Estimation of threshold time series models using efficient jump MCMC

Elena Goldman*
Department of Finance and Economics
Lubin School of Business, Pace University
New York, NY 10038
e-mail: egoldman@pace.edu, tel: 212-618-6516

and

Terence D. Agbeyegbe
Department of Economics
Hunter College, CUNY
695 Park Avenue
New York, NY 10021
e-mail: tagbeyeg@hunter.cuny.edu, tel: 212-772-5405

April 7, 2005

Abstract

This paper shows how a Metropolis-Hastings algorithm with efficient jump can be constructed for the estimation of multiple threshold time series of the U.S. short term interest rates. The results show that interest rates are persistent in a lower regime and exhibit weak mean reversion in the upper regime. For model selection and specification several techniques are used such as marginal likelihood and information criteria, as well as estimation with and without truncation restrictions imposed on thresholds.

*Authors are indebted to Professor Hiroki Tsurumi from Rutgers University for his invaluable comments.

1 Introduction

Many time series exhibit non-linear dynamics that can be characterized by regime changes in persistence and volatility. Recently many papers used Threshold Autoregressive (TAR) models in economic and financial time series to model the dynamics of short-term interest rates, real exchange rates, unemployment rate, stock prices, production, and inventories.

In spite of popularity of TAR models the problem of efficient estimation of thresholds is still unsolved. The likelihood function is discontinuous in the thresholds and has multiple peaks in case of more than two regimes. There is enormous literature using classical methods where a threshold is estimated via grid search or residual scatter plots (see e.g. Tong and Lim (1980), Tong (1983, 1990), Tsay (1989, 1998), Hansen (1997), Lanne and Saikkonen (2002), etc.). We face a problem in estimating simultaneously multiple thresholds by these methods since accurate estimation with many dimensions of grid search takes a long time. To solve this problem, e.g. Gonzalo and Pitarakis (2002) used sequential estimation on a grid of points and found that their method works well only when regimes samples are equally spaced. The Bayesian analogue of a grid search is a Griddy Gibbs sampler within MCMC and it was done e.g. in Phann, Schotman and Tscherig (1996). Other Bayesian methods were suggested for a two-regime simple autoregressive specification with one threshold by Geweke and Terui (1993) using Monte Carlo integration for the case of noninformative priors. This method requires numerical multiple integration. Chen and Lee (1995) used a Metropolis algorithm within a Gibbs sampler and noted that their procedure avoids sophisticated integration or use of plots, which may result in imprecise values of estimated threshold parameter. However, they admit that the MCMC algorithm that they suggested is restrictive to a simple two-regime TAR model and much remains to be done for a general TAR model with multiple regimes. Koop and Potter (1999) extended Monte Carlo integration method suggested in Geweke and Terui (1993) for a three regime autoregressive model with informative priors. Forbes, Kalb and Kofman (1999) extended numerical integration for multiple thresholds using Rao-Blackwellized estimates of marginal posterior densities. In this paper we propose a Metropolis-Hastings algorithm with efficient jumps and estimate multiple thresholds jointly in a general class of ARMA-GARCH models. Our model can be extended to other models with many parameters. This algorithm is more efficient than existing methods; it avoids numerical integration or grid search and achieves convergence in reasonable computational time.

In this paper we analyze behavior of the U.S. short-term interest rates using a threshold model. The short-term interest rate is a key variable in many finance models, including term structure models and models that use risk-free rate as an input. Switches in regimes for short-term interest rates were studied among others by Hamilton (1988), Cai (1994), Gray (1996), etc. using Markov switching model, Lanne and Saikkonen (2002) and Phann, Schotman and Tscherig (1996) used threshold models. Hamilton (1988) analyzed the term structure of interest rates in response to FED's change in monetary policy in October 1979, which resulted in a period of a high level and high volatility of interest rates. He suggested to use a two-regime Markov-switching model instead of a constant coefficient autoregressive model for short rates since the former provided a better description of the univariate process for the short rate and was more consistent with historical correlation between long and short rates. Phann, Schotman and Tscherig (1996) (PST) modeled the dynamics of short-term interest rates as an autoregressive threshold model with autoregressive parameters and volatility depending on the level of interest rates. They noted that the traditional linear models of interest rate dynamics, emphasized in Chan, Karolyi, Longstaff and Saunders (CKLS) (1992), are not able to explain the empirical regularity that long term rates are not as volatile as short-term rates. The general stochastic differential equation used by CKLS to model the adjustment process of short-term interest rate can be represented by equation (1),

$$dr(t) = (\mu + \beta r(t))dt + \sigma r(t)^{\lambda} dW \tag{1}$$

where r(t) is the interest rate level at time t, t is time, W is a standard Brownian motion and μ, β, σ, and λ are parameters.

In this model $\mu + \beta r(t)$ is the drift and $\sigma^2 r(t)^{2\lambda}$ is the variance of unexpected interest rate changes. The β is interpreted as the mean reversion parameter-it measures the speed of mean reversion in rate levels; σ^2 as the volatility parameter and λ as the elasticity of volatility with respect to the level of interest rates. At higher λs, the volatility is more sensitive to interest rates. Since the CKLS model is inadequate in explaining the empirical regularity that long term rates are not as volatile as short-term rates PST considered a non-linear framework and obtained results that are consistent with this empirical fact.

