Discussion Paper Series

No.43

An Extension of the Markov-Switching Model with Time-Varying Transition Probabilities: Bull-Bear Analysis of the Japanese Stock Market

Akifumi Isogai, Satoru Kanoh, and Toshifumi Tokunaga

November 2004

Hitotsubashi University Research Unit for Statistical Analysis in Social Sciences

A 21st-Century COE Program

Institute of Economic Research
Hitotsubashi University
Kunitachi, Tokyo, 186-8603 Japan
http://hi-stat.ier.hit-u.ac.jp/

An Extension of the Markov-Switching Model with Time-Varying Transition Probabilities: Bull-Bear Analysis of the Japanese Stock Market

Akifumi Isogai
MTB Investment Technology Institute Co., Ltd.,
2-5-6 Shiba, Minato-ku, Tokyo, 105-0014, Japan
(isogai@mtec-institute.co.jp)
Satoru Kanoh
Institute of Economic Research, Hitotsubashi University,
2-1 Naka, Kunitachi City, Tokyo 186-8603, Japan
(kanoh@ier.hit-u.ac.jp)
Toshifumi Tokunaga
Department of Business Administration, Nanzan University,
18 Yamazato-cho, Shouwa-ku, Nagoya 466-8673, Japan
(tokunaga@nanzan-u.ac.jp)

first draft: November 1, 2004

Abstract

This paper attempts to extend the Markov-switching model with time-varying transition probabilities (TVTP). The transition probabilities in the conventional TVTP model are functions of exogenous variables that are time-dependent but with constant coefficients. In this paper the coefficient parameters that express the sensitivities of the exogenous variables are also allowed to vary with time. Using data on Japanese monthly stock returns, it is shown that the explanatory power of the extended model is superior to conventional models.

KEY WORDS: Gibbs sampling, Kalman filter, Marginal likelihood, Market dynamics, Time-varying sensitivity.

1. Introduction

The classification of market conditions - for example, in the stock market, but also in foreign exchange markets or the economy overall, i.e. the business cycle - into bull and bear markets has been broadly accepted in both the academic and the practical world. The term "bull market" refers to a medium- and long-term upward trend in the market and "bear market" refers to a downward trend. Despite the apparent simplicity of these terms, their precise definitions in the analysis of the stock market are far from clear. One way to define the terms is by using stock prices levels. Lunde and Timmermann (2004), for example, define bull and bear markets as situations when a stock price index exceeds locally defined thresholds. On the other hand, Biscarri and Gracia (2002) search the turning points of bull and bear markets by checking if the current stock price index is at a local maximum (minimum) within a certain time span. Both these definitions of bull and bear markets are computationally simple; however, the determination of the thresholds in the former and the sample span in the latter inevitably will be arbitrary.

A more rigorous method to determine bull and bear markets is to use the Markov-switching model (MSM) proposed by Hamilton (1989). The stock price index is considered to result from one of two unobservable regimes - the bull market and the bear market - and those regimes are estimated using this model. The switching mechanism is expressed by a Markov model with constant transition probabilities. Schaller and van Norden (1997) and Maheu and McCurdy (2000) apply the model to analyze bull-bear conditions in the U.S. stock market. The Markov-switching model has also been used to identify periods of expansion and contraction in the business cycle.

The basic MSM assumes that transition probabilities are constant, which in practice is rarely the case. Diebold, Lee, and Weinbach (1994) and Filardo (1994) therefore developed an MSM with time-varying transition probabilities (TVTP). The transition probabilities are expressed as functions of the exogenous variables, and the probabilities vary over time through the time dependency of the exogenous variables. Applying the model to exchange rates, Diebold, Lee, and Weinbach (1994), for example, argue that the likelihood of a revaluation can be expected to depend on economic fundamentals. Filardo (1994), meanwhile, examining business cycles, considers time-varying transition probabilities that are functions of leading economic indicators. Since the publication of these studies, the approach has been used in a large number of studies on the business cycle and exchange rate movements, but only a few have applied it to the analysis of stock prices. The paper by Schaller and van Norden (1997) is the first application of MSM-TVTP to the U.S. stock market and Maheu and McCurdy (2000) use the duration of the market condition as an exogenous variable.

This paper has two objectives. The first is to extend the TVTP model (MSM-TVTP with "time-varying sensitivities") to allow the sensitivities of the exogenous variables determining the transition probabilities to vary over time; we do this by building on the time-varying logit model developed by Kanoh and Li (1990). The second objective of the paper is to apply the model to the Japanese stock market and analyze the market dynamics. The model is expected to detect changes of the sensitivities to the exogenous variables over time. Recently, Lunde and Timmermann (2004) proposed a model with time-varying sensitivities in which bull and bear hazard rates depend on the duration of the market condition and the interest rate under the assumption that the states are observable. The TVTP with time-varying sensitivities in this paper is constructed without the assumption of observable bull-bear conditions and those conditions are to be identified within the model.

Some of the other characteristics of the model are as follows. First, the proposed model nests several MSMs, in particular, the MSM with constant transition probabilities by Hamilton (1989), and the MSM-TVTP by Diebold, Lee, and Weinbach (1994) and Filardo (1994). This enables us to evaluate the importance of the extension of the model step-by-step. Second, though the time-varying sensitivity is described by a state space model, it is not possible to employ the maximum likelihood estimation using the Kalman filtering algorithm because of the nonlinearity of the model. The estimation and evaluation of the results are conducted using Gibbs sampling, while the comparisons of the models are made using marginal likelihood following Chib (1995).

The remainder of this paper is organized as follows. In Section 2, the basic framework of the models is presented; moreover, the estimation procedure and the method of comparing models are summarized. Section 3 describes the Japanese stock market data and other market-related data used as exogenous variables in the estimation. Section 4 reports the estimation results of the MSM with constant transition probabilities, the MSM-TVTP by Diebold, Lee, and Weinbach (1994) and Filardo (1994), and the MSM-TVTP with time-varying sensitivities. Section 5 concludes.

2. Basic framework of the MSMs

The section starts with the description of the basic framework of the models. Consider the following model for a stock return, R_t.

$$(1-\phi(L))(R_t-\mu_{S_t})=\epsilon_t,\quad \epsilon_t\sim N(0,\sigma^2_{S_t}),\tag{1}$$

$$\mu_{S_t}=(1-S_t)\mu_0+S_t\mu_1,\quad \mu_0>\mu_1,\quad S_t=\{0,1\},\tag{2}$$

$$\sigma^2_{S_t}=(1-S_t)\sigma^2_0+S_t\sigma^2_1,\quad \sigma^2_1=(1+h)\sigma^2_0,\tag{3}$$

where $\phi(L)=\phi_1L+\phi_2L^2+\cdots+\phi_pL^p$ and L is a lag operator. The state or regime S_t is an unobserved random variable that takes zero or one. The evolution of S_t is governed by a Markov process with the following probabilities:

$$\begin{cases} p_t^{(11)} \equiv \Pr[S_t = 1|S_{t-1} = 1, Z_{t-1}] = \Pr[S_t^* \geq 0|S_{t-1} = 1, Z_{t-1}], \\ p_t^{(00)} \equiv \Pr[S_t = 0|S_{t-1} = 0, Z_{t-1}] = \Pr[S_t^* < 0|S_{t-1} = 0, Z_{t-1}], \end{cases} \tag{4}$$

$$S_t^* = f(S_{t-1}, Z_{t-1}; \theta) + u_t; \quad u_t \sim N(0,1), \tag{5}$$

where S_t^* is a latent variable, Z_t is a vector of exogenous variables, and θ is a parameter vector specifying the function f. Note that $p_t^{(01)} \equiv \Pr[S_t = 0|S_{t-1} = 1, Z_{t-1}] = 1 - p_t^{(11)}$, $p_t^{(10)} \equiv \Pr[S_t = 1|S_{t-1} = 0, Z_{t-1}] = 1 - p_t^{(00)}$.

Let Θ be the parameter vector to be estimated:

$$\Theta = \{\underbrace{\mu_0, \mu_1}_{\mu}, \underbrace{\phi_1, ..., \phi_p}_{\Phi}, \underbrace{\sigma_0^2, \sigma_1^2}_{\Sigma}, \underbrace{S_1, ..., S_T}_{\tilde{S}_T}, \underbrace{S_1^*, ..., S_T^*}_{\tilde{S}_T^*}, \theta\}. \tag{6}$$

In order to estimate Θ, the Gibbs sampling technique is employed and the following steps from (a) to (f) are iterated. $(k = 1, 2, ..., M, M+1, ..., M+N)$

(a) Generate $\mu^{(k+1)}$ from $g\left(\mu | \Phi^{(k)}, \Sigma^{(k)}, \tilde{S}_T^{(k)}, \tilde{S}_T^{*(k)}, \theta^{(k)}\right)$,

(b) Generate $\Phi^{(k+1)}$ from $g\left(\Phi | \mu^{(k+1)}, \Sigma^{(k)}, \tilde{S}_T^{(k)}, \tilde{S}_T^{*(k)}, \theta^{(k)}\right)$,

(c) Generate $\Sigma^{(k+1)}$ from $g\left(\Sigma | \mu^{(k+1)}, \Phi^{(k+1)}, \tilde{S}_T^{(k)}, \tilde{S}_T^{*(k)}, \theta^{(k)}\right)$,

(d) Generate $\tilde{S}_T^{(k+1)}$ from $g\left(\tilde{S}_T \middle| \mu^{(k+1)}, \Phi^{(k+1)}, \Sigma_T^{(k+1)}, \tilde{S}_T^{*(k)}, \theta^{(k)}\right)$,

(e) Generate $\tilde{S}_T^{*(k+1)}$ from $g\left(\tilde{S}_T^* \middle| \mu^{(k+1)}, \Phi^{(k+1)}, \Sigma_T^{(k+1)}, \tilde{S}_T^{(k+1)}, \theta^{(k)}\right)$,

(f) Generate $\theta^{(k+1)}$ from $g\left(\theta | \mu^{(k+1)}, \Phi^{(k+1)}, \Sigma_T^{(k+1)}, \tilde{S}_T^{(k+1)}, \tilde{S}_T^{*(k+1)}\right)$.

where g denotes an appropriate density function whose specification is explained below.

