Power of a LR Test for Regime Shifts: A Simulation Study

Takeshi Otsu

Abstract

This paper reexamines a finite property of the standardized likelihood ratio test statistic designed for a regime switching model by Hansen (1992). The main findings are as follows. First, we find that the power of the test is not excellent but moderate when we increase number of sample sets generated by a random generator and use a finer grid in the simulation. Secondly, the size of the test is larger than the nominal size for a finer grid point. It suggests the over-rejection of the null hypothesis, and, specifically, the rejection frequency is twice as much as the nominal size of 5%. Finally, the over-rejection of the null hypothesis disappears and the power of the test reduces when the Bartlett kernel is introduced into the covariance estimator of the test statistic.

Key words: Likelihood Ratio test, Monte Carlo simulation, Markov switching

JEL classification: C24

1 Introduction

The concept of 'regime shifts' is frequently used to describe changes in the state of economy. On the empirical front, it is materialized by nonlinear econometric models, such as thresholds models (Tong, 1983) and Markov-switching models (Hamilton, 2016). These models attempt to describe discrete changes or nonlinearities observed in economic and financial data.

A difficulty comes into play when we conduct statistical testing for these models. It arises from the fact that some parameters are unidentified under the null hypothesis. This is the problem of unidentified nuisance parameters, which was analyzed by Davies (1977, 1987). The existence of unidentified nuisance parameters makes it impossible to consistently estimate model parameters under the null hypothesis and to apply the standard asymptotic distributional theory to test statistics because the likelihood function has multiple local optima and some of the elements of the score vector (the first-order derivatives) are identically zeros under the null hypothesis.

Following Davies (1977), Hansen (1996) extended the empirical process theory to a wide class of estimation problems and test statistics, and Hansen (1992) used it to derive a bound for the asymptotic distribution of the standardized likelihood ratio statistic for Markov switching models. Garcia (1998) examined the asymptotic distribution of the likelihood ratio test statistic, assuming the score is not identically zero under the null hypothesis. But, the power of such a test is unknown. Cho and White (2007) proposed a quasi-likelihood ratio (QLR) test and derives its asymptotic null distribution. Qu and Zhuo (2021) used a higher-order approximation to refine the asymptotic null distribution of the QLR statistic.

Although Hansen (1992) showed some evidence that the proposed likelihood ratio test had a good power, the computation requirements did not allow for good enough simulation exercises to examine the finite sample distribution. In the literature, this statistic is rarely used. Thus, this paper reexamines the power of the standardized likelihood ratio test in a finite sample.

The main findings are as follows. First, we find that the power of the test is not excellent but moderate when we increase number of sample sets generated by a random generator and use a finer grid in the simulation. Secondly, the size of the test is larger than the nominal size for a finer grid. It suggests the over-rejection of the null hypothesis, and, specifically, the rejection frequency is twice as much as the nominal size of 5%. Finally, the over-rejection of the null hypothesis disappears and the power of the test reduces when the Bartlett kernel is introduced into the covariance estimator of the test statistic.

The rest of the paper is organized as follows. In section 2, we briefly discuss the testing hypotheses to be studied and review the likelihood ratio test statistic proposed by Hansen (1992). Section 3 explains specification for the simulation and examines its results. The final section is allocated to discussion.

2 Testing Hypothesis and Test Statistic

2.1 Testing Hypothesis

Suppose we have two different states of economy. Let y_t an economic time-series variable at time t ($t=1, 2, \cdots, n$). Then, we have a simple regime-switching model as follows:

$$y_t=\mu+\mu_d S_t+\varepsilon_t, \tag{1}$$

$$\varepsilon_t \sim N(0, \sigma^2), \tag{2}$$

where μ, μ_d, and σ are all unknown constant parameters, and S_t denotes the economic state at time t, taking either 1 or 0. The transition between states is assumed to be governed by a first-order Markov process that is independent of ε_t. Then, the transition probabilities are:

$$P[S_t=1|S_{t-1}=1]=p, \tag{3}$$

$$P[S_t=0|S_{t-1}=0]=q. \tag{4}$$

When μ_d is zero, this model reduces to one-state model. Therefore, it is interesting to conduct the following hypothesis testing to see if the two-state model is appropriate.

$$H_0 : \mu_d = 0, \quad H_1 : \mu_d \neq 0, \tag{5}$$

where H_0 denotes the null hypothesis and H_1 the alternative one. One might use the conventional t-statistic to test, but it does not have the standard null distribution. Neither do other conventional statistics, such as the likelihood ratio or the chisquare statistics. This is because p and q are unidentified under the null hypothesis. That is, we cannot find unique estimates for these parameters to maximize the likelihood function. Further, the scores with respect to μ_d, p and q are identically zeros under the null hypothesis. Then, the standard distributional theory is inapplicable. To circumvent these problems, Hansen (1992) proposed a standardized likelihood ratio test statistic and resorted to Monte Carlo simulations to compute its p-values.

