Simulation-Based Inference for Markov Switching Models

Gabriel Rodriguez Rondon* Jean-Marie Dufour†

Abstract
Markov switching models have wide applications in economics, finance, and other fields. Most studies focusing on identifying the number of regimes in a Markov switching model have been limited to testing the null hypothesis of only one regime (i.e., a linear model with no switching) against an alternative hypothesis with two regimes. Even in such simple cases, this type of problem raises issues of nonstandard asymptotic distributions, identification failure, and nuisance parameters. In this paper, we propose Monte Carlo test methods [Dufour (2006)] which deal transparently with these distributional issues, even allowing for finite-sample inference. The procedure is applied to likelihood ratio statistics. The tests circumvent the issues plaguing conventional hypothesis testing. This also allows one to deal with non-stationary processes, models with non-Gaussian errors and multivariate settings, which have received little attention in the literature. An important contribution of this paper is the Maximized Monte Carlo Likelihood Ratio Test (MMC-LRT), which is an identifications-robust valid test procedure both in finite samples and asymptotically. Further, the methods proposed are applicable to more general settings where a null hypothesis with M_0 regimes is tested against an alternative with $M_0 + m$ regimes where both $M_0 \geq 1$ and $m \geq 1$. This allows one to compare different Markov switching models and Hidden Markov Models. Simulation results suggest the proposed tests are able to control the level of the test and have good power.

Key Words: Hypothesis testing, Monte Carlo tests, Likelihood ratio, Markov switching, Hidden Markov Model, Nonlinearity, Regimes

1. Introduction

Markov-switching models (MSM) were first introduced by Goldfeld and Quandt (1973) and later popularized by Hamilton (1989) as an alternative approach to modelling U.S. GNP growth. These models allow one to treat a series as a nonlinear process where the nonlinearity arises from discrete shifts. The process before and after a shift can be described as two separate regimes, and Hamilton (1989) describes these regimes as episodes where the behavior of the series is significantly different. Using U.S. GNP growth as an example, one regime can characterize a period of positive growth, while the other represents a period of negative growth due to recessions. Due to this flexibility, they have since been widely used in macroeconomics and finance. For example, MSMs have been applied to the identification of business cycles [Chauvet, 1998; Chauvet and Hamilton, 2006; Chauvet et al., 2002; Diebold and Rudebusch, 1996; Hamilton, 1989; Kim and Nelson, 1999; Qin and Qu, 2021], interest rate dynamics (Garcia and Perron, 1996), financial markets (Marcucci, 2005), conditional heteroskedasticity models [Augustyniak, 2014; Gray, 1996; Haas et al., 2004; Hamilton and Susmel, 1994; Klaassen, 2002], conditional correlations (Pelletier, 2006) and identification of structural VAR models [Herwartz and Lütkepohl, 2014; Lanne et al