At the end of the $(M+N)$ th iteration, the first M results are dropped and the average values of the remaining N results are calculated as the estimates.

$$E[\Theta] = \frac{1}{N} \sum_{k=M+1}^{M+N} \Theta^{(k)}. \tag{7}$$

The generation of the estimates in steps (a) to (e) is summarized below. In step (f), the generation of θ depends on the formulation of f and, therefore, the details will be explained in the next section.

First, in step (a), $\mu' = [\mu_0 \ \mu_1]$ is obtained from the posterior distribution (9) using the prior distribution (8).

prior distribution: $\quad \mu|\Sigma \sim N(m_{0_\mu}, \Sigma_{0_\mu})_{I[\mu_0 > \mu_1]}, \quad (8)$

posterior distribution: $\quad \mu|\Sigma \sim N(m_{1_\mu}, \Sigma_{1_\mu})_{I[\mu_0 > \mu_1]}, \quad (9)$

where

$$
\begin{aligned}
m_{1_\mu} &= (\Sigma_{0_\mu}^{-1} + \Sigma^{-1} A'_\mu A_\mu)^{-1} (\Sigma_{0_\mu}^{-1} m_{0_\mu} + \Sigma^{-1} A'_\mu Y_\mu), \\
\Sigma_{1_\mu} &= (\Sigma_{0_\mu}^{-1} + \Sigma^{-1} A'_\mu A_\mu)^{-1}, \\
A'_\mu &= \begin{bmatrix} 1-\phi(L) & \cdots & 1-\phi(L) \\ (1-\phi(L))S_{p+1} & \cdots & (1-\phi(L))S_T \end{bmatrix}, \\
Y'_\mu &= [(1-\phi(L))R_{p+1} \;\cdots\; (1-\phi(L))R_T],
\end{aligned}
$$

and $I[\mu_0 > \mu_1]$ is an indicator function used to denote that the constraint $\mu_0 > \mu_1$ is satisfied.

Then, in step (b), $\Phi' = [\phi_1 \;\cdots\; \phi_p]$ is generated from the posterior distribution (11) using the prior distribution (10).

prior distribution: $$\Phi|\mu \sim N(m_{0_\phi}, \Sigma_{0_\phi})_{I_{S(\phi)}}, \tag{10}$$

posterior distribution: $$\Phi|\mu \sim N(m_{1_\phi}, \Sigma_{1_\phi})_{I_{S(\phi)}}, \tag{11}$$

where

$$
\begin{aligned}
m_{1_\phi} &= (\Sigma_{0_\phi}^{-1} + A'_\phi A_\phi)^{-1} (\Sigma_{0_\phi}^{-1} m_{0_\phi} + A'_\phi Y_\phi), \\
\Sigma_{1_\phi} &= (\Sigma_{0_\phi}^{-1} + A'_\phi A_\phi)^{-1}, \\
A'_\phi &= \begin{bmatrix} R_p^* & \cdots & R_{T-2}^* \\ \vdots & & \vdots \\ R_1^* & \cdots & R_{T-p}^* \end{bmatrix}, \\
Y'_\phi &= [R_{p+1}^* \;\cdots\; R_T^*],\; R_t^* = (R_t - \mu_{S_t})/\sigma_{S_t},
\end{aligned}
$$

and I_{S_ϕ} is an indicator function used to denote that the roots of $\phi(L) = 0$ lie outside the unit circle.

In step (c), σ_0^2 and $h^*(= 1 + h)$ are generated from the posterior distributions (14) and (15) using the prior densities (12) and (13), respectively.

prior distribution: $$\sigma_0^2|\mu, \Phi, h \sim IG\left(\frac{\nu_0}{2}, \frac{\delta_0}{2}\right), \tag{12}$$

$$h^*|\mu, \Phi \sim IG\left(\frac{\nu_{h0}}{2}, \frac{\delta_{h0}}{2}\right), \tag{13}$$

posterior distribution: $$\sigma_0^2|\mu, \Phi, h \sim IG\left(\frac{\nu_1}{2}, \frac{\delta_1}{2}\right), \tag{14}$$

$$h^*|\mu, \Phi \sim IG\left(\frac{\nu_{h1}}{2}, \frac{\delta_{h1}}{2}\right), \tag{15}$$

where

$$
\begin{aligned}
&\nu_1 = \nu_0 + (T - p), \qquad \delta_1 = \delta_0 + Y'_0 Y_0, \\
&\nu_{h1} = \nu_{h0} + T_h, \qquad \delta_{h1} = \delta_{h0} + Y'_1 Y_1, \\
&Y'_0 = [\epsilon_{p+1}^* \;\cdots\; \epsilon_T^*],\; \epsilon_t^* = (1-\phi(L))(R_t - \mu_{S_t})/(1 + S_t h), \\
&T_h = \text{count}(Y_1),\; Y'_1 = [\epsilon_{p+1}^{**} \;\cdots\; \epsilon_T^{**}], \\
&\epsilon_t^{**} = \begin{cases} (1-\phi(L))(R_t - \mu_{S_t})/\sigma_0, & \text{if } S_t = 1, \\ \textit{No sampling}, & \text{if } S_t = 0. \end{cases}
\end{aligned}
$$

T_h is the number of non-zero elements of Y_1.

Step (d) follows the method developed by Albert and Chib (1993).

$$P(S_t|\tilde{R}_T, S_{-t}) \propto \begin{cases} P(S_t|S_{t-1})P(S_{t+1}|S_t)P(R_t, ..., R_p|\tilde{S}_p) \prod_{k=p+1}^{t+p} f(R_k|\tilde{R}_{k-1}, \tilde{S}_k), \text{ if } t \leq p, \\ P(S_t|S_{t-1})P(S_{t+1}|S_t)\prod_{k=t}^{t+p} f(R_k|\tilde{R}_{k-1}, \tilde{S}_k), \text{ if } p+1 \leq t \leq T-p+1, \\ P(S_t|S_{t-1})P(S_{t+1}|S_t)\prod_{k=t}^{T} f(R_k|\tilde{R}_{k-1}, \tilde{S}_k), \text{ if } T-p \leq t \leq T, \end{cases} \quad (16)$$

where $\tilde{R}_t = \{R_1, ..., R_t\}$, $\tilde{S}_t = \{S_1, ..., S_t\}$, and $S_{-t} = \{S_1, ..., S_{t-1}, S_{t+1}, ..., S_T\}$.

In step (e), for a given formulation of f, S_t^* is obtained by sampling u_t from a truncated normal distribution in such a way that S_t^* satisfies the inequalities given in (4).

The comparison of the different models $(\mathcal{M}_1, ..., \mathcal{M}_I)$ is conducted based on marginal likelihood following Chib (1995), i.e.,

$$\ln \hat{m}(Y_T|\mathcal{M}_i) = \ln g(Y_T|\Theta_i^*, \mathcal{M}_i) + \ln \pi(\Theta_i^*|\mathcal{M}_i) - \ln \hat{\pi}(\Theta_i^*|Y_T, \mathcal{M}_i), \quad i = 1, 2, ..., I, \quad (17)$$

where Θ_i^* is the estimated parameter vector in the model $\mathcal{M}_i$; and $g(Y_T|\Theta_i^*, \mathcal{M}_i)$, $\pi(\Theta_i^*|\mathcal{M}_i)$ and $\hat{\pi}(\Theta_i^*|Y_T, \mathcal{M}_i)$ are the likelihood function, the prior density and the posterior density, respectively, for any model i $(\mathcal{M}_i)$.

3. Data description and summary statistics

The Tokyo Stock Price Index (TOPIX) is used as an indicator of movements in the Japanese stock market (Figure 1(a)). The TOPIX is published by the Tokyo Stock Exchange and is one of the major value-weighted indices used in Japan. The sample period is from January 1970 to December 2003 for a total of 408 observations. The following three variables, which are often claimed to influence the stock market, are used as exogenous variables in Z: (1) TOR: the turnover ratio (calculated as the aggregate trading value / aggregate market value of the shares listed in the First Section of the Tokyo Stock Exchange), (2) DPR: the dividend price ratio (calculated as the average dividend / average price of the share listed on the First Section of the Tokyo Stock Exchange), and (3) ONR: the short-term interest rate (overnight call money rate). The movements of these variables during the sample period are depicted in (Figures 1(b), 1(c), and 1(d).

Baker and Stein (2002), for example, analyzing the U.S. stock market, suggest that market liquidity indices, such as the turnover ratio, are good proxies for market sentiment. Schaller and van Norden (1997) examine whether the price/dividend ratio (the inverse of the DPR) influences the transition probabilities in a MSM using U.S. stock data. Lunde and Timmermann (2004) analyze how hazard

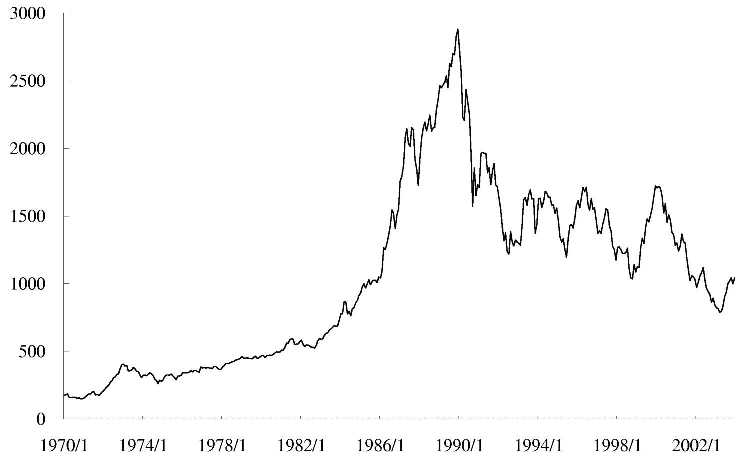

Figure 1(a): TOPIX(monthly values)