This paper shows how a Metropolis-Hastings algorithm with efficient jump can be applied to the estimation of multiple threshold time series models. The algorithm that we suggest is more efficient than existing methods for threshold models. The model is then estimated for the US 3 month Treasury bill rate. Our first result is that the interest rates are persistent in a lower regime and exhibit weak mean reversion in the upper regime. Second, volatility is more persistent in the lower regime than in the upper regime, while the level of volatility is higher when interest rates are high. Third, allowing for both GARCH effect and level effect (dependence of volatility on the level of interest rate) in TAR model results in smaller elasticity λ estimates, significant GARCH effects, and smaller standard errors of threshold parameter estimates. These results have important implications for measuring risk-free rate and for the term structure of interest rates. The empirical analysis to be presented corroborates earlier results obtained in Gray (1996) using Markov-switching model and Phann, Schotman and Tscherig (1996) using TAR model. However, our specification is more general, since we allow multiple regimes and thresholds, ARMA, GARCH, and all parameters are flexible to change with regime. The model we suggest generalizes other models available in the literature so that CKLS, GARCH, ARMA, and multiple threshold models for interest rates are nested in our model as special cases.

For model selection we used several techniques such as marginal likelihood and Bayesian information criteria, as well as estimation with and without truncation restrictions imposed on thresholds. For testing nonlinearity we suggest to use posterior distribution of difference in parameters in different regimes.

The plan of the paper is as follows. Section 2 presents our non-linear framework for multiple regime modelling and Metropolis-Hastings algorithm. Section 3 presents empirical estimates of the model applied to the U.S. T-bill rate. Section 4 concludes.

2 The Threshold Model

Our generalized threshold model nests the threshold autoregressive model (TAR or SETAR), ARMA, GARCH, and CKLS models. Volatility of interest rates is assumed to have both level and GARCH effects, following earlier interest rate models suggested by Brenner (1996) and Koedijk (1997), who found that if GARCH is omitted it results in a biased estimate of level effect. In what follows, we will let y_t represent the level of interest rate at time t and define new parameters, for different regimes, extending equation (1) to the following threshold model.

$$\begin{aligned}
\Delta y_t &= \gamma_1^j + \gamma_2^j y_{t-1} + |y_{t-1}|^{\lambda^j} u_t \qquad (2) \\
u_t &= \frac{\Theta^j(B)}{\Phi^j(B)}\,\epsilon_t \qquad (3) \\
\epsilon_t &\sim N(0, \sigma_t^2) \qquad (4) \\
\sigma_t^2 &= \alpha_0^j + \sum_{i=1}^{a} \alpha_i^j \epsilon_{t-i}^2 + \sum_{i=1}^{s} \beta_i^j \sigma_{t-i}^2 \qquad (5)
\end{aligned}$$

$$\alpha_0^j > 0,\ \alpha_i^j \geq 0,\ (j = 1, \cdots, a),\ \beta_i^j \geq 0,\ (i = 1, \cdots, s)$$

where y_t belongs to regime j given by

$$\begin{cases} j = 1 & y_{t-d} < r_1 \\ j = 2 & r_1 \leq y_{t-d} < r_2 \\ ... & ... \\ j = k+1 & y_{t-d} \geq r_k \end{cases} \qquad (6)$$

We assume in the model above that there are k thresholds $r_1, r_2, ..., r_k$ and $k+1$ regimes for y_t.

$\Theta(B)$ and $\Phi(B)$ are q-th order and p-th order polynomials in the backward shift operator B respectively:

$$\Theta(B) = 1 + \theta_1 B + \cdots + \theta_q B^q \quad \text{and} \quad \Phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p.$$

Each parameter in this model $\{\gamma^j, \phi^j, \theta^j, \alpha^j, \beta^j, \lambda^j\}$ takes $k+1$ values depending on the regime j where y_{t-d} belongs. The delay parameter d will be selected together with the orders (p, q, a, s) of ARMA(p, q)-GARCH(a, s) process in each regime to fit the best model according to a Bayesian information criterion based on estimated marginal likelihood.

Let the prior pdf of the parameters $\gamma^{'}$, $\phi^{'}$, $\theta^{'}$, $\alpha^{'}$ and $\beta^{'}$ be given by

$$\pi(\gamma,\phi,\theta,\alpha,\beta) = \prod_{i=1}^{k+1} \mathrm{N}(\gamma_{0i},\Sigma_{\gamma_i}) \times \mathrm{N}(\phi_{0i},\Sigma_{\phi_i}) \times \mathrm{N}(\theta_{0i},\Sigma_{\theta_i}) \times \mathrm{N}(\alpha_{0i},\Sigma_{\alpha_i})$$
$$\times \mathrm{N}(\beta_{0i},\Sigma_{\beta_i}) \times I\left(0 \leq \lambda^{(i)} \leq \lambda^{up}\right) \times I\left(r_i \in [r^{(i)}_{\text{low}}, r^{(i)}_{\text{up}}]\right)$$

We assume a uniform prior for parameters λ and r with necessary constraints imposed on these parameters: (i) $\lambda \geq 0$ and less than certain upper bound[1]; (ii) each r_i is constrained so that minimum $\delta\%$ of observations are within each regime. We need sufficient sample size in each regime in order to estimate all parameters of the model. We use $\delta = 5\%$ of total number of observations as a minimum sample size in each regime.[2]

Given threshold values $r_1, ..., r_k$ the samples $\{y_t\}$ and $\{x_t\}$ are separated into regimes given by (6). We will denote $y^j_t = y_t$ and $x^j_t = x_t$ if $\{y_t,\ x_t\}$ belong to regime j.

