Modified Quasi-Likelihood Ratio Test for Regime Switching[*]

Hiroyuki Kasahara
Vancouver School of Economics
University of British Columbia
hkasahar@mail.ubc.ca

Tatsuyoshi Okimoto
Graduate School of
International Corporate Strategy
Hitotsubashi University
tokimoto@ics.hit-u.ac.jp

Katsumi Shimotsu
Faculty of Economics
University of Tokyo
shimotsu@e.u-tokyo.ac.jp

October 1, 2013

Abstract

In this paper, we propose a modified quasi-likelihood ratio test of the null hypothesis of one regime against the alternative of two regimes in Markov regime switching models. The asymptotic distribution of the proposed test statistic is a simple function of Gaussian random variables, and the inference is no more complicated than in the standard case. Our simulations show that the proposed test has good finite sample size and power that are comparable to the quasi-likelihood ratio test of Cho and White (2007). We apply our test to stock returns and Japanese policy functions.

Keywords: Markov regime switching; mixture model; modified likelihood ratio statistic; penalty term.

JEL Classification Numbers: C12, C13, C22.

1

1 Introduction

The Markov regime-switching model is very attractive for modeling the dynamics of economic and financial time series because it can capture many important features, such as structural changes, nonlinearity, high persistence, fat tails, leptokurtosis, and asymmetric dependence. Thus, the regime-switching models have been prevalent in economics and finance since Hamilton (1989) proposed it to describe the business cycle. For instance, Evans and Wachtel (1993) use it to analyze the inflation regime and the sources of inflation uncertainty. Hamilton and Susmel (1994) introduce the regime-switching-ARCH model to capture the volatility clustering in stock market more accurately than the GARCH model does. Similarly, Gray (1996) develops the regime-switching-GARCH model to examine the conditional distribution of interest rates. The regime-switching framework is also adopted to identify the monetary policy regime as in Sims and Zha (2006) and Inoue and Okimoto (2008). In addition, Ang and Bekaert (2002) and Okimoto (2008) employ the regime-switching model to distinguish the bear and bull regimes in international equity markets. Lastly, the regime-switching structure may be critical for evaluating asset price and for understanding the term structures of interest rates as emphasized by Dai et al. (2007).

Despite its popularity in economics and finance, few statistical tests test the null hypothesis of a single regime against the alternative hypothesis of two (or more) regimes. This is because the standard regularity conditions for the likelihood ratio test are violated by problems such as a nuisance parameter identified only under the alternative and a parameter on the boundary of the parameter space. Hansen (1992) proposes a test that treats the log-likelihood as an empirical process indexed by the parameters and takes the supremum of the log-likelihood over the nuisance parameters not identified under the null hypothesis. Cho and White (2007) derive the asymptotic distribution of the quasi-likelihood ratio (QLR) statistic by rewriting the model as a two-component mixture model, thereby ignoring the temporal dependence of the regimes. Carrasco et al. (2013) propose an information matrix-type test for parameter stability. However, these tests are computationally intensive, requiring either bootstrapping or simulating a functional of a Gaussian process, because their asymptotic distribution depends on the model as well as on the parameter space.

Building on the idea of Chen et al. (2001), in this paper, we propose a modified quasi-likelihood ratio (MQLR) statistic for testing the null hypothesis of a single

regime against the alternative hypothesis of two regimes for regime switching models with a scalar regime-specific parameter. The MQLR statistic is easy to compute and has a simple asymptotic null distribution. The asymptotic distribution is either the square of $\max\{0, N(0,1)\}^2$ or a function of a two-dimensional Gaussian random variable, and which asymptotic distribution applies to the model at hand can easily be verified. Consequently, the critical values can be easily obtained, and the inference is no more complicated than in the standard case.

We also investigate the finite sample performance of the MQLR test using Monte Carlo simulations. Our simulation results indicate that the MQLR test has similar finite sample size and power properties to the quasi-likelihood ratio test of Cho and White (2007), which requires extensive simulations for obtaining critical values.

The remainder of this paper is organized as follows. Section 2 presents the MQLR test and analyzes its asymptotic distribution. Section 3 reports simulation results. Section 4 reports empirical applications of the MQLR test to stock returns and Japanese policy functions. The technical assumptions and proofs are collected at the end of the paper in the mathematical appendix. All limits below are taken as $n \to \infty$, unless stated otherwise. Let := denote "equals by definition."

2 Modified quasi-log-likelihood test

In this section, we propose the MQLR statistic and derive its asymptotic distribution.

2.1 Model and quasi-log-likelihood

We consider the following regime switching model:

$$X_t|\mathcal{F}_{t-1} \sim \begin{cases} F(\cdot|X^{t-1}; \theta_0^*, \theta_1^*), & \text{if } S_t = 1, \\ F(\cdot|X^{t-1}; \theta_0^*, \theta_2^*), & \text{if } S_t = 2, \end{cases}$$

where $X^{t-1} := (X'_{t-1}, \ldots, X'_1)$, S_t follows a first-order Markov process, and $\mathcal{F}_{t-1}$ is the σ-algebra generated by $(X'_{t-1}, \ldots, X'_1, S_t, \ldots, S_1)$. $\theta_0 \in \mathbb{R}^{r_0}$ is a parameter common across regimes, and θ_1 and θ_2 are scalar-valued regime-specific parameters, with superscript $*$ denoting its true value.

Recently, Cho and White (2007) consider testing the null hypothesis of a single regime in this model. Because analyzing the likelihood function of this model is

3

difficult, Cho and White (2007) rewrite it as a two-term mixture model and analyze the resulting quasi-log-likelihood function. Define the conditional density of X_t as

$$f_t(\theta^j) := f(X_t | X^{t-1}; \theta^j), \quad \text{where } \theta^j := (\theta_0, \theta_j).$$

Define $\Pr(S_t = 1)$ in the stationary distribution of S_t as $\pi := \Pr(S_t = 1)$. Then, the quasi-log-likelihood function of the data can be written as a mixture model:

$$L_n(\pi, \theta) := \sum_{t=1}^{n} \ell_t(\pi, \theta), \quad \ell_t(\pi, \theta) := \log(\pi f_t(\theta^1) + (1 - \pi) f_t(\theta^2)). \tag{1}$$

Let θ_* be the unknown true value of θ_j under the null of a single regime. Partition the null of a single regime as follows:

$$H_0 : \{\pi^* = 1 \text{ and } \theta_1 = \theta_*\} \cup \{\theta_1 = \theta_2 = \theta_*\} \cup \{\pi^* = 0 \text{ and } \theta_2 = \theta_*\}$$
$$= H_{01} \cup H_{02} \cup H_{03},$$

H_{01}, H_{02}, and H_{03} violate standard assumptions of the validity of the likelihood ratio test. Under H_{01} and H_{03}, the true parameter value, π^*, lies on the boundary of the parameter space $[0, 1]$, and either θ^1 or θ^2 are not identified. Under H_{02}, the parameter π^* is not identified.

Cho and White (2007) derive the asymptotic distribution of the quasi-log-likelihood ratio (QLR) statistic by considering $H_{01} \cup H_{03}$ and H_{02} separately. The asymptotic distribution of the QLR statistic under $H_{01} \cup H_{03}$ is given by

$$\sup_{\theta_2 \in \Theta_*(\epsilon)} (\min[0, \mathcal{G}(\theta_2)])^2, \tag{2}$$

where $\Theta_*(\epsilon)$ is a subset of the parameter space for θ_2, and $\mathcal{G}(\theta_2)$ is a Gaussian process whose variance depends on the model.[1] By contrast, the asymptotic distribution of the QLR statistic under H_{02} is relatively simple. If the second-order derivative of the quasi-log-likelihood function differs from zero, the asymptotic distribution is

$$\max[0, G_0]^2, \quad G_0 \sim N(0, 1). \tag{3}$$

---

[1]We do not give a precise definition of $\Theta_*(\epsilon)$ because it is involved. See Cho and White (2007, p. 1679).