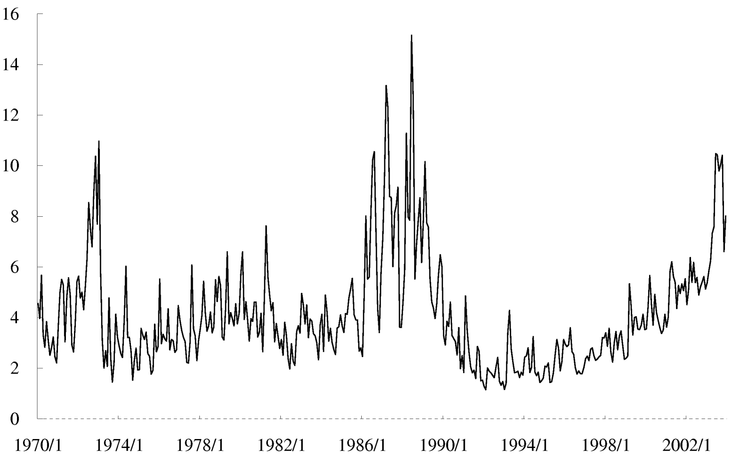

Figure 1(b): TOR

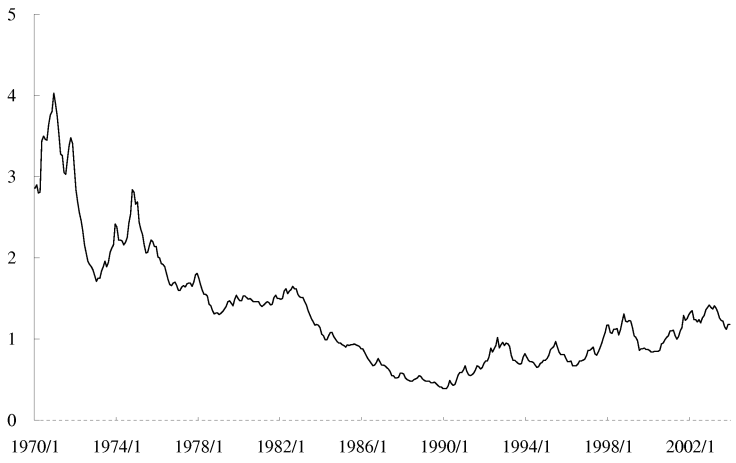

Figure 1(c): DPR

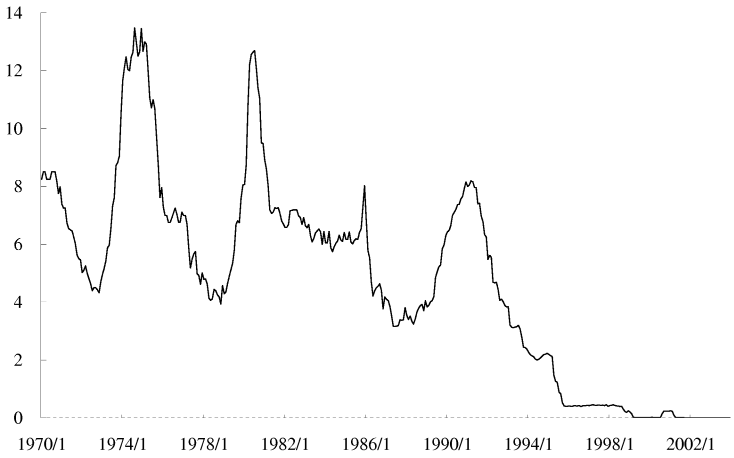

Figure 1(d): ONR

Table 1: Summary statistics

	TOPIX	(1) TOR	(2) DPR	(3) ONR
	%	%	%/month	%/year
Average	0.573	4.10	1.31	4.72
Stdev.	5.173	2.18	0.73	3.45
Minimum	-20.417	1.15	0.39	0.001
(date)	(90/9)	(92/2)	(89/12)	(01/10 -)
25%	-2.409	2.64	0.78	1.06
Median	0.635	3.55	1.14	4.85
75%	3.725	4.95	1.60	6.98
Maximum	18.153	15.16	4.03	13.48
(date)	(90/10)	(88/6)	(70/12)	(74/8)

rates depend on the interest rate using U.S. stock data. As can be seen, all of these studies concentrate on the U.S. stock market. In this paper, we use these variables and examine their effect on the transition probabilities using data for the Japanese stock market.

Table 1 lists the summary statistics for these variables. The average monthly return of the TOPIX is slightly positive. The maximum change is observed between 1990/9 and 1990/10. The maximum of the turnover ratio and the minimum of the dividend price ratio are recorded in the late 1980s when the stock price index approached its peak (see Figure 1(a)). In contrast, the maxima of the dividend price ratio and the overnight call money rate are recorded in the first half of the 1970s when stock prices were at their lowest during the sample period. The minima of the turnover ratio and the overnight call money rate are recorded after stock prices reached their peak. Since 2001/10, the overnight call money rate has not changed over time and stayed at the lowest level in history. In the following analysis, the explanatory variables are standardized.

4. Empirical Results

In the estimation, the parameters of the prior distributions in models (1)-(3), which describe the return process, are set as $m'_{0_\mu} = [0\ 0], \Sigma_{0_\mu} = I, m'_{0_\phi} = [0 \cdots 0], \Sigma_{0_\phi} = I, \nu_0 = \nu_{h0} = 1, \sigma_0 = \sigma_{h0} = 1$, where I is a unit matrix. The dropped draws M and the remaining draws N are 3,000 and 7,000, respectively.

4.1 The Markov-switching model with constant transition probabilities

Let us express function (5) as follows:

$$f(S_{t-1}, Z_{t-1}; \theta) = \gamma_0(1 - S_{t-1}) + \lambda_0 S_{t-1}, \quad \text{where } \theta = \{\gamma_0, \lambda_0\}. \tag{18}$$

This is the model proposed by Hamilton (1989). The estimates of the parameters $\theta' = [\gamma_0\ \delta_0]$ are obtained from the following posterior distribution (20) using the prior distribution (19),

prior distribution: $\theta | \tilde{S}_T, \tilde{S}_T^* \sim N(m_{0_\theta}, \Sigma_{0_\theta}),$ (19)

posterior distribution: $\theta | \tilde{S}_T, \tilde{S}_T^* \sim N(m_{1_\theta}, \Sigma_{1_\theta}),$ (20)

where

$$m_{1_\theta} = (\Sigma_{0_\theta}^{-1} + A'_\theta A_\theta)^{-1}(\Sigma_{0_\theta}^{-1} m_{0_\theta} + A'_\theta Y_\theta),$$

$$\Sigma_{1_\theta} = (\Sigma_{0_\theta}^{-1} + A'_\theta A_\theta)^{-1},$$

$$A'_\theta = \begin{bmatrix} 1 - S_1 & \cdots & 1 - S_{T-1} \\ S_1 & \cdots & S_{T-1} \end{bmatrix}, \; Y'_\theta = [S_2^* \ \cdots \ S_T^*].$$

Table 2 shows the estimation results from different settings of the model. The transition probabilities listed in the last two rows are calculated as:

$$\begin{cases} p_t^{(11)} = \Pr[u_t \geq -\delta_0] = 1 - \Psi(-\delta_0), & \forall t, \\ p_t^{(00)} = \Pr[u_t < -\gamma_0] = \Psi(-\gamma_0), & \forall t. \end{cases} \tag{21}$$

$\Psi(\cdot)$ refers to the cumulative density function of the standard normal distribution.

The first column reports the results when the monthly return of the TOPIX follows an AR(0) process that takes two states for the conditional mean (mean switching model). The second column is for the case when the AR(0) process takes two states for the conditional variance (variance switching model). The third column is for the combination of the two cases when the AR(0) process takes two states for the conditional mean and variance (mean-variance switching model). Among these three cases, the third one displays the highest marginal likelihood. The estimates of the parameters in the third case are similar to those in the second case, implying that the shift in the variance rather than

Table 2: Comparing MSMs with constant transition probabilities

AR(p)	$p=0$	$p=0$	$p=0$	$p=1$
μ_{S_t}	$\mu_0+\mu_1 S_t$	μ	$\mu_0+\mu_1 S_t$	$\mu_0+\mu_1 S_t$
$\sigma^2_{S_t}$	σ^2	$S_t\sigma_0^2+(1+S_t)\sigma_1^2$	$S_t\sigma_0^2+(1+S_t)\sigma_1^2$	$S_t\sigma_0^2+(1+S_t)\sigma_1^2$
$\ln\hat{m}$	-552.23	-542.52	-540.95	-543.86
ϕ	—	—	—	0.064 (0.052)
μ_0	0.253 (0.209)	0.035 (0.041)	0.086 (0.048)	0.090 (0.053)
μ_1	-0.662 (0.389)	—	-0.131 (0.082)	-0.140 (0.090)
σ_0^2	0.922 (0.099)	0.256 (0.043)	0.267 (0.046)	0.269 (0.052)
σ_1^2	—	1.331 (0.132)	1.347 (0.129)	1.336 (0.137)
γ_0	-0.581 (0.858)	-1.694 (0.278)	-1.640 (0.296)	-1.597 (0.276)
δ_0	0.403 (0.780)	2.093 (0.308)	1.996 (0.291)	1.963 (0.288)
$p^{(00)}$	0.688	0.949	0.942	0.938
$p^{(11)}$	0.627	0.977	0.972	0.970

in the mean is an important factor. The last column presents the case when an AR(1) process is assumed in the mean-variance switching model. Table 2 also shows that the marginal likelihood value decreases and the estimated AR parameter is near zero and insignificant. Therefore, the addition of the AR term does not improve the performance of the model. Based on these results, the extension of the Markov-switching model below uses the AR(0) specification.

A closer look at the estimation results of the mean-variance switching model reveals that state 0 has a positive mean and a smaller variance, while state 1 has a negative mean and a larger variance. Thus, if market conditions are classified into two states, one of these is a rise in stock prices following a stable trend, while the other is a decline in stock prices characterized by fluctuations. This classification coincides with that in the study by Maheu and McCurdy (2000) on the U.S. market. Therefore, bull-bear conditions in the Japanese market have the same characteristics as those in the U.S. market. The probability of staying in the state is 0.942 in the case of state 0 and 0.972 in the case of state 1, indicating that the transition between the states is not likely to occur frequently.