Let us rescale variables y^j_t and x^j_t by the level heteroscedasticity component $|y_{t-1}|^{\lambda^j}$, $(j = 1, ..., k+1)$:

$$\begin{aligned} \widetilde{y}^j_t &= y^j_t / |y_{t-1}|^{\lambda^j} \\ \widetilde{x}^j_t &= x^j_t / |y_{t-1}|^{\lambda^j} \end{aligned} \tag{7}$$

The posterior distribution is given by

$$p(\gamma,\phi,\theta,\alpha,\beta|data) = \pi(\gamma,\phi,\theta,\alpha,\beta) \prod_{j=1}^{k+1} \prod_{t\in T_j} \frac{1}{\sigma_t |y_{t-1}|^{\lambda^j}} \phi\left(\frac{\widetilde{y}^j_t - g(Z_t)}{\sigma_t}\right) \tag{8}$$

[1] which in practice is never greater than 1.5. Some models make stationarity of variance restriction $\lambda < 1$. However, most studies that used CKLS model (1) without GARCH effect found that λ is between 1 and 1.5 for US T-bills

[2] δ depends on number of observations, if the sample size is small higher δ is recommended. For example, Koop and Potter (1999) used 15% as a minimum sample size in each regime.

where for every $t \in T_j = \{t : r_{j-1} \leq y_{t-d} \leq r_j\}$

$$
\begin{aligned}
e_t &= \widetilde{y}_t^j - \widetilde{x}_t^j \gamma^j \\
\epsilon_t &= \widetilde{y}_t^j - g(Z_t) \\
g(Z_t) &= \widetilde{x}_t \gamma^j - \sum_{i=1}^{p} \phi_i^j e_{t-i} - \sum_{i=1}^{q} \theta_i^j \epsilon_{t-i} \\
\\
\sigma_t^2 &= \alpha_0^j + \sum_{i=1}^{r} \alpha_i^j \epsilon_{t-i}^2 + \sum_{i=1}^{s} \beta_i^j \sigma_{t-i}^2
\end{aligned}
$$

Phann, Schotman and Tschernig (1996, PST) considered a two regime threshold autoregressive model with CKLS heteroscedasticity for the U.S. 3 month T-bill rates. Their model is a special case of our model if parameters of GARCH and ARMA are set to zero and level effect λ is not allowed to change with regime. We also allow possibility of more than two regimes in our model. In section 3 we show that a GARCH component is important since as we found the interest rate data shows strong conditional heteroscedasticity. Moreover, the persistence of volatility changes with regime exhibiting higher persistence in a lower regime. We allow more flexibility for the model error (ARMA(p, q)-GARCH(a, s) rather than white noise) where the orders p, q, a, s in each regime[3] are selected using a Bayesian information criterion. The delay parameter d and the threshold variable are also part of the model selection. PST use the threshold variable y_{t-1} with $d = 1$. Tsay (1998) suggests to use a stationary z_{t-d} variable as a threshold. It could be modelled e.g. as Δy_{t-d} as is suggested in Hansen (1997). Another choice of a threshold variable could be a moving average $\Sigma_{i=0}^{d} |y_{t-i}|/(d+1)$ used in Tsay (1998) and in Frances and Van Dyck (2000). We have chosen the threshold variable as y_{t-d}, which makes intuitive sense for interest rate models, as regimes are separated by high and low interest rates. In the

[3]Orders can be different in different regimes, so we have choice of p, q, a, s in each regime separately

next section we explain an efficient Metropolis-Hastings algorithm for the threshold model that works much faster compared with estimation on a grid of points, as is suggested in PST and other literature.

2.1 Estimation of the Threshold Model

The way we estimate threshold parameters $r = (r_1, ..., r_k)$ is different from the existing methods in the literature. We employ an efficient Metropolis jumping rule (see e.g. Gelman, Carlin and Rubin (2004)), which solves the problem of efficient simultaneous estimation of multiple thresholds.

We describe the algorithm for the case of two and three regimes. For higher number of regimes the estimation procedure is similar. We will estimate parameters in blocks: (i) regression parameters, γ; (ii) AR coefficients ϕ, (iii) MA coefficients, θ; (iv) ARCH coefficients, α, (v) GARCH coefficients β, (vi) threshold parameters r, (vii) volatility parameters λ.

We describe the procedure below.

(i) Choose initial values for $\gamma,\ \phi,\ \theta,\ \alpha,\ \beta,\ \ r$, and λ. Start from crude estimates of mean or mode of the posterior distribution. Generally it is hard to find a good approximation for mean or mode of threshold distribution because of unconventional shape of the likelihood function. For starting values we simply divide the sample into regimes with equal samples. In case we have two regimes we use the mean value of y_t as a starting point $r^{(0)}$, if there are three regimes we divide sample into three equal subsamples to find starting values $r_1^{(0)}$ and $r_2^{(0)}$.

Let the started points be denoted by $\gamma^{(0)},\ \phi^{(0)},\ \theta^{(0)},\ \alpha^{(0)},\ \beta^{(0)},\ r^{(0)}$, and $\lambda^{(0)}$. Given initial values $r_1^{(0)},\ r_2^{(0)}$, and d the samples $\{y_t\}$ and $\{x_t\}$ are separated into regimes based on y_{t-d}:

$$\begin{cases} j = 1 & y_{t-d} < r_1 \\ j = 2 & r_1 \le y_{t-d} < r_2 \\ j = 3 & y_{t-d} \ge r_2 \end{cases}$$

(ii) Once the data is separated into regimes given thresholds r and level λ the model is transformed into arranged ARMA-GARCH model,[4] estimation of which using MCMC is standard (see e.g. Nakatsuma (2000) who used independence chains or Goldman and Tsurumi (2005) who used random walk algorithm). We draw parameters of ARMA-GARCH in each regime block by block using random walk Markov Chain, e.g. parameters of a regression block are drawn for all regimes separately (using arranged by regimes data) and then are accepted or rejected jointly in one block. The acceptance rates are controled by multiplying the variance of the proposal density with a scaling constant.

