

Siddiqui, S. (1998), "A Qualitative Threshold Model of Daily Exchange Rate Movements," *Economics Letters*, 59, 243-248.

Terasvirta, T. (1990), "Generalizing Threshold Autoregressive Models," Discussion Paper 90-44, University of California, San Diego, Department of Economics.

Tong, H. (1983), "Threshold Models in Nonlinear Time Series Analysis," New York: Springer-Verlag.

—— (1990), "Nonlinear Time Series Analysis: A Dynamical System Approach," New York: Oxford University Press.

Table 1: Summary Statistics relating to weekly first-differences in one month Treasury bill yields reported in annulized percentage terms. The sample period is May 1986 to May 1996, a total of 523 observations. The data are plotted in Fig.1.

Statistics	Mean	Variance	Skewness	Kurtosis	$\text{Corr}(dr_t, r_{t-1})$
Estimate	-0.0015	0.1637	0.1018	4.4288	-0.1256

Table 2: Estimation results (n=423). The numbers in the paranthese are the t-statistics.

parameters	Constant	GARCH	DTGARCH	GRS	STGARCH
mean ϕ_{01}	0.1637 (2.7423)	0.0879 (2.0353)	0.0083 (0.1468)	0.3797 (2.1174)	0.3665 (1.3376)
ϕ_{11}	-0.0315 (2.0816)	-0.0180 (-1.9885)	-0.0198 (-1.6863)	-0.0560 (-1.6786)	-0.0466 (-0.9818)
ϕ_{02}	-	-	0.1213 (2.0915)	0.0179 (0.3313)	0.0586 (1.3556)
ϕ_{12}	-	-	-0.0189 (-1.5963)	-0.0080 (-0.7551)	-0.0184 (-2.1138)
variance α_0	0.4010 (32.341)	-	-	-	-
α_{11}	-	0.1866 (3.1018)	0.1295 (2.4216)	0.1740 (0.8390)	0.2805 (1.1361)
α_{31}	-	0.0699 (4.7313)	0.0514 (2.6342)	0.2318 (9.4048)	0.2616 (5.7921)
α_2	-	0.6446 (5.6048)	0.4746 (5.8207)	0.1585 (3.0697)	0.0982 (1.1695)
α_{12}	-	-	0.4507 (3.1176)	0.3349 (3.5439)	0.1485 (1.6133)
α_{32}	-	-	0.1065 (11.158)	0.0000 (0.0000)	0.0736 (4.2764)
probability β_{0p}	-	-	-	-7.3982 (-0.9452)	-2.3580 (-3.1508)
β_{1p}	-	-	-	5.0187 (1.0098)	0.6328 (0.6053)
β_{0q}	-	-	-	0.7610 (2.1565)	1.9659 (2.7019)
β_{1q}	-	-	-	-1.2576 (-1.5172)	-3.6627 (-3.0441)

Table 3: Summary statistics for various specifications

	Constant	GARCH	DTGARCH	GRS	STGARCH
# parameters	3	5	9	13	13
Log-likelihood	-264.14	-198.47	-172.80	-152.27	-145.63
AIC	-267.14	-203.47	-181.80	-165.27	-158.63
Schwarz	-273.53	-214.12	-200.97	-192.96	-186.31
MSE	0.1644	0.1514	0.1455	0.1453	0.1331
MAE	0.2060	0.1899	0.1787	0.1859	0.1783

The second row reports L^*, the maximum value achieved for the log of the likelihood function.
AIC is calculated as $L^* - k$, k is the number of parameters.
Schwarz is calculated as $L^* - (k/2) \cdot \ln(n)$.
We use two loss functions to measure the in-sample-forecasting: mean squared error (MSE) and mean absolute error (MAE).

$$MSE = T^{-1} \sum (\hat{u}_t^2 - h_t^2)^2,$$
$$MAE = T^{-1} \sum |\hat{u}_t^2 - h_t^2|.$$