Figure 2 displays the probability for the market to be in state 1. It is instructive to look at this figure together with Figure 1(a) depicting the movement of the TOPIX. Thus, Figure 2 suggests that in the period from 1975 to 1985, the stock market was mainly in state 1, and a look at Figure 1(a) reveals that, indeed, during this period the TOPIX followed an upward trend with small fluctuations. In contrast, for the period after 1990, the graphs show that market conditions were characterized by state 1, i.e. a downward trend displaying large fluctuations. The period between 1986 and 1989 is, however, classified as state 0 in spite of the abrupt upward trend observed in Figure 1(a). This classification is due to the large fluctuations of the stock price index during the period. It is found from the estimated figures that the classification of the period is more affected by the variance than by the mean value.

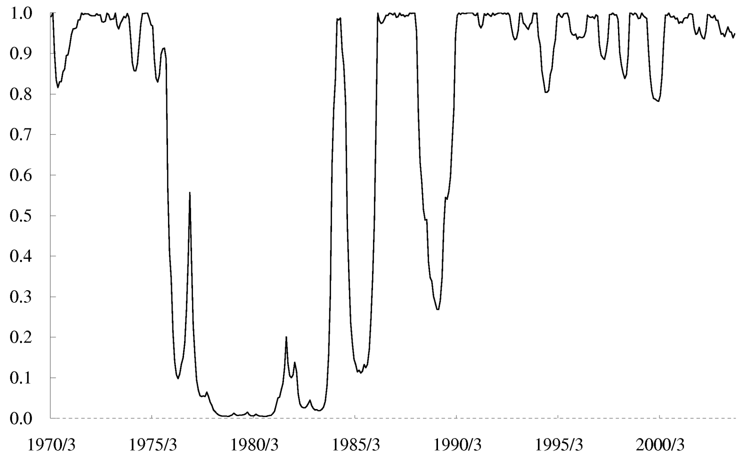

Figure 2: Probability of state 1, mean-variance switching model

4.2 The Markov-switching model with time-varying transition probabilities (MSM-TVTP)

4.2.1 The MSM-TVTP with constant sensitivities and its application to the Japanese market data

Having outlined the basic model with constant transition probabilities, we now introduce a model with time-varying transition probabilities. The model builds on work by Diebold, Lee, and Weinbach (1994) and Filardo (1994), who extended the MSM to include time-varying probabilities that are expressed as functions of exogenous variables. Because the coefficients of the exogenous variables in such models are constant, these are termed as MSM-TVTP with constant sensitivities.

One study applying the TVTP model empirically to U.S. stock market data is that by Schaller and van Norden (1997). Using the maximum likelihood method, the study found that, on the one hand, the price/dividend ratio had little effect on the transition probability when the stock market was following a low-variance upward trend. On the other hand, the ratio did have a substantial impact on the probability when the stock market was on a downward trend with high variance. However, as the coefficients on the price/dividend ratio were insignificant, the overall evidence on the link between the ratio and the transition probability was inconclusive. Here, in order to examine the time dependency of the transition probabilities, the TVTP model is applied to the Japanese stock market using the variables listed in Table 1 and the bubble dummy ("D90") that takes 0 until 1989/12 and 1 after 1990/1.

The model is formulated as follows:

$$f(S_{t-1}, Z_{t-1}; \theta) = \left(\gamma_0 + \sum_{k=1}^{K} \gamma_k z_{k,t-1}\right)(1 - S_{t-1}) + \left(\lambda_0 + \sum_{k=1}^{K} \lambda_k z_{k,t-1}\right) S_{t-1}, \tag{22}$$

where $\theta' = [\gamma_0, \gamma_1, ..., \gamma_K, \lambda_0, \lambda_1, ..., \lambda_K]$.

The estimates of the parameters are obtained in step (f) of the previous section from the following posterior distribution (24) using the prior distribution (23),

$$\textbf{prior distribution:} \qquad \theta|\tilde{S}_T, \tilde{S}_T^* \sim N(m_{0_\theta}, \Sigma_{0_\theta}), \tag{23}$$

$$\textbf{posterior distribution:} \qquad \theta|\tilde{S}_T, \tilde{S}_T^* \sim N(m_{1_\theta}, \Sigma_{1_\theta}), \tag{24}$$

where

$$m_{1_\theta} = (\Sigma_{0_\theta}^{-1} + A'_\theta A_\theta)^{-1}(\Sigma_{0_\theta}^{-1} m_{0_\theta} + A'_\theta Y_\theta),$$

$$\Sigma_{1_\theta} = (\Sigma_{0_\theta}^{-1} + A'_\theta A_\theta)^{-1},$$

$$A'_\theta = [a'_{\theta_1}, ..., a'_{\theta_{T-1}}] = \begin{bmatrix} 1 - S_1 & \cdots & 1 - S_{T-1} \\ (1 - S_1)z_{1,1} & \cdots & (1 - S_{T-1})z_{1,T-1} \\ \vdots & & \vdots \\ (1 - S_1)z_{K,1} & \cdots & (1 - S_{T-1})z_{K,T-1} \\ S_1 & \cdots & S_{T-1} \\ S_1 z_{1,1} & \cdots & S_{T-1} z_{1,T-1} \\ \vdots & & \vdots \\ S_1 z_{K,1} & \cdots & S_{T-1} z_{K,T-1} \end{bmatrix}, \; Y'_\theta = [S_2^* \; \cdots \; S_T^*].$$

The transition probabilities are calculated as:

$$\begin{cases} p_t^{(11)} = \Pr\left[u_t \geq -\lambda_0 - \sum_{k=1}^{K} \lambda_k z_{k,t-1}\right] = 1 - \Psi\left(-\lambda_0 - \sum_{k=1}^{K} \lambda_k z_{k,t-1}\right), & \forall t, \\ p_t^{(00)} = \Pr\left[u_t < -\gamma_0 - \sum_{k=1}^{K} \gamma_k z_{k,t-1}\right] = \Psi\left(-\gamma_0 - \sum_{k=1}^{K} \gamma_k z_{k,t-1}\right), & \forall t. \end{cases} \tag{25}$$

Table 3 summarizes the values of the marginal likelihood for the combination of the variables in Table 1 and the dummy variable. Here, $m'_{0_\theta} = [0 \; \cdots \; 0]$ and $\Sigma_{0_\theta} = I$ in the prior distribution. Panel A shows the results when only one of the three variables or the dummy variable is used as the explanatory variable. The marginal likelihood shows that none of the three variables plays an important role in the TVTP model with constant sensitivities. The marginal likelihood values are less than the one for the mean-variance switching model. The thin line in Figure 7(a) indicates the probability of remaining in the state of low returns with high variance ($p_t^{(00)}$) and that in Figure 7(b) indicates the probability of remaining in the state of high returns with low variance ($p_t^{(11)}$). The figures show that $p_t^{(00)}$ varies substantially with the DPR but $p_t^{(11)}$ is hardly affected by the DPR. These findings are consistent

Table 3: Comparing TVTP models with constant sensitivities

Panel A. $K=1$		Panel B. $K=2$	
Z_t	$\ln \hat{m}$	Z_t	$\ln \hat{m}$
{TOR}	-545.46	{TOR, DPR}	-548.18
{DPR}	-547.67	{TOR, ONR}	-544.16
{ONR}	-543.12	{DPR, ONR}	-546.46
{D90}	-536.96	{TOR, D90}	-538.22
		{DPR, D90}	-542.88
		{ONR, D90}	-538.16

with those obtained by Schaller and van Norden (1997). Panel B in Table 3 shows the results when two out of the three variables plus the dummy are used in combination. When the dummy is not used, no effects resulting from the combination of two variables are observed.

To summarize the above, the models of time-varying transition probabilities through exogenous variables cannot explain the dynamics of the Japanese stock market. However, the model with the dummy variable distinguishing the periods before and after the burst of the bubble has some explanatory power. The marginal likelihood, when the bubble dummy is used, generally become larger than that of the mean-variance switching model except when combined with the DPR variable. Note that the model only with the dummy, i.e. without any market-related variables, has the largest marginal likelihood among the TVTP models with constant sensitivities. These results suggest it is necessary to consider the nature of the time dependency of the transition probabilities more carefully.

4.2.2 A new model: a TVTP model with time-varying sensitivities

The explanatory power of the models with the dummy variable suggests that the sensitivities to the market-related variables may change over time. In order to capture the time dependency of the sensitivities to the market-related variables, the sensitivities in the TVTP model are formulated with a time series structure under the assumption that the impact of market-related variables on the transition probability may differ depending on economic conditions. The extended model is as follows:

$$f(S_{t-1}, Z_{t-1}; \theta_t) = \left(\gamma_{0,t} + \sum_{k=1}^{K} \gamma_{k,t} z_{k,t-1}\right)(1 - S_{t-1}) + \left(\lambda_{0,t} + \sum_{k=1}^{K} \lambda_{k,t} z_{k,t-1}\right) S_{t-1}, \quad (26)$$

where

$$\theta_t' = [\gamma_{0,t}, \gamma_{1,t}, ..., \gamma_{K,t}, \lambda_{0,t}, \lambda_{1,t}, ..., \lambda_{K,t}]$$

and

$$\begin{cases} \gamma_{k,t} = \omega_{\gamma_k} + \rho_{\gamma_k}\ \gamma_{k,t-1} + \eta_{\gamma_k,t}, & \eta_{\gamma_k,t} \sim N(0, \sigma_{\gamma_k}^2), \quad k = 0, 1, ..., K, \\ \lambda_{k,t} = \omega_{\lambda_k} + \rho_{\lambda_k}\ \lambda_{k,t-1} + \eta_{\lambda_k,t}, & \eta_{\lambda_k,t} \sim N(0, \sigma_{\lambda_k}^2), \quad k = 0, 1, ..., K. \end{cases}$$

Note that the coefficients of the explanatory variables follow an AR(1) process.