(iii) **Threshold parameters**

We found that standard random walk Metropolis algorithm with constant variance wanders between thresholds and it is very hard to control the acceptance rate; as a result it takes long time to converge. Alternatively using estimation on a grid of points between upper and lower bounds (griddy Gibbs sampling) is also not efficient (in our model with single threshold and 200 grid points MCMC with griddy Gibbs takes 20 times longer than using efficient jump MCMC, explained below). Estimation with number of regimes higher than 2 using griddy Gibbs is impractical.

We suggest to use the following procedure with efficient jump.

Let the superscript, (i), denote the i-th draw. Each threshold parameter $r_j^{(i)}$ $(1 \leq j \leq k)$ can be drawn either in a separate block given other threshold parameters, or all thresholds $\{r^{(i)} = (r_1^{(i)}, ..., r_k^{(i)})\}$ could be drawn in one block, in the latter case acceptance rate will be lower.

[4] We construct arranged ARMA-GARCH model in a similar way as Tsay (1989) and others constructed arranged autoregressive model sorting data y_t by regimes.

We generate

$$r_j^{(i)} \sim N(r_j^{(i-1)}, stdr_j^{(i-1)})$$

where $stdr_j^{(i-1)}$ is initially selected as a constant C_0, such that the proposal normal distribution covers all threshold parameter space (e.g, C_0=half-distance between upper and lower bound for each regime). After sufficient number of draws mmm we set $stdr_j^{(i-1)}$ equal to the standard deviation of the sample of accepted draws $\{r_j^{(l)},\ l = 1, ..i - 1\}$ multiplied by a scaling constant C. The variance of the proposal density is therefore proportional to the variance matrix estimated from the simulation.

$$stdr^{(i)} = C * stdr(\{r^{(l)}\}), \qquad l = n_0, ..i - 1$$

Variance is adjusted using a scaling constant C, so that acceptance rate is reasonable. Gelman et al. (2004) suggest optimal acceptance rate of 44% for 1 parameter and 23% for many parameters.

If $r_j^{(i)}$ does not satisfy the condition

$$r_j^{low} < r_j^{(i)} < r_j^{up}$$

where r_j^{low} and r_j^{up} are defined so that regimes below and above $r_j^{(i)}$ have at least δ % of observations, then generate $r_j^{(i)}$ again until it falls within upper and lower bounds. We set $\delta = 5\%$.

We accept $r^{(i)} = (r_1^{(i)}, ..., r_k^{(i)})\}$ with probability

$$a_6 = \min\left\{\frac{p(\gamma^{(i)}, \phi^{(i)}, \theta^{(i)}, \alpha^{(i)}, \beta^{(i)}, r^{(i)}, \lambda^{(i-1)}|\text{data})}{p(\gamma^{(i)}, \phi^{(i)}, \theta^{(i)}, \alpha^{(i)}, \beta^{(i-1)}, r^{(i-1)}, \lambda^{(i-1)}|\text{data})},\ 1\right\}.$$

Otherwise set $r^{(i)} = r^{(i-1)}$.

Alternatively, one can construct a separate block and acceptance rate for each threshold.

(iv) **Elasticity λ blocks**

Draws of $\lambda_j^{(i)}$ are done similarly to draws of $r_j^{(i)}$ using efficient jump with the exception that we set the constraint: $\lambda_j^{(i)} > 0$. In this block we use the same procedure as Qian, Ashizawa and Tsurumi (2005) for estimation of CKLS parameter λ in a linear model (1).[5]

Generate

$$\lambda_j^{(i)} \sim N(\lambda_j^{(i-1)}, stdl_j^{(i-1)})$$

where $stdl_j^{(i-1)}$ is initially selected as a constant and after sufficient number of draws mmm we set $stdl_j^{(i-1)}$ proportional to the standard deviation of the sample of accepted draws $\{\lambda_j^{(l)},\ l = 1, ..i-1\}$.

If $\lambda_j^{(i)} > 0$ we accept $\lambda_j^{(i)}$ with probability

$$a_7 = \min\left\{\frac{p(\gamma^{(i)}, \phi^{(i)}, \theta^{(i)}, \alpha^{(i)}, \beta^{(i)}, r^{(i)}, \lambda_j^{(i)}, \lambda_{-j}^{(i)}|\text{data})}{p(\gamma^{(i)}, \phi^{(i)}, \theta^{(i)}, \alpha^{(i)}, \beta^{(i-1)}, r^{(i)}, \lambda^{(i-1)}, \lambda_{-j}^{(i)}|\text{data})},\ 1\right\}.$$

Otherwise set $\lambda_j^{(i)} = \lambda_j^{(i-1)}$. given previous draws of other parameters and λ's in other regimes $(\lambda_{-j}^{(i)})$. After all λ's are drawn we set $\lambda^{(i)} = (\lambda_1^{(i)}, ..., \lambda_3^{(i)})$.

Alternatively, all λ_j can be drawn in one block.

As with any standard MCMC procedure, we make N draws of the parameters in each of the blocks, and we burn the first m draws. Out of the remaining $N - m$ draws, we keep every h-th draw. We check convergence by testing that the draws attain mean and covariance stationarity.

2.2 Choice of the model

Estimation of a threshold model involves choice of (i) number of regimes, (ii) the threshold variable (e.g. y_{t-d}, $\Delta(y_{t-d})$, or $\Sigma_{i=0}^{d}|y_{t-i}|/(d+1)$) and the

[5] The focus of this paper is not on CKLS model, but we use it as one of the features of interest rate models. The main contribution of this paper is efficient jump for estimation of threshold parameters that eliminates problems associated with other algorithms.

delay parameter d, (iii) orders of ARMA(p, q)-GARCH(r, s) process in each regime.