2.2 The Standardized Likelihood Ratio Test Statistic

We provide a schematic explanation for the standardize likelihood ratio test statistic proposed by Hansen (1992). Let $l_t(\cdot)$ a log-transformed probability density of the model consisting of eq.(1) to eq.(4). Then, the log-likelihood function can be written in the form:

$$\mathcal{L}_n(\alpha, \theta) = \sum_{t=1}^{n} l_t(\alpha, \theta), \tag{6}$$

where n is the sample size, $\alpha = (\mu_d, p, q)$ and $\theta = (\mu, \sigma)$. Note that θ is identified under H_0 in (5). For a maximum likelihood estimator, θ can be written as

$$\theta(\alpha) = \underset{\theta}{\mathrm{argmax}} \lim_{n \to \infty} \frac{1}{n} E\mathcal{L}_n(\alpha, \theta). \tag{7}$$

Therefore, the concentrated log-likelihood function is given by

$$\mathcal{L}_n^C(\alpha) = \mathcal{L}_n(\alpha, \theta(\alpha)). \tag{8}$$

To test H_0 in (5), consider the following likelihood ratio function:

$$LR_n(\alpha_A) = \mathcal{L}_n^C(\alpha_A) - \mathcal{L}_n^C(\alpha_N), \tag{9}$$

where α_N is α under H_0 and α_A is α under H_1. The test statistic proposed by Hansen (1992) is given by the supremum of eq.(9):

$$LR_n = \sup_{\alpha_A} LR_n(\alpha_A). \tag{10}$$

To study a bound of the asymptotic distribution of this statistic, consider the following decomposition for any α:

$$LR_n(\alpha) = E[LR_n(\alpha)] + Q_n(\alpha), \tag{11}$$

where $E[LR_n(\alpha)]$ is the mean and $Q_n(\alpha)$ is the deviation from the mean. Note:

$$Q_n(\alpha) = \sum_{t=1}^{n} q_t(\alpha), \tag{12}$$

where

$$\begin{aligned} q_t(\alpha) = {} & l_t(\alpha,\ \theta(\alpha)) - l_t(\alpha_N,\ \theta(\alpha_N)) \\ & - E[l_t(\alpha,\ \theta(\alpha)) - l_t(\alpha_N,\ \theta(\alpha_N))]. \end{aligned} \tag{13}$$

If an empirical process central limit theorem holds, the $Q_n(\alpha)$ weakly converges to a mean zero Gaussian process, $Q(\alpha)$:

$$\frac{1}{\sqrt{n}} Q_n(\alpha) \Longrightarrow Q(\alpha), \tag{14}$$

as $n \to \infty$. For different values of α, the covariance function is given by

$$K(\alpha_1,\ \alpha_2) = \sum_{k=-\infty}^{\infty} E[q_t(\alpha_1) q_{t-k}(\alpha_2)], \tag{15}$$

and the associated variance function is

$$V(\alpha)=K(\alpha,\ \alpha). \tag{16}$$

Let $\hat{\theta}(\alpha)$ a consistent estimator of $\theta(\alpha)$ in eq.(7): it is assumed that $\hat{\theta}(\alpha)$ is consistent for $\theta(\alpha)$ at rate $\sqrt{n}$, uniformly in α. Further, we assume that we have a stochastic order relation for the following Euclidean metric:

$$\sup_{\alpha}\| \mathcal{L}_n^C(\alpha,\ \theta(\alpha))-\mathcal{L}_n^C(\alpha,\ \hat{\theta}(\alpha))\|=O_n(1). \tag{17}$$