Cho and White (2007) combine these two asymptotic distributions and derive the asymptotic distribution of the QLR statistics. For example, combining (2) and (3) provides the distribution

$$\max\{\max[0, G_0]^2, \ \sup_{\theta_2}(\min[0, \mathcal{G}(\theta_2)])^2\}.$$

Practical implementation of the QLR statistic has, however, several difficulties. First, the QLR statistic has a model-dependent null distribution because $\mathcal{G}(\theta_2)$ depends on the model. Second, the asymptotic null distribution also depends on the parameter space over which the supremum is taken in (2). Third, taking the supremum over the parameter space is often difficult and requires elaborate simulations to obtain precise critical values. Consequently, obtaining reliable critical values for the QLR statistic requires extensive model-dependent simulations.

2.2 Modified quasi-log-likelihood

We propose a modified QLR (MQLR) statistic that substantially reduces the implementation problem associated with the QLR statistic. Following Chen et al. (2001), our approach adds a penalty term to the quasi-log-likelihood function. By doing so, the MQLR statistic can clear the inference problem that occurs when π is on the boundary and θ_j is not identified. With an appropriate choice of the penalty term, the modified quasi-maximum likelihood estimator of π is bounded away from 0 and 1, and the asymptotic distribution of the MQLR statistic is a simple function of standard normal random variables.

Because the model is symmetric with respect to (π, θ_1) and $(1 - \pi, \theta_2)$, we assume $\pi \in [0, 1/2]$ without the loss of generality. We consider the following modified quasi-log-likelihood function that adds a penalty term to the quasi-log-likelihood function provided in (1):

$$PL_n(\pi, \theta) := L_n(\pi, \theta) + C \log(2\pi), \tag{4}$$

where $L_n(\pi, \theta)$ is defined in (1) and C is a positive non-random number. Because the function $\log(2\pi)$ takes the value $-\infty$ when $\pi = 0$, the penalty term bounds π away from 0. Consequently, we do not need to consider the cases H_{01} and H_{03}, and the asymptotic distribution of the test statistic is substantially simplified.

Chen et al. (2001) introduce the addition of a penalty term that forces the esti-

mates of π away from 0 and 1 in the context of testing a finite mixture model. Chen and Kalbfleisch (2005), Chen (1998), and Chen et al. (2001, 2004) apply the modified likelihood function with a similar penalty term to test the number of components in various finite mixture models.

Define the unrestricted modified quasi-maximum likelihood estimator (MQMLE) as

$$(\hat{\pi}, \hat{\theta}_0, \hat{\theta}_1, \hat{\theta}_2) := \underset{(\pi, \theta_0, \theta_1, \theta_2) \in [0, 1/2] \times \Theta_0 \times \Theta_* \times \Theta_*}{\arg\max} PL_n(\pi, \theta),$$

where Θ_0 and Θ_* are the parameter space of θ_0 and θ_j $(j = 1, 2)$, respectively. Define the restricted MQMLE imposing H_0 as

$$(\hat{\theta}_0^n, \hat{\theta}_1^n) := \underset{(\theta_0, \theta_1) \in \Theta_0 \times \Theta_*}{\arg\max} \sum_{t=1}^{n} \log(f_t(\theta^1)),$$

where the superscript n is used to denote the null-imposing quasi-maximum likelihood estimator. The technical assumptions in this paper are essentially the same as those in Cho and White (2007).

The following proposition shows that $\hat{\pi}$ is bounded away from 0 in probability.

Proposition 1. *Suppose that Assumptions 1–5 and either Assumption 6 or 7 in the mathematical appendix hold, and that H_0 holds. Then, $\log(2\hat{\pi}) = O_p(1)$.*

Using Proposition 1, we can analyze the asymptotic distribution of the MQLR statistic.

2.3 Asymptotic distribution of the MQLR statistic: nonzero second-order derivative case

Define the MQLR statistic as

$$MQLR_n := PL_n(\hat{\pi}, \hat{\theta}_0, \hat{\theta}_1, \hat{\theta}_2) - PL_n(1/2, \hat{\theta}_0^n, \hat{\theta}_1^n, \hat{\theta}_1^n).$$

As shown in Cho and White (2007), the asymptotic behavior of the quasi-log-likelihood ratio depends on whether the quasi-log-likelihood function has zero second-order derivatives under the null. The following proposition establishes the asymptotic distribution of the MQLR statistic when the quasi-log-likelihood function has nonzero second-order derivatives under the null.

Proposition 2. *Suppose that Assumptions 1–6 in the mathematical appendix hold and that H_0 holds. Then, $MQLR_n \to_d \max[0, G_0]^2$, where $G_0 \sim N(0, 1)$.*

In this case, the asymptotic distribution of the MQLR statistic can be easily tabulated. The 90%, 95%, and 99% critical values are 1.64, 2.71, and 5.41, respectively.

2.4 Asymptotic distribution of the MQLR statistic: zero second-order derivative case

The asymptotic distribution of the MQLR statistic differs from that given in Proposition 2 when the quasi-log-likelihood function has zero second-order derivatives under the null. This occurs when the second-order derivative is a linear function of the first-order derivatives. Namely, for each θ^1 and for some nonzero $(\alpha', \beta)' \in \mathbb{R}^{r_0+1}$,

$$f_t^{(0,2)}(\theta^1) = \alpha f_t^{(1,0)}(\theta^1) + \beta f_t^{(0,1)}(\theta^1) \tag{5}$$

where $f_t^{(i,j)}(\theta^1) := \nabla_{\theta_0}^i \nabla_{\theta_1}^j f_t(\theta_0, \theta_1)$ with $\theta^1 = (\theta_0, \theta_1)$. The degeneracy (5) occurs, for example, when $f_t(\theta_0, \theta_1)$ is the normal density with type-specific mean θ_1 and a common but unknown variance θ_0. Under (5), the quasi-log-likelihood function must be expanded eight times to obtain its asymptotic distribution.

The following proposition establishes the asymptotic distribution of the MQLR statistic in the zero-second order derivative case. Define

$$\begin{aligned}
r_{t*}^{(i,j)} &:= \nabla_{\theta_0}^i \nabla_{\theta_1}^j f_t(\theta_0^*, \theta_*)/f_t(\theta_0^*, \theta_*), \\
s_{t*} &:= r_{t*}^{(0,4)} - 6\beta r_{t*}^{(0,3)} - 6\alpha' r_{t*}^{(1,2)} + 6\alpha' r_{t*}^{(1,1)}\beta + 3\alpha' r_{t*}^{(2,0)}\alpha, \tag{6} \\
r_{t*}^{(1)} &:= \nabla_{\theta^1} f_t(\theta_0^*, \theta_*)/f_t(\theta_0^*, \theta_*).
\end{aligned}$$

$r_{t*}^{(1)}$ corresponds to the score of the one-component model. Collect the variances and covariances of $(r_{t*}^{(0,3)}, s_{t*}, r_{t*}^{(1)})$ as

$$\begin{bmatrix} C_{33} & C_{3s} & C_{31} \\ C_{s3} & C_{ss} & C_{s1} \\ C_{13} & C_{1s} & C_{11} \end{bmatrix} := \begin{bmatrix} E[(r_{t*}^{(0,3)})^2] & E[r_{t*}^{(0,3)} s_{t*}] & E[r_{t*}^{(0,3)} r_{t*}^{(1)'}] \\ E[s_{t*} r_{t*}^{(0,3)}] & E[s_{t*}^2] & E[s_{t*} r_{t*}^{(1)'}] \\ E[r_{t*}^{(1)} r_{t*}^{(0,3)}] & E[r_{t*}^{(1)} s_{t*}] & E[r_{t*}^{(1)} r_{t*}^{(1)'}] \end{bmatrix},$$

and define

$$\Omega^{(3)} := C_{33} - C_{31}C_{11}^{-1}C_{13}, \quad \Omega^{(s)} := C_{ss} - C_{s1}C_{11}^{-1}C_{1s}, \quad \Omega^{(3,s)} := C_{3s} - C_{31}C_{11}^{-1}C_{1s}.$$

Proposition 3. *Suppose that Assumptions 1–5 and 7 in the mathematical appendix hold and that H_0 holds. Then, $MQLR_n \to_d \max[G_0^2, \max[0, G_*]^2]$, where $(G_0, G_*) \sim N(0, V)$, where the diagonal elements of V equal 1 and $cov(G_0, G_*) = \Omega^{(3,s)}/[\Omega^{(3)}\Omega^{(s)}]^{1/2}$.*

Although the asymptotic distribution of the MQLR statistic is more complicated than the nonzero second-order derivative case, the inference is still substantially easier than in the QLR statistic. Because the matrices $\Omega^{(3,s)}$, $\Omega^{(3)}$, and $\Omega^{(s)}$ can be consistently estimated from the data, the asymptotic distribution of the MQLR statistic can be easily simulated.