We use several criteria, like significance of coefficients, marginal likelihood, and a modified Bayesian information criterion (MBIC) discussed in Goldman and Tsurumi (2005). This criterion is a Bayesian analogue of Akaike information criterion given by

$$AIC(p, q, r, s, d, n_{regimes}) = -2\Sigma_{j=1}^{n_{regimes}} ln(L_j(p_j, q_j, r_j, s_j, d)) + 2(\nu + 1)$$

where $ln(L_j())$ is a log-likelihood function for regime j and ν are degrees of freedom. For the case of $n_{regimes} = 2 :\ \ \nu = 2 * k + p_1 + p_2 + q_1 + q_2 + a_1 + a_2 + s_1 + s_2 + 1$, where k is the dimension of x_t and is assumed to be the same in two regimes, $p_1, p_2, q_1, q_2, a_1, a_2, s_1, s_2$ are orders of ARMA-GARCH parameters in regimes 1 and 2 correspondingly and 1 degree of freedom is given to the choice of delay parameter d. In a similar way we can find ν for higher number of regimes.

The modified bayesian information criterion is given by:

$$MBIC = -2 \ln\, m(x) + 2(\nu + 1)$$

where the marginal likelihood $m(x)$ is computed by the Laplace-Metropolis estimation and evaluated at either posterior mean or mode.[6]

For the choice of number of regimes we find the smallest MBIC as well as perform sensitivity analysis where estimation of thresholds r_j is done without restriction $r^{low\,(j)} < r_j < r^{up\,(j)}$. We look at the sensitivity of posterior densities of r_j to imposing δ % restriction. If r_j is closer than δ % of observations toone of the neighbouring thresholds, or either upper or lower bound for the sample $\{y_t\}$ we suspect that number of regimes is less than was assumed.

[6]Alternatively one can use Chib and Jeliazkov (2001) estimator of marginal likelihood.

To test whether the dynamics changes with regime we look at posterior distributions of differences in parameters of interest in upper and lower regimes. If there is considerable difference in some parameter's distributions it supports the hypothesis of non-linearity of series y_t.

Finally, the formal test is based on minimizing the modified Bayesian information criterion when models with 1 regime, 2 regimes, and 3 regimes are compared.

3 Empirical estimation of interest rate dynamics

In this section we present the results of estimation of the generalized threshold model (2)-(5) for the U.S. short-term interest rates. The 3 month T-bill monthly data for the period 1962-2005 was downloaded from the Federal Reserve Board of Governor's website.[7] As was mentioned we have chosen the threshold variable as y_{t-d}, which makes sense for interest rate models, as regimes are separated by high and low interest rates. Using model selection criteria identified in the previous section we found that the best model has 2 regimes compared to 3 regimes or 1 regime. We estimated model with and without restrictions of minimum $\delta = 5\%$ of observations in each regime and results were identical for the two-regime model. The best threshold variable turned out to be y_{t-1} with delay parameter $d = 1$.

All results of estimation are given in Table 1 and Figures 1-4. The data are presented in Figure 1(a). We can identify the period of 1979-1982 which is characterized by high levels of interest rates and high volatility.[8] The posterior distribution of the estimated threshold is given in Figure 1(b) and the mean value of estimated threshold distribution (10%) is shown on the

[7] http://federalreserve.gov/releases/h15/data.htm

[8] Many authors included the period of FED's experiment 1979-1982 as part of the sample. For example, PST considered the period 1962-1990.

graph with data separating y_t into two regimes; the upper regime coincides with historical period of change in monetary policy.

Table 1: Threshold ARMA model with CKLS-GARCH volatility

	Regime 1: $(y_{t-1} < r)$			Regime 2: $(y_{t-1} \geq r)$		
	mean	(std)	Corr	mean	(std)	Corr
γ_1	0.047	(0.030)	0.447	2.405	(1.545)	0.491
γ_2	-0.006	(0.007)	0.442	-0.208	(0.131)	0.507
ϕ_1	0.507	(0.138)	0.918	-0.123	(0.432)	0.886
θ_1	-0.129	(0.166)	0.927	0.740	(0.403)	0.876
α_0	0.001	(0.000)	0.771	0.029	(0.024)	0.777
α_1	0.142	(0.030)	0.796	0.328	(0.215)	0.657
β_1	0.835	(0.030)	0.883	0.329	(0.229)	0.826
λ	0.391	(0.130)	0.799	0.638	(0.137)	0.791
ρ=max AR root	0.507	(0.138)	0.918	0.396	(0.212)	0.665
r	10.034	(0.246)	0.315			

Notes: (1)ARMA(1,1)–GARCH(1,1) model for each regime was selected based on MBIC, significance of parameters and convergence patterns.
(2)Estimated model has n1=483 (93%) observations in the lower regime and n2=34 (7%) observations in the upper regime.
(3)The figures in parentheses are posterior standard deviations.
(4) Corr is the first order autocorrelation of the MCMC draws.
(5)ρ is the maximum absolute value of the inverse roots of the AR parameters

Some of the results given in Table 1 are similar to results obtained in PST for the period 1962-1990 (p. 161 Table 2), while certain results are different as we describe below.[9] One of the similar results is that the interest rate follows a unit root process in the lower regime ($y_{t-1} < r$) and the process has slow mean reversion once the interest rate is above the threshold. The same results was obtained earlier by Gray (1996) who used a Markov-switching

[9]We also estimated our model with the same sample as PST (1962-1990) and results are very similar to the full sample 1962-2005 results presented in the text.