Let $\hat{q}_t(\alpha)$ the estimator of $q_t(\alpha)$ in eq.(13) associated with $\hat{\theta}(\alpha)$, and it is used to compute $\hat{Q}_n(\alpha)$ via eq.(12). Then, eq.(17) implies

$$\sup_{\alpha}\| Q_n(\alpha)-\hat{Q}_n(\alpha)\|=O_n(1). \tag{18}$$

The sample analogue of eq.(15) can be written as

$$\begin{aligned}\hat{K}_n(\alpha_1,\ \alpha_2)=&\sum_{t=1}^{n}\hat{q}_t(\alpha_1)\hat{q}_t(\alpha_2)\\&+\sum_{k=1}^{M}w_{kM}\sum_{t=k+1}^{n}(\hat{q}_{t-k}(\alpha_1)\hat{q}_t(\alpha_2)+\hat{q}_t(\alpha_1)\hat{q}_{t-k}(\alpha_2)),\end{aligned} \tag{19}$$

where w_{kM} is the Bartlett kernel with a bandwidth of M:

$$w_{kM}=1-\frac{k}{M+1}. \tag{20}$$

When we set $\alpha_1=\alpha_2$ in eq.(19) and divide its both sides by n to obtain the estimator of variance: $\hat{V}_n(\alpha)/n$.

Although the mean function in eq.(11) is unknown, we know from eq.(9) that, under the H_0,

$$E[LR_n(\alpha)]\leq 0. \tag{21}$$

This implies

$$\frac{1}{\sqrt{n}}LR_n(\alpha)\leq\frac{1}{\sqrt{n}}Q_n(\alpha). \tag{22}$$

As assumed before, the right hand side weakly converges to $Q(\alpha)$. Therefore, taking $\sup$ of the both sides in eq.(22), we can obtain a bound for the asymptotic distribution of the statistic in eq.(10). Hansen (1992) argued that the bound has excessively strong tendency not to reject the null hypothesis: it is too conservative in practice. To reduce the over-conservative tendency, it is proposed to standardize the statistic so that the variance is same for all values of α.

For the sample analogue of eq.(22) with the estimator, $\hat{\theta}(\alpha)$, we have

$$\frac{1}{\sqrt{n}}\widehat{LR_n}(\alpha) \leq \frac{1}{\sqrt{n}}\hat{Q}_n(\alpha). \tag{23}$$

The right hand side weakly converges to $Q(\alpha)$ under the conditions of eq.(14) and eq. (18). The standardized version is obtained by dividing both sides by $\sqrt{\hat{V}_n(\alpha)/n}$:

$$\frac{\widehat{LR_n}(\alpha)}{\sqrt{\hat{V}_n(\alpha)}} \leq \frac{\hat{Q}_n(\alpha)}{\sqrt{\hat{V}_n(\alpha)}}. \tag{24}$$

Hansen (1992) assumes:

$$\frac{\hat{Q}_n(\alpha)}{\sqrt{\hat{V}_n(\alpha)}} \Longrightarrow \frac{Q(\alpha)}{\sqrt{V(\alpha)}}. \tag{25}$$

Define the standardized likelihood ratio function as

$$\widehat{LR_n^*}(\alpha) = \frac{\widehat{LR_n}(\alpha)}{\sqrt{\hat{V}_n(\alpha)}}. \tag{26}$$

Then, the standardized likelihood ratio statistic is

$$\widehat{LR_n^*} = \sup_{\alpha} \widehat{LR_n^*}(\alpha). \tag{27}$$

Similarly, we define

$$\hat{Q}_n^*(\alpha) = \frac{\hat{Q}_n(\alpha)}{\sqrt{V_n(\alpha)}}, \tag{28}$$

and

$$Q^*(\alpha) = \frac{Q(\alpha)}{\sqrt{V(\alpha)}}. \tag{29}$$

Together with eq.(24) and eq.(25), we have

$$\widehat{LR_n^*} \le \sup_\alpha \hat{Q}_n^*(\alpha) \implies \sup_\alpha Q^*(\alpha). \tag{30}$$

Finally, we have the following bound for the asymptotic distribution of the statistic in eq.(27):

$$P\left\{\widehat{LR_n^*} \ge x\right\} \le P\left\{\sup_\alpha \hat{Q}_n^*(\alpha) \ge x\right\}$$
$$\implies P\left\{\sup_\alpha Q^*(\alpha) \ge x\right\}. \tag{31}$$

The distribution of $\sup_\alpha Q^*(\alpha)$ is generally non-standard, but it is completely characterized by its covariance function:

$$K^*(\alpha_1, \alpha_2) = \frac{K(\alpha_1, \alpha_2)}{\sqrt{V(\alpha_1)}\sqrt{V(\alpha_2)}}. \tag{32}$$

The consistent estimators of the components on the right hand side are given by eq.(19). Therefore, it is possible to obtain the approximate distribution of $\sup_\alpha Q^*(\alpha)$ from the empirical distribution of the random draws, $\sup_\alpha \hat{Q}_n^*(\alpha)$.