The estimation of the parameters in the AR(1) process is carried out using the Kalman filtering algorithm. For a discussion of Kalman filtering and Gibbs sampling, refer to Carter and Korn (1994). Recently, Lunde and Timmermann (2004) estimated a model with time-varying sensitivities in which bull and bear hazard rates depend on the duration of the market condition and the interest rate. Their model assumes that the states are observable and the estimation is conducted by maximizing a penalized log likelihood function. Instead in this paper the TVTP with time-varying sensitivities is constructed without the assumption of observable bull-bear conditions; instead, those conditions are to be identified within the model.

In order to estimate p_t^{11} and p_t^{00} in (25), $\{\gamma_{k,1},...,\gamma_{k,T},\lambda_{k,1},...,\lambda_{k,T},\omega_{\gamma_k},\rho_{\gamma_k},\omega_{\lambda_k},\rho_{\lambda_k},\sigma^2_{\gamma_k},\sigma^2_{\lambda_k}\}$, $k=1,...,K$, are obtained using the following procedure.

First, θ_t $(t=1,...,T)$ is generated by

$$p(\tilde{\theta}_T|\tilde{S}_T,\tilde{S}_T^*) = p(\theta_T|\tilde{S}_T,\tilde{S}_T^*)\prod_{t=1}^{T-1} p(\theta_t|\theta_{t+1},\tilde{S}_t,\tilde{S}_t^*), \tag{27}$$

where $p(\cdot|\cdot)$ denotes an appropriate conditional density and $\tilde{\theta}_T = [\theta_1,....,\theta_T]$. (27) implies the sampling of $\tilde{\theta}_T$ is conducted by first generating θ_T from $p(\theta_T|\tilde{S}_T,\tilde{S}_T^*)$ and then θ_t from $p(\theta_t|\theta_{t+1},\tilde{S}_t,\tilde{S}_T^*)$ from $t=T-1$ to $t=1$ recursively.

$\theta_T|\tilde{S}_T,\tilde{S}_T^*$ and $\theta_t|\theta_{t+1},\tilde{S}_{t+1},\tilde{S}_T^*$ in (27) follow

$$\theta_T|\tilde{S}_T,\tilde{S}_T^* \sim N(\theta_{T|T},P_{T|T}), \tag{28}$$

$$\theta_t|\theta_{t+1},\tilde{S}_{t+1},\tilde{S}_{t+1}^* \sim N(\theta_{t|t,\theta_{t+1}},P_{t|t,\theta_{t+1}}) \quad t=1,2,...,T-1, \tag{29}$$

where

$$\begin{aligned}
\theta_{T|T} &= E(\theta_T|\tilde{S}_T,\tilde{S}_T^*),\\
P_{T|T} &= \mathrm{Cov}(\theta_T|\tilde{S}_T,\tilde{S}_T^*),\\
\theta_{t|t,\theta_{t+1}} &= E(\theta_t|\tilde{S}_t,\tilde{S}_t^*,\theta_{t+1}) = E(\theta_t|\theta_{t|t},\theta_{t+1}),\\
P_{t|t,\theta_{t+1}} &= \mathrm{Cov}(\theta_t|\tilde{S}_t,\tilde{S}_t^*,\theta_{t+1}) = \mathrm{Cov}(\theta_t|\theta_{t|t},\theta_{t+1}).
\end{aligned}$$

Here $\theta_{T|T}$ and $P_{T|T}$ in (28) are obtained from the following Kalman filtering algorithm. For t $(= 1,...,T)$,

$$\begin{aligned}
\theta_{t|t-1} &= \omega + \rho\theta_{t-1|t-1},\\
P_{t|t-1} &= \rho P_{t-1|t-1}\rho' + Q,
\end{aligned}$$

where $\omega = [\omega_{\gamma_0},\omega_{\gamma_1},...,\omega_{\gamma_K},\omega_{\lambda_0},\omega_{\lambda_1},...,\omega_{\lambda_K}]'$, ρ and Q are diagonal matrices whose diagonal elements are $[\rho_{\gamma_0},\rho_{\gamma_1},...,\rho_{\gamma_K},\rho_{\lambda_0},\rho_{\lambda_1},...,\rho_{\lambda_K}]$ and $[\sigma^2_{\gamma_0},\sigma^2_{\gamma_1},...,\sigma^2_{\gamma_K},\sigma^2_{\lambda_0},\sigma^2_{\lambda_1},...,\sigma^2_{\lambda_K}]$, respectively. S_t^* is generated conditionally on S_t and S_{t-1} by :

$$S_t^* = a_{\theta_t}\theta_{t|t-1} + u_t, \quad u_t \sim \begin{cases} N(0,1)_{[u_t \geq -a_{\theta_t}\theta_{t|t-1}]}, & \text{if } S_t = 1, \\ N(0,1)_{[u_t < -a_{\theta_t}\theta_{t|t-1}]}, & \text{if } S_t = 0. \end{cases}$$

By letting

$$\eta_{t|t-1} = S_t^* - a_{\theta_t}\theta_{t|t-1},$$

$$f_{t|t-1} = a_{\theta_t}P_{t|t-1}a'_{\theta_t} + R,$$

$\theta_{t|t}$ and $P_{t|t}$ are obtained as:

$$\theta_{t|t} = \theta_{t|t-1} + K_t\eta_{t|t-1},$$

$$P_{t|t} = P_{t|t-1}K_t a_{\theta_t}P_{t|t-1},$$

where $K_t = P_{t|t-1}a'_{\theta_t}f_{t|t-1}^{-1}$ and R is the variance of the observation error that is normalized to be 1. In the above, $\theta_{t|t-1} = E(\theta_t|\tilde{S}_{t-1}, \tilde{S}_{t-1}^*)$, $\theta_{t|t} = E(\theta_t|\tilde{S}_t, \tilde{S}_t^*)$, $P_{t|t-1} = \mathrm{Cov}(\theta_t|\tilde{S}_{t-1}, \tilde{S}_{t-1}^*)$, $P_{t|t} = \mathrm{Cov}(\theta_t|\tilde{S}_t, \tilde{S}_t^*)$.

The backward estimation of $\theta_{t|t,\theta_{t+1}}$ and $P_{t|t,\theta_{t+1}}$ $(t = 1, 2, ..., T-1)$ is carried out by

$$\theta_{t|t,\theta_{t+1}} = \theta_{t|t} + P_{t|t}\rho'(\rho P_{t|t}\rho' + Q)^{-1}(\theta_{t+1} - \omega - \rho\theta_{t|t}),$$

$$P_{t|t,\theta_{t+1}} = P_{t|t} - P_{t|t}\rho'(\rho P_{t|t}\rho' + Q)^{-1}\rho P_{t|t}.$$

Then, $\Phi'_{\gamma_k} = [\omega_{\gamma_k} \ \rho_{\gamma_k}]$ in the $\gamma_{k,t}$-series is obtained from the posterior distribution given below.

prior distribution: $\quad \Phi_{\gamma_k}|\tilde{\gamma}_{k,T} \sim N(m_{0_{\gamma_k}}, \Sigma_{0_{\gamma_k}})_{I_{S(\gamma_k)}},$ (30)

posterior distribution: $\quad \Phi_{\gamma_k}|\tilde{\gamma}_{k,T} \sim N(m_{1_{\gamma_k}}, \Sigma_{1_{\gamma_k}})_{I_{S(\gamma_k)}},$ (31)

where

$$m_{1_{\gamma_k}} = (\Sigma_{0_{\gamma_k}}^{-1} + A'_{\gamma_k}A_{\gamma_k})^{-1}(\Sigma_{0_{\gamma_k}}^{-1}m_{0_{\gamma_k}} + A'_{\gamma_k}Y_{\gamma_k}),$$

$$\Sigma_{1_{\gamma_k}} = (\Sigma_{0_{\gamma_k}}^{-1} + A'_{\gamma_k}A_{\gamma_k})^{-1},$$

$$A'_{\gamma_k} = [\gamma_{k,1} \ \cdots \ \gamma_{k,T-1}], \ Y'_{\phi} = [\gamma_{k,2} \ \cdots \ \gamma_{k,T}],$$

$$\tilde{\gamma}_{k,T} \equiv [\gamma_{k,1}, ..., \gamma_{k,t}, ..., \gamma_{k,T}],$$

and $I_{S_{\gamma_k}}$ is an indicator function used to denote that the roots of $\gamma_k(L) = 0$ lie outside the unit circle.

The generation of $\Phi'_{\gamma_k} = [\omega_{\gamma_k} \ \rho_{\gamma_k}]$ is iterated from $k = 0$ to $k = K$. Then, after replacing the $\gamma_{k,t}$-series with the $\lambda_{k,t}$-series, the generation of $\Phi'_{\lambda_k} = [\omega_{\lambda_k} \ \rho_{\lambda_k}]$ is iterated from $k = 0$ to $k = K$ under the similar prior and posterior distributions of $\Phi'_{\lambda_k} = [\omega_{\lambda_k} \ \rho_{\lambda_k}]$.

Lastly, $\sigma_{\gamma_k}^2$ is obtained from the posterior distribution given below.

prior distribution: $\quad \sigma_{\gamma_k}^2|\omega_{\gamma_k}, \rho_{\gamma_k}, \tilde{\gamma}_{k,T} \sim IG\left(\frac{\nu_{\gamma_{k,0}}}{2}, \frac{\delta_{\gamma_{k,0}}}{2}\right),$ (32)

$$\textbf{posterior distribution:} \quad \sigma^2_{\gamma_k}|\omega_{\gamma_k}, \rho_{\gamma_k}, \tilde{\gamma}_{k,T} \sim IG\left(\frac{\nu_{\gamma_{k,1}}}{2}, \frac{\delta_{\gamma_{k,1}}}{2}\right), \tag{33}$$

where

$$\nu_{\gamma_{k,1}} = \nu_{\gamma_{k,0}} + (T-1), \quad \delta_{\gamma_{k,1}} = \delta_{\gamma_{k,0}} + Y_0'Y_0,$$

$$Y_0' = [\eta_2^* \ \cdots \ \eta_T^*], \ \eta_t^* = \gamma_{k,t} - \omega_{\gamma_k} - \rho_{\gamma_k}\gamma_{k,t-1}.$$

The generation of $\sigma^2_{\gamma_k}$ is iterated from $k = 0$ to $k = K$. Then, after replacing the $\gamma_{k,t}$-series with the $\lambda_{k,t}$-series, the generation of $\sigma^2_{\lambda_k}$ is iterated from $k = 0$ to $k = K$ under the similar prior and posterior distributions of $\sigma^2_{\lambda_k}$.

