

Lanne, Markku; Saikkonen, Pentti

Working Paper

Modeling the US short-term interest rate by mixture autoregressive processes

SFB 373 Discussion Paper, No. 2000,76

Provided in Cooperation with:

Collaborative Research Center 373: Quantification and Simulation of Economic Processes,
Humboldt University Berlin

Suggested Citation: Lanne, Markku; Saikkonen, Pentti (2000) : Modeling the US short-term interest rate by mixture autoregressive processes, SFB 373 Discussion Paper, No. 2000,76, Humboldt University of Berlin, Interdisciplinary Research Project 373: Quantification and Simulation of Economic Processes, Berlin,
<https://nbn-resolving.de/urn:nbn:de:kobv:11-10048024>

This Version is available at:

<https://hdl.handle.net/10419/62194>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Modeling the U.S. Short-Term Interest Rate by Mixture Autoregressive Processes*

Markku Lanne
and
Pentti Saikkonen

University of Helsinki

Abstract

A new kind of mixture autoregressive model with GARCH errors is introduced and applied to the U.S. short-term interest rate. According to the diagnostic tests developed in the paper and further informal checks the model is capable of capturing both of the typical characteristics of the short-term interest rate: volatility persistence and the dependence of volatility on the level of the interest rate. The model also allows for regime switches whose presence has been a third central result emerging from the recent empirical literature on the U.S. short-term interest rate. Realizations generated from the estimated model seem stable and their properties resemble those of the observed series closely. The drift and diffusion functions implied by the new model are in accordance with the results in much of the literature on continuous-time diffusion models for the short-term interest rate, and the term structure implications agree with historically observed patterns.

⁰ *Financial support from the Deutsche Forschungsgemeinschaft through the SFB 373, and the Yrjö Jahnsson Foundation is gratefully acknowledged. The second author also thanks the Alexander von Humboldt Foundation for financial support under a Humboldt research award. Part of this research was done while the authors were visiting the Institute of Statistics and Econometrics at the Humboldt University in Berlin.

1 Introduction

The default-free short-term interest rate is a key variable in many economic and financial models, and much effort has been devoted to building empirical time series models that satisfactorily describe its dynamic behavior. Most of this work concerns U.S. interest rates that are characterized by volatility persistence and the dependence of the volatility on the level of the interest rate (see e.g. Brenner, Harjes and Kroner, 1996). Unfortunately, many models presented in the literature fail to fit both of these properties. Volatility persistence is adequately modeled by different kinds of GARCH models, whereas probably the best known model incorporating the latter feature is the generalized continuous-time model of Chan, Karolyi, Longstaff and Sanders (1992). Recently Koedijk, Nissen, Schotman and Wolf (1997), and Brenner et al. (1996) have suggested combined level-GARCH models that to some extent seem to capture both effects.

Much of the interest in building models of the short-term interest rate stems from the fact that these models are used in pricing interest-rate sensitive assets and in risk management. Typically these applications are based on Monte Carlo simulation where the empirical model of the short-term interest rate is (a part of) the data generating process (DGP). Another application requiring a model that provides a close approximation to the true data, is the estimation of continuous-time short-rate models and stochastic volatility models by simulation-based methods such as the efficient method of moments (EMM). This was recently emphasized by Andersen and Lund (1997), who showed that the dynamics of the level-GARCH models mentioned above are excessively erratic at the estimated parameter values, despite the fact that they pass most of the conventional diagnostic tests. Therefore, these models may not serve as good auxiliary models (score generators) in the EMM. Andersen and Lund (1997) suggested level-EGARCH models with Gaussian and Student- t distributed error terms that produce realizations with the range satisfactorily corresponding to that of the observed series. The recent finite-sample simulation results of Andersen, Chung and Sørensen (1999) confirm that a suitable parametric auxiliary model is critical for the EMM estimation of a stochastic volatility model even though this parametric model can be augmented with nonparametric terms. They also found that for the efficiency of the EMM in small samples, parsimony is critical. It can probably be more easily achieved by having a sufficient parametric model than by augmenting a poor model with a large number of nonparametric terms.

In addition to volatility persistence and level effects, another feature that has recently been claimed to characterize especially U.S. interest rate data, is the presence of regime switches which the above mentioned models do not explicitly take into account. Much of the literature applying regime switching models to U.S. short-term interest rates has been inspired by the exceptional three-year period (from November 1979 to October 1982) when the Federal Reserve ceased targeting interest rates causing increases in both the level and volatility. With the exception of Pfann, Schotman and Tschernig (1996) who considered threshold autoregressive (TAR) models,

Markov switching (MS) models have typically been fitted to interest rate data. Cai (1994), Gray (1996), and Ang and Bekaert (1998) present models that are extensions of Hamilton's (1989) MS model allowing for autoregressive conditional heteroskedasticity. All of these papers provide convincing evidence for the presence of multiple regimes in the U.S. interest rates.

While there thus seems to be evidence of regime switching in the U.S. short-term interest rates, the commonly employed MS specification may not be optimal for modeling these series. First, as Ang and Bekaert (1998) point out, a usual problem with such empirical models is the unsatisfactory regime classification. Typically the *ex ante* probability of a regime is around one half in a two-regime model. Second, these models have also found to be difficult to estimate, with problems arising from nonuniqueness of the maximum of the likelihood function and the unboundedness problem causing the conditional variance in one regime to approach zero as the conditional variance in the other regime approaches infinity (see Gray, 1995). Finally, the estimated MS models, as well as all the other models referred to above, have produced residuals that are either autocorrelated at virtually all lags or conditionally heteroskedastic. Moreover, Ang and Bekaert (1998) noticed that the MS models are rather bad at matching the unconditional moments of the interest rate series. Probably these problems explain why these models are not more commonly used.

In this paper we provide evidence that the dynamics of the postwar U.S. short-term interest rate can be captured by a new kind of model which is a mixture of linear autoregressive models with GARCH errors. This model is an extension of the recent mixture autoregressive (MAR) models introduced by Le, Martin and Raftery (1996), and Wong and Li (1999a,b, 2000). In the simple special case of no GARCH type heteroskedasticity our MAR model can be obtained from the TAR model by adding an unobservable white noise term to the threshold parameters and thereby making them time variant. In the same way as the TAR model our MAR model is therefore capable of allowing for regime switches in the data. However, in this respect it is considerably more flexible than the TAR model because it is not bound to fixed threshold parameters, and in this sense it can also be thought of as a smooth version of the TAR model. In fact, as far as the conditional expectation is concerned, our MAR model is entirely similar to the smooth transition autoregressive (STAR) model which has previously been used to introduce smoothness to TAR models. In general, the conditional distribution of the MAR model is very different from that of the STAR model or TAR model, however.

Our MAR model also explicitly contains probabilities of regime switches commonly referred to as mixing proportions in mixture models. The mixing proportions are functions of a past value of the process and considerably simpler than the transition probabilities which are their counterparts in MS models. This feature, which is perhaps the most distinctive difference between our MAR model and the MS model, also implies that the MAR model can be readily extended to allow for GARCH type heteroskedasticity. In our experience this relative simplicity translates into numerically unproblematic estimation compared to MS models even when no GARCH type

heteroskedasticity is present.

In Section 5 we present estimation results of different specifications of the MAR model with weekly postwar U.S. three-month Treasury bill rate data. To be able to assess the adequacy of a fitted MAR model, diagnostic tests on the conditional mean and variance, based on the general approach of Wooldridge (1990), are developed. It turns out that to capture both the volatility persistence and level effects, a MAR model with GARCH type heteroskedasticity is required. Diagnostic tests reveal no misspecification of our preferred MAR-GARCH model with the exception of some unmodeled autocorrelation at higher lags. The model also seems to be economically tenable in the sense that the *ex ante* probabilities implied by the model are able to clearly single out the high volatility periods reported in the previous literature. Further informal checks based on simulation experiments also lend support to the model, indicating that it matches the observed data well. Contrary to the results concerning MS models in the previous literature, our preferred MAR model seems to provide clear regime classification (according to the regime classification measure of Ang and Bekaert (1998)). Our model is also more parsimonious than the generalized regime switching model of Gray (1996) and has superior forecasting performance.

The plan of the paper is as follows. In Section 2 our basic MAR model and some extensions, including the combination of the MAR model and GARCH model, are presented. The model is also compared to some other multi-regime models that have been used to model interest rates in the previous literature. In Section 3 maximum likelihood (ML) estimation of the parameters of the model is discussed, and in Section 4 the diagnostic checks are introduced. Section 5 presents the empirical results, and Section 6 concludes.

2 Mixture Autoregressive Models

2.1 Special Case

To illustrate the general models to be introduced in this paper, we first consider a simple special case given by

$$y_t = (\nu_1 + b_1 y_{t-1}) I(y_{t-1} < c + \eta_t) + (\nu_2 + b_2 y_{t-1}) I(y_{t-1} \geq c + \eta_t) + \sigma \varepsilon_t \quad (1)$$

where the random variables η_t and ε_s are independent for all t and s with $\eta_t \sim NID(0, \sigma_\eta^2)$ and $\varepsilon_t \sim NID(0, 1)$, $I(\cdot)$ is the indicator function, and, in addition to the variance σ_η^2 , also ν_i , b_i ($i = 1, 2$), c and $\sigma > 0$ are unknown parameters. When $\sigma_\eta^2 = 0$ the random variable η_t can be dropped out of (1) and the model becomes a special case of a homoskedastic TAR model. We refer to Tong (1990) for a comprehensive account of TAR models. An essential feature of this TAR model is that it is composed of two AR(1) models with different conditional means depending on whether y_{t-1} is below or above the value of the threshold parameter c which divides the sample space into two regimes, the lower regime and the upper regime. The conditional means in the lower and upper regimes are $\nu_1 + b_1 y_{t-1}$ and $\nu_2 + b_2 y_{t-1}$,

respectively, and the change between these two conditional means is abrupt, occurring always when y_{t-1} crosses the value of the threshold parameter c .

Now suppose that $\sigma_\eta^2 > 0$. A major difference between this case and the preceding TAR case is that, even if the value of the threshold parameter c is assumed known, it is not known which one of the two regimes applies and, given y_{t-1} , switching from one regime to another is randomized. For instance, if the lower regime has been in operation for some time until the point of time $t - 2$ and then $y_{t-1} \geq c$ occurs switching to the upper regime does not occur if the random variable η_t takes a value sufficiently far below zero. Similarly, a sufficiently large value of η_t implies a switch to the upper regime even when $y_{t-1} < c$ occurs. Thus, as far as regime switches are concerned, model (1) is considerably more flexible than the corresponding TAR model obtained by assuming $\sigma_\eta^2 = 0$. Flexibility of this kind appears useful because in some cases regime switches based on strict rules can make conventional TAR models difficult to apply. For instance, the occurrence of a regime switch may also depend on randomly changing external factors and not only on whether y_{t-1} is strictly below or above the value of the threshold parameter c . Therefore, it seems reasonable to make the constant threshold parameter used in conventional threshold models time variant. The idea used in model (1) provides a way to do this. This kind of flexibility is likely to be relevant for interest rate applications, as implied by the conclusions of Pfann et al. (1996) who found the dependence on a single threshold variable triggering regime switches a major limitation of their TAR models.

The random variable η_t may also be interpreted as representing uncertainty economic agents have when they make their decisions which can result in a regime switch. This uncertainty may reflect the fact that different economic agents have different views of the state of economy. As a result, realizations of the process y_t may vary above and below the threshold during several points of time before they finally switch to a new regime or draw away from the threshold without leaving the current regime.