GARCH model. The pdfs of persistence parameter γ_2 in two regimes are presented in Figure 2(a) and the posterior distribution of the difference in persistence for two regimes ($\gamma_2^{(1)} - \gamma_2^{(2)}$) is given in Figure 2(b). We see that in the lower regime (regime 1) the pdf is tightly distributed around zero. In the upper regime (regime 2), the distribution is centered around -0.21 and is more disperse. The posterior distribution of the difference shows that 85% highest posterior density interval (HPDI) corresponds to positive values of differences in γ_2, i.e. persistence is higher in lower regime. Using any reasonable significance level, e.g. the 95% highest posterior density interval, we clearly do not reject the null hypothesis of a unit root for γ_2 in the lower regime. Although for the upper regime using the 95% HPDI we would also fail to reject unit root (95% HPDI is (-0.468, 0.060)), the variance of this distribution is high and the upper limit of the 95% HPDI is very close to zero. Using, for example 87% HPDI (-0.402, -0.002) we would reject the unit root hypothesis for the upper regime.

Let us test stationarity of interest rates. In each regime the null hypothesis of a unit root is given by either:

$$H_0 : \gamma_2 = 0 \text{ versus } H_1 : \gamma_2 < 0 \tag{9}$$

or

$$H_0 : \rho = 1 \text{ versus } H_1 : \rho < 1 \tag{10}$$

where ρ is the maximum absolute value of the inverse roots of the AR parameters in the error term u_t. Rejection of both null hypotheses above implies stationarity. However, if we find a unit root in the lower regime, but mean reversion in the upper regime, it would still imply stationarity since interest rates exhibit mean reversion once they are higher than the threshold.[10]

[10] For a formal definition of stationarity of a two-regime TAR model see, e.g. Frances and Van Dyck (2000). They explain that a unit root in lower regimes and stationarity in

Table 1 shows that the max absolute value of inverted roots of AR parameters (ρ) in the error term is less than 1, so the error term is stationary in both regimes. Overall, we conclude that interest rates are weakly stationary, since the unit root hypothesis in the upper regime was rejected using 87% HPDI, but was not rejected with 95% HPDI.

The posterior pdf of the threshold parameter presented in Figure 1 (b) with mean value $r = 10.03$ (see Table 1) is close to PST but the standard error is significantly smaller. Since we included a GARCH component (which turns out to be significant in the model) we increased the precision of the threshold estimator.

Figure 4(a) shows the posterior pdf of a GARCH parameter ab, which measures the persistence of volatility in each regime:

$$ab = \Sigma_{i=1}^{l}(\alpha_i + \beta_i) \tag{11}$$

where $l = max(a, s)$, $\alpha_i = 0$ for $i \geq a$ and $\beta_i = 0$ for $i \geq s$. The posterior pdf for ab can be used to test a null hypotheses whether a GARCH component is present in the data. The null and alternative hypotheses are given by: $\mathrm{H}_0 : ab = 0$ (*i.e.* the error process does not have GARCH) versus $\mathrm{H}_1 : ab > 0$ (*i.e.* the error process has GARCH). In Figure 4(a) we show the pdfs for ab in each regime and find ab significant in both regimes. The persistence in the lower regime is close to one, while in the upper regime it is less than one (using any HPDI). We also show the posterior distribution of the difference in the persistence in volatility in lower and upper regimes and find lower persistence in the upper regime using 85% HPDI. If we look at a constant term in the GARCH model we find that when interest rates fall they become less volatile (if we compare magnitudes of α_0 in regimes

the upper regime implies stationarity of non-linear time series.

1 and 2 from Table 1), however at lower interest rates volatility is more persistent. The property that level of volatility is proportional to interest rates is well-known, however change in persistence is an interesting result. We may conclude that overall volatility process is stationary, since it is mean-reverting when interest rates are high and volatility is high.

Finally, we find two similar distributions for level heteroscedasticity parameter λ in two regimes (presented in Figure 3). HPDI for both distributions include $\lambda = 0.5$. Our estimates of λ are smaller than estimates given in CKLS and PST and are consistent with results of other studies that omitting the GARCH component results in overestimating the level effect in volatility of interest rates.

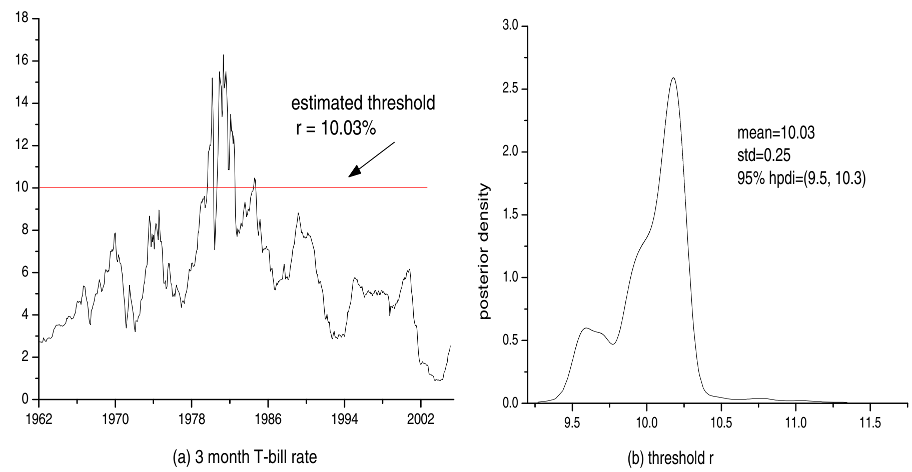

Figure 1: (a) U.S. three-month T-bill rate, monthly data, 1962-2005; (b) Posterior pdf of the threshold parameter r