The transition probabilities are calculated as:

$$\begin{cases} p_t^{(11)} = \Pr\left[u_t \geq -\lambda_{0,t} - \sum_{k=1}^{K} \lambda_{k,t} z_{k,t-1}\right] = 1 - \Psi\left(-\lambda_{0,t} - \sum_{k=1}^{K} \lambda_{k,t} z_{k,t-1}\right), & \forall t, \\ p_t^{(00)} = \Pr\left[u_t < -\gamma_{0,t} - \sum_{k=1}^{K} \gamma_{k,t} z_{k,t-1}\right] = \Psi\left(-\gamma_{0,t} - \sum_{k=1}^{K} \gamma_{k,t} z_{k,t-1}\right), & \forall t. \end{cases} \tag{34}$$

In the following experiment, for the sake of simplicity, it is assumed, as in Lunde and Timmermann (2004), that the sensitivities in the equations specifying the transition probabilities follow a random walk. The parameters of the prior distributions are set for $k = 0, 1, ..., K$ as $m'_{0_{\gamma_k}} = [0 \ \cdots \ 0], \Sigma_{0_{\gamma_k}} = I, m'_{0_{\lambda_k}} = [0 \ \cdots \ 0], \Sigma_{0_{\lambda_k}} = I, \nu_{\gamma_{k,0}} = \nu_{\lambda_{k,0}} = 1,$ and $\delta_{\gamma_{k,0}} = \delta_{\lambda_{k,0}} = 1$.

The estimation results of the experiment are presented in Table 4. They show that the likelihood values increase by a large margin in all of the three cases and the differences are as great as 10. However, the values of the marginal likelihood in the three models do not differ much. The estimation results are similar to those presented in Table 2, with state 0 showing a positive mean and a smaller variance and state 1 having a negative mean and a larger variance. In other words, the observed pattern is in line with the distinction of bull and bear markets proposed by Maheu and McCurdy (2000), suggesting that their criteria can also be applied to the Japanese market.

Figures 3 to 5 display the transitions of the sensitivities in the three models. It can be seen that the sensitivities change greatly during the sample period in all three models.

Figures 6 to 8 display the probabilities of remaining in each state with each variable used as an explanatory variable. The thick lines express the probabilities for the model with time-varying sensitivities and the thin lines express those for constant sensitivities. While the probabilities for the model with constant sensitivities move within a narrow range near 1.0, the probabilities for the model with time-varying sensitivities fluctuate widely between 0 and 1. This is particularly the case with $p_t^{(11)}$. Notice that such large fluctuations of the transition probabilities are detected only by the MSM-TVTP with time-varying sensitivities. This suggests that the time dependency of the sensitivities must be taken into consideration in describing market dynamics. Reflecting the differences in the movements of the variables, particularly after 1989, the movements of the various probabilities of remaining in the same state show very different patterns. The probabilities in Figure 6 display irregular fluctuations that also reflect the irregular movements of the variable TOR.

The three panels in Figure 9 show the probabilities of the market being in state 1 (bear) for the three explanatory variables. The three figures show more or less similar patterns with some differences after 1989. The similarity comes as no surprise given the closeness of the marginal likelihood values.

Though it would be possible to identify the reasons underlying the different patterns in relation to the movement of each variable, this is an empirical matter that would distract from the theoretical concerns of this paper. Below, the model with TOR is briefly addressed. The reason for this choice is that even though market sentiments play an important role in determining market dynamics, the relationship represents an underinvestigated phenomenon and so far, no one has attempted to examine it using the TVTP-MSM.

Turning to Figure 6, conspicuous differences in the transition probabilities can be observed when compared with the other models. For example, in the late 1980s, the probabilities of the bull market continuing remains relatively high and the probabilities of the bear market continuing swing widely between 0 and 1. In Figure 9(a), there is a considerable drop in the probability of being in state 1 in the late 1980s which is not observed in the other two cases. Figure 9(b) and 9(c) depicting the DPR and ONR models show a large dip in the probability around 1995 which does not show up in the TOR model. As TOR is regarded as a good proxy for market sentiments (e.g. Baker and Stein (2002)), these features shown by the model with TOR may be considered to reflect investor sentiments at that time. A possible explanation is as follows. In the late 1980s, the market participants' optimistic sentiments drove market behavior. This is confirmed in (34) by the fact that large positive values of z_{t-1} combined with large negative values of γ_{1t} make $p_t^{(00)}$ close to 1. When they are combined with moderate negative values of λ_{1t}, $p_t^{(11)}$ becomes volatile. On the other hand, market participants' pessimistic sentiments around 1995 intensified the possibility of a bear market. This is also confirmed by the fact that large negative values of z_{t-1} combined with large negative γ_{1t}'s make $p_t^{(00)}$ nearly 0. Note that λ_{0t} takes large positive values during this period and keeps $p_t^{(11)}$ close to 1. Since the absolute values of λ_{1t} are small, z_{t-1} does not affect $p_t^{(11)}$. Further, in Figure 6(a), it can be noticed that the probability of the bull market continuing deviates from zero after 1999 while the probabilities stay at zero for the models with DPR and ONR. The increase in the probability of the bull market continuing in the TOR model reflects the abrupt increase in TOR and suggests that investors' confidence in the continuation of the bull market has grown in recent years in parallel with the recovery of the Japanese economy.

Table 4: Comparing TVTP models with time-varying sensitivities

	const. only	(1) TOR	(2) DPR	(3) ONR
$\ln \hat{m}$	-536.46	-533.84	-533.35	-534.61
μ_0	0.081 (0.050)	0.082 (0.052)	0.080 (0.048)	0.084 (0.048)
μ_1	-0.123 (0.083)	-0.134 (0.091)	-0.133 (0.088)	-0.130 (0.084)
σ_0^2	0.237 (0.046)	0.265 (0.055)	0.274 (0.052)	0.245 (0.043)
σ_1^2	1.321 (0.121)	1.387 (0.133)	1.419 (0.140)	1.367 (0.131)
$\sigma_{\gamma_0}^2$	1.013 (0.921)	1.099 (1.037)	5.326 (4.085)	2.757 (2.971)
$\sigma_{\gamma_1}^2$	—	3.467 (3.661)	6.276 (10.08)	2.056 (2.725)
$\sigma_{\lambda_0}^2$	1.005 (0.907)	0.607 (0.329)	3.736 (5.621)	1.463 (0.913)
$\sigma_{\lambda_1}^2$	—	0.702 (0.605)	1.516 (1.626)	1.459 (1.307)

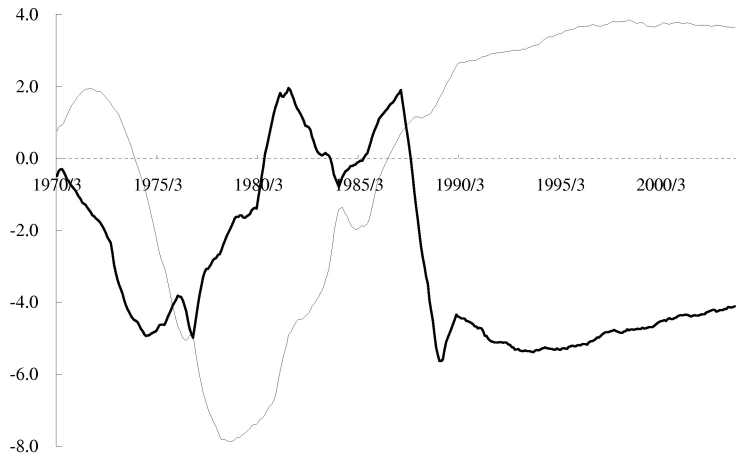

Figure 3(a): $\gamma_{0,t}$ (thin line), $\gamma_{1,t}$ (thick line) at state 0, TVTP model with time-varying sensitivities (TOR)

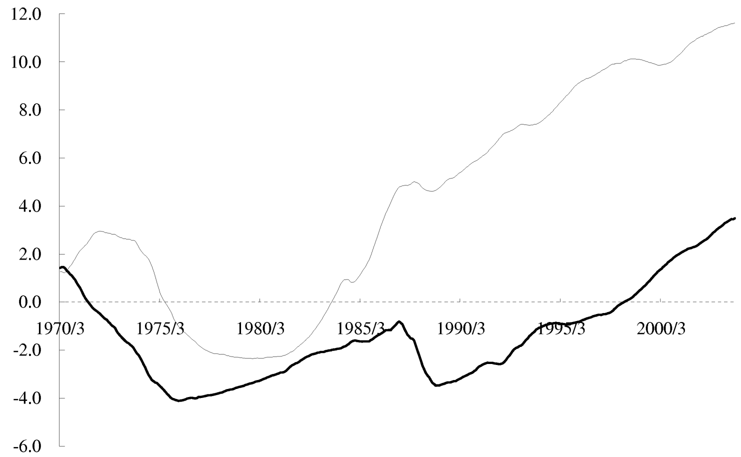

Figure 3(b): $\lambda_{0,t}$ (thin line), $\lambda_{1,t}$ (thick line) at state 1, TVTP model with time-varying sensitivities (TOR)