3 Simulations

In this section, we examine the finite sample properties of the MQLR statistic using several Monte Carlo simulations. We consider three cases: Case 1: Constant change model; Case 2: Coefficient change model; and Case 3: Variance change model. Case 1 belongs to the zero second-order derivative case, and Cases 2 and 3 belong to the nonzero second-order derivative case. We explain each case in detail sequentially. Regarding the constant C in the penalty term, Chen et al. (2001) recommend using $C = 1$. We set $C = 1$ in simulations and empirical applications unless stated otherwise.

In Case 1, we consider the same simulation environment as Cho and White (2007), and, thus, our results are comparable with theirs. The model for $X_t|\mathcal{F}_t$ is an AR(1) model with possible Markov regime-switching in the constant term and a common unknown variance, $\pi \cdot N(\theta_1 + \theta_{01}X_{t-1}, \sigma^2) + (1 - \pi) \cdot N(\theta_2 + \theta_{01}X_{t-1}, \sigma^2)$, where $\theta_0 = (\theta_{01}, \sigma^2)'$ with its true value $\theta_0^* = (\theta_{01}^*, \sigma^{2*})'$. For the size simulation, X_t follows an AR(1) process, $X_t = 0.5X_{t-1} + u_t$ with $u_t \sim$ i.i.d. $N(0, 1)$. For the power simulation, X_t is generated by $X_t = -\beta_* \cdot \mathbf{1}_{\{S_t=1\}} + \beta_* \cdot \mathbf{1}_{\{S_t=2\}} + 0.5X_{t-1} + u_t$ with $u_t \sim$ i.i.d. $N(0, 1)$ and $P(S_t = 1|S_{t-1} = 1) = P(S_t = 2|S_{t-1} = 2) = p^*$, where $p^* \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$ and $\beta_* \in \{0.2, 0.4, 0.6, 0.8, 1.0\}$.

The normal density with mean μ and variance σ^2, $f(x; \mu, \sigma^2)$, has the property $\nabla_{\mu\mu} f(x; \mu, \sigma^2) = 2\nabla_{\sigma^2} f(x; \mu, \sigma^2)$. Consequently, Case 1 belongs to the zero second-

order derivative case, and the asymptotic distribution of the MQLR statistic is given by Proposition 3 with $\alpha = (0, 2)$ and $\beta = 0$, such that

$$
\begin{aligned}
r_{t*}^{(1)} &= [\nabla_\mu \log(f(X_t; \mu, \sigma^2)), \nabla_\mu \log(f(X_t; \mu, \sigma^2))X_{t-1}, \nabla_{\sigma^2} \log(f(X_t; \mu, \sigma^2))]', \\
r_{t*}^{(0,3)} &= \nabla_{\mu\mu\mu} f(X_t; \mu, \sigma^2)/f(X_t; \mu, \sigma^2), \\
s_{t*} &= \frac{\nabla_{\mu\mu\mu\mu} f(X_t; \mu, \sigma^2) - 12\nabla_{\mu\mu\sigma^2} f(X_t; \mu, \sigma^2) + 12\nabla_{\sigma^2\sigma^2} f(X_t; \mu, \sigma^2)}{f(X_t; \mu, \sigma^2)},
\end{aligned}
\tag{7}
$$

evaluated at $(\mu, \sigma^2) = (\theta_* + \theta_{01}^* X_{t-1}, \sigma^{2*})$. As shown in the following proposition, G_* and G_0 are mutually independent in Case 1. The 90th, 95th, and 99th percentile of the asymptotic distribution of the MQLR statistic are 3.33, 4.51, and 7.35, respectively, and the p-values can be easily simulated.

Proposition 4. *Suppose* $f(X_t|X^{t-1}, \pi, \theta_0, \theta_1, \theta_2) = \pi \cdot f(X_t; \theta_1 + \theta_{01}X_{t-1}, \sigma^2) + (1 - \pi) \cdot f(X_t; \theta_2 + \theta_{01}X_{t-1}, \sigma^2)$, *where* $\theta_0 = (\theta_{01}, \sigma^2)'$ *and* $f(x; \mu, \sigma^2)$ *is the normal density with mean* μ *and variance* σ^2. *Then, under the assumptions of Proposition 3,* $MQLR_n \to_d \max[G_0^2, \max\{0, G_*\}^2]$, *where* G_0 *and* G_* *are two independent* $N(0, 1)$ *random variables.*

Table 1 reports results of a size simulation for Case 1 computed using 5000 replications. The numbers in parentheses are the results reported in Table II of Cho and White (2007). The MQLR test is comparable to the QLR test of Cho and White (2007) with a reasonable size for sample size $n \geq 100$. When $n = 50$, both tests reject the null too often, but are acceptable given such a small sample size.

We also examined the finite sample size of the MQLR test under $C = 0.5$ and $C = 2$. The results are reported in Table 2. When $C = 0.5$, the MQLR test is oversized, whereas, when $C = 2$, it is undersized. Chen et al. (2001) note that their modified likelihood ratio test is not sensitive to the values of C, but our results do not corroborate theirs. Further investigation on the optimal choice of C is an important future research area.

Results of power simulation for Case 1 are shown in Table 3. The numbers in parentheses are the results reported in Table III of Cho and White (2007). When $\beta_* \leq 0.6$, the power of the MQLR test is similar to the power of the QLR test regardless of the value of p^*. When $\beta_* \geq 0.8$, the power of the MQLR test depends on π^*, and the MQLR test has less power when π^* is close to 0 or 1. The QLR test also has less power when π^* is close to 0 or 1, even though the relationship between

9

its power and π^* is different from that of the MQLR test.

When π is close to 0 or 1, the MQLR test may lose power against the QLR test, because the QLR test tests both $H_{01} : \theta_1 = \theta_2$ and $H_{02} : \pi(1 - \pi) = 0$. Table 4 compares the power of the MQLR test and the QLR test at the 5% nominal level for Case 1 when p^* is close to 0 and 1. When p^* is small and $\beta_* \geq 0.6$, the MQLR test has a stronger power than the QLR test. The QLR test tends to be more powerful than the MQLR test when p^* is large or β_* is small. Overall, neither the MQLR test nor the QLR test dominates the other.

To investigate the finite sample properties of the MQLR statistic in more depth, we consider the coefficient change model as Case 2. The model for $X_t | \mathcal{F}_t$ is $\pi \cdot N(\theta_{01} + \theta_1 X_{t-1}, \sigma^2) + (1 - \pi) \cdot N(\theta_{01} + \theta_2 X_{t-1}, \sigma^2)$, where $\theta_0 = (\theta_{01}, \sigma^2)'$. Case 2 belongs to the nonzero second-order derivative case, and the asymptotic distribution of the MQLR statistic is given as $\max\{0, N(0, 1)\}^2$. For the size simulation, we use $X_t = 0.5X_{t-1} + u_t$ with $u_t \sim$ i.i.d. $N(0, 1)$. For the power simulation, we set the DGP as $X_t = 1 + (0 \cdot \mathbf{1}_{\{S_t=1\}} + \beta_* \cdot \mathbf{1}_{\{S_t=2\}})X_{t-1} + u_t$ with $P(S_t = 1|S_{t-1} = 1) = P(S_t = 2|S_{t-1} = 2) = p^*$, where $\beta_* \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$ and $p^* \in \{0.1, 0.2, 0.3, 0.4, 0.5\}$.