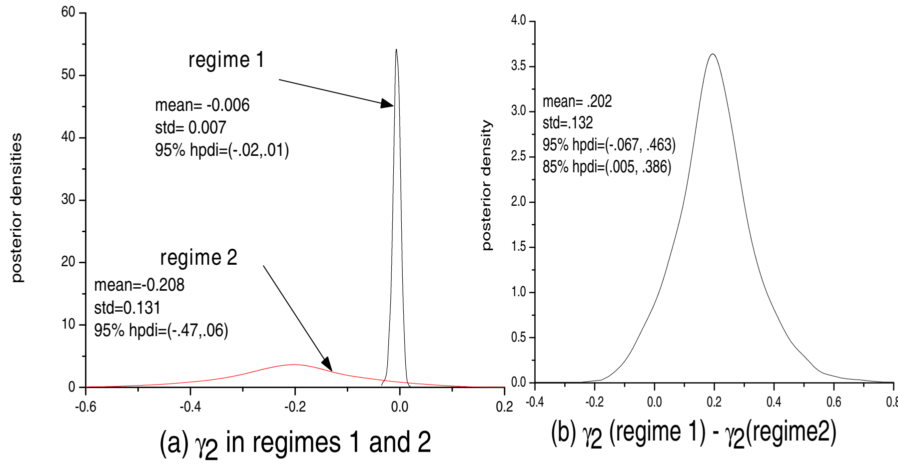

Figure 2: (a) Posterior pdfs of persistence parameter γ_2 in regimes 1 and 2; (b) Posterior pdf of the difference in persistence parameters.

4 Conclusion

We applied MCMC algorithm with efficient jump for a general class of threshold time series models with multiple regimes that works much faster than estimation on a grid of points and is simpler than numerical integration techniques suggested in literature.

We used as example U.S. T-bills data and found that the data can be characterized by two regimes; the interest rate data process follows a unit root in a lower regime and a weak mean reversion process above a certain threshold; the volatility is more persistent in lower regime; the modification to allow for GARCH error results in more efficient threshold parameter estimates and lower level effects in volatility.

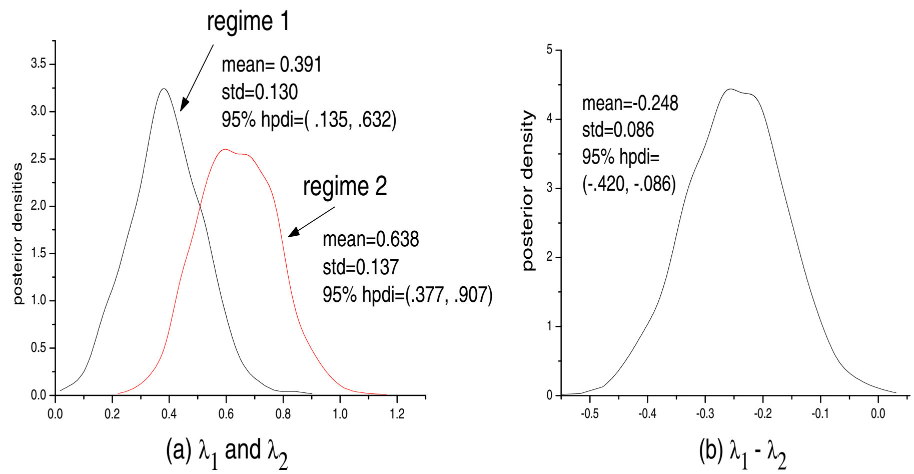

Figure 3: (a) Posterior pdfs of elasticity of volatility λ in regimes 1 and 2; (b) Posterior pdf of the difference in elasticity of volatility parameters.

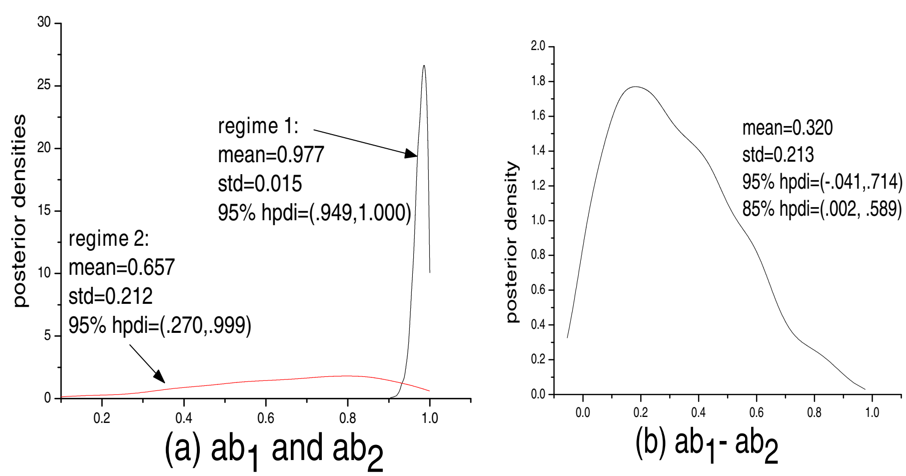

Figure 4: (a) Posterior pdfs of persistence in volatility ab in regimes 1 and 2; (b) Posterior pdf of the difference in persistence of volatility parameters.

References

Ball, C.A and W.N. Torous, "Unit roots and estimation of interest rate dynamics," *Journal of Empirical Finance*, 3, (1996), pp. 215-238.