That the regime switches implied by model (1) are not based on a strict rule can be thought of as a smoothness property. To illustrate this, let $E_{t-1}(\cdot)$ signify the conditional expectation with respect to the information set $\{y_{t-j}, j \geq 1\}$ and conclude from (1) that

$$E_{t-1}(y_t) = (\nu_1 + b_1 y_{t-1}) [1 - \Phi((y_{t-1} - c)/\sigma_\eta)] + (\nu_2 + b_2 y_{t-1}) \Phi((y_{t-1} - c)/\sigma_\eta) \quad (2)$$

where $\Phi(\cdot)$ denotes the cumulative distribution function of the standard normal distribution. Thus, the conditional expectation depends on the smooth function $\Phi(\cdot)$ instead of the (discontinuous) indicator function obtained in the case of the TAR model. In fact, the conditional expectation is identical to that obtained from the STAR model with the transition function defined by $\Phi(\cdot)$. The STAR model corresponding to (1) can be defined by replacing $I(y_{t-1} < c + \eta_t)$ and $I(y_{t-1} \geq c + \eta_t)$ by $1 - \Phi((y_{t-1} - c)/\sigma_\eta)$ and $\Phi((y_{t-1} - c)/\sigma_\eta)$, respectively.¹ In terms of the conditional expectation model (1) introduces a similar smoothing to the TAR model as the

¹Although the cumulative distribution function of the normal distribution has also been considered in STAR models (see Chan and Tong, 1986, and Tong, 1990) the logistic function has been a

STAR model. However, as will be seen below, the stochastic properties of model (1) and the STAR model are generally rather different. The same is also true for their smoothness properties because the regime switches implied by model (1) are abrupt while those implied by the STAR model are slow unless the transition function $\Phi(\cdot)$ increases steeply. As far as our interest rate application is concerned, it is interesting to note that according to the results of Pfann et al. (1996) the STAR model fails to provide useful smoothness to the TAR model for the U.S. short-term interest rate. Their estimated transition function was so steep that the STAR model was virtually indistinguishable from the corresponding TAR model.

Unless otherwise stated we shall henceforth assume that $\sigma_\eta^2 > 0$. We shall also assume that equation (1) defines a stationary process. In the case of TAR and STAR models conditions which ensure stationarity are known (see Meyn and Tweedie, 1993, Chan and Tong, 1986, and Tong, 1990, p. 183) but for model (1) or its extensions to be presented below such conditions are not available. Therefore, we suggest taking a practical approach to this issue and examining the stability of model (1) and its generalizations by simulation (cf. Gallant, Rossi and Tauchen, 1993, for using simulation for this purpose).

Given the past values y_{t-j} , $j \geq 1$, the conditional distributions obtained for y_t from the TAR and STAR models discussed above are normal. The conditional means were already discussed and the conditional variance is σ^2 in both cases. However, in the case of model (1) a nonnormal conditional distribution with nonconstant variance is obtained. Let $f_{t-1}(\cdot)$ be the conditional density function of y_t given y_{t-j} , $j \geq 1$, and let $\phi(\cdot)$ signify the density function of the standard normal distribution. Then, using the independence of η_t and past values of y_t we can write the conditional joint density of y_t and η_t given y_{t-j} , $j \geq 1$, as the product of the marginal density of η_t and conditional density of y_t given both y_{t-j} , $j \geq 1$, and η_t . After this, $f_{t-1}(y_t)$ is obtained by integration and, using the indicator functions in (1) to divide the range of integration suitably, we find that model (1) implies

$$\begin{aligned} f_{t-1}(y_t) &= \frac{1}{\sigma} \frac{1}{\sigma_\eta} \int_{y_{t-1}-c}^{\infty} \phi((y_t - \nu_1 - b_1 y_{t-1})/\sigma) \phi(\eta/\sigma_\eta) d\eta \\ &\quad + \frac{1}{\sigma} \frac{1}{\sigma_\eta} \int_{-\infty}^{y_{t-1}-c} \phi((y_t - \nu_2 - b_2 y_{t-1})/\sigma) \phi(\eta/\sigma_\eta) d\eta. \end{aligned}$$

This gives

$$\begin{aligned} f_{t-1}(y_t) &= \frac{1}{\sigma} \phi((y_t - \nu_1 - b_1 y_{t-1})/\sigma) [1 - \Phi((y_{t-1} - c)/\sigma_\eta)] \\ &\quad + \frac{1}{\sigma} \phi((y_t - \nu_2 - b_2 y_{t-1})/\sigma) \Phi((y_{t-1} - c)/\sigma_\eta). \end{aligned} \tag{3}$$

more common choice (see Granger and Teräsvirta, 1993, and the references therein). However, as far as the general ideas of the STAR model are concerned, there is no essential difference between these two choices.

Thus, the conditional distribution resulting from (1) is a mixture of two normal distributions which are the same as the conditional distributions obtained from the two regimes of the corresponding TAR model. To put this another way, (1) defines a mixture of two AR(1) models. Given the value of y_{t-1} , the AR(1) model of the lower regime applies with probability $1 - \Phi((y_{t-1} - c)/\sigma_\eta)$ and the AR(1) model of the upper regime applies with probability $\Phi((y_{t-1} - c)/\sigma_\eta)$. If the value of y_{t-1} is close to the value of the threshold parameter c (as measured by the standard deviation σ_η) there is a fair probability to end up in any of the two regimes. However, when the value of y_{t-1} is far from the value of the threshold parameter c , the probability of a regime switch is negligible. This gives a probabilistic description of the flexibility of model (1) compared to the corresponding TAR model in which these two probabilities are always zero or one.

Based on the above discussion, we call model (1) a mixture autoregressive model and use the abbreviation MAR. The probabilities $1 - \Phi((y_{t-1} - c)/\sigma_\eta)$ and $\Phi((y_{t-1} - c)/\sigma_\eta)$ are called mixing proportions. Previous published papers on mixture autoregressive models are Le, Martin and Raftery (1996) and Wong and Li (2000). The models studied in these papers differ from our model in that the mixing proportions are assumed invariant in time. A generalization to the case of time variant mixing proportions is given by Wong and Li (1999b). Our MAR model may also be compared to the autoregressive MS model discussed for example by Hamilton (1994, Chapter 22.4). In the first order case corresponding to (1) this model assumes that the parameters in the conditional mean change according to an unobservable two state Markov chain independent of the error term ε_t . Instead of a Markov chain parameter changes in our MAR model are modeled by a past value of the process y_t and the unobservable white noise process η_t . In the standard MS model the transition probabilities of the Markov chain are not determined by any observable variables but extensions which relax this feature are available. From our point of view the most relevant extension in this direction is due to Gray (1996) whose GRS model is similar to our model in that the transition probabilities depend on a lagged value of the process. The Markov chain structure used by Gray (1996) to model parameter changes makes his model more complicated than ours, however.

As the above discussion has already revealed, the conditional distribution implied by the MAR model (1) is very different from that obtained from the so far commonly used TAR and STAR models. In addition to the conditional mean given by (2) the conditional variance can also be readily obtained from the conditional distribution (3) or from equation (1). By straightforward calculation it can be seen that the conditional variance is

$$Var_{t-1}(y_t) = \sigma^2 + [\nu_1 - \nu_2 + (b_1 - b_2)y_{t-1}]^2 \Phi((y_{t-1} - c)/\sigma_\eta) [1 - \Phi((y_{t-1} - c)/\sigma_\eta)]. \quad (4)$$

Thus, unlike in the previously discussed TAR and STAR models, the MAR model (1) implies conditional heteroskedasticity even though the innovation ε_t is homoskedastic. Conditional heteroskedasticity is related to regime switches. It is prominent when the

value of y_{t-1} is close to the value of the threshold parameter c and diminishes when the distance of y_{t-1} from c gets large. This reflects the uncertainty in the regime switches caused by the random external factor η_t .

For most economic time series, including interest rates, the MAR model (1) is probably not adequate, but various extensions are straightforward to obtain. For instance, the assumption that two mixtures in (3) have the same variance could be relaxed so that the parameter σ would be replaced by σ_1 in the first component distribution and by σ_2 in the second one. Then σ^2 in the conditional variance (4) would be $\sigma_1^2[1 - \Phi((y_{t-1} - c)/\sigma_\eta)] + \sigma_2^2\Phi((y_{t-1} - c)/\sigma_\eta)$. The number of mixtures can also be larger than two and more lags can be used in the component autoregressive models. These extensions are discussed in more detail in the next section. It may also be worth noting that instead of the normal distribution assumed in (1) other distributions could in principle be considered without any difficulty.

2.2 Mixture Autoregressive Model

An obvious way to generalize the MAR model of the previous section is to replace the conditional distribution in (3) by a more general mixture of normal distributions with each mixture based on a conventional Gaussian autoregressive model. Instead of y_{t-1} we now assume that the mixing proportions depend on a general lagged value y_{t-d} so that, if the number of component distributions is m , the mixing proportions are given by

$$\pi_{i,t-d} = \begin{cases} 1 - \Phi((y_{t-d} - c_1)/\sigma_\eta), & i = 1 \\ \Phi((y_{t-d} - c_{i-1})/\sigma_\eta) - \Phi((y_{t-d} - c_i)/\sigma_\eta), & i = 2, \dots, m-2 \\ \Phi((y_{t-d} - c_{m-1})/\sigma_\eta), & i = m \end{cases} \quad (5)$$

where $d \geq 1$ is a delay parameter and $c_1 < \dots < c_{m-1}$ are threshold parameters. Our generalization of the conditional distribution in (3) can thus be written as

$$f_{t-1}(y_t) = \sum_{i=1}^m \frac{1}{\sigma_i} \phi((y_t - \nu_i - b_{i1}y_{t-1} - \dots - b_{ip}y_{t-p})/\sigma_i) \pi_{i,t-d}. \quad (6)$$

Of course, it is also possible to arrive at this conditional distribution by extending equation (1) to

$$y_t = \sum_{i=1}^m (\nu_i + b_{i1}y_{t-1} + \dots + b_{ip}y_{t-p} + \sigma_i \varepsilon_t) I(c_i + \eta_t \leq y_{t-d} < c_{i-1} + \eta_t) \quad (7)$$

where $I(c_i + \eta_t \leq y_{t-d} < c_{i-1} + \eta_t)$ should be interpreted as $I(c_1 + \eta_t \leq y_{t-d})$ for $i = 1$ and $I(y_{t-d} < c_{m-1} + \eta_t)$ for $i = m$. We call this model an m -component mixture autoregressive model of order p and delay d or a $\text{MAR}(m, p, d)$ model. Of course, an obvious generalization is obtained by letting the orders vary in the component models which only amounts to replacing p in (6) and (7) by p_i . In this case the abbreviation $\text{MAR}(m, p_1, \dots, p_m, d)$ is used. On the other hand, various restricted

versions of the general $MAR(m, p, d)$ model may be considered. For instance, one may assume homoskedasticity or that the autoregressive parameters b_{i1}, \dots, b_{ip} are the same for all $i = 1, \dots, m$. The latter restriction, which implies that the nonlinearity in the conditional mean is only due to level shifts, is found to provide an adequate description for the interest rate data in the application of this paper.