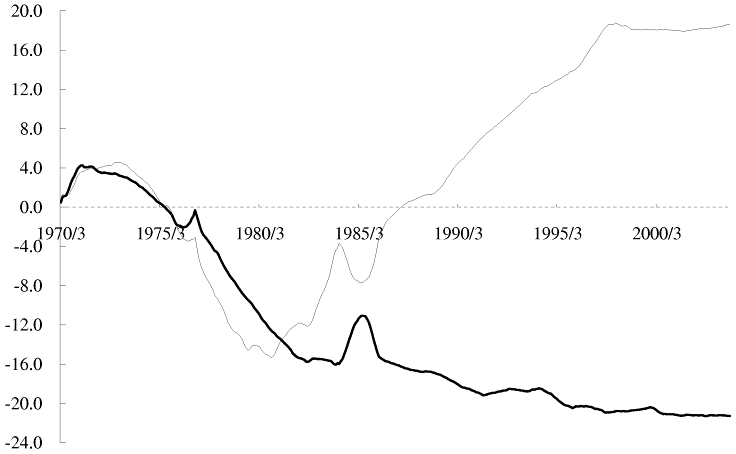

Figure 4(a): $\gamma_{0,t}$ (thin line), $\gamma_{1,t}$ (thick line) at state 0, TVTP model with time-varying sensitivities (DPR)

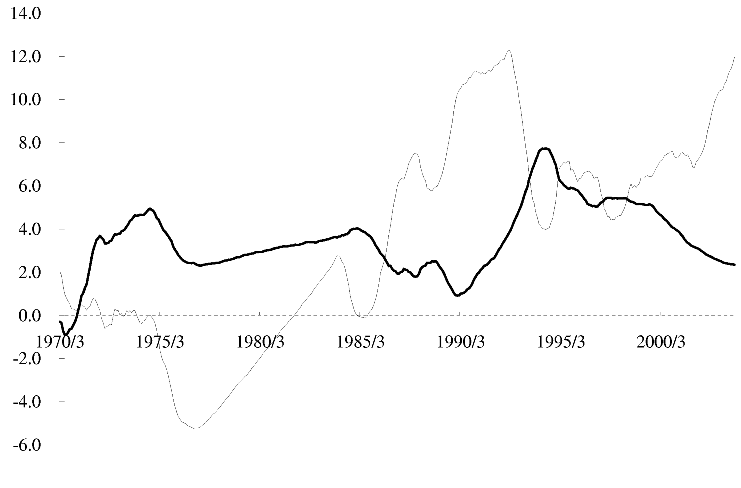

Figure 4(b): $\lambda_{0,t}$ (thin line), $\lambda_{1,t}$ (thick line) at state 1, TVTP model with time-varying sensitivities (DPR)

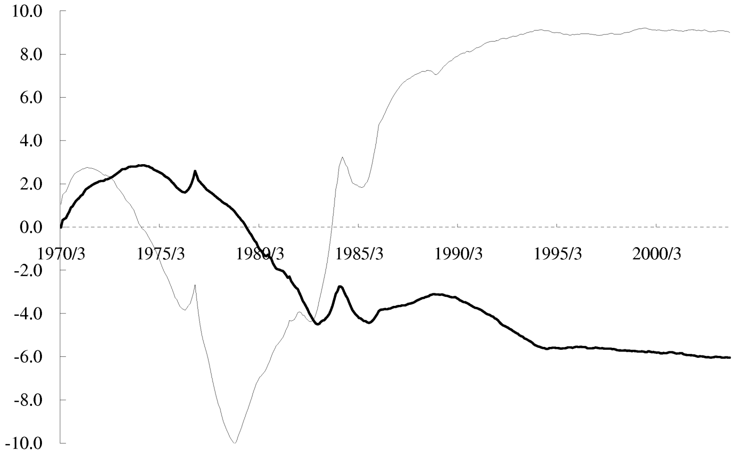

Figure 5(a): $\gamma_{0,t}$ (thin line), $\gamma_{1,t}$ (thick line) at state 0, TVTP model with time-varying sensitivities (ONR)

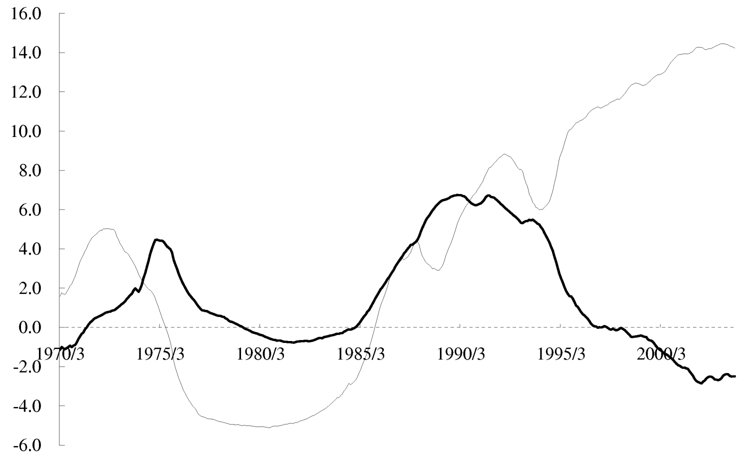

Figure 5(b): $\lambda_{0,t}$ (thin line), $\lambda_{1,t}$ (thick line) at state 1, TVTP model with time-varying sensitivities (ONR)

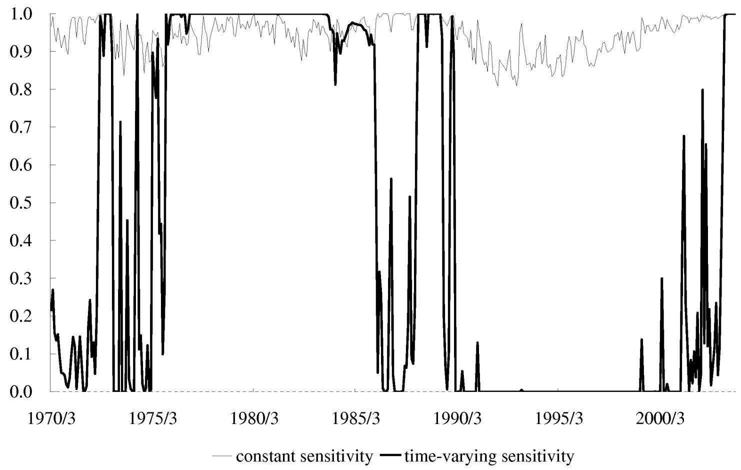

Figure 6(a): Transition probability $p_t^{(00)}$, MSM-TVTP (TOR)

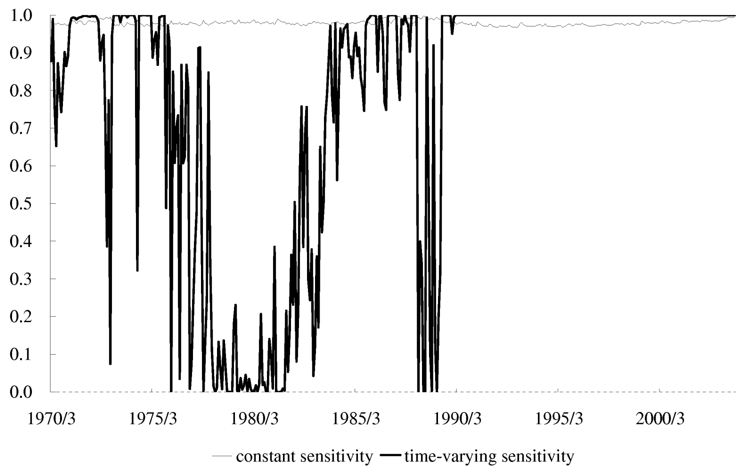

Figure 6(b): Transition probability $p_t^{(11)}$, MSM-TVTP (TOR)

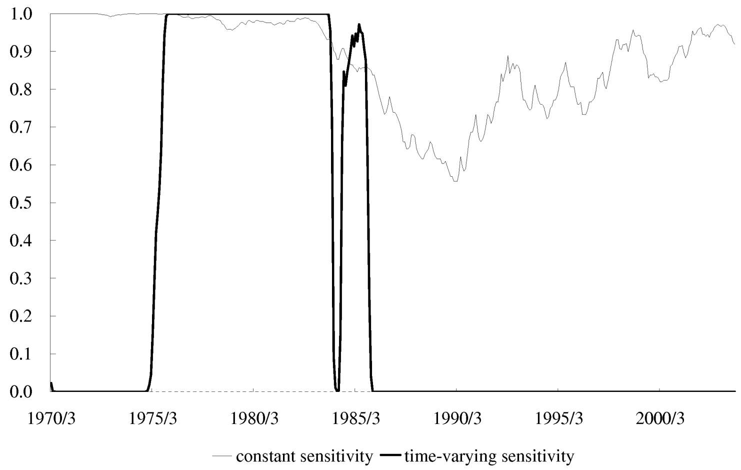

Figure 7(a): Transition probability $p_t^{(00)}$, MSM-TVTP (DPR)

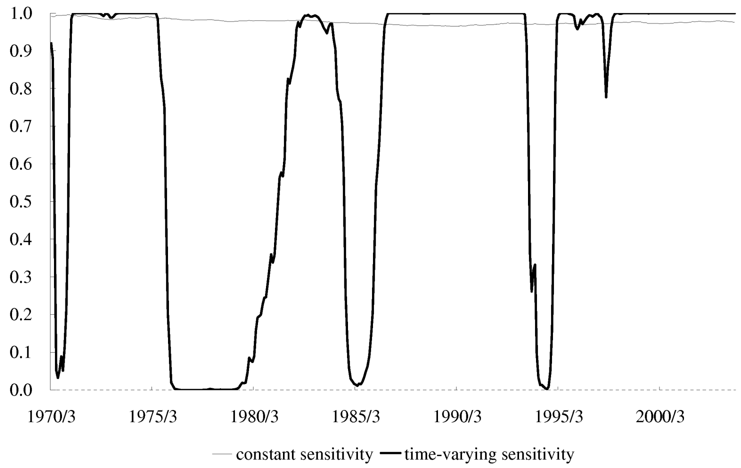

Figure 7(b): Transition probability $p_t^{(11)}$, MSM-TVTP (DPR)