Tables 5 and 6 report the size and power simulation results for Case 2.[2] The MQLR test has a good size even for a small n. Further, the MQLR test has a good power, which does not substantially depend on π^*.

The last case considered in this paper is a variance change model (Case 3). In this case, the null model is the same as before, whereas the alternative model is $\pi \cdot N(\theta_{01} + \theta_{02}X_{t-1}, \theta_1^2) + (1 - \pi) \cdot N(\theta_{01} + \theta_{02}X_{t-1}, \theta_2^2)$, where $\theta_0 = (\theta_{01}, \theta_{02})'$. For the power simulation, the DGP is formulated as $X_t = 0.5X_t + u_t$ with $u_t \sim N(0, 1^2 \cdot \mathbf{1}_{\{S_t=1\}} + \beta_*^2 \cdot \mathbf{1}_{\{S_t=2\}})$, $P(S_t = 1|S_{t-1} = 1) = P(S_t = 2|S_{t-1} = 2) = p^*$, where $\beta_* \in \{1.1, 1.2, 1.3, 1.4, 1.5\}$ and $p^* \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$.

Table 7 reports the results of the size simulation for Case 3. The MQLR test overrejects slightly for $n \leq 100$ but performs well otherwise. Table 8 shows that the MQLR test has good power properties for all values of p^*.

In summary, our proposed MQLR test has very good finite sample size and power properties in many cases. The simulation results demonstrate the attractiveness of the MQLR test because of the simplicity of its procedure.

---

[2]For Cases 2 and 3, we provide only the results of the MQLR test, because very intensive simulations are needed to obtain the critical values of the QLR statistic.

4 Empirical Applications

In this section, we apply our MQLR test to stock returns and Japanese policy functions.

4.1 Stock returns

Our stock return application is based on Ang and Bekaert (2002) and Okimoto (2008). Both studies use a regime switching model to examine the differences in the characteristics between bear and bull markets in international equity markets. Specifically, they assume the existence of two regimes, and test the equivalence of the expected return and volatilities across the regimes. Following these studies, we test regime switching on expected return and volatility by considering the following simple model for the stock return:

$$r_t = \pi \cdot N(\mu_1, \sigma_1^2) + (1 - \pi) \cdot N(\mu_2, \sigma_2^2). \tag{8}$$

This model corresponds to both studies' models used to describe the marginal behavior of one country's stock returns. Ang and Bekaert (2002) use data from the United States, the United Kingdom, and Germany. They reject the equivalence of expected returns across regimes only for the United States at the 10% significance level, and reject the equivalence of volatility across regimes for all three countries. By contrast, Okimoto (2008) rejects the equivalence of expected returns across regimes and volatilities across regimes for all G6 countries.

We use the monthly MSCI stock indices from Bloomberg for a sample period from 1983:1 to 2012:9 . The countries examined are the same as in Okimoto (2008): Canada, France, Germany, Japan, the United Kingdom, and the United States. When testing the null of no regime switching in expected returns, namely, $\mu_1 = \mu_2$, we assume $\sigma_1 = \sigma_2$ and use critical values computed from the asymptotic distribution given by Proposition 4. When testing the null of no regime switching in volatility, namely, $\sigma_1 = \sigma_2$, we assume $\mu_1 = \mu_2$ and use the critical values from $\max\{0, N(0,1)\}^2$. Table 9 reports the results of the MQLR test. We reject the null of no regime switching in expected returns at the 10% significance level for all countries except for Japan. Thus, our results for expected returns are more consistent with those of Okimoto (2008). In contrast, the null of no regime switching in volatility is rejected for all

countries as in Ang and Bekaert (2002) and Okimoto (2008).

4.2 Japanese policy functions

Our second empirical application is based on Doi et al. (2011, hereafter DHO). DHO consider regime switching for several Japanese policy functions to evaluate the sustainability of Japanese fiscal policy. The first function estimated by DHO is a Bohn (1998)-type fiscal policy function given by

$$x_t = \alpha + \beta b_{t-1} + \gamma_1 GVAR_t + \gamma_2 gap_t + \rho_1 x_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim \text{i.i.d. } N(0, \sigma^2), \qquad (9)$$

where x is the ratio of primary surplus to GDP, b is the debt-to-GDP ratio, $GVAR$ is the temporary deviation from the trend level of government expenditure divided by GDP, and gap is the output gap measured by the deviation from the trend divided by GDP.[3] As Bohn (1998) emphasized, the primary surplus should respond positively to the debt-to-GDP ratio, namely, $\beta > 0$, for the fiscal policy to be sustainable. Expanding the analysis of Bohn (1998) and its application to Japan by Doi and Ihori (2009), DHO consider regime switching for the equation (9) and find that while two distinct regimes exist, neither of them are sustainable. To show the appropriateness of regime switching, DHO use the Markov switching criterion proposed by Smith et al. (2006), but conduct no tests for regime switching. Therefore, we apply the MQLR test to the equation (9). Table 10 reports the results in which we test the null of no regime switching in the coefficient of *each regressor* while assuming the other parameters are identical across two regimes. Critical values are computed from the asymptotic distribution given by Proposition 4 for a constant term and from $\max\{0, N(0,1)\}^2$ for the other parameters. The null of no regime switching is rejected for all coefficients except for GVAR. Our results are broadly consistent with the results of DHO, whose estimates of the coefficient of GVAR do not significantly differ across regimes.

Another fiscal policy function examined by DHO is that of Davig and Leeper (2007) given by

$$\tau_t = \alpha + \beta b_{t-1} + \gamma_1 gap_t + \gamma_2 g_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma^2), \qquad (10)$$

---

[3]DHO use the method proposed by Hodrick and Prescott (1997) to calculate the trend level of variables.

where τ is the tax revenue-to-GDP ratio, and g is the government expenditure-to-GDP ratio. According to the terminology originally developed by Leeper (1991), an "active" policy is not constrained by the level of government debt, which implies that an active policy is characterized by a positive coefficient on the debt-to-GDP ratio ($\beta > 0$). For the US data, Davig and Leeper (2007) find that the fiscal policy alternates between the "active" and "passive" phase, while DHO find two distinct active regimes for Japan with $\beta < 0$ for one regime and $\beta = 0$ for the other. To confirm the existence of regime switching in the equation (10) for Japan, we conduct the MQLR test for each coefficient. As seen in Table 11, the results demonstrate that the regime switching is significant for all coefficients except for a constant term. This is consistent with DHO, because their estimates of the constant term do not significantly differ across regimes.

DHO also estimate the monetary policy function expressed by the following Taylor rule:

$$call_t = \alpha + \beta \pi_t + \gamma_1 gap_t + \gamma_2 ex_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma^2). \tag{11}$$

Here, $call$ is the quarterly average overnight call rate (nominal), π is the inflation rate measured as the rate of change of the GDP deflator from four quarters ago, and ex is the deviation of the real effective exchange rate from its Hodrick-Prescott trend. Monetary policy is said to be active if $\beta > 1$ and inactive (i.e., passive) if $\beta < 1$. DHO demonstrate that there is a permanent regime change after the third quarter of 1995; however, the Japanese monetary policy is inactive for both regimes. We test the regime switching in the equation (11) to see whether we can obtain additional evidence of the regime switching. The results reported in Table 12 illustrate that the regime switching is significant for the constant and exchange rate, suggesting that the low interest policy introduced by the Bank of Japan in 1995 is one of the main causes of a regime change. In addition, the insignificance of the regime switching in the inflation implies that there is no significant change in the monetary policy reaction to the inflation. These results are broadly consistent with those of DHO.