3 Monte Carlo Simulation

To conduct the Monte Carlo simulation, we first generate samples under the null hypothesis. Since the model follows independent and identically distributed (i.i.d.) normal distribution, we generate samples of i.i.d. normal observations, using the random generator of the SFMT Mersenne-Twister 19937 (GAUSS software). To evaluate a finite sample property, the sample size is set to 131, same as in Hansen (1992). We randomly generate 1131 sample points, and burn out the first and the last 500 samples. Then, we compute the maximum likelihood (ML) estimates of the model under the null hypothesis to obtain the estimate of $\mathcal{L}_n^C(\alpha_N)$.

To compute the estimate of $\mathcal{L}_n^C(\alpha_A)$, we estimate $\theta(\alpha)$ of the model under the alternative hypothesis for each combination of $\alpha=(\mu_d, p, q)$ using some grids of values. The estimate of $LR_n(\alpha_A)$ is computed for each combination, and divided by its standard deviation (eq.(26)). Then, the maximum value is selected for $\widehat{LR_n^*}$ (eq.(27)). The set of the grid points is same as that of Hansen (1992): 20 grid points for μ_d, that is, the range of $[0.1, 2]$ in steps of 0.1, and two sets of grid points for p and q:

GP6 (6 grid points) : 0.15 to 0.90 in steps of 0.15;

GP8 (8 grid points) : 0.12 to 0.89 in steps of 0.11.

To study size and power of the standardized likelihood ratio ($\widehat{LR_n^*}$) statistic, we need to obtain draws from the empirical distribution of $\sup_\alpha Q^*(\alpha)$. Following Hansen (1992, 1996), we use the expression below:

$$\widetilde{LR_n^*}(\alpha)=\frac{\sum_{k=0}^{M}\sum_{t=1}^{n}\hat{q}_t(\alpha)u_{t+k}}{\sqrt{M+1}\sqrt{\hat{V}_n(\alpha)}}, \tag{33}$$

where $u_t(t=1, 2, \cdots, n+M)$ is a random sample of $N(0, 1)$ variables. The process of eq.(33) would approximately give rise to the process of $\hat{Q}_n^*(\alpha)$. We use 1000 Monte Carlo samples to calculate asymptotic p-values associated with the $\widehat{LR_n^*}$ statistic. We set M at the values from 0 to 4 throughout simulations in this section.

To begin with, we examine size of the test, that is, the frequency of rejections under the null hypothesis in eq.(5). We generate 50 sets of samples of length 131, consisting of i.i.d normal observations. For each sample set, we compute the $\widehat{LR_n^*}$ statistic and its associated p-value. If the p-value is smaller than 0.20, 0.10, or 0.05, the null hypothesis is rejected at the 20, 10, or 5 per cent level, respectively.

We count the frequencies of rejections out of 50 trials. Selected results are presented. Table 1 shows a result closest to that of Hansen (1992, 1996) in our experiments. Table 2, however, shows that the simulation results vary a lot. Then, we increase the number of sample sets to 100 in Table 3, 200 in Table 4, and 500 in Table 5. We confirm that the simulation results give less variation as we have more sample sets to compute the rejection frequencies. This experiment suggests that at least 500 sample sets will be necessary to draw a conclusion. The results in these tables indicate the actual size of the test is close to their nominal values of 10% and 5% levels and slightly smaller for 20% level.

To examine the power of the test, we need specify a model under the alternative hypothesis. Hansen (1992) used the estimates of Hamilton (1989) model, which is described as:

$$y_t = \mu + \mu_d S_t + u_t, \tag{34}$$

$$\phi(L) u_t = \varepsilon_t, \tag{35}$$

$$\varepsilon_t \sim N(0, \sigma^2), \tag{36}$$

$$\phi(L) = 1 - \sum_{j=1}^{K} \phi_j L^j, \quad L^k y_t \equiv y_{t-k} \tag{37}$$

where K is set to 4, and μ, μ_d, σ and ϕ_j are all unknown constant parameters. The difference from the model in the previous section is in the error term that follows an autoregressive (AR) process as in eq.(35). S_t, p, and q are defined same as in the previous section. We replicate the maximum-likelihood (ML) estimates in Table 6.