Basically, the discussion given to motivate and interpret the simple MAR model of the previous section also applies to the general model defined above. In the same way we can also derive the first and second conditional moments. The conditional mean is

$$E_{t-1}(y_t) = \sum_{i=1}^m (\nu_i + b_{i1}y_{t-1} + \dots + b_{ip}y_{t-p}) \pi_{i,t-d} \quad (8)$$

and, by straightforward calculation, one can see that the conditional variance can be written as

$$Var_{t-1}(y_t) = \sum_{i=1}^m \sigma_i^2 \pi_{i,t-d} + \sum_{i=1}^m \left[(\nu_i - \bar{\nu}) + (b_{i1} - \bar{b}_1)y_{t-1} + \dots + (b_{ip} - \bar{b}_p)y_{t-p} \right]^2 \pi_{i,t-d} \quad (9)$$

where $\bar{\nu} = \sum_{i=1}^m \nu_i \pi_{i,t-d}$ and $\bar{b}_j = \sum_{i=1}^m b_{ij} \pi_{i,t-d}$, ($j = 1, \dots, p$). Thus, the conditional mean is a weighted average of the means of the m component distributions with weights given by the mixing proportions. A similar weighted average of the conditional variances of the component distributions appears in the conditional variance. However, this is not the only source of conditional heteroskedasticity. The conditional variance also contains another component which measures the variability of the conditional means of the composite distributions. This component is due to uncertainty in regime switches and it is prominent when y_{t-d} is close to a threshold.

2.3 Extensions with Autoregressive Conditional Heteroskedasticity

It is plausible that the conditional heteroskedasticity inherent in the MAR model of the previous section does not adequately describe most financial time series. This claim is also supported by the fact that the TAR models Pfann et al. (1996) estimated for the three-month U.S. Treasury bill rate, had conditionally heteroskedastic residuals. It is straightforward to extend the general MAR model of the previous section to allow for ARCH type heteroskedasticity in the component models. A simple way to define such a model is to make the variances of the component distributions in (6) dependent on t according to an ARCH or GARCH process. Thus, instead of (6) one might consider

$$f_{t-1}(y_t) = \sum_{i=1}^m \frac{1}{\sigma_{it}} \phi((y_t - \nu_i - b_{i1}y_{t-1} - \dots - b_{ip}y_{t-p})/\sigma_{it}) \pi_{i,t-d} \quad (10)$$

where σ_{it} is obtained from the GARCH(r, q) process

$$\sigma_{it}^2 = \sigma_i^2 + \beta_{i1}\sigma_{i,t-1}^2 + \dots + \beta_{ir}\sigma_{i,t-r}^2 + \alpha_{i1}u_{i,t-1}^2 + \dots + \alpha_{iq}u_{i,t-q}^2 \quad (11)$$

where $u_{i,t} = y_t - \nu_i - b_{i1}y_{t-1} - \dots - b_{ip}y_{t-p}$. As an obvious extension of equation (7) we now have

$$y_t = \sum_{i=1}^m (\nu_i + b_{i1}y_{t-1} - \dots + b_{ip}y_{t-p} + \sigma_{it}\varepsilon_t) I(c_i + \eta_t \leq y_{t-d} < c_{i-1} + \eta_t). \quad (12)$$

Given the discussion of the previous section, conditional first and second moments can also be easily obtained in this case. Of course, the conditional mean is still given by (8) and the conditional variance is obtained from (9) by replacing σ_i^2 by σ_{it}^2 .

In principle, it would be possible to consider the case where the orders of both the $\text{MAR}(m, p, d)$ model and related $\text{GARCH}(r, q)$ model dependent on the index i . However, the resulting model would become rather complex and from a practical point of view, special cases obtained by parameter restrictions are likely to be of greater interest. For instance, the assumption that the dynamics of the GARCH process (11) are the same for each regime may often be plausible so that instead of (11) it would be reasonable to consider its restricted version

$$\sigma_{it}^2 = \sigma_i^2 + \beta_1 \sigma_{i,t-1}^2 + \dots + \beta_r \sigma_{i,t-r}^2 + \alpha_1 u_{i,t-1}^2 + \dots + \alpha_q u_{i,t-q}^2. \quad (13)$$

This model implies that only the sizes of the component variances vary between different regimes but their dynamics do not. This special case is found quite reasonable in the application of this paper. An even more restricted special case is obtained by assuming that $\sigma_1^2 = \dots = \sigma_m^2$ so that conditional heteroskedasticity is the same for all regimes.

Finally, it may be worth noting that, since the conditional variance obtained from (11) is determined by quantities depending only on the current regime, our MAR-GARCH model avoids the path dependence problem which makes parameter estimation in the MS-GARCH models of Cai (1994) and Susmel and Hamilton (1994) virtually impossible. In this respect our MAR model is similar to the GRS model of Gray (1996) but, as already discussed, it is simpler and the simplicity is even more pronounced in models which include a GARCH component.

3 Parameter Estimation

After specifying the orders of the MAR model (including the delay d) and related GARCH model (if needed) it is in principle straightforward to estimate the parameters by ML. For simplicity, we shall assume that these orders are the same for every component model. Let θ be the parameter vector containing the parameters in the $\text{MAR}(m, p, d)$ - $\text{GARCH}(r, q)$ model defined by the conditional distribution (10) or some special case thereof obtained by zero restrictions or equality restrictions. In the most general case considered here the parameter vector θ thus contains the parameters $\nu_i, b_{i1}, \dots, b_{ip}, \sigma_i^2, \beta_{i1}, \dots, \beta_{ir}, \alpha_{i1}, \dots, \alpha_{iq}, (i = 1, \dots, m), c_1, \dots, c_{m-1}$ and σ_η^2 . A necessary condition for meaningful estimation is that the parameter vector θ is identifiable. It is known that in finite mixtures of normal distributions identifiability

holds unless some of the component distributions are redundant (see e.g. Yakowitz and Spragins, 1968). Unless otherwise stated, we shall assume that this is not the case.

Now, suppose we have the observed time series y_{-l+1}, \dots, y_T where $l = \max\{p, d\}$ and y_{-l+1}, \dots, y_0 are treated as fixed initial values. Making the dependence of the conditional distribution (10) explicit on the parameter vector θ we can write the (conditional) log-likelihood function of the data as

$$l_T(\theta) = \sum_{t=1}^T \log f_{t-1}(y_t; \theta)$$

where $f_{t-1}(y_t; \theta)$ is obtained from equations (10), (11) and (5). Although the log-likelihood function looks complicated its value can be computed straightforwardly for any given value of the parameter vector θ . Thus, numerical methods can be used to find ML estimates. The likelihood function is clearly twice continuously differentiable so that, assuming that the observed series is generated by a stationary and ergodic process, it is reasonable to apply standard large sample results to construct statistical tests and confidence intervals. In particular, approximate standard errors can be obtained from the diagonal elements of the matrix $-\left(\partial^2 l_T(\hat{\theta}) / \partial \theta \partial \theta'\right)^{-1}$ where $\hat{\theta}$ signifies the ML estimate of θ . More general Wald tests as well as likelihood ratio tests with conventional asymptotic χ^2 distributions can also be obtained. Of course, these test procedures cannot be used to test hypotheses which imply that the number of component models m can be reduced because under such hypotheses unidentifiability of nuisance parameters makes the testing problem nonstandard (see Davies, 1977). An explicit example of this is given by the hypothesis $\nu_1 = \nu_2$ and $b_1 = b_2$ in (1). Under this hypothesis the parameters c and σ_η^2 vanish from the model and therefore cannot be identified. Note that the related multi-regime models such as MS, TAR and STAR models also suffer from similar unidentifiability problems. In general, lack of identifiability makes all standard large sample inference procedures invalid. Thus, since the number of component models is generally unknown in practice it would be of great interest to have test procedures which could be used to help finding a correct number of component models. In particular, it would be of interest to have a test procedure which could be used to check whether a standard AR(p) model with possibly GARCH errors would be adequate. Attempts to solve these testing problems lies outside the scope of this paper although significant progress has recently been made in this area by Andrews (1993), Andrews and Ploberger (1994), and Hansen (1996) among others. In the application of the paper we shall employ less formal procedures to choose the number of component models which in most cases is likely to be at most three.

4 Diagnostic Checks

After a specified time series model has been fitted to data diagnostic checks should be applied to reveal possible misspecifications. When parameter estimation can be carried out by least squares, it is possible to compute the residuals of the fitted model and investigate whether their properties are similar to those of the corresponding theoretical errors. In this context, statistical tests and visual inspection of the residual series are typically employed. In the MAR models introduced in Section 2 the use of residual based diagnostic checks is less straightforward because it is not quite obvious how residuals can be obtained. Ideally one would like to have an empirical counterpart of the error term ε_t in (12). However, even though ε_t can be easily solved from equation (12) the solution depends on the process η_t and is therefore not observable. Instead of trying to base diagnostic checks on an empirical counterpart of the error term ε_t we are therefore content with using residuals which can only be used to check whether the specified model is capable of describing the first two conditional moments of the observed process. The theoretical counterparts of these residuals, also mentioned by Wong and Li (2000), are given by

$$e_t = \frac{y_t - E_{t-1}(y_t)}{\sqrt{Var_{t-1}(y_t)}}$$

where the conditional expectation $E_{t-1}(y_t)$ and the conditional variance $Var_{t-1}(y_t)$ are obtained from the MAR model (12) or its relevant special case in the way explained in Section 2. Thus, e_t depends on the parameters of the specified model and, when there is need to be explicit about this dependence, the notation $e_t(\theta)$ is used. A correct specification of the model implies that e_t is a martingale difference sequence with unit variance and any significant autocorrelation or heteroskedasticity indicates inadequacies in the conditional mean or variance of the specified model. It should be noted, however, that even in the case of a correctly specified model e_t is not normally distributed. Since e_t is not observable in practice it has to be replaced by an empirical counterpart. An obvious way to do this is to replace the unknown parameters in $E_{t-1}(y_t)$ and $Var_{t-1}(y_t)$ by their ML estimators discussed in the previous section. This gives the series $\hat{e}_t = e_t(\hat{\theta})$.