5 Mathematical Appendix

The assumptions are essentially the same as those in Cho and White (2007).

Assumption 1.

(i) The observable random variables $\{X_t \in \mathbb{R}^d\}_{t=1}^n$ are generated as a sequence of strictly stationary β-mixing random variables such that for some $c > 0$ and $\rho \in [0,1)$, the β-mixing coefficient, β_τ, is at most $c\rho^\tau$.

(ii) The sequence of unobserved regime indicators, $\{S_t \in \{1,2\}\}_{t=1}^n$, is generated as a first-order Markov process such that $Pr(S_t = j|S_{t-1} = i) = p_{ij}^*$ with $p_{ii}^* \in [0,1]$ $(i, j = 1, 2)$.

(iii) The given $\{X_t\}$ is a Markov regime-switching process: for some $\theta^* := (\theta_0^*, \theta_1^*, \theta_2^*) \in \mathbb{R}^{r_0+2}$,

$$X_t|\mathcal{F}_{t-1} \sim \begin{cases} F(\cdot|X^{t-1}; \theta_0^*, \theta_1^*), & \text{if } S_t = 1, \\ F(\cdot|X^{t-1}; \theta_0^*, \theta_2^*), & \text{if } S_t = 2, \end{cases}$$

where $\mathcal{F}_{t-1} := \sigma(X^{t-1}, S^t)$ is the σ-algebra generated by $(X_{t-1}', \ldots, X_1', S_t, \ldots, S_1)$. The conditional cumulative distribution function (CDF) of $X_t|\mathcal{F}_{t-1}$, $F(\cdot|X^{t-1}; \theta_0^*.\theta_j^*)$, has a probability density function (PDF) $f(\cdot|X^{t-1}; \theta_0^*, \theta_j^*)$ $(j = 1, 2)$. Further, for $(p_{11}^*, p_{22}^*) \in [0,1) \times [0,1) \backslash \{(0,0)\}$, θ^* is unique in $\mathbb{R}^{r_0+2}$.

Assumption 2.

(i) A model for $f(\cdot|X^{t-1}; \theta_0^*.\theta_j^*)$ is $\{f(\cdot|X^{t-1}; \theta^j) : \theta^j := (\theta_0, \theta_j) \in \tilde{\Theta}\}$, where $\tilde{\Theta} := \Theta_0 \times \Theta_* \in \mathbb{R}^{r_0+1}$, and Θ_0 and Θ_* are convex and compact sets in $\mathbb{R}^{r_0}$ and $\mathbb{R}$, respectively. Further, for each $\theta^j \in \tilde{\Theta}$, $f(\cdot|X^{t-1}; \theta^j)$ is a measurable PDF with CDF $F(\cdot|X^{t-1}; \theta^j)$ $(j = 1, 2)$.

(ii) For every $x \in \mathbb{R}^d$, $f(x|X^{t-1}; \cdot)$ is eight times continuously differentiable on $\tilde{\Theta}$ almost surely.

Assumption 3. (θ_0^*, θ_*) maximizes $n^{-1}E[\sum_{t=1}^n \log(f_t(\theta^1))]$ uniquely in the interior of $\tilde{\Theta}$.

Assumption 4. For all $(\pi, \theta) \in [0,1] \times \Theta$, $n^{-1}E[\sum_{t=1}^n \ell_t(\pi, \theta)]$ exists and is finite.

Assumption 5.

(i) There exists a sequence of positive, strictly stationary, and ergodic random variables, $\{M_t\}$, such that (a) for some $\delta > 0$, $E[M_t^{1+\delta}] < \Delta < \infty$, (b) $\sup_{(\pi,\theta) \in [0,1] \times \Theta} ||\nabla_{(\pi,\theta)}\ell_t(\pi, \theta)\nabla_{(\pi,\theta)}\ell_t(\pi, \theta)'||_\infty \le M_t$, and (c) $\sup_{(\pi,\theta) \in [0,1] \times \Theta} ||\nabla_{(\pi,\theta)}^2\ell_t(\pi, \theta)||_\infty \le M_t$.

(ii) There exists a sequence of positive, strictly stationary, and ergodic random variables, $\{M_t\}$, such that for some $\delta > 0$, $E[M_t^{1+\delta}] < \Delta < \infty$;

$\sup_{\theta^1 \in \tilde{\Theta}} |\nabla_{i_1} \cdots \nabla_{i_k} f_t(\theta^1)/f_t(\theta^1)|^4 \leq M_t$; $\sup_{\theta^1 \in \tilde{\Theta}} |\nabla_{i_1} \cdots \nabla_{i_\ell} f_t(\theta^1)/f_t(\theta^1)|^2 \leq M_t$; $\sup_{\theta^1 \in \tilde{\Theta}} |\nabla_{\theta_1}^8 f_t(\theta^1)/f_t(\theta^1)| \leq M_t$; $\sup_{\theta^1 \in \tilde{\Theta}} |\nabla_{j_1} \nabla_{\theta_1}^7 f_t(\theta^1)/f_t(\theta^1)| \leq M_t$; $\sup_{\theta^1 \in \tilde{\Theta}} |\nabla_{j_1} \nabla_{j_2} \nabla_{\theta_1}^6 f_t(\theta^1)/f_t(\theta^1)| \leq M_t$, where $k = 1, 2, 3, 4$, $\ell = 5, 6, 7$, $i_1, \ldots, i_7 \in \{\theta_{01}, \theta_{02}, \ldots, \theta_{0r_0}, \theta_1\}$, and $j_1, j_2 \in \{\theta_{01}, \theta_{02}, \ldots, \theta_{0r_0}\}$.

Define $B(\pi, \theta) := E[\nabla_{(\pi,\theta)} \ell_t(\pi, \theta) \nabla_{(\pi,\theta)} \ell_t(\pi, \theta)']$, and let $\lambda_{\max}(\cdot)$ and $\lambda_{\min}(\cdot)$ respectively denote the maximum and the minimum eigenvalues of a given matrix. Define $r_t^{(0,j)}(\theta_*) := \nabla_{\theta^1}^j f_t(\theta_0^*, \theta_*)/f_t(\theta_0^*, \theta_*)$ and $r_t^{(u)}(\theta_2) := [1 - r_t(\theta_2), r_t^{(1)}(\theta_2)']'$, where $r_t^{(1)}(\theta_2) := \nabla_{\theta^1} f_t(\theta_0^*, \theta_2)/f_t(\theta_0^*, \theta_*)$. Define s_{t*} as in (6).

Assumption 6. *For each* $(\pi^*, \theta_0^*, \theta_1^*, \theta_2^*)$, $\lambda_{\min}(B(\pi^*, \theta_0^*, \theta_1^*, \theta_2^*)) \geq 0$ *such that (a) if* $\lambda_{\min}(B(\pi^*, \theta_0^*, \theta_1^*, \theta_2^*)) > 0$, *then* $\lambda_{\max}(B(\pi^*, \theta_0^*, \theta_1^*, \theta_2^*)) < \infty$ *or (b) if* $\lambda_{\min}(B(\pi^*, \theta_0^*, \theta_1^*, \theta_2^*)) = 0$, *then for each* $\theta_2 \neq \theta_*$ *and* $\theta_2' \neq \theta_*$, $\lambda_{\min} C^{(u)}(\theta_2, \theta_2') > 0$, $\lambda_{\max} C^{(u)}(\theta_2, \theta_2') < \infty$, *where*

$$C^{(u)}(\theta_2, \theta_2') := \begin{bmatrix} E[r_t^{(0,2)}(\theta_*)^2] & E[r_t^{(0,2)}(\theta_*)r_t^{(u)}(\theta_2')']' \\ E[r_t^{(u)}(\theta_2)r_t^{(0,2)}(\theta_*)] & E[r_t^{(u)}(\theta_2)r_t^{(u)}(\theta_2')'] \end{bmatrix}.$$