According to Hansen (1992), the AR parameters were set to zero for the alternative model in the simulation. But, the zero AR parameters do not maximize the likelihood function with other estimates in Table 6. This does not satisfy the theoretical requirement in eq.(7). In experiments, we computed power of the test with the AR parameters set to zero, keeping other estimates in Table 6. Then, the power was 100% for the tests of the nominal sizes from 20% to 5%, which is a quite different result from that of Hansen (1992). Thus, to keep the theoretical consistency, we use the estimates of μ_d, p and q in Table 6 for the simulation and recalculate their associated ML estimates of μ and σ, setting $\phi_j(j=1, 2, 3, 4)$ to zero.

Table 7 and 8 show the results from 50 sets of samples. We obtain a similar result for the nominal size of 20% to that of Hansen (1992), but very different results for other nominal sizes in Table 7. In addition, the results vary quite a lot as seen in Table 8. Then, we increase the number of trials to 500 in Table 9. We find that the power is more than 80% at the nominal size of 20%, around 70% to 80% for 10% nominal size, and from 55% to 70% for 5% nominal size. Therefore, we do not come to a conclusion that the test has excellent power as in Hansen (1992). The power is moderate at the nominal sizes of 10% and 5%, which are the significance levels frequently used in the literature. This implies that the test will have a moderate discriminatory power in practice. We note that the power tends to be stronger as the bandwidth (M) is greater.

In Table 10, we use a finer grid for p and q, that is, 8 grid points (GP8). The powers of the tests are lower than those in Table 9. The sizes show that the tests tend to over-reject the null hypothesis relative to the nominal sizes. The actual sizes show that the null hypothesis would be rejected twice as high as at the nominal size of 5%. That is, the tests tend to over-reject the null.

Finally, we evaluate the effect of the estimator of the covariance function in eq. (16) on the simulation results. Instead of eq. (33), we use the following expression to approximate the empirical distribution of $Q^*(\alpha)$:

$$\widetilde{LR}_n^*(\alpha)_B = \frac{\sum_{k=0}^{M}\sum_{t=1}^{n}\sqrt{w_{kM}}\,\hat{q}_t(\alpha)\,u_{t+k}}{\sqrt{\hat{V}_n(\alpha)}}, \tag{38}$$

where w_{kM} is the Bartlett kernel, given by eq.(20). This ensures the positive semi-definite consistent estimate of covariance matrix in a finite sample (Newey and West, 1987). Table 11 shows that introduction of the Bartlett kernel makes the actual sizes of the tests lower than the nominal sizes, and the over-rejection of the null hypothesis disappears with 8 grid-point simulation. It reduces the power of the test, however. When the bandwidth (M) is 1, the actual size (3.6%) is close to the nominal size of 5%, but the power is low, around 31%. The formulation of the covariance estimator has substantial influence on the size and the power of the test.

4 Discussion

This paper reexamines the standardized likelihood ratio test statistic designed for a regime switching model by Hansen (1992). The main findings are as follows. First, we find that the power of the test is not excellent but moderate when we increase number of sample sets generated by a random generator and use a finer grid point in the simulation. Secondly, the size of the test is larger than the nominal size for a finer grid. This suggests the over-rejection of the null hypothesis. Specifically, the rejection frequency is twice as much as the nominal size of the 5%. Finally, the over-rejection of the null hypothesis disappears and the power of the test reduces when the Bartlett kernel is introduced into the covariance estimator of the test statistic.

A couple of caveats are in order. First, we do not examine the size and the power of the test when a model follows an autoregressive process from eq.(34) to eq.(37). Hansen (1992, 1996) examined autoregressive processes, different from those discussed in the previous section, and found that the power of the test reduced. This suggests that model specifications would matter. Secondly, a set of finer grid points might give rise to more accurate results. In general, the finer the grid, the more accurate the empirical distribution of the test statistic. In the paper, we only used 6 and 8 grid points for p and q, and 20 points for μ_d. Finally, choice of the covariance estimator of the test statistic and its bandwidth also affects its finite property. These are subjects for future research.