In addition to visual inspection of the series \hat{e}_t misspecification tests for checking whether a specified MAR model can adequately describe the first two conditional moments of the process, can be derived by applying the general approach of Wooldridge (1990). The idea of these tests is to (incorrectly) treat e_t as Gaussian white noise and obtain misspecification tests which are robust to this (incorrect) distributional assumption. As already indicated in the previous section, standard tests cannot be obtained for checking the specified number of component models. In what follows we shall therefore assume that this feature is correctly specified. First, consider testing the adequacy of the specified lag length p against the alternative $p + s$ where s is a chosen positive integer. This test is primarily designed to check the adequacy of

the conditional mean which in its most general form is specified by (8). It should be noted, however, that in the MAR-GARCH model the specified lag length also affects the conditional variance given by (9) with σ_i^2 possibly replaced by σ_{it}^2 in (11) or (13). Other aspects of the conditional variance are here assumed to be correctly specified. Under the alternative the previous expressions of the conditional mean and variance are modified in an obvious way. The conditional mean becomes

$$E_{t-1}(y_t) = \sum_{i=1}^m (\nu_i + b_{i1}y_{t-1} + \dots + b_{i,p+s}y_{t-p-s}) \pi_{i,t-d}$$

so that, if $\vartheta_1 = [b_{1,p+1} \dots b_{1,p+s} \dots b_{m,p+1} \dots b_{m,p+s}]'$, the null hypothesis is $\vartheta_1 = 0$. Let $\mu_{1t}(\theta, \vartheta_1) = E_{t-1}(y_t)$ and $\omega_{1t}^2(\theta, \vartheta_1) = \text{Var}_{t-1}(y_t)$ be the conditional mean and variance of y_t of the unrestricted model and, for convenience, set $\hat{\omega}_{1t}^2 = \omega_{1t}^2(\hat{\theta}, 0)$. Furthermore, define

$$\hat{\varphi}_{1t} = \begin{bmatrix} \hat{\omega}_{1t}^{-1} \partial \mu_{1t}(\hat{\theta}, 0) / \partial \theta' \\ \frac{1}{\sqrt{2}} \hat{\omega}_{1t}^{-2} \partial \omega_{1t}^2(\hat{\theta}, 0) / \partial \theta' \end{bmatrix}$$

and

$$\hat{\lambda}_{1t} = \begin{bmatrix} \hat{\omega}_{1t}^{-1} \partial \mu_{1t}(\hat{\theta}, 0) / \partial \vartheta_1' \\ \frac{1}{\sqrt{2}} \hat{\omega}_{1t}^{-2} \partial \omega_{1t}^2(\hat{\theta}, 0) / \partial \vartheta_1' \end{bmatrix}.$$

Explicit expressions of these partial derivatives as well as those required in the test procedure to be described shortly will not be given because some of them are very complicated. In the applications of the paper all partial derivatives have been computed numerically. Our test procedure for the specification of the lag length can now be obtained by following Example 3.3 of Wooldridge (1990). It consists of two least squares regressions. In the first one $\hat{\lambda}_{1t}$ is regressed on $\hat{\varphi}_{1t}$ to yield the residuals $\ddot{\lambda}_{1t}$. In the second least squares regression 1 is regressed on $\left[\hat{e}_t \quad (1/\sqrt{2})(\hat{e}_t^2 - 1) \right] \ddot{\lambda}_{1t}$ ($t = 1, \dots, T$) and the uncentered R_1^2 is computed. The test statistic with an asymptotic χ_{ms}^2 distribution under the null hypothesis is then given by TR_1^2 . We shall also use this test with the restriction $b_{ij} = b_j$ for all $i = 1, \dots, m$. In that case the definitions of the vectors $\hat{\varphi}_{1t}$ and $\hat{\lambda}_{1t}$ are modified in an obvious way and the degrees of freedom in the limiting χ^2 distribution are replaced by s .

Now consider testing the adequacy of the conditional variance implied by a specified MAR model. In addition to the number of component models it is here assumed that the conditional mean is correctly specified. Thus, $\mu_{2t}(\theta) = E_{t-1}(y_t)$ is given by (8) whether the null hypothesis holds or not. We shall explicitly only consider the case where the null hypothesis states that the conditional variance is given by (9) with σ_i^2 replaced by the restricted GARCH process (13). Under the alternative (13) is extended to

$$\sigma_{it}^2 = \sigma_i^2 + \beta_1 \sigma_{i,t-1}^2 + \dots + \beta_r \sigma_{i,t-r}^2 + \alpha_1 u_{i,t-1}^2 + \dots + \alpha_{q+s} u_{i,t-q-s}^2$$

so that the null hypothesis implies that the parameter vector $\vartheta_2 = [\alpha_{p+1} \dots \alpha_{q+s}]'$ is zero. Let $\omega_{2t}^2(\theta, \vartheta_2) = \text{Var}_{t-1}(y_t)$ be the conditional variance of the unrestricted

model and denote $\hat{\omega}_{2t}^2 = \omega_{1t}^2(\hat{\theta}, 0)$. The test procedure can again be obtained by following Example 3.3 of Wooldridge (1990). Instead of $\hat{\varphi}_{1t}$ and $\hat{\lambda}_{1t}$ we now have

$$\hat{\varphi}_{2t} = \begin{bmatrix} \hat{\omega}_{2t}^{-1} \partial \mu_{2t}(\hat{\theta}) / \partial \theta' \\ \frac{1}{\sqrt{2}} \hat{\omega}_{2t}^{-2} \partial \omega_t^2(\hat{\theta}, 0) / \partial \theta' \end{bmatrix}$$

and

$$\hat{\lambda}_{2t} = \begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}} \hat{\omega}_{2t}^{-2} \partial \omega_t^2(\hat{\theta}, 0) / \partial \vartheta_2' \end{bmatrix},$$

respectively. The first step is again to run the least squares regression of $\hat{\lambda}_{2t}$ on $\hat{\varphi}_{2t}$ and obtain the residuals $\check{\lambda}_{2t}$. In the second least squares regression 1 is regressed on $\begin{bmatrix} 0 & (1/\sqrt{2})(\hat{e}_t^2 - 1) \end{bmatrix} \check{\lambda}_{2t}$ ($t = 1, \dots, T$) and the uncentered R_2^2 is computed. The test statistic with an asymptotic χ_s^2 distribution under the null hypothesis is then obtained as TR_2^2 . Similarly, the GARCH specification with regime-dependent coefficients (11) can be tested against the specification with constant coefficients (13). The test statistic has an asymptotic χ_{mr}^2 distribution under the null hypothesis.

In the same way as equation (13) is augmented with further lags of $u_{i,t}^2$ above, it can be augmented with other variables that are thought to affect the conditional variance in order to test for the adequacy of the specification. For instance, an asymmetric specification allowing the α 's to depend on the signs of the lagged u_{it} 's may be relevant in some applications. A test can be derived by adding $I(u_{i,t-1} > 0)u_{i,t-1}^2, \dots, I(u_{i,t-q} > 0)u_{i,t-q}^2$ to (13) and testing the null hypothesis that their coefficients are zero along the lines described above. The test statistic has an asymptotic χ^2 null distribution with mq and q degrees of freedom in models with regime-dependent and constant ARCH coefficients, respectively.

Following Example 3.3 of Wooldridge (1990) it is also possible to obtain an overall test for the lag length and conditional variance but we shall not consider this test. It may be worth noting that the above tests rely on the consistency of the ML estimator $\hat{\theta}$ so that violations from the distributional assumption used in the MAR model may make these tests invalid.

In addition to the diagnostic checks discussed above we shall evaluate the performance of a fitted MAR model by using several informal procedures suggested in the recent literature. First, because the model can be used as a data generating process in Monte Carlo simulation it is important to check that it is capable of reproducing realizations close enough to the observed series. To this end we simulate long realizations to find out whether the values typically are in the range observed in the data as suggested by Andersen and Lund (1997). Furthermore we shall compare the unconditional moments implied by the different model specifications and the data. Second, a common problem with multi-regime models is that they do not clearly classify observations into regimes, but for instance in a two-regime model the conditional probability of each regime stays around a half most of the time. This may suggest misspecification, and therefore we compute the regime switching performance

measures recently suggested by Ang and Bekaert (1998). Third, the forecasting performance of the model is examined, not only because in practice this kind of models are often used for forecasting, but also because bad out-of-sample forecasting performance may suggest overparametrization to which multi-regime models are naturally prone. Finally, because general conditions for stationarity of our MAR models are not available, stability is checked by means of various Monte Carlo simulation procedures. Moreover, using the approximations suggested by Stanton (1997) the implications of the estimated MAR models for continuous-time dynamics are checked and compared to results in previous literature.

5 Empirical Results

In this section we present an application of the MAR model to the short-term U.S. Treasury bill rate. More precisely, the data set consists of the weekly three-month U.S. Treasury bill secondary market rates covering the weeks from the full January 1954 to September 1999 (2386 observations). The data are extracted from the H.15 release of the Federal Reserve System. A somewhat shorter series from the same source was analyzed by Andersen and Lund (1997). The series is displayed in Figure 1. This is probably the longest weekly series of U.S. risk-free rate available, and using as long a sample as possible may be important in identifying the regimes of the MAR model.

5.1 Estimated MAR Models and Diagnostic Checks

We consider three different MAR model specifications. The first is a MAR(2,2,1) model with no ARCH type heteroskedasticity while the other two are two- and three-regime MAR models with a GARCH(1,1) component. The delay parameter d is chosen to be 1 throughout, because with weekly data it is not conceivable that further lags would affect the regime classification. In all the models the autoregressive coefficients are restricted to be constant across the regimes, and this restriction cannot be rejected at the 5 % significance level for any of the models. It turns out that the model with no ARCH type component is easily rejected by the diagnostic tests, whereas discriminating between the other two models is not possible on the basis of these tests alone, but other (informal) criteria must be considered. These criteria clearly indicate that the three-regime model is required to adequately model the interest rate data. The parameter estimates and asymptotic standard errors are presented in Table 1 and the results of the diagnostic tests in Table 2. All the computations are performed with the GAUSS CML library, using the Berndt, Hall, Hall and Hausman (1974) algorithm. The standard errors are obtained from the inverse of the final Hessian matrix.

With the exception of the level parameters μ_2 and μ_3 in all the models, and the variance parameters σ_1 and σ_2 in the MAR(2,2,1)-GARCH(1,1) model, all the pa-

rameters are accurately estimated and significant at the 5 % (assuming asymptotic normality). Insignificant level parameters were also reported by Gray (1996), and Ang and Bekaert (1998) in MS specifications of the U.S. short-term interest rate. Although this finding suggests the presence of regime switching only in the conditional variance, our simulation experiments indicate that models with constant level parameters fail to produce realizations resembling the observed series. The insignificance of the variance parameters in the two-regime MAR-GARCH model suggests that three regimes are required to adequately model the series. In all the models the sum of the autoregressive coefficients is almost unity reflecting the high persistence of the interest rate series. Because of the nonlinearity of the models this does not necessarily indicate that instability is likely to be a problem. The sums of the GARCH coefficients are 0.995 and 0.963 in the two- and three-regime MAR-GARCH models, respectively. For the three-regime model this deviates considerably more from unity than in the GARCH specifications without regime switching in the previous literature. Employing standard inference the restrictions in both the MAR(2,2,1) and the two-regime MAR-GARCH model would be rejected against the three-regime MAR-GARCH model. However, as discussed in Section 3, the standard likelihood ratio test statistics do not have the usual asymptotic χ^2 null distribution, and hence model selection has to be based on diagnostic tests and some informal criteria.

The residuals of the MAR(3,2,1)-GARCH(1,1) model are plotted in Figure 2. The residuals of the two-regime model look the same, while the residuals of the MAR(3,2,1) model are clearly heteroskedastic. The large negative outlier corresponds to the stock market crash 1987; the model with a dummy for the week of the crash was also estimated, but the results were essentially the same. According to the diagnostic tests in Table 2 higher-order autocorrelation in the series is not adequately modeled by any of the models. In the previously considered empirical models of the short-term U.S. interest rates unmodeled low-order autocorrelation has also been a problem. The MAR(2,2,1) model with no ARCH type heteroskedasticity is clearly misspecified suggesting that an ARCH component is indeed required to model the persistence in variance. A GARCH(1,1) specification seems to be sufficient, and the constancy of the GARCH coefficients across the regimes cannot be rejected at the 5% level.