Assumption 7. *For each* $(\pi^*, \theta_0^*, \theta_1^*, \theta_2^*)$, $\lambda_{\min}(B(\pi^*, \theta_0^*, \theta_1^*, \theta_2^*)) \geq 0$ *such that (a) if* $\lambda_{\min}(B(\pi^*, \theta_0^*, \theta_1^*, \theta_2^*)) > 0$, *then* $\lambda_{\max}(B(\pi^*, \theta_0^*, \theta_1^*, \theta_2^*)) < \infty$ *or (b) if* $\lambda_{\min}(B(\pi^*, \theta_0^*, \theta_1^*, \theta_2^*)) = 0$, *then for each* $\theta_2 \neq \theta_*$ *and* $\theta_2' \neq \theta_*$, $\lambda_{\min} C^{(v)}(\theta_2, \theta_2') > 0$ *and* $\lambda_{\max} C^{(v)}(\theta_2, \theta_2') < \infty$, *where*

$$C^{(v)}(\theta_2, \theta_2') := \begin{bmatrix} E[s_{t*}^2] & E[s_{t*} r_t^{(0,3)}(\theta_*)] & E[s_{t*} r_t^{(u)}(\theta_2')'] \\ E[r_t^{(0,3)}(\theta_*) s_{t*}] & E[r_t^{(0,3)}(\theta_*)^2] & E[r_t^{(0,3)}(\theta_*) r_t^{(u)}(\theta_2')'] \\ E[r_t^{(u)}(\theta_2) s_{t*}] & E[r_t^{(u)}(\theta_2) r_t^{(0,3)}(\theta_*)] & E[r_t^{(u)}(\theta_2) r_t^{(u)}(\theta_2')'] \end{bmatrix}.$$

5.1 Proof of Proposition 1

The proof uses a similar argument to that of Lemma 1 in Chen et al. (2001). Let

$$\begin{aligned} QLR_n(\pi, \theta) &:= [L_n^*(\pi, \theta) - L_n^*(1/2, \hat{\theta}_0^n, \hat{\theta}_1^n)], \\ MQLR_n(\pi, \theta) &:= [PL_n^*(\pi, \theta) - L_n^*(1/2, \hat{\theta}_0^n, \hat{\theta}_1^n)], \end{aligned} \tag{12}$$

be the quasi-log-likelihood ratio and modified quasi-log-likelihood ratio with respect to the maximized one-component log-likelihood, respectively. From (4), we have $MQLR_n(\pi, \theta) = QLR_n(\pi, \theta) + C \log(2\pi)$. Let $(\hat{\pi}^q, \hat{\theta}^q) := \arg\max_{\pi,\theta} L_n^*(\pi, \theta) = \arg\max_{\pi,\theta} QLR_n(\pi, \theta)$ denote the maximum quasi-likelihood estimator (QMLE) of (π, θ).

First, observe that

$$0 \leq MQLR_n(\hat{\pi}, \hat{\theta}) \leq QLR_n(\hat{\pi}, \hat{\theta}) \leq QLR_n(\hat{\pi}^q, \hat{\theta}^q), \tag{13}$$

where the second and third inequalities hold because $C \log(2\pi) \leq 0$ and $(\hat{\pi}^q, \hat{\theta}^q)$ is the QMLE. Cho and White (2007, Theorem 6) show that $QLR_n(\hat{\pi}^q, \hat{\theta}^q) = O_p(1)$. Hence, it follows from (13) that $MQLR_n(\hat{\pi}, \hat{\theta}) = O_p(1)$.

Second, because $(\hat{\pi}^q, \hat{\theta}^q)$ is the QMLE and $MQLR_n(\pi, \theta) = QLR_n(\pi, \theta) + C \log(2\pi)$, we have $0 \leq MQLR_n(\hat{\pi}, \hat{\theta}) - C \log(2\hat{\pi}) \leq QLR_n(\hat{\pi}^q, \hat{\theta}^q)$. Because $MQLR_n(\hat{\pi}, \hat{\theta}) = O_p(1)$ and $QLR_n(\hat{\pi}^q, \hat{\theta}^q) = O_p(1)$, we have $C \log(2\hat{\pi}) = O_p(1)$, giving the stated result. $\square$

5.2 Proof of Proposition 2

We prove the stated result by showing $Z_n + o_p(1) \leq MQLR_n(\hat{\pi}, \hat{\theta}) \leq Z_n + o_p(1)$, where Z_n is a random variable that converges to $\max[0, G_0]^2$ in distribution.

Define $QLR_n(\pi, \theta)$ and $MQLR_n(\pi, \theta)$ as in the proof of Lemma 1. Let $\hat{\theta}_{1/2}$ be the maximizer of $MQLR_n(\pi, \theta)$ while fixing $\pi = 1/2$, and let $\hat{\theta}_{1/2}^q$ be the maximizer of $QLR_n(\pi, \theta)$ while fixing $\pi = 1/2$. Because $\log(2\pi) = 0$ when $\pi = 1/2$, we have $\hat{\theta}_{1/2} = \hat{\theta}_{1/2}^q$. Therefore, $MQLR_n(\hat{\pi}, \hat{\theta}) \geq MQLR_n(1/2, \hat{\theta}_{1/2}) = QLR_n(1/2, \hat{\theta}_{1/2}^q)$. On the other hand, from (13) in the proof of Lemma 1, we have $MQLR_n(\hat{\pi}, \hat{\theta}) \leq QLR_n(\hat{\pi}, \hat{\theta})$. Therefore, $QLR_n(1/2, \hat{\theta}_{1/2}^q) \leq MQLR_n(\hat{\pi}, \hat{\theta}) \leq QLR(\hat{\pi}, \hat{\theta})$.

Observe that $\hat{\pi}$ is bounded away from 0 in probability because $C \log(2\hat{\pi}) = O_p(1)$ from Lemma 1. Consequently, from Theorem 4(a) of Cho and White (2007), both $QLR_n(\hat{\pi}, \hat{\theta})$ and $QLR_n(1/2, \hat{\theta}_{1/2}^q)$ converge to the same limit Z_n that is asymptotically distributed as $\max[0, G_0]^2$. Therefore, $MQLR_n(\hat{\pi}, \hat{\theta}) \rightarrow_d \max[0, G_0]^2$ follows. $\square$

5.3 Proof of Proposition 3

The proof is the same as the proof of Proposition 2, except it uses Theorem 5 of Cho and White (2007) in place of Theorem 4(a) of Cho and White (2007). $\square$

5.4 Proof of Proposition 4

For a standard normal density $\phi(x)$, its n-th derivative is $(-1)^n H_n(x)\phi(x)$, where $H_k(x)$ is the Hermite polynomial of degree k, e.g., $H_1(x) = x$, $H_2(x) = x^2 - 1$,

$H_3(x) = x^3 - 3x$, $H_4(x) = x^4 - 6x^2 + 3$. Using this property and that $\nabla_{\mu\mu} f(x; \mu, \sigma^2) = 2\nabla_{\sigma^2} f(x; \mu, \sigma^2)$, the scores $(r_{t*}^{(1)}, r_{t*}^{(0,3)}, s_{t*})$ defined in (7) can be written as

$$r_{t*}^{(1)} = \left[\frac{H_1(\varepsilon_t)}{\sigma^*}, \frac{H_1(\varepsilon_t) X_{t-1}}{\sigma^*}, \frac{H_2(\varepsilon_t)}{2\sigma^{*2}} \right]', \quad r_{t*}^{(0,3)} = \frac{1}{\sigma^{3*}} H_3(\varepsilon_t), \quad s_{t*} = -\frac{2}{\sigma^{*4}} H_4(\varepsilon_t),$$

where $\varepsilon_t = (X_t - \theta_* - \theta_{01}^* X_{t-1})/\sigma^* \sim N(0, 1)$. Because ε_t and X_{t-1} are independent and Hermite polynomials of different orders are orthogonal, $r_{t*}^{(1)}$, $r_{t*}^{(0,3)}$, and s_{t*} are mutually uncorrelated. Consequently, $C_{31} = 0$, $C_{1s} = 0$, $\Omega^{(3,s)} = C_{3s} = 0$, and G_* and G_0 are mutually independent. $\square$

References

Ang, A. and Bekaert, G. (2002), "International Asset Allocation with Regime Shifts," *Review of Financial Studies*, 15, 1137–1187.