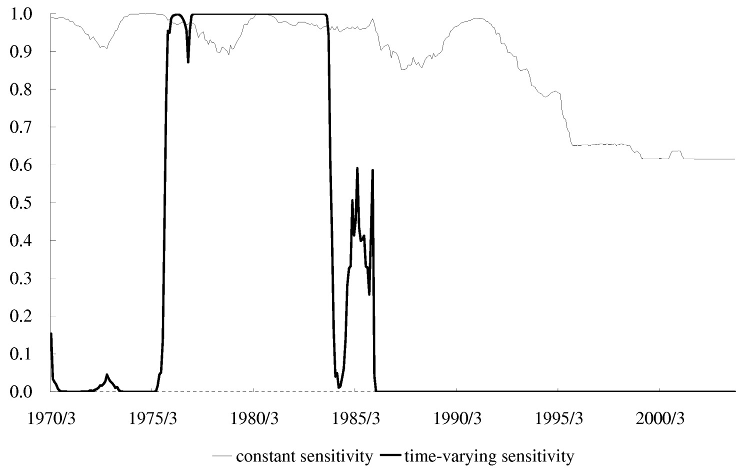

Figure 8(a): Transition probability $p_t^{(00)}$, MSM-TVTP (ONR)

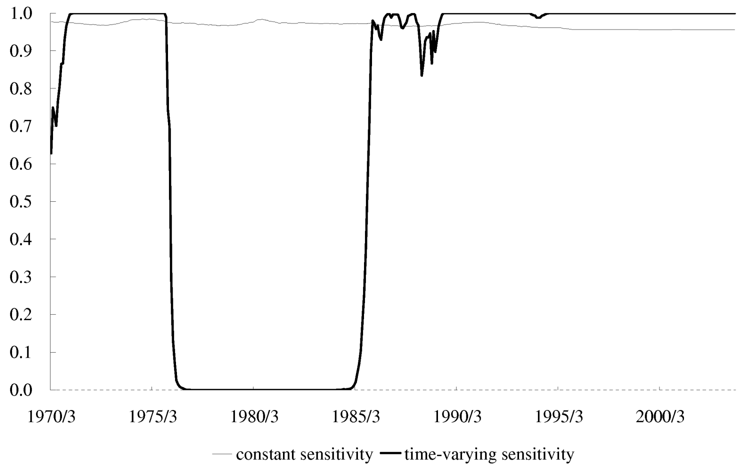

Figure 8(b): Transition probability $p_t^{(11)}$, MSM-TVTP (ONR)

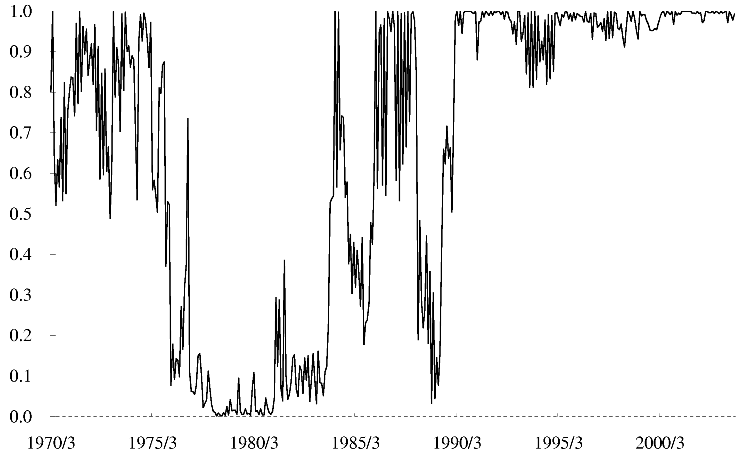

Figure 9(a): Probability of state 1, MSM-TVTP with time-varying sensitivities (TOR)

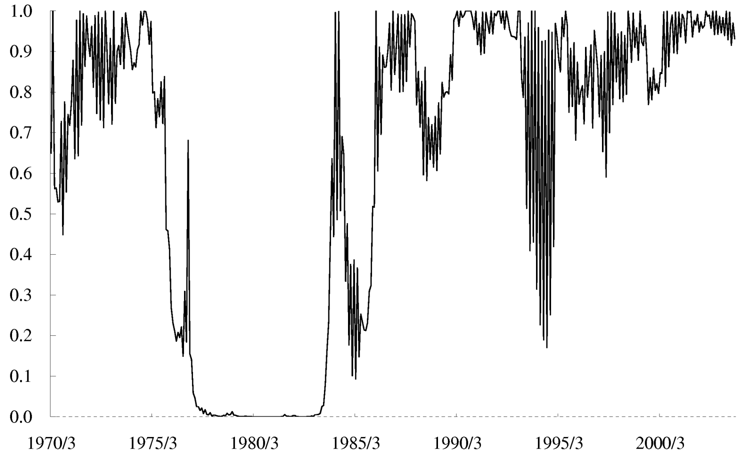

Figure 9(b): Probability of state 1, MSM-TVTP with time-varying sensitivities (DPR)

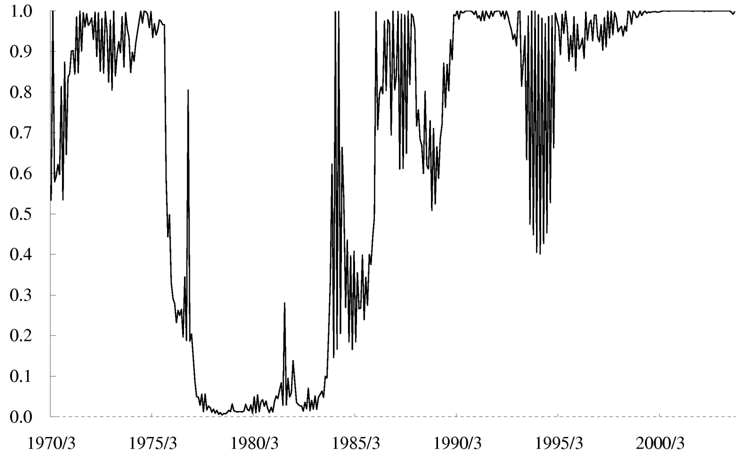

Figure 9(c): Probability of state 1, MSM-TVTP with time-varying sensitivities (ONR)

5. Conclusion

The purpose of this paper was to extend the Markov-switching model with time-varying transition probabilities (MSM-TVTP) proposed by Diebold et al. (1994) and Filardo (1994) by allowing the coefficients of the exogenous variables to vary over time. Since the state probabilities were nonlinear functions of the sensitivity parameters, the estimation was conducted employing the Gibbs sampling technique.

Using Japanese stock market data spanning a period of 34 years, we found that the marginal likelihood values improved by a large margin when compared with conventional MSMs. The estimation results also showed that the sensitivities changed considerably over the sample period and the estimated transition probabilities displayed large fluctuations between 0 and 1. These characteristics could be detected only by the MSM-TVTP with time-varying sensitivities. These results suggest that the time dependency of sensitivities is an important factor in explaining market dynamics. While these findings are on the Japanese market, it would be interesting to try the MSM-TVTP with time-varying sensitivities to study the dynamics of the U.S. market.

The empirical analysis suggested that the distinction between bull and bear markets proposed by Maheu and McCurdy (2000) also applies to the Japanese market. When the turnover ratio was used as an explanatory variable, the estimated patterns of the state probabilities and of the transition probabilities could be interpreted based on investor sentiments, which is compatible with the hypothesis that market sentiment affects market dynamics.

As Maheu and McCurdy (2000) points out, the duration of the state can be an important factor as a determinant of the transition probabilities. However, to our knowledge, empirical research on the Japanese stock market has shown the duration of a state as a determinant of transition probabilities to be insignificant, e.g. Shibata (2004). A possible extension of the empirical research presented in this paper would be to allow the sensitivities of the transition probabilities to the duration of states to be time-varying.

References

Albert, J. H., and Chib, S. (1993), "Bayes Inference via Gibbs Sampling of Autoregressive Time Series Subject to Markov Mean and Variance Shifts," *Journal of Business & Economic Statistics*, 11, 1 - 15.

Biscarri, J. G. and Gracia, F. P. (2002), "Bulls and Bears: Lessons from Some European Countries," mimeo, University of Navarre.

Baker, M., and Stein, J., C. (2002), "Market Liquidity as a Sentiment Indicator", Discussion Paper, Harvard Institute of Economic Research.

Carter, C. K., and Korn, R. (1994), "On Gibbs Sampling for State Space Models," *Biometrika*, 81, 541 - 553.

Chib, S. (1995), "Marginal Likelihood from the Gibbs Output," *Journal of the American Statistical Association*, 90, 1313 - 1321.

Diebold, Francis X., Joon-Haeng Lee, and Gretchen C. Weinbach (1994), "Regime Switching with Time-varying Transition Probabilities," in *Nonstationary Time Series Analysis and Cointegration*, ed. C. Hargreaves, Oxford University Press, Oxford.

Filardo, A., J. (1994), "Business Cycle Phases and Their Transition Dynamics," *Journal of Business & Economic Statistics*, 12, 299 - 308.

Filardo, A. J., and Gordon, S. F. (1998), "Business Cycle Durations," *Journal of Econometrics*, 85, 99 - 123.

Hamilton, J. D. (1989), "A New Approach to the Economic Analysis of Non-stationary Time Series and the Business Cycle," *Econometrica*, 57, 357 - 384.

Kanoh, S., and Li, Z. D. (1990), "A Method of Exploring the Mechanism of Inflationary Expectations Based on Qualitative Survey Data," *Journal of Business & Economic Statistics*, 8, 395 - 403.

Kim, C. J., and Nelson, C. R. (1998), "Business Cycle Turning Points, A New Coincident Index, and Tests of Duration Dependence Based on a Dynamic Factor Model with Regime Switching," *Review of Economics and Statistics*, 80, 188-201.

Kim, C. J., and Nelson, C. R. (1999), *State-Space Models with Regime Switching*, MIT Press.

Lunde, A., and Timmermann, A. (2004), "Duration Dependence in Stock Prices: An Analysis of Bull and Bear Markets," *Journal of Business & Economic Statistics*, 22, 253 - 273.

Maheu, M., and McCurdy, T. H. (2000), "Identifying Bull and Bear Markets in Stock Returns," *Journal of Business & Economic Statistics*, 18, 100 - 112.

Schaller, H., and van Norden, S. (1997), "Regime Switching in Stock Market Returns," *Applied Financial Economics*, 7, 177 - 191.

Shibata, M. (2004), "The stock market cycle and its duration dependence," unpublished manuscript.