In addition to the tests for the adequacy of the lag length in the conditional mean and variance some other diagnostic checks of interest for this particular series were computed (cf. Brenner et al. (1996)). Whether the models capture the stylized fact of level effect (conditional variance growing with the level of the interest rate) is tested by testing the significance of the lagged level in the conditional variance equation, and all the models clearly pass this test. We also tested for the significance of a dummy variable for the period of the "new operating procedures" of the Federal Reserve (November 1979–September 1982) when the volatility of the interest rate series was considerably higher than in the rest of the sample. Because the simple MAR model without GARCH type conditional heteroskedasticity does not capture the volatility persistence in the series, it fails to model this structural break while the other two models seem to be adequate in this respect. Finally, the adequacy of the GARCH

specifications against asymmetric alternatives is tested by allowing the coefficient of the unscaled squared residual u_{t-1}^2 in the conditional variance equation to be different for positive and negative u_{t-1} . The fact that this kind of constancy of the ARCH coefficient cannot be rejected in either model lends support to the simple GARCH specification against asymmetric alternatives such as EGARCH.

5.2 Dynamic Properties

Because there are no results available on the stability of the different MAR models, the stability of the estimated models is examined by simulation. To this end Gallant et al. (1993) suggested computing the conditional mean and volatility profiles with a large range of initial values, each profile obtained by computing the mean at every point t ($t = 1, \dots, T$) of the respective moment from a large number of realizations. By over-plotting all these profiles we obtain a profile bundle whose shape should reveal potential excessive dependence on initial conditions. If the bundle retains its thickness, instability of the model is suspected, while if it gets narrower over the sequence, this lends support to stability. The conditional moment profiles of the AR(2)-GARCH(1,1) and the two- and three-regime MAR-GARCH models were computed based on 100 different sets of consecutive initial values randomly drawn from the interest rate series. The conditional mean profile bundles of none of the models indicated instability while the conditional volatility bundles depicted in Figure 3 seem to differ in this respect between the models. The AR-GARCH model is included to illustrate the profile bundles in a case where instability in the conditional volatility is known to prevail (the sum of the parameter estimates in the conditional variance function is unity), and this is indeed confirmed by its conditional volatility bundle. The volatility bundle of the two-regime MAR-GARCH model does not seem to get narrower, either, while that of the three-regime model clearly dampens albeit slowly. Hence these results lend support to the three-regime model against the two-regime model.

Since the main motivation of building models of the short-term interest rate is the use of these models as DGPs in Monte Carlo simulation in asset pricing applications, it is important to check that the models are capable of reproducing realizations resembling the observed series. As was recently demonstrated by Andersen and Lund (1997), various (level-)-AR-GARCH specifications of the short-term interest rate are excessively erratic at their estimated parameter values, typically producing realizations with values far beyond the observed range. Andersen and Lund (1997) checked their level-EGARCH model for the three-month U.S. Treasury yield by simulating a long realization of 100,000 observations from the estimated model, dividing this realization into 100 subsamples of 1000 observations each and recording the maximum of each subsample. For their model the simulated rate never exceeded 75% and only a single sample had rates higher than 37%. Moreover, the numbers of samples with the maximum rate over 25% were 7 and 4, respectively, for models with Gaussian and Student- t error terms. For the level-AR-GARCH model, rates in excess of 100%

were common, and in 17 subsamples the maximum exceeded 1000%. We conducted a similar experiment, and for the three-regime MAR-GARCH model the maximum rate never exceeded 35% with only one subsample having rates in excess of 25%, at best. It must be emphasized, though, that somewhat different results are easily obtained by repeating this exercise with a different seed for the random number generator. Still, typically only three or four subsamples had rates in excess of 25% and the rate never exceeded 60%.

Further evidence on the capability of the different models to match the characteristics of the data is provided in Table 3 which presents unconditional moments. They are computed as sample moments from simulated realizations of 100,000 observations. The excess skewness and kurtosis in relation to the normal distribution in the observed data is best matched by the three-regime MAR-GARCH model, whereas the two-regime and MAR(3,2,1) model imply much higher values of these moments. The unconditional variance of the two-regime model is also excessively high.

To get an idea of how well the MAR-GARCH model captures the high volatility periods reported in the previous literature the time series of conditional volatility computed as the square root of the conditional variance (9) is plotted in Figure 4. In accordance with the previous findings, three periods of high volatility stand out, namely the OPEC oil crisis 1973–1974, the Fed’s policy change 1979–1982 and the 1987 stock market crash. These same periods are also visible in Figure 5 that plots the ex ante conditional probabilities (5) of the preferred MAR-GARCH model. The conditional probability of the upper regime deviates considerably from zero only in the period of the Fed experiment 1979–1982. Thus it seems that the third regime is required only to model this exceptional period. The plots for the middle and lower regimes are almost mirror images of each other. The higher conditional probability of the middle regime and smaller probability of the lower regime at the oil crisis can clearly be seen. While there are distinct spikes at the 1987 stock market crash, this period does not stand out in the probability plots as clearly as in Figure 4.

The conditional probabilities can also be used to compute regime classification measures that give information on the ability of the model to distinguish between the regimes. A good model would classify regimes such that the conditional ex ante probabilities were close to zero or one. If the probabilities in a two-regime model are in general close to a half, the regime classification performance of the model is inferior, indicating potential misspecification. We computed the *RCM* measure of Ang and Bekaert (1998),

$$RCM(m) = 100m^2 \frac{1}{T} \sum_{t=1}^T \left(\prod_{i=1}^m \pi_{i,t-d} \right)$$

where m is the number of regimes. This measure can take values between 0 and 100, with 0 indicating perfect regime classification. For our preferred three-regime MAR-GARCH model this measure equals 1.62 while for the two-regime MAR-GARCH model the figure is 15.38. For univariate MS models of the monthly U.S. three-month

interest rate, covering the period from January 1972 to August 1996, Ang and Bekaert (1998) got values of the *RCM* measure ranging between 23.29 and 37.44. Thus these measures seem to favor the MAR model, and, in particular, the three-regime MAR-GARCH specification.

5.3 Forecasting Performance

Because forecasting volatility is one of the main practical uses of the models of the short-term interest rate, it is important to provide evidence of the forecasting performance. Moreover, out-of-sample forecasting exercises can also be motivated as complementary checks of potential overparametrization in addition to diagnostic tests and the informal evaluation procedures. It is conceivable that the relative advantage of the MAR-GARCH model would be greater at long horizons, but for simplicity, we concentrate on the one-period horizon where the forecasts from the regime-switching models can be computed in closed form. Following Gray (1996), we compared the forecasts from the regime-switching models to those from a constant-variance autoregressive model and an AR-GARCH model. The evaluation of volatility forecasts is difficult because the actual variance is not observable but must also be based on the estimated model. This complicates the comparison of measures of accuracy such as the root mean squared error (*RMSE*). Therefore, we also report the following commonly used R^2 measure,

$$R^2 = 1 - \frac{\sum_{t=1}^T (V_t - \hat{V}_t)^2}{\sum_{t=1}^T V_t^2},$$

where V_t is the actual variance (given by the model) computed as $(y_t - \hat{E}_{t-1}y_t)^2$ and \hat{V}_t is its forecast. $\hat{E}_{t-1}y_t$ denotes the estimate of the conditional expectation of the interest rate y_t from the estimated model. The numerator of the latter term is $T \cdot MSE$, so that of two models having the same *MSE* the R^2 is larger for the one that implies the larger actual variance on average. Hence, this measure relates the forecast errors to the average magnitude of the variance implied by the model. Notice that R^2 can also take a negative value if the unexplained sum of squares in the numerator is higher than the total sum of squares in the denominator. This could happen, for instance, in an AR model when the variance in the estimation period is much higher than in the forecasting period, so that the constant \hat{V}_t always tends to be larger than V_t .

In addition to the two- and three-regime MAR-GARCH models, measures of forecast accuracy are provided for the AR, AR-GARCH and the GRS (Gray, 1996) models. We considered two different estimation periods to highlight the forecasting performance of the models in different kinds of situations. In the first case the estimation period (January 1954–December 1978) does not include the Fed experiment while it is included in the out-of-sample forecasting period. The expectation is that multi-regime models do not necessarily perform better out of sample here

because the estimation period has no observations of the exceptional regime in the out-of-sample period. In the second case the estimation period (January 1954–March 1981) includes the first half of the period of the ‘new operating procedures’ of the Fed. Hence one would expect the multi-regime models to perform better both in and out of sample. The models were estimated with data from the entire estimation period and all the forecasts were computed keeping these parameter values fixed. The adequacy of the estimated models was tested using the Ljung-Box test for autocorrelation in the squared residuals, and apart from the AR model, the null of no autocorrelation could not be rejected for any of the models at conventional significance levels.²

The root mean squared prediction errors and the R^2 measures of the one-period ahead forecasts are presented in Table 4. As far as the $RMSE$ ’s are concerned, one of the MAR-GARCH models is in each case superior, and the differences between the two- and three-regime models are small. However, as discussed above, the R^2 is probably a more sensible measure of forecast accuracy. According to this measure the differences are minor in the sample but in the out-of-sample period the two-regime model clearly seems to be superior. Since the $RMSE$ ’s of these two models are quite close to each other, the three-regime model must produce a much smaller conditional variance, especially in the first experiment. This suggests that at least for forecasting one period ahead the three-regime model may be overparametrized. Especially in the first case this observation is not surprising since the estimation period excludes the exceptional period of the Fed experiment, suggesting that the two-regime model might be adequate. Still, it is conceivable that the three-regime model would be superior for forecasting several periods ahead provided the estimation period had observations of this regime so that its parameters could be properly estimated. According to the R^2 measures the GRS and AR-GARCH models are not much inferior, either. The length of the out-of-sample period may be one explanation for the relatively good behavior of the AR-GARCH model that does not take regime switching into account. Although it produces large forecasting errors during the period of the new operating procedures, their number is negligible in relation to the entire forecasting period.

5.4 Comparison with Continuous-Time Models

Much of the recent empirical literature on the short-term U.S. interest rate deals with continuous-time models. In order to compare the properties of our MAR-GARCH model with those of the models presented in this literature, its implications on the parameters of a typical continuous-time model must first be derived. To this end we employ Stanton’s (1997) first-order approximations of the drift and diffusion functions of the following diffusion process,

$$dy_t = \mu(y_t)dt + \sigma(y_t)dW_t,$$

where W_t is a standard Brownian motion. These approximations are based on the conditional expectation and variance conditional on the lagged interest rate y_{t-1} . Stanton

²The estimation and test results are not reported here, but they are available upon request.

(1997) estimated these nonparametrically. Because for our MAR(3,2,1)-GARCH(1,1) model the conditional expectation (8), and variance (9) depend also on y_{t-2} , y_{t-3} and y_{t-4} , these formulas cannot be used directly. Instead, we employed the following simulation procedure. First, a long realization of the model is simulated. For each observation, y_{t-1} , in the realization the approximations of the drift and diffusion are computed and recorded. The values of the lagged interest rates, y_{t-1} , are then divided into bins of width 0.25, and within each bin the average drift and diffusion are calculated. A realization of length 500,000,000 was needed to obtain smooth plots.