Table 4: Out-of-sample forecasting

	Constant	GARCH	DTGARCH	GRS	STGARCH
OMSE	0.0370	0.0218	0.0256	0.0244	0.0217
OMAE	0.1716	0.1023	0.0943	0.1192	0.0903
TMSE	0.1660	0.1527	0.1469	0.1457	0.1340
TMAE	0.2282	0.1999	0.1858	0.1986	0.1818
R^2	0.1293	0.2108	0.2295	0.2199	0.2398

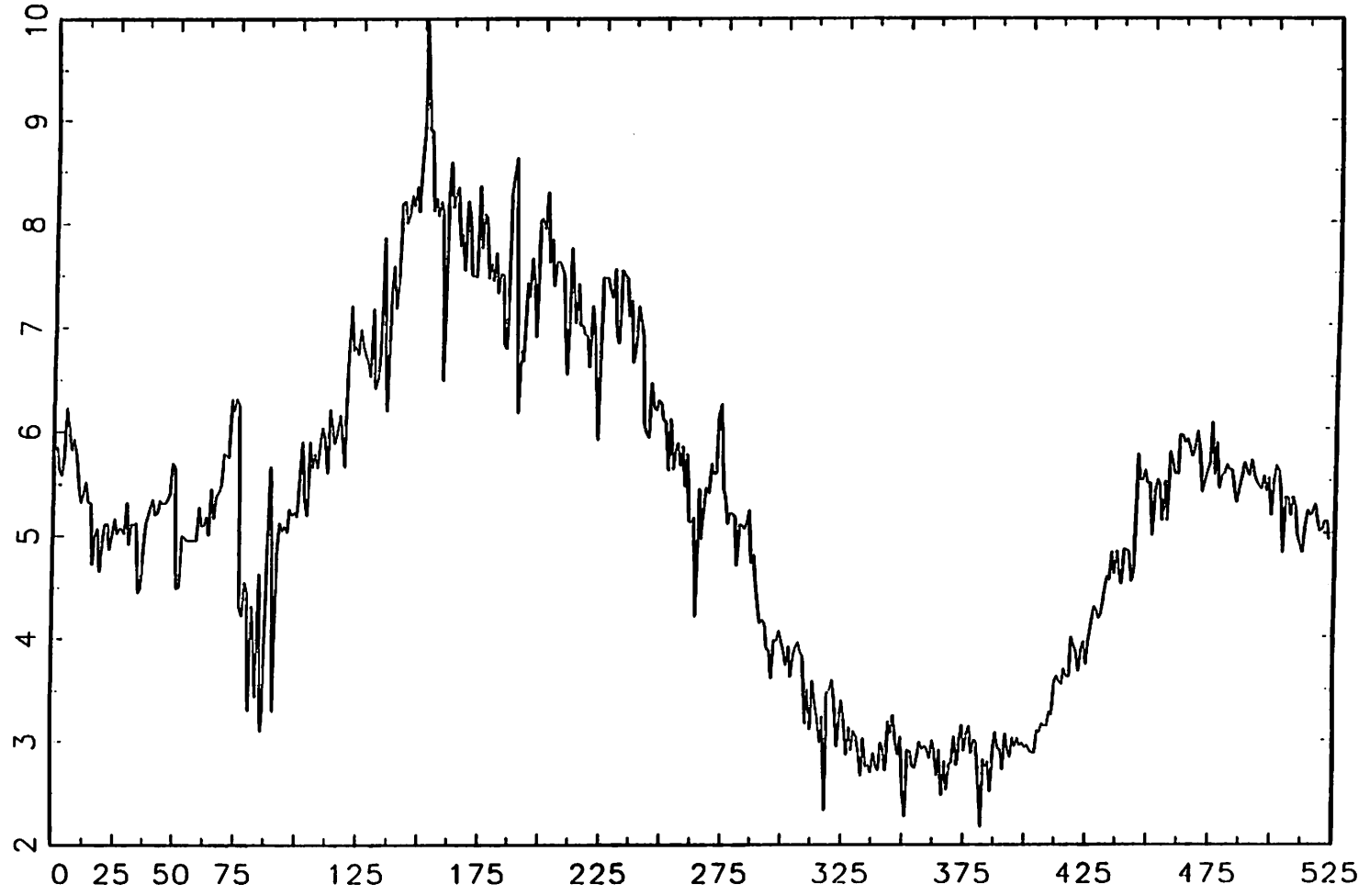Figure 1: Short Rates, 1986–1996
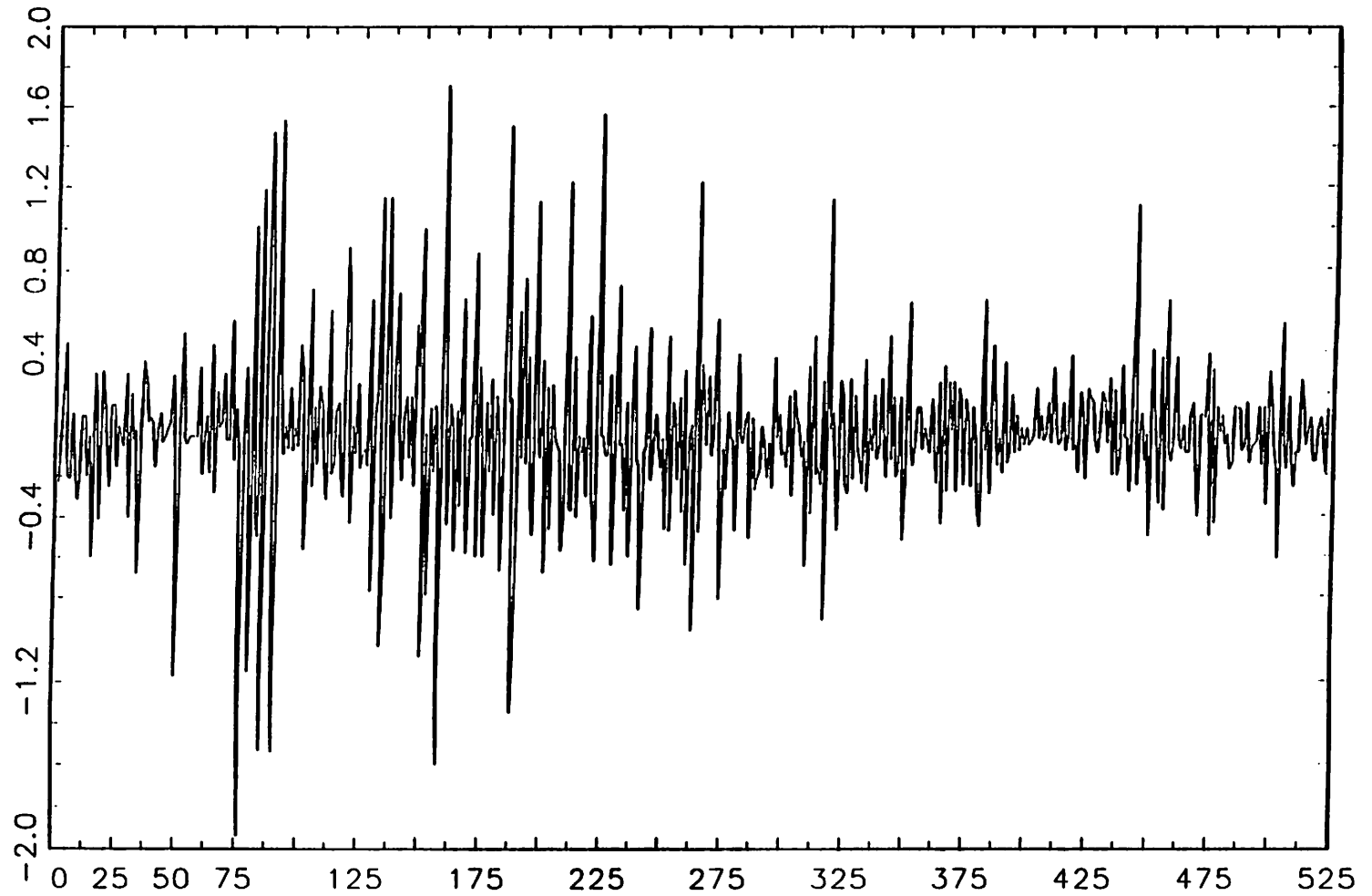
Figure 2: Change of Short Rates, 1986−1996