Ball, C.A and W.N. Torous, "The stochastic volatility of short-term interest rates: some international evidence," *Journal of Finance*, 54, (1999), pp. 2339-2359.

Bliss, R.J., and D.C. Smith, "The Elasticity of Interest rate Volatility: Chan, Karolyi, Longstaff and Sanders Revisited", Federal Reserve Bank of Atlanta, Working Paper 97:13a (1998)

Bollerslev, T., 1986, Generalized autoregressive conditional heteroscedasticity, Journal of Econometrics, 31, 307–327.

Brennan, M. J. and E.S. Schwartz, "Analyzing Convertible Bonds," *Journal of Financial and Quantitative Analysis*, 15, (1980), pp. 907-929.

Brenner, R.J., R. H. Harjes and K.F. Kroner, "Another Look at Models of the Short-Term Interest Rates," *Journal of Financial and Quantitative Analysis*, 31:1, (1996), pp. 85-107.

Broze, L. O. Scaillet and J. Zakoian, "Testing for Continuous-Time Models of Short-term Interest Rate," *Journal of Empirical Finance*, 2, (1995), pp. 199-223.

Broze, L. O. Scaillet and J. Zakoian, "Quasi-Indirect Inference for Diffusion Processes," *Econometric Theory*, 14:2, (1998), pp. 161-186.

Chan, K.C, G. A. Karolyi, F.A. Longstaff and A.B. Sanders, "An Empirical Comparison of Alternative Models of the Short-Term Interest Rate," *Journal of Finance*, 48, (1992) pp. 1209-1227.

Chen, C.W.S. and Lee,J.C., 1995, Bayesian inference of threshold autoregressive models, Journal of Time Series Analysis, 16,483-492.

Chib, S., 1995, Marginal likelihood from the Gibbs output, Journal of the American Statistical Association, 90, 1313–1321.

Chib, S. and I. Jeliazkov (2001) "Marginal Likelihood from the Metropolis-Hastings Output," *Journal of the American Statistical Association*, 96, 270–281.

Cowels, M.K. and B.P. Carlin, 1996, Markov Chain Monte Carlo convergence diagnostics: a comparative review, Journal of the American Statistical Association, 91, 883-904.

Cox, J., J. Ingersoll, and S. Ross, "An Analysis of Variable Rate Loan Contracts," *Journal of Finance*, 35, (1980) pp. 389-403.

Cox, J., J. Ingersoll, and S. Ross, "A Theory of the Term Structure of Interest Rates," *Econometrica*, 53, (1985) pp. 385-407.

Dahlquist, M. and S.F. Gray, "Regime-switching and interest rates in the European Monetary system," *Journal of International Economics*, 50, (2000), pp. 399-419.

Geweke, J. and Terui, N., 1993, Bayesian threshold autoregressive models for nonlinear time series, Journal of Time Series Analysis, 14, 441-454.

Goldman, E., and H. Tsurumi, 2005, Bayesian Analysis of a Doubly-Truncated Regression Model with ARMA-GARCH Error, *Studies in Nonlinear Dynamics & Econometrics*, forthcoming.

Gray, S.F., "Modelling the conditional distribution of interest rates as a regime-switching process," *Journal of Financial Economics*, 42, (1996), pp. 27-62.

Heidelberger, P. and Welch,P., 1983, Simulation run length control in the presence of an initial transient, Operations Research, 31, 1109-1144.

Koop, G., and S.M. Potter, 1999, Dynamic Asymmetries in U.S. Unemployment, Journal of Business and Economic Statistics, 17, 298-312.

Lanne and Saikkonen, 2002, Threshold autoregressions for Strongly Autocorrelated Time Series, Journal of Business and Economic Statistics, 20, 282-289.

Nabeya, S. and K. Tanaka, 1988, Asymptotic theory of a test of constancy of regression coefficients against the random walk alternative, Annals of Statistics, 218–235.

Nakatsuma, T., 2000, Bayesian analysis of ARMA-GARCH models: A Markov chain sampling approach, Journal of Econometrics, 95, 57–69.

Nowman, K.B., "Gaussian Estimation of Single-Factor Continuous-Time Models of The Term Structure of Interest Rates," *Journal of Finance*, 52, (1997), pp. 1695-1706.

Phann, G.A., Schotman, P.C., Tschering, R., 1996, Non-linear interest rate dynamics and implications for the term structure, Journal of Econometrics, 74, 149-176.

Ploberger, W., Kr'́amer, W. and K.Kontrus, 1989, A new test for structural stability in the linear regression model, Journal of Econometrics, 40, 307-318.

Qian, X., Ashizawa, R., and Tsurumi, H., 2005, Estimation of short term interest rate models: comparison of Bayesian and non-Bayesian estimators, forthcoming in the Communications in Statistics.

Tong, H., 1990, Nonlinear time series: a dynamic system approach, Oxford, U.K.: Oxford University Press.

Tsay, R.S., 1989, Testing and modelling threshold autoregressive processes, Journal of the American Statistical Association, 84,231-240.

Tsay, R.S., 1998, Testing and modelling multivariate threshold models, Journal of the American Statistical Association, 93, 1188-1202.

Tsurumi, H., 2000a, Bayesian statistical computations of nonlinear financial time series models: a survey with illustrations, Asia-Pacific Financial Markets, 1-29.

Vasicek, O., "An Equilibrium Characterization of The Term Structure," *Journal of Financial Economics*, 5, (1977), pp. 177-188.

Yu, J. and Phillips, P.C.B., "A Gaussian approach for continuous-time models of the short-term interest rate," *Econometrics Journal*, 4, (2001), pp. 211-225.