The drift and diffusion functions are plotted in Figure 6. The drift is similar in shape to those obtained by Aït-Sahalia (1996), Stanton (1997), Jiang (1998) and Ang and Bekaert (1998)³ all using different methods and data sets. At low interest rate levels the drift is close to zero indicating near unit root behavior, and turns clearly negative at about 8% with the mean reversion getting stronger with the interest rate level. In contrast to the results of Stanton (1997) and Jiang (1998), however, there is no marked abrupt change in the steepness for rates beyond 15%. The diffusion function is, in general, upward sloping. It bears a close resemblance to those obtained in the studies mentioned above in that at low levels the curve is almost constant but gets increasingly steeper with the interest rate level. At low interest rate levels the estimated diffusion function takes somewhat lower and at high levels somewhat higher values than Stanton's (1997). One explanation to this might be the recent observation of Chapman and Pearson (2000) that Stanton's (1997) estimator tends to give too smooth diffusion function estimates, overestimating at low and underestimating at high interest rate levels. Stanton (1997) also provides 95% bootstrapped confidence bands for the estimated drift and diffusion functions. While such confidence bands could in principle be obtained for the approximations based on our model, their computation would be extremely burdensome because for each bootstrap sample the nonlinear MAR-GARCH model would have to be estimated. Therefore, we must be content with the observation that the estimated drift in Figure 6 lies clearly within the 95% confidence band of Stanton's (1997) drift function estimate, and our parametric model is expected to produce a tighter confidence band, provided the model is adequate. The fact that our diffusion function estimate does not entirely lie within Stanton's (1997) 95% confidence band is likely to follow from the results of Chapman and Pearson (2000) mentioned above.

5.5 Illustration: Bond Pricing

To demonstrate one practical application of the MAR-GARCH model, we show how it can be used to value zero-coupon bonds with different maturities. This also gives information on the term structure implications of the model and facilitates comple-

³Aït-Sahalia (1996) modeled a daily series of the 7-day Eurodollar deposit spot rate covering the period from 1 June 1973 to 25 February 1995. Stanton (1997) and Jiang (1998) considered a daily series of the U.S. three-month Treasury bill yield covering the periods from January 1965 to July 1995 and from January 1962 to January 1996, respectively.

mentary comparisons with the previous literature.

Bond pricing can be based on the risk-neutral valuation principle and in practice conducted using Monte Carlo simulation. The general idea is to simulate a large number of realizations of (the discretization of) the risk-adjusted interest rate process

$$d\tilde{y}_t = [\mu(\tilde{y}_t) - \lambda(\tilde{y}_t)] dt + \sigma(\tilde{y}_t) dW_t \quad (14)$$

over a period starting at $y_t = y_0$ at date t and finishing at date T , the maturity of the bond. An estimate of the price of a zero-coupon bond at date t with payoff \$1 at time T ,

$$P_t(T) = E_t \left[\exp \left(- \int_t^T \tilde{y}_s ds \right) \right], \quad (15)$$

is then obtained as the average of the expression in the brackets in (15) over the simulated realizations. In addition to the drift and diffusion functions computed in Section 5.4, this approach requires knowledge of the market price of interest rate risk $\lambda(\tilde{y}_t)$, i.e. the excess return the investor requires to bear each extra unit of risk. It is well known that with an one-factor interest rate model the absence of arbitrage requires the risk premium on any asset to be proportional to the standard deviation of its return. The calculation of the market price of interest rate risk can, therefore, be based on any two assets, and, following Stanton (1997) we have chosen to estimate it using data on the three- and six-month interest rates. In order to apply Stanton's (1997) approximation (his formula (51)) we need an estimate of the conditional expectations of the difference of the holding period returns between the six- and three-month rates and the diffusion function of the six-month rate. These were computed from separate MAR-GARCH models for these variables with the lagged three-month rate as the threshold variable ⁴, using the simulation procedure described in Section 5.4. The estimated market price of interest rate risk function is depicted in Figure 7. Although it lies well within the 95% confidence band of Stanton's (1997) estimate (except when the short rate is very close to 20%), the shapes of the two estimates are quite different. Stanton (1997) obtains the result that $\lambda(\cdot)$ is very close to zero when the short rate is 15% or less and then decreases steeply at high short-rate levels. A similar conclusion emerges from the study of Jiang (1998) who also uses nonparametric methods. Our estimate of $\lambda(\cdot)$, on the other hand, does not show any abrupt change at high short rate levels, and the steepest decline takes place between 2 and 6% levels of the short rate. Both Stanton's (1997) and Jiang's (1998) estimates are very inaccurate at high short rate levels which makes direct comparisons difficult. Because of the computational difficulties already discussed in Section 5.4 no confidence bands are reported in Figure 7.

Table 5 gives the estimated bond prices for maturities of one, two and three years. In addition to the prices based on the risk-adjusted process (14), those assuming λ to

⁴The results in this section are based on data from the period from January 1959 to September 1999, the availability of data on the six-month rate. The estimated models are not reported to save space, but the results are available upon request.

be zero are presented. This allows for direct comparison with the results of Stanton (1997) who also reported both figures. In each case a rising yield curve is obtained. When y_0 is 1% the differences between the prices obtained assuming the estimated $\lambda(\cdot)$ and $\lambda = 0$ are minor albeit growing with the maturity. However, when y_0 equals 5%, the differences are substantial already for the one-year bond. These results agree, in general, with those of Stanton (1997); the differences are especially small for the $\lambda = 0$ case. On the other hand, when the respective estimates of the market price of risk are included in the valuation, there are differences for maturities over one year. When $y_0 = 1\%$ our estimates imply lower and when $y_0 = 5\%$, higher interest rates than Stanton's (1997). This is due to the fact that our estimate of the market price of risk starts getting larger (in absolute value) at lower short-rate levels than his.

The term structure implications of the MAR models are further illustrated in Figure 8 which depicts the simulated yield curves corresponding to some short-rate levels and up to 15 years of maturity, computed from the bond prices obtained as describe above. These curves are, in general, consistent with the historical U.S. term structures in two respects. First, they are steeper at the short end and get flatter towards the long end of the maturity spectrum. Second, the higher is the short-rate level, the flatter is the yield curve.

6 Conclusion

The presence of multiple regimes is a central result emerging from the recent empirical literature on the U.S. short-term interest rate. It can, however, be argued that the suggested multi-regime MS and TAR models let alone GARCH type models are not optimal for modelling this time series. In addition to numerical problems in estimation many of these have been shown to produce excessively erratic realizations which makes them unsuitable for the typical applications of pricing interest-rate sensitive assets and continuous-time model estimation by simulation-based methods. In this paper we introduce a new kind of model which is a mixture of linear autoregressive models with GARCH errors. It is also shown that robust diagnostic tests can easily be derived applying the approach of Wooldridge (1990).

The MAR model is applied to weekly data on the U.S. three-month Treasury bill rate. According to diagnostic tests as well as some informal checks a MAR with a GARCH component seems to provide a good description of the U.S. short-term interest rate process. Specifically, the realizations generated from the estimated model seem stable and their properties resemble those of the observed series closely. The out-of-sample forecasting performance of the MAR model is also superior compared to some alternatives in the previous literature. Using the approach of Stanton (1997) the MAR model can also be used to compute approximations of the parameters of the continuous-time diffusion model on which much of the interest in the empirical interest rate literature has recently concentrated. The drift and diffusion functions implied by the MAR model turned out to be, in general, in accordance with the results in this literature. Using additional data on the six-month interest rate, an approximation of

the market price of interest rate risk is also estimated and the implied term structure patterns are found to accord with those historically observed.

References

- Aït-Sahalia, Y. (1996). Testing continuous-time models of the spot interest rate. *Review of Financial Studies* 9, 385-426.
- Andersen, T.G. and J. Lund (1997). Estimating continuous-time stochastic volatility models of the short-term interest rate. *Journal of Econometrics* 77, 343-377.
- Andersen, T.G., H.-J. Chung and B.E. Sørensen (1999). Efficient method of moments estimation of a stochastic volatility model: A Monte Carlo study. *Journal of Econometrics* 91, 61-87.
- Andrews, D.K.W. (1993). Tests for parameter instability and structural change with unknown change point. *Econometrica* 61, 821-856.
- Andrews, D.K.W. and W. Ploberger (1994). Optimal tests when a nuisance parameter is present only under the alternative. *Econometrica* 62, 1383-1414.
- Ang, A. and G. Bekaert (1998). Regime switches in interest rates. NBER Working Paper 6508.
- Berndt, E.K., B.H. Hall, R.E. Hall and J.A. Hausman (1974). Estimation and inference in non-linear structural models. *Annals of Economic and Social Measurement* 4, 653-665.
- Brenner, R.J., R.H. Harjes and K.F. Kroner (1996). Another look at models of the short-term interest rate. *Journal of Financial and Quantitative Analysis* 31, 85-107.
- Cai, J. (1994). A Markov model of switching-regime ARCH. *Journal of Business and Economic Statistics* 12, 309-316.
- Chan, K.C., G.A. Karolyi, F.A. Longstaff and A.B. Sanders (1992). An empirical comparison of alternative models of the short-term interest rate. *Journal of Finance* 47, 1209-1227.
- Chan, K.S. and H. Tong (1986). On estimating thresholds in autoregressive models. *Journal of Time Series Analysis* 7, 179-194.
- Chapman, D.A. and N.D. Pearson (2000). Is the short rate drift actually non-linear? *Journal of Finance* 55, 355-388.

- Davies, R.B. (1977). Hypothesis testing when a nuisance parameter is present only under the alternative. *Biometrika* 61, 247-254.
- Gallant, A.R., P.E. Rossi and G. Tauchen (1993). Nonlinear dynamic structures. *Econometrica* 61, 871-907.
- Granger, C.W.J. and T. Teräsvirta (1993). *Modelling Non-linear Economic Relationships*. Oxford University Press. Oxford.
- Gray, S.F. (1995). An analysis of conditional regime-switching models. Manuscript, Graduate School of Business, Stanford University.
- 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. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica* 57, 357-384.
- Hamilton, J. D. (1994). *Time Series Analysis*. Princeton University Press. Princeton.
- Hamilton, J.D. and R. Susmel (1994). Autoregressive conditional heteroskedasticity and changes in regime. *Journal of Econometrics* 64, 307-333.
- Hansen, B.E. (1996). Inference when a nuisance parameter is present only under the alternative. *Econometrica* 64, 413-430.
- Jiang, G.J. (1998). Nonparametric modeling of U.S. interest rate term structure dynamics and implications on the prices of derivative securities. *Journal of Financial and Quantitative Analysis* 33, 465-497.
- Koedijk, K.G., F.G.J.A. Nissen, P.C. Schotman and C.C.P. Wolff (1997). The dynamics of short-term interest rate volatility reconsidered. *European Finance Review* 1, 105-130..
- Le, N.D., R. D. Martin, and A.E. Raftery (1996). Modeling flat stretches, bursts, and outliers in time series using mixture transition distribution models. *Journal of the American Statistical Association* 91, 1504-1515.
- Meyn, S.P. and R.L. Tweedie (1993). *Markov Chains and Stochastic Stability*. Springer-Verlag. Berlin, Heidelberg, New York.
- Pfann, G.A., P.C. Schotman and R. Tschernig (1996). Nonlinear interest rate dynamics and implications for the term structure. *Journal of Econometrics* 74, 149-176.
- Stanton, R. (1997). A nonparametric model of term structure dynamics and the market price of interest rate risk. *Journal of Finance* 52, 1973-2002.

Tong, H. (1990). Non-linear Time Series: A Dynamical System Approach. Oxford University Press. Oxford.

Wong, C.S. and W.K. Li (1999a). On a mixture autoregressive conditional heteroscedastic model. Research Report 225 (Department of Statistics and Actuarial Science, University of Hong Kong).

Wong, C.S. and W.K. Li (1999b). On a generalized mixture autoregressive model. Research Report 242 (Department of Statistics and Actuarial Science, University of Hong Kong).

Wong, C.S. and W.K. Li (2000). On a mixture autoregressive model. Journal of the Royal Statistical Society, Series B

Wooldridge, J.F. (1990). A unified approach to robust, regression-based specification tests. Econometric Theory 6, 17-43.