Bohn, H. (1998), "The Behavior of U.S. Public Debt and Deficits," *Quarterly Journal of Economics*, 113, 949–963.

Carrasco, M., Hu, L., and Ploberger, W. (2013), "Optimal Test for Markov Switching Parameters," University of Montreal.

Chen, H., Chen, J., and Kalbfleisch, J. D. (2001), "A Modified Likelihood Ratio Test for Homogeneity in Finite Mixture Models," *Journal of the Royal Statistical Society, Series B*, 63, 19–29.

— (2004), "Testing for a Finite Mixture Model with Two Components," *Journal of the Royal Statistical Society, Series B*, 66, 95–115.

Chen, J. (1998), "Penalized Likelihood-ratio Test for Finite Mixture Models with Multinomial Observations," *Canadian Journal of Statistics*, 26, 583–559.

Chen, J. and Kalbfleisch, J. D. (2005), "Modified Likelihood Ratio Test in Finite Mixture Models with a Structural Parameter," *Journal of Statistical Planning and Inference*, 129, 93–107.

Cho, J. S. and White, H. (2007), "Testing for Regime Switching," *Econometrica*, 75, 1671–1720.

Dai, Q., Singleton, K., and Yang, W. (2007), "Regime Shifts in a Dynamic Term Structure Model of U.S. Treasury Bond Yields." *Review of Financial Studies*, 20, 1669–1706.

Davig, T. and Leeper, E. M. (2007), "Fluctuating Macro Policies and the Fiscal Theory," *NBER Macroeconomics Annual 2006*, 21, 247–298.

Doi, T., Hoshi, T., and Okimoto, T. (2011), "Japanese Government Debt and Sustainability of Fiscal Policy," *Journal of the Japanese and International Economies*, 25, 414–433.

Doi, T. and Ihori, T. (2009), *The Public Sector in Japan: Past Developments and Future Prospects.*, Cheltenham, UK: Edward Elgar.

Evans, M. and Wachtel, P. (1993), "Were Price Changes During the Great Depression Anticipated? : Evidence from Nominal Interest Rates," *Journal of Monetary Economics*, 32, 3–34.

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

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

Hamilton, J. and Susmel, R. (1994), "Autoregressive Conditional Heteroskedasticity and Changes in Regime." *Journal of Econometrics*, 64, 307–333.

Hansen, B. (1992), "The Likelihood Ratio Test Under Non-Standard Conditions: Testing the Markov Switching Model of GNP," *Journal of Applied Econometrics*, 7, 61–82.

Hodrick, R. and Prescott, E. C. (1997), "Postwar U.S. Business Cycles: An Empirical Investigation," *Journal of Money, Credit, and Banking*, 29, 1–16.

Inoue, T. and Okimoto, T. (2008), "Were There Structural Breaks in the Effect of Japanese Monetary Policy?: Re-evaluating the Policy Effects of the Lost Decade," *Journal of the Japanese and International Economies*, 22, 320–342.

Leeper, E. M. (1991), "Equilibria under "Active" and "Passive" Monetary and Fiscal Policies," *Journal of Monetary Economics*, 27, 129–147.

Okimoto, T. (2008), "New Evidence of Asymmetric Dependence Structures in International Equity Markets," *Journal of Financial and Quantitative Analysis*, 43, 787–815.

Sims, C. and Zha, T. (2006), "Were There Regime Switches in U.S. Monetary Policy?" *American Economic Review*, 96, 54–81.

Smith, A., Naik, P. A., and Tsai, C.-L. (2006), "Markov-Switching Model Selection using Kullback-Leibler Divergence," *Journal of Econometrics*, 134, 553–577.

Table 1: Size of the MQLR test in Case 1

Sample Size	Nominal Levels (%)			
	10.00	7.50	5.00	2.50
50	13.70 (13.53)	11.32 (10.56)	8.40 (7.90)	4.98 (4.20)
100	12.00 (11.60)	9.50 (9.16)	7.16 (6.73)	4.34 (3.20)
200	10.60 (10.00)	8.20 (7.50)	5.88 (5.30)	3.18 (2.43)
500	9.40 (10.20)	7.18 (7.53)	4.96 (5.43)	2.60 (2.33)

The numbers in parentheses are from Table II of Cho and White (2007). Model for $X_t|\mathcal{F}_{t-1} : \pi \cdot N(\theta_1 + \theta_0 X_{t-1}, \sigma^2) + (1 - \pi) \cdot N(\theta_2 + \theta_0 X_{t-1}, \sigma^2)$. Number of replications: $5,000$; DGP: $X_t = 0.5X_{t-1} + u_t$ and $u_t \sim$ i.i.d. $N(0, 1)$.

Table 2: Size of the MQLR test in Case 1 for different values of C

	$C = 0.5$				$C = 2$			
	Nominal Levels (%)				Nominal Levels (%)			
Sample Size	10.00	7.50	5.00	2.50	10.00	7.50	5.00	2.50
50	16.62	13.92	10.54	6.26	11.42	9.32	7.10	4.12
100	14.18	11.42	8.68	5.26	9.38	7.62	5.54	3.54
200	13.04	10.56	7.64	4.60	8.14	6.42	4.80	2.58
500	12.90	10.24	7.36	4.44	7.78	6.42	4.68	2.84

Model for $X_t|\mathcal{F}_{t-1} : \pi \cdot N(\theta_1 + \theta_0 X_{t-1}, \sigma^2) + (1 - \pi) \cdot N(\theta_2 + \theta_0 X_{t-1}, \sigma^2)$. Number of replications: $5,000$; DGP: $X_t = 0.5X_{t-1} + u_t$ and $u_t \sim$ i.i.d. $N(0, 1)$.

Table 3: Power of the MQLR test and QLR test in Case 1 (in percent; 5% nominal level)

	β_*				
p^*	0.20	0.40	0.60	0.80	1.00
0.1	5.94 (5.76)	7.16 (6.63)	11.20 (7.06)	21.06 (7.13)	37.02 (7.30)
0.3	5.40 (5.60)	7.00 (6.06)	12.90 (9.50)	34.28 (23.93)	75.38 (62.16)
0.5	5.64 (6.03)	6.82 (6.83)	14.30 (11.33)	43.56 (35.20)	89.58 (85.10)
0.7	6.08 (6.16)	6.86 (6.33)	12.48 (9.46)	32.28 (26.80)	70.60 (68.83)
0.9	5.68 (5.53)	6.62 (6.53)	8.00 (8.13)	8.82 (14.83)	8.06 (29.26)

The numbers in parentheses are from Table III of Cho and White (2007). Model for $X_t|\mathcal{F}_{t-1} : \pi \cdot N(\theta_1 + \theta_0 X_{t-1}, \sigma^2) + (1 - \pi) \cdot N(\theta_2 + \theta_0 X_{t-1}, \sigma^2)$. Number of observations: 500; Number of replications: $5,000$; DGP: $X_t = -\beta_* \cdot \mathbf{1}_{\{S_t=1\}} + \beta_* \cdot \mathbf{1}_{\{S_t=2\}} + 0.5X_{t-1} + u_t$; $P(S_t = 1|S_{t-1} = 1) = P(S_t = 2|S_{t-1} = 2) = p^*$ and $u_t \sim$ i.i.d. $N(0, 1)$.