Acknowledgements **A part of this work is financially supported by** *Tokubetsu Kenkyu Josei* **(Seijo University) under the title of「経済モデルの状態数に関する実証分析」.**

References

Cho, J. S. and White, H., 2007, "Testing for regime switching," Econometrica, 75, 1671-1720.

Davies, R. B., 1977, "Hypothesis testing when a nuisance parameter is present only under the alternative," Biometrika, 64, 247-254.

Davies, R. B., 1987, "Hypothesis testing when a nuisance parameter is present only under the alternative," Biometrika, 74, 33-34.

Garcia, R., 1998, "Asymptotic null distribution of the likelihood ratio test in Markov switching models," International Economic Review 39(3), 763-788.

Hamilton, J. D., 1989, "A new approach to the economic analysis of nonstationary time series and the business cycle," Econometrica 57, 357-384.

Hamilton, J. D., 2016. "Macroeconomic Regimes and Regime Shifts," Ch. 3., In Handbook of Macroeconomics, Volume 2A, edited by John B. Taylor and Harald Uhlig. Amsterdam: Elsevier, 163-201.

Hansen, B. E., 1992, "The likelihood ratio test under nonstandard conditions: Testing the Markov switching model of GNP," Journal of Applied Econometrics 7, S61-S82. Erratum, 1996, 11, 195-198.

Hansen, B. E., 1996, "Inference when a nuisance parameter is not identified under the null hypothesis," Econometrica, 64, 413-430.

Newey, W. K., and K. D. West, 1987, "A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix," Econometrica 55, 703-708.

Qu, Z. and F. Zhuo, 2021, “Likelihood ratio-based tests for Markov regime switching,” The Review of Economic Studies 88(2), 937-968.

Tong, H., 1983, Threshold Models in Non-linear Time Series Analysis. New York: Springer.

Table 1 Rejection Frequency under the Null: 50 trials

Size(%)	Hansen*			Replication 1		
Bandwidth	Nominal Size (%)			Nominal Size (%)		
M	20	10	5	20	10	5
0	12	4	0	12	6	6
1	14	8	0	14	8	6
2	16	8	2	14	8	6
3	14	8	2	14	10	6
4	14	6	2	14	10	6

Note: *Hansen (1992, Erratum: 1996). 6 grid points.

Table 2 Rejection Frequency under the Null: 50 trials, variation

Size (%)	Replication 2			Replication 3			Replication 4		
Bandwidth	Nominal Size (%)			Nominal Size (%)			Nominal Size (%)		
M	20	10	5	20	10	5	20	10	5
0	6	2	0	20	10	2	24	14	6
1	10	2	0	20	10	4	24	12	6
2	10	4	2	22	8	4	24	14	8
3	12	4	2	22	8	4	22	16	8
4	12	6	2	24	10	4	22	16	10

Note: 6 grid points for p and q.

Table 3 Rejection Frequency under the Null: 100 trials, variation

Size (%)	Replication 1			Replication 2			Replication 3		
Bandwidth	Nominal Size (%)			Nominal Size (%)			Nominal Size (%)		
M	20	10	5	20	10	5	20	10	5
0	10	8	5	13	10	5	19	10	7
1	15	9	5	14	7	4	22	14	7
2	14	7	5	15	10	4	21	14	8
3	16	7	6	16	10	5	21	15	8
4	16	7	6	15	10	6	21	15	10

Note: 6 grid points for p and q.

Table 4　Rejection Frequency under the Null: 200 trials, variation

Size (%)	Replication 1			Replication 2			Replication 3		
Bandwidth	Nominal Size (%)			Nominal Size (%)			Nominal Size (%)		
M	20	10	5	20	10	5	20	10	5
0	13.5	5.0	3.0	17.0	8.5	3.5	20.5	13.0	7.5
1	16.0	5.0	3.0	19.0	10.5	3.5	21.0	14.0	8.0
2	16.5	5.5	3.5	19.5	12.0	4.5	21.5	14.0	9.5
3	17.0	5.5	3.5	21.0	12.5	4.5	22.0	13.5	10.0
4	16.5	6.0	4.0	21.5	12.5	5.0	23.0	15.0	10.5

Note: 6 grid points for p and q.