Yakowitz, S.J. and J.D. Spragins (1968). On the identifiability of finite mixtures. The Annals of Mathematical Statistics 39, 209-214.

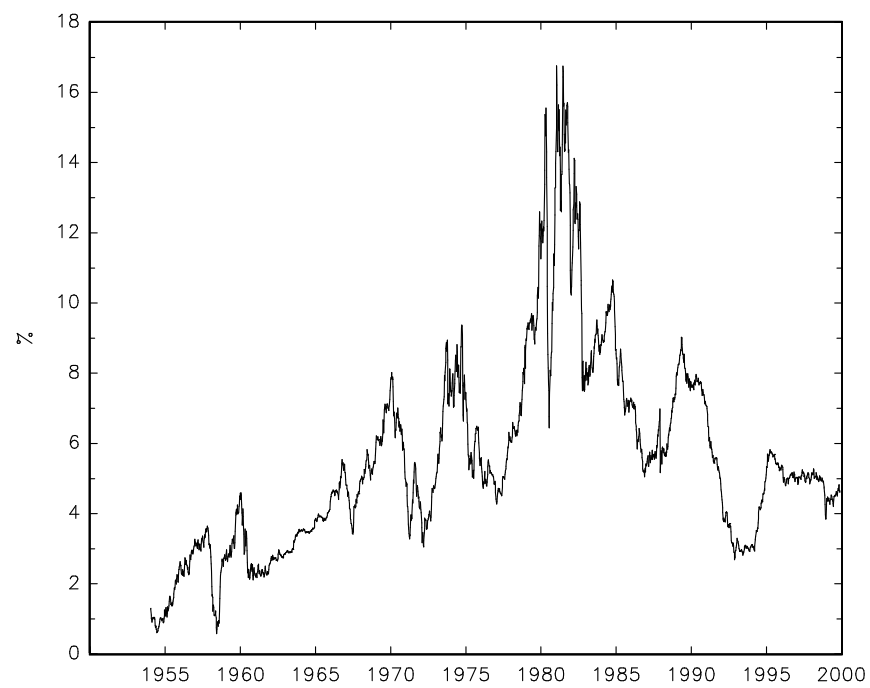


Figure 1: The weekly observations of the three-month U.S. Treasury bill rate covering the period from January 1954 to September 1999.

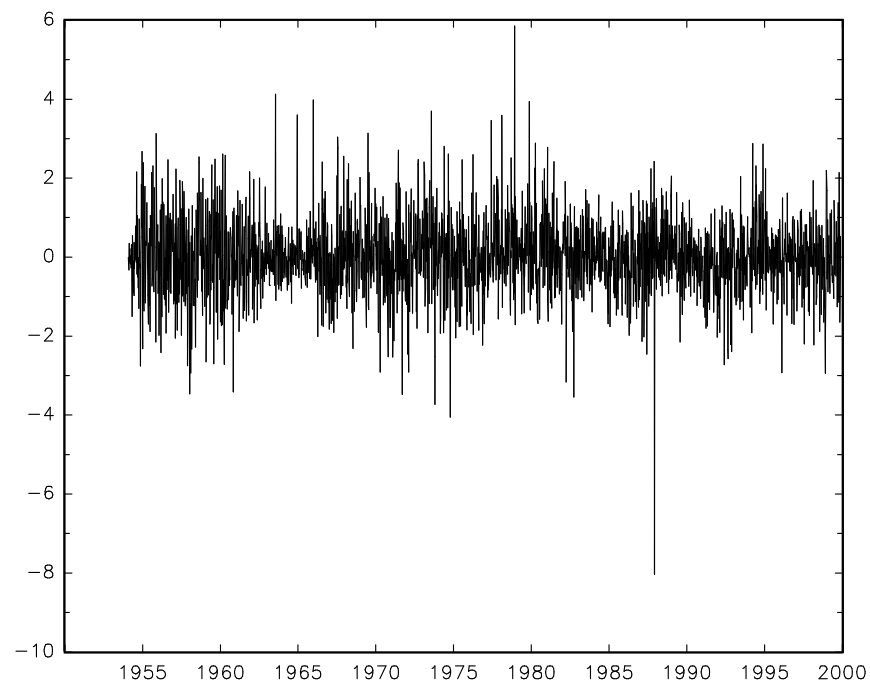


Figure 2: The residuals of the $\text{MAR}(3,2,1)\text{-GARCH}(1,1)$ model for the three-month U.S. Treasury bill rate.

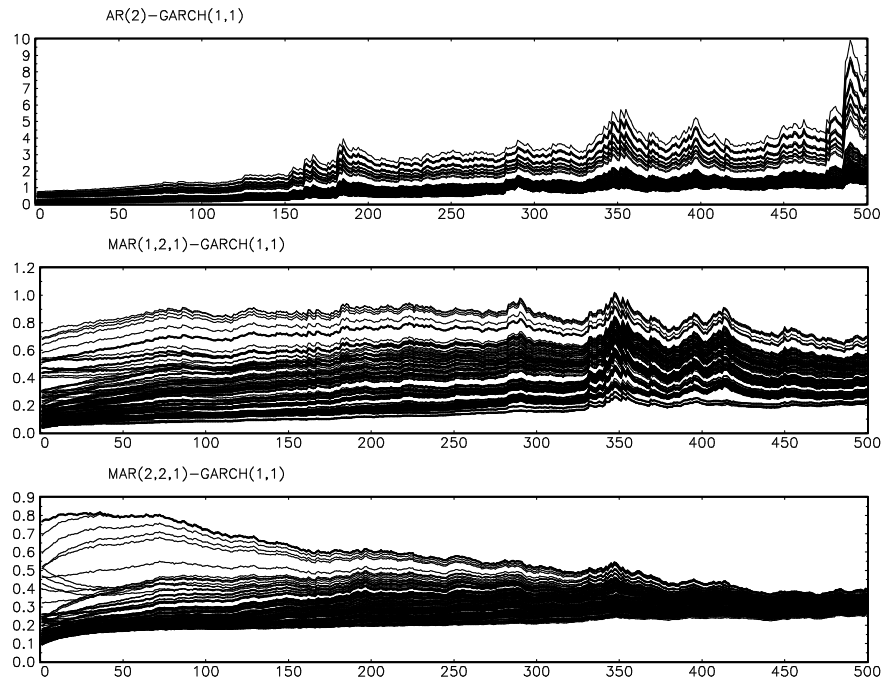


Figure 3: Profile bundles of the conditional volatility for the AR(2)-GARCH(1,1), MAR(2,2,1)-GARCH(1,1) and MAR(3,2,1)-GARCH(1,1) models of the three-month U.S. Treasury bill rate.

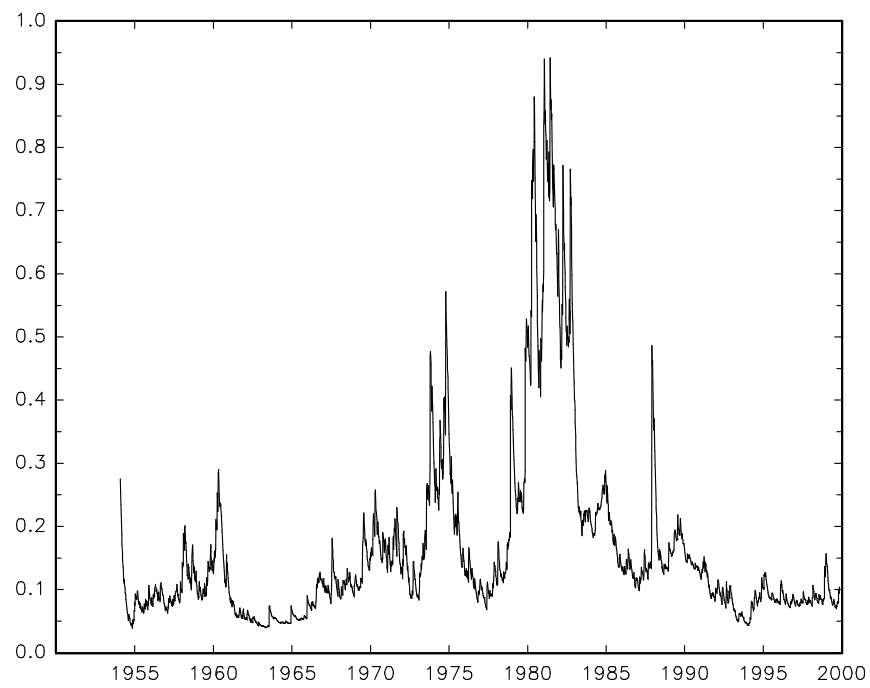


Figure 4: Conditional volatility of the U.S. three-month interest rate series computed from the $MAR(3,2,1)$ -GARCH(1,1) model.

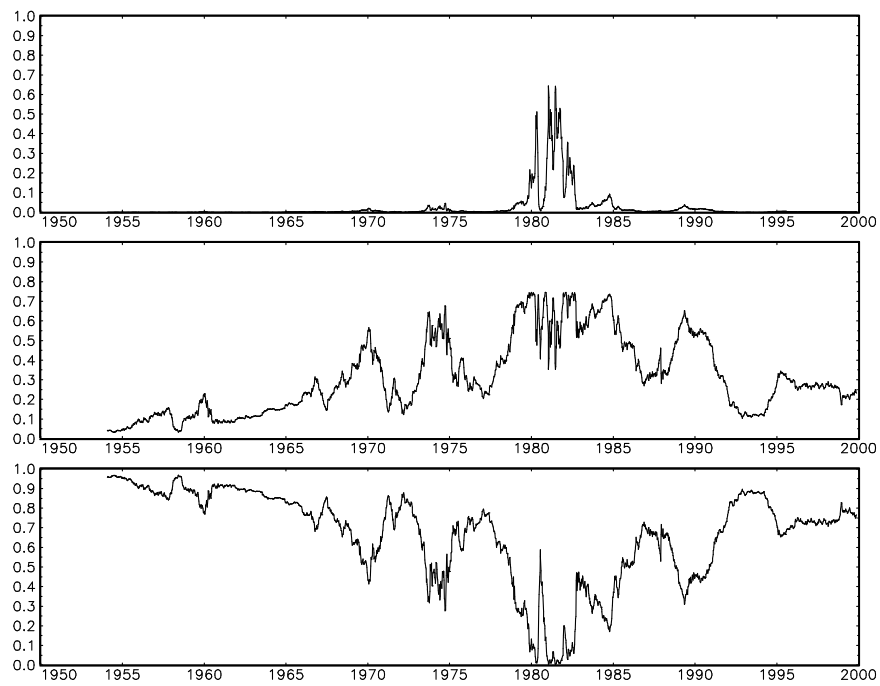


Figure 5: The ex ante conditional probabilities of the upper, middle and lower regimes implied by the MAR(3,2,1)-GARCH(1,1) model for the three-month U.S. Treasury bill rate.