Table 4: Power of the MQLR test and QLR test in Case 1 for small and large p^* (in percent; 5% nominal level)

| | MQLR test | | | | | QLR test | | | | |
| | β_* | | | | | β_* | | | | |
p^*	0.20	0.40	0.60	0.80	1.00	0.20	0.40	0.60	0.80	1.00
0.01	5.84	6.32	9.86	15.68	21.96	6.56	6.56	7.96	12.46	17.04
0.05	5.62	6.66	10.20	17.56	27.62	6.56	6.04	8.46	13.42	21.10
0.95	5.46	5.62	6.46	6.14	4.78	6.24	6.10	6.46	6.40	6.48
0.99	5.82	6.08	6.18	5.96	5.04	6.44	6.24	6.56	6.52	6.80

Model for $X_t|\mathcal{F}_{t-1} : \pi \cdot N(\theta_1 + \theta_0 X_{t-1}, \sigma^2) + (1 - \pi) \cdot N(\theta_2 + \theta_0 X_{t-1}, \sigma^2)$. In the QLR test, θ_1 and θ_2 are restricted to be in the interval $[-2, 2]$. Number of observations: 500; Number of replications: 5,000; DGP: $X_t = -\beta_* \cdot \mathbf{1}_{\{S_t=1\}} + \beta_* \cdot \mathbf{1}_{\{S_t=2\}} + 0.5 X_{t-1} + u_t$; $P(S_t = 1|S_{t-1} = 1) = P(S_t = 2|S_{t-1} = 2) = p^*$ and $u_t \sim$ i.i.d. $N(0,1)$.

Table 5: Size of the MQLR test in Case 2

| | Nominal Levels (%) | | | |
Sample Size	10.00	7.50	5.00	2.50
50	10.20	8.28	5.90	3.66
100	9.06	7.14	5.26	2.92
200	8.74	6.86	5.26	3.02
500	9.26	7.12	5.02	2.58

Model for $X_t|\mathcal{F}_{t-1} : \pi \cdot N(\theta_{01} + \theta_1 X_{t-1}, \sigma^2) + (1 - \pi) \cdot N(\theta_{01} + \theta_2 X_{t-1}, \sigma^2)$. Number of replications: 5,000; DGP: $X_t = 0.5 X_{t-1} + u_t$ and $u_t \sim$ i.i.d. $N(0,1)$.

Table 6: Power of the MQLR test in Case 2 (in percent; 5% nominal level)

| | β_* | | | | |
p^*	0.10	0.20	0.30	0.40	0.50
0.1	6.98	11.92	24.04	45.84	73.86
0.3	7.44	12.22	24.64	50.60	82.18
0.5	7.26	11.64	25.64	53.60	85.48
0.7	7.52	11.72	24.86	52.38	84.16
0.9	7.30	12.50	23.94	45.34	72.02

Model for $X_t|\mathcal{F}_{t-1} : \pi \cdot N(\theta_{01} + \theta_1 X_{t-1}, \sigma^2) + (1 - \pi) \cdot N(\theta_{01} + \theta_2 X_{t-1}, \sigma^2)$. Number of observations: 500; Number of replications: 5,000; DGP: $X_t = 1 + (0 \cdot \mathbf{1}_{\{S_t=1\}} + \beta_* \cdot \mathbf{1}_{\{S_t=2\}}) X_{t-1} + u_t$; $P(S_t = 1|S_{t-1} = 1) = P(S_t = 2|S_{t-1} = 2) = p^*$ and $u_t \sim$ i.i.d. $N(0,1)$.

21

Table 7: Size of the MQLR test in Case 3

Sample Size	Nominal Levels (%)			
	10.00	7.50	5.00	2.50
50	14.64	12.52	10.02	6.72
100	11.94	9.78	7.32	4.44
200	10.22	8.20	5.66	3.44
500	9.68	7.22	4.86	2.64

Model for $X_t | \mathcal{F}_{t-1} : \pi \cdot N(\theta_{01} + \theta_{02} X_{t-1}, \theta_1^2) + (1 - \pi) \cdot N(\theta_{01} + \theta_{02} X_{t-1}, \theta_2^2)$. Number of replications: $5,000$; DGP: $X_t = 0.5 X_{t-1} + u_t$ and $u_t \sim$ i.i.d. $N(0,1)$.

Table 8: Power of the MQLR test in Case 3 (in percent; 5% nominal level)

p^*	β_*				
	1.10	1.20	1.30	1.40	1.50
0.1	7.36	12.50	22.34	37.04	55.70
0.3	6.94	12.00	22.14	37.26	55.18
0.5	7.28	12.36	22.46	36.78	55.18
0.7	7.54	12.10	22.02	37.08	54.70
0.9	7.30	12.30	21.58	36.18	53.40

Model for $X_t | \mathcal{F}_{t-1} : \pi \cdot N(\theta_{01} + \theta_{02} X_{t-1}, \theta_1^2) + (1 - \pi) \cdot N(\theta_{01} + \theta_{02} X_{t-1}, \theta_2^2)$. Number of observations: 500; Number of replications: $5,000$; DGP: $X_t = 0.5 X_t + u_t$;
$u_t \sim N(0, 1^2 \cdot \mathbf{1}_{\{S_t=1\}} + \beta_*^2 \cdot \mathbf{1}_{\{S_t=2\}})$; $P(S_t = 1 | S_{t-1} = 1) = P(S_t = 2 | S_{t-1} = 2) = p^*$.

Table 9: Results of the MQLR test on stock return data

	Canada	France	Germany	Japan	U.K.	U.S.
Expected return	28.8	15.9	32.7	0.00	3.92	19.1
p-value	< 0.01	< 0.01	< 0.01	0.50	0.07	< 0.01
Volatility	30.8	10.4	27.8	4.66	7.81	11.2
p-value	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01

The rows Expected return and Volatility report the value of the MQLR statistic for testing the null of no regime switching in expected return and volatility, respectively. The p-values of the MQLR statistic in Expected return are based on the asymptotic distribution in Case 1 in Section 3. The p-values of the MQLR statistic in Volatility are based on $\max\{0, N(0,1)\}^2$.

Table 10: Results of the MQLR test on Bohn (1998) type fiscal policy function

	Constant	b_{t-1}	$GVAR_t$	gap_t	x_{t-1}
MQLR statistic	4.42	3.67	0.00	6.55	4.86
p-value	0.05	< 0.01	0.50	< 0.01	< 0.01

The dependent variable is the primary balance divided by GDP. b_{t-1} is the average of the debt-to-GDP ratios for the last four quarters. GVAR is the deviation of the government expenditure from its trend divided by GDP. gap_t is the output gap measured as the deviation from the Hodrick-Prescott trend. x_{t-1} is the lagged dependent variable. The MQLR test tests the null of no regime switching in the coefficient of each regressor while assuming that other parameters are the same across two regimes. p-values are computed from the asymptotic distribution given by Proposition 4 for a constant term and from $\max\{0, N(0,1)\}^2$ for the other parameters.

Table 11: Results of the MQLR test on Davig and Leeper (2007) type fiscal policy function

	Constant	b_{t-1}	gap_t	g_t
MQLR statistic	3.16	4.70	4.25	5.04
p-value	0.11	< 0.01	< 0.01	< 0.01

The dependent variable is the tax revenue divided by GDP. b_{t-1} is the average of the debt to GDP ratios for the last four quarters. gap_t is the output gap measured as the deviation from the Hodrick-Prescott trend. g_t is the government expenditure divided by GDP. The MQLR test is conducted in the same manner as in Table 10.

Table 12: Results of the MQLR test on monetary policy function

	Constant	π_t	gap_t	ex_t
MQLR statistic	8.2	0.01	0.86	4.75
p-value	< 0.01	0.50	0.19	< 0.01

The dependent variable is the average overnight call rate. π_t is the inflation rate measured as the rate of change of GDP deflator from four quarters ago. gap_t is the output gap measured as the deviation from the Hodrick-Prescott trend. ex_t is the deviation of the real effective exchange rate from its Hodrick-Prescott trend. The MQLR test is conducted in the same manner as in Table 10.