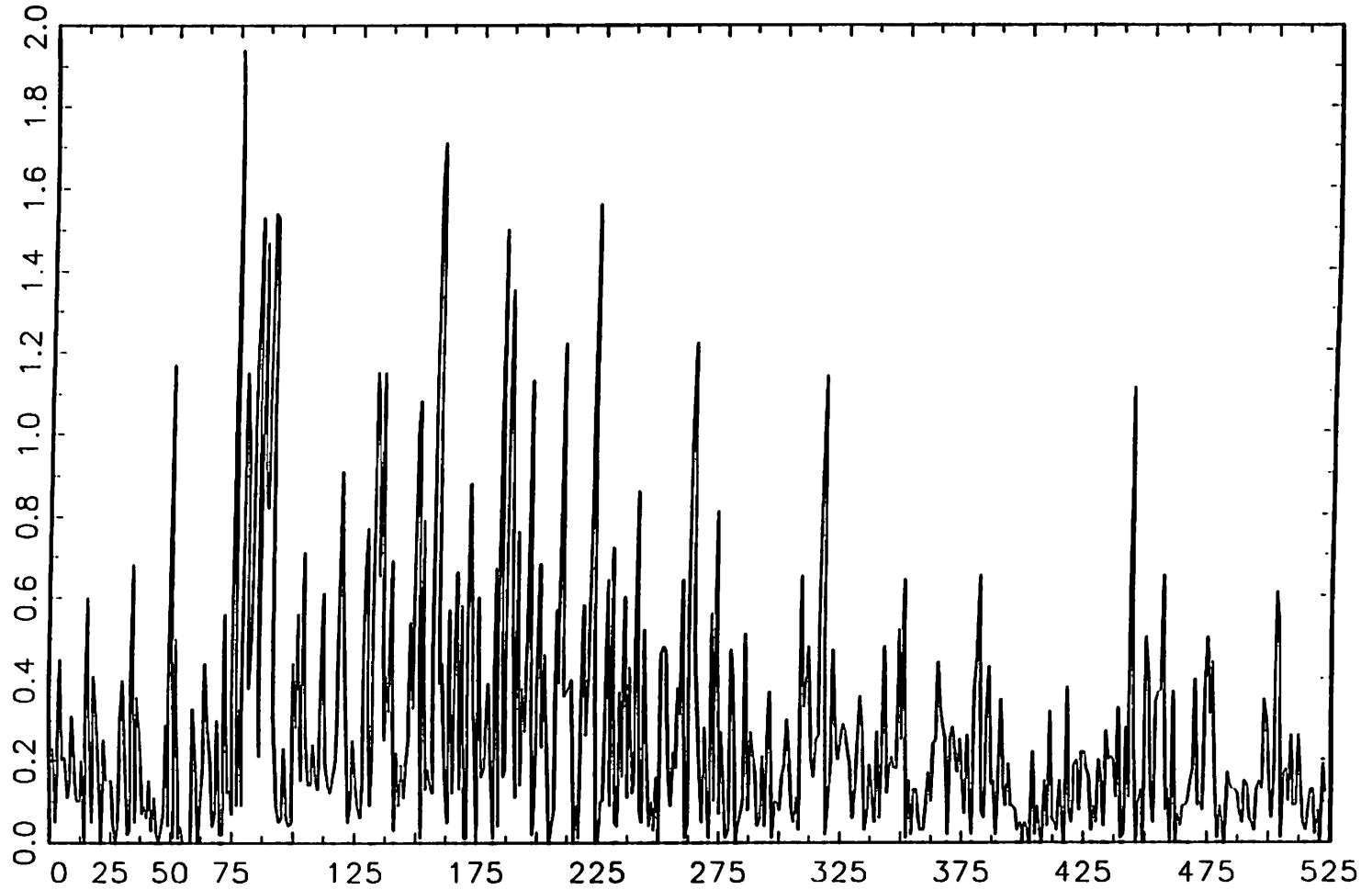Figure 3: Absolute value of Change of Short Rates, 1986−1996