Table 5　Rejection Frequency under the Null: 500 trials, variation

Size (%)	Replication 1			Replication 2			Replication 3		
Bandwidth	Nominal Size (%)			Nominal Size (%)			Nominal Size (%)		
M	20	10	5	20	10	5	20	10	5
0	14.0	7.4	2.4	15.6	9.6	5.2	17.4	9.8	4.2
1	16.8	8.4	3.4	17.6	11.0	7.0	19.0	11.2	5.0
2	17.2	9.2	3.4	19.0	11.0	6.8	19.6	10.8	5.4
3	18.4	9.4	3.8	19.8	11.2	7.0	20.2	10.8	5.4
4	18.2	8.8	4.2	20.0	12.2	7.4	21.4	10.8	5.4

Note: 6 grid points for p and q.

Table 6　Estimates of Markov Switching Model: $K=4$

Parameter	Estimate	Standard Error*
μ	-0.35889	0.44941
μ_d	1.52240	0.44647
ϕ_1	0.01349	0.15913
ϕ_2	-0.05751	0.20484
ϕ_3	-0.24701	0.14413
ϕ_4	-0.21290	0.13249
σ	0.76900	0.09154
p	0.90407	0.90407
q	0.75463	0.10040
Log-Likelihood value: -183.669		# of obs.: 131

Note: The dependent variable is the rate of change of US real GNP, 1952:2 to 1984: 4. See Hamilton (1989).
* Heteroskedastic consistent estimates.

Table 7 Rejection Frequency under the Alternative: 50 trials

Power (%)	Hansen*			Replication 1		
Bandwidth	Nominal Size (%)			Nominal Size (%)		
M	20	10	5	20	10	5
0	86	80	74	86	70	58
1	86	80	74	84	74	66
2	86	76	74	90	78	70
3	86	76	74	94	80	70
4	86	76	74	94	80	70

Note: *Hansen (1992, Erratum: 1996). 6 grid points.

Table 8 Rejection Frequency under the Alternative: 50 trials, variation

Power (%)	Replication 2			Replication 3			Replication 4		
Bandwidth	Nominal Size (%)			Nominal Size (%)			Nominal Size (%)		
M	20	10	5	20	10	5	20	10	5
0	72	58	40	82	74	54	90	76	60
1	78	58	52	88	76	56	92	78	64
2	84	66	52	88	78	66	90	82	72
3	84	66	54	88	78	66	92	86	70
4	88	66	56	88	80	64	96	84	74

Note: 6 grid points for p and q.

Table 9 Rejection Frequency under the Alternative: 500 trials, variation

Power (%)	Replication 1			Replication 2			Replication 3		
Bandwidth	Nominal Size (%)			Nominal Size (%)			Nominal Size (%)		
M	20	10	5	20	10	5	20	10	5
0	82.8	70.2	55.4	84.6	73.6	60.4	85.6	72.4	59.4
1	85.8	75.0	63.0	87.0	77.2	65.0	88.0	77.4	65.6
2	87.2	77.0	64.8	88.4	79.4	67.8	88.8	80.2	67.4
3	87.6	78.2	66.6	89.6	80.2	69.4	90.0	81.6	69.0
4	88.4	78.8	68.0	90.2	81.2	71.4	90.2	82.0	63.8

Note: 6 grid points for p and q.

Table 10　Rejection Frequency : 500 trials with GP8

Bandwidth	Null (Size) Nominal Size (%)			Alternative (Power) Nominal Size (%)		
M	20	10	5	20	10	5
0	22.8	15.4	9.8	80.2	64.4	50.6
1	24.6	16.4	11.4	83.4	70.2	57.8
2	26.0	15.6	11.2	85.0	71.8	60.2
3	25.8	16.6	12.0	85.4	75.6	61.6
4	26.6	17.4	11.8	86.2	76.2	61.6

Note: 8 grid points for p and q.

Table 11　Rejection Frequency : 500 trials with GP8, Bartlett kernel

Bandwidth	Null (Size) Nominal Size (%)			Alternative (Power) Nominal Size (%)		
M	20	10	5	20	10	5
0	24.2	14.8	9.2	83.2	66.4	52.0
1	10.6	6.2	3.6	65.0	44.4	31.4
2	5.6	2.6	1.6	49.8	30.2	16.6
3	2.8	1.2	0.6	38.0	17.6	8.6
4	2.0	1.0	0.2	28.0	11.0	5.0

Note: 8 grid points for p and q. Bartlett kernel for covariance weights.