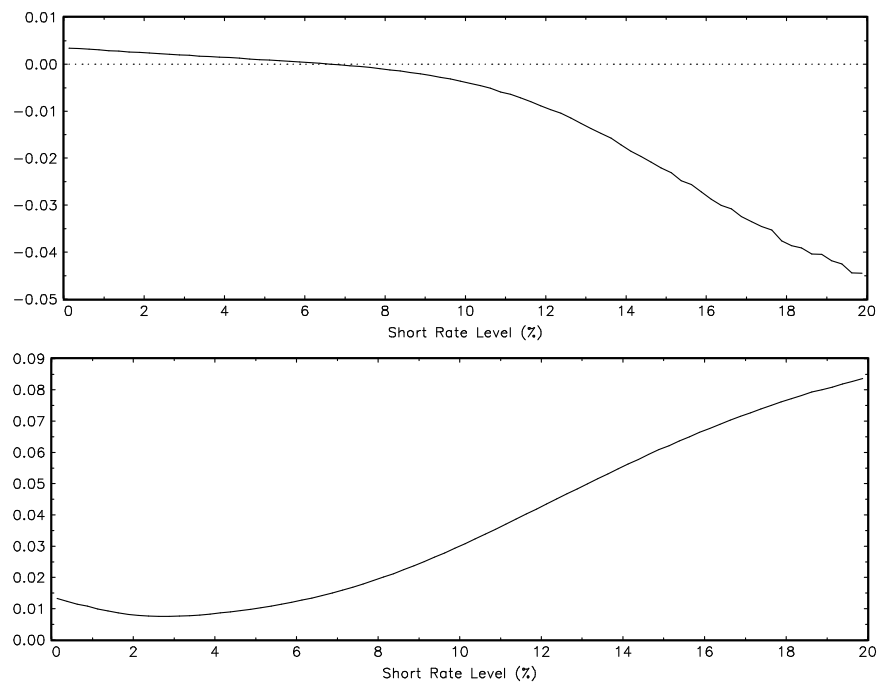


Figure 6: Drift (upper panel) and diffusion functions (lower panel) implied by the MAR(3,2,1)-GARCH(1,1) model of the three-month U.S. Treasury bill rate.

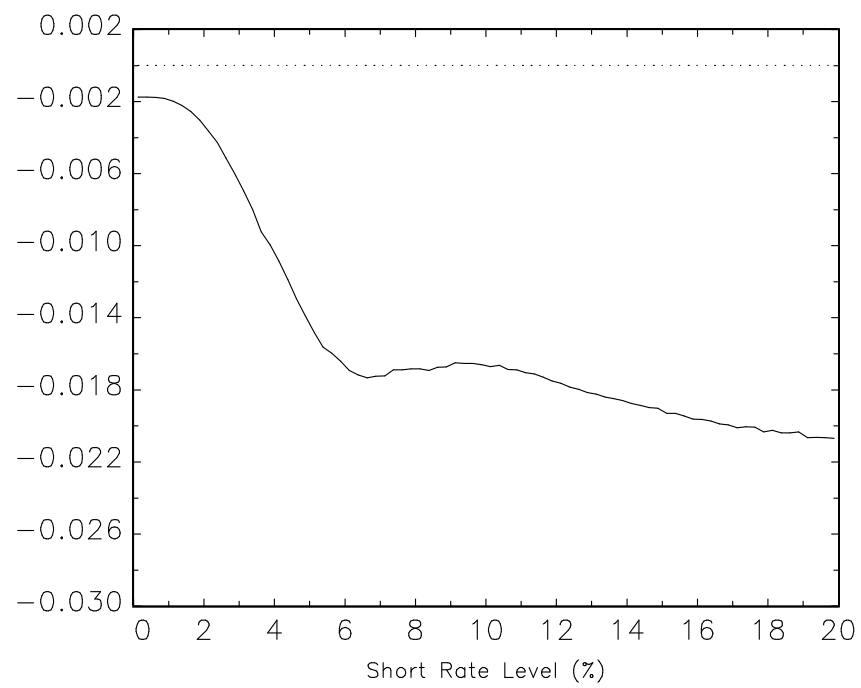


Figure 7: Market price of interest rate risk implied by the MAR models.

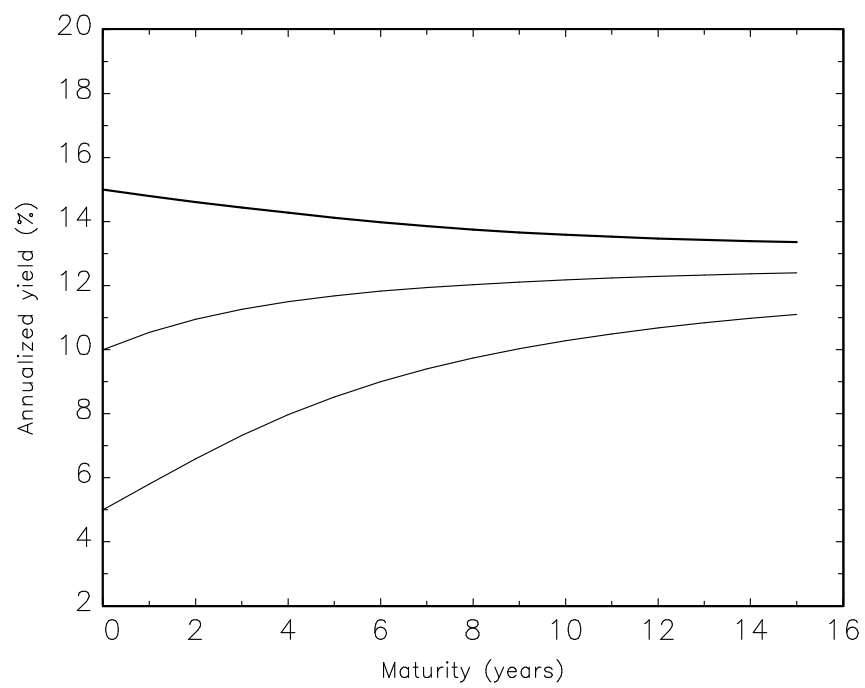


Figure 8: Term structures for some short-rate levels.

Table 1: Estimation results of the different MAR models.

	Model		
	MAR(3,2,1)	MAR(2,2,1)- GARCH(1,1)	MAR(3,2,1)- GARCH(1,1)
μ_1	0.012 (0.005)	0.012 (0.005)	0.009 (0.005)
μ_2	0.019 (0.011)	-0.020 (0.057)	0.012 (0.012)
μ_3	0.004 (0.052)		-0.093 (0.138)
c_1	5.273 (0.446)	14.920 (2.811)	7.236 (0.776)
c_2	11.551 (0.413)		15.437 (1.342)
b_1	1.213 (0.020)	1.242 (0.022)	1.237 (0.021)
b_2	-0.215 (0.020)	-0.244 (0.022)	-0.238 (0.021)
σ_1	0.004 (0.001)	0.00002 (0.0002)	0.00002 (0.00001)
σ_2	0.026 (0.004)	0.019 (0.024)	0.002 (0.0004)
σ_3	0.398 (0.056)		0.079 (0.030)
σ_η	8.386 (1.171)	22.720 (8.110)	12.950 (4.100)
α_1		0.122 (0.021)	0.115 (0.016)
β_1		0.873 (0.022)	0.848 (0.020)
Log likelihood	1476.073	1720.475	1762.149

The figures in the parentheses are standard errors computed from the inverse of the final Hessian matrix.

Table 2: Diagnostic checks of the different MAR models.

	Model		
	MAR(3,2,1)	MAR(2,2,1)- GARCH(1,1)	MAR(3,2,1)- GARCH(1,1)
Conditional mean			
AR(3) ^a	0.096	0.088	0.082
AR(4)	0.235	0.110	0.135
AR(5)	0.025	0.0003	0.001
General MAR(-GARCH) ^b	0.149	0.178	0.090
Conditional variance			
ARCH(1) ^c	0.005		
GARCH(1,2) ^d		0.615	0.738
GARCH(1,3)		0.484	0.254
GARCH(1,4)		0.160	0.126
General GARCH(1,1) ^e		0.095	0.102
Level effect ^f	0.977	0.426	0.470
Structural Break ^g	0.002	0.082	0.094
Asymmetry ^h		0.223	0.194

All figures are marginal significance levels. The numerical derivatives required in the calculation of the test statistics are computed by the gradp routine in GAUSS.

^aIn the tests labeled AR(l) the alternative model is the corresponding MAR (-GARCH) model with l lags in the conditional mean equation.

^bThe alternative model is the MAR (-GARCH) model with regime-dependent autoregressive coefficients.

^cThe alternative model is the MAR-ARCH(1) model.

^dIn the tests labeled GARCH(1, l) the alternative model is the corresponding MAR-GARCH model with l lags of the squared error term in the conditional variance.

^eThe alternative model is the MAR-GARCH(1,1) model with regime-dependent GARCH coefficients.

^fThe alternative model is the MAR(-GARCH) model with the lagged level of the interest rate entering the conditional variance equation.

^gThe alternative model is the MAR(-GARCH) model with a dummy variable for the period of the "new operating procedures" of the Fed (November 1979 – September 1982) entering the conditional variance equation.

^hThe alternative model is the MAR-GARCH model with u_{t-1}^2 in the conditional variance equation allowed to take a different coefficient when u_{t-1} is positive.

Table 3: Comparison of unconditional moments.

	Population	Model		
		MAR(3,2,1)	MAR(2,2,1)-	MAR(3,2,1)-
			GARCH(1,1)	GARCH(1,1)
Mean	5.508	5.669	4.876	5.221
Variance	7.863	6.367	47.755	5.552
Skewness	1.125	2.171	4.467	1.150
Kurtosis	4.745	11.497	40.809	6.777

The moments implied by the different models are computed as sample moments from a simulated realization of 100,000 observations of the respective model.

Table 4: Forecasting performance of different models of the one-month U.S. Treasury bill rate January 1970–April 1994.

Model	In Sample	
	January 1954– December 1978	January 1954– March 1981
AR(1)	0.0877 (0.0660)	0.1711 (0.0617)
AR(1)-GARCH(1,1)	0.0971 (−0.1656)	0.1593 (0.1926)
GRS	0.0890 (0.1449)	0.1666 (0.2660)
MAR(2,2,1)-GARCH(1,1)	0.0836 (0.1540)	0.1511 (0.2889)
MAR(3,2,1)-GARCH(1,1)	0.0817 (0.1635)	0.1512 (0.2925)
	Out of Sample	
	January 1979– September 1999	April 1982– September 1999
AR(1)	0.2660 (0.0342)	0.2014 (0.0348)
AR(1)-GARCH(1,1)	0.2431 (0.1675)	0.1985 (0.0661)
GRS	0.2687 (0.1775)	0.2088 (0.0706)
MAR(2,2,1)-GARCH(1,1)	0.2383 (0.2154)	0.1976 (0.0952)
MAR(3,2,1)-GARCH(1,1)	0.2622 (0.0715)	0.2032 (0.0518)

The figures are root mean squared prediction errors of one period ahead forecasts. The figures in the parenthesis are the corresponding R^2 values. The parameters are estimated using the in-sample period data and held fixed over the out-of-sample period. GRS is the generalized regime-switching model of Gray (1996).

Table 5: Bond prices implied by the MAR(3,2,1)-GARCH(1,1) model.

		Bond Price	
		Estimated Market	Zero Market
y_0		Price of Risk	Price of Risk
1%	1 year	0.9876	0.9887
	2 years	0.9707	0.9755
	3 years	0.9484	0.9610
5%	1 year	0.9436	0.9508
	2 years	0.8765	0.9032
	3 years	0.8027	0.8575

The figures are prices of a zero-coupon bond with payoff \$1 at maturity. They are obtained as averages of the values of the bond price (15) based on 1,000 simulated realizations of the risk-adjusted short-rate process (14) with initial value y_0 . The discretization uses 500 time periods per week, and the antithetic variate approach is employed to reduce variability.