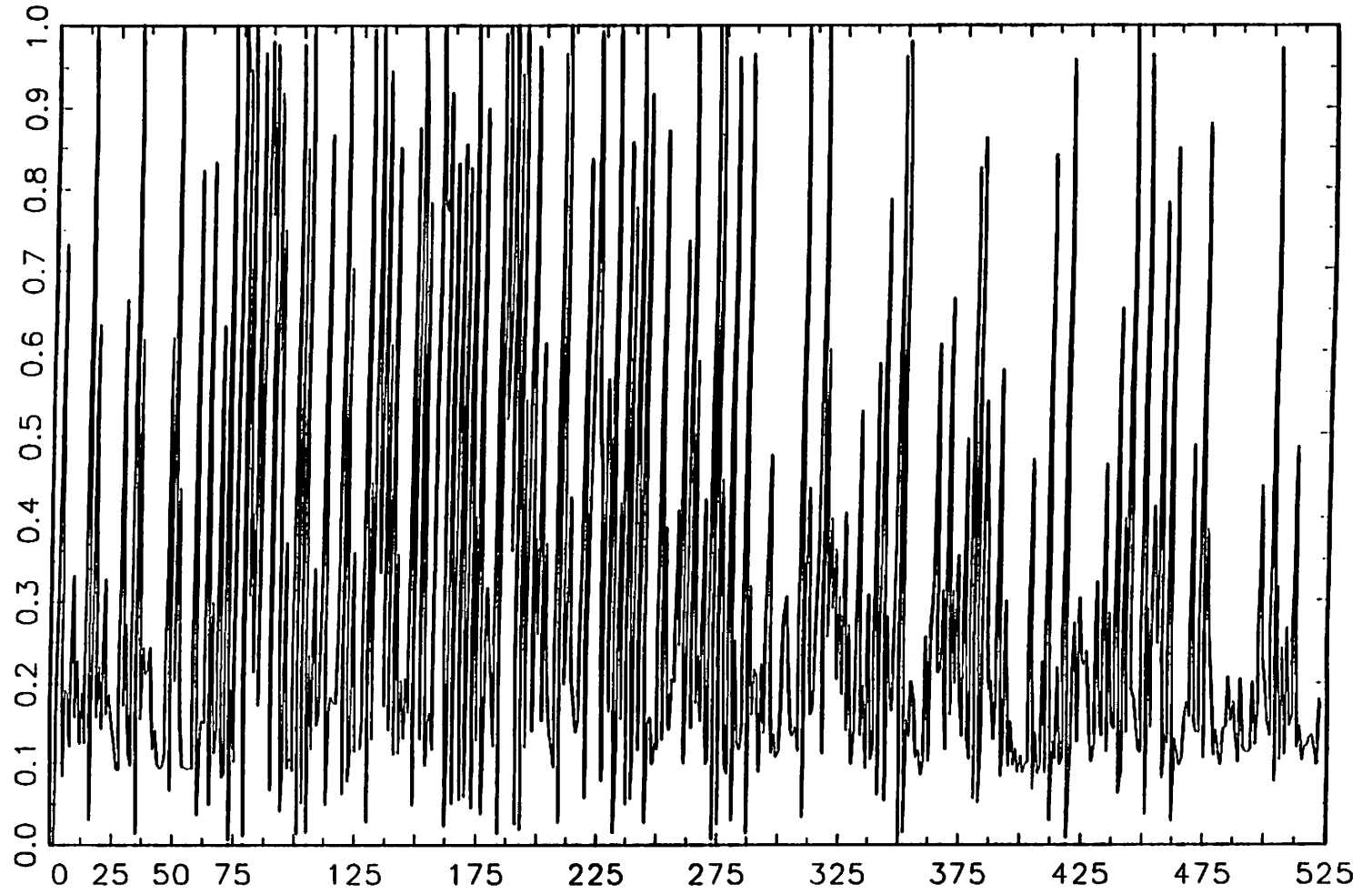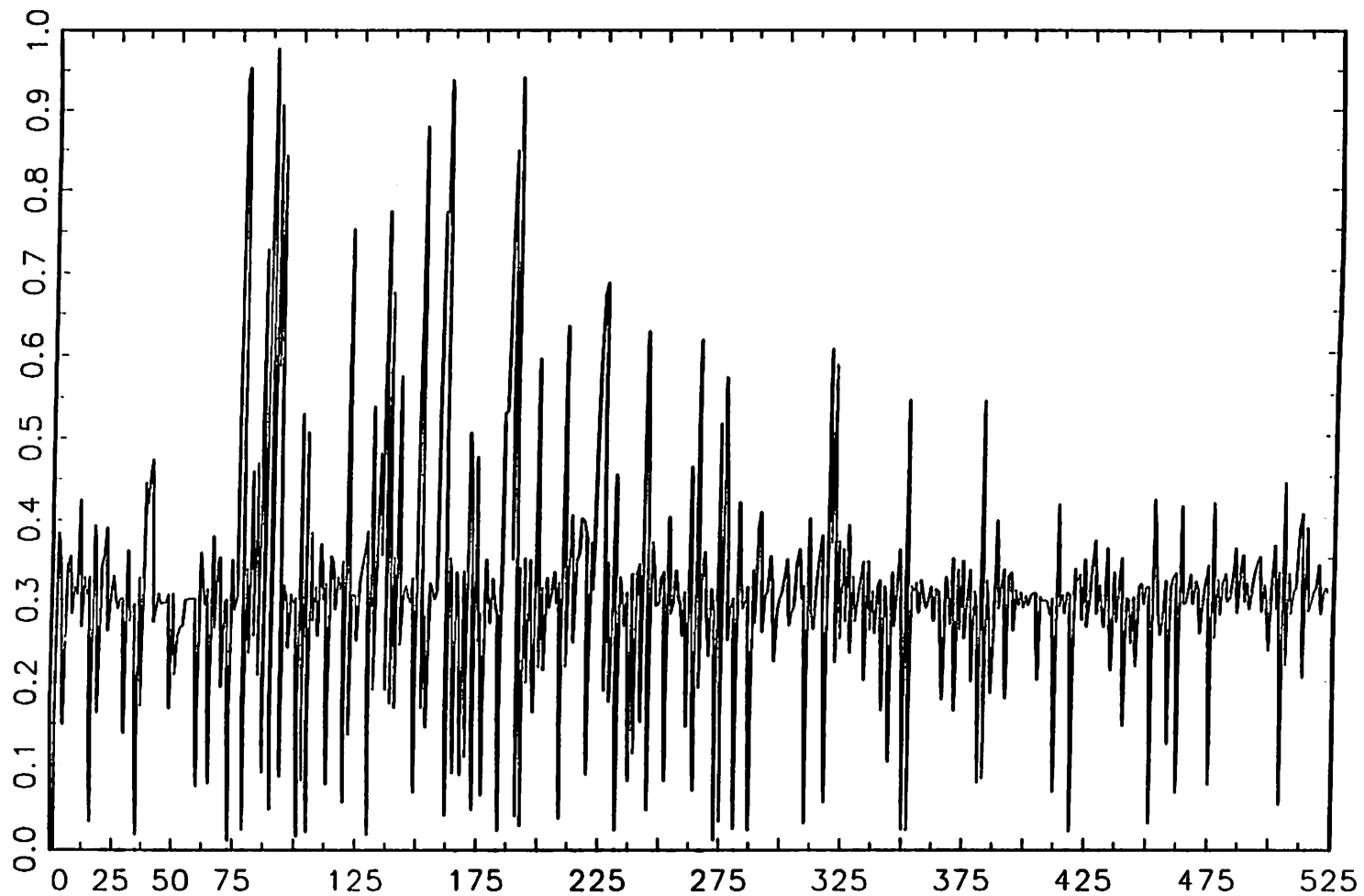Figure 4: Filter Probability of State 1 in GRS

Figure 5: Ex-post Probability of State 1 in GRS

Figure 6: Probability of State 1 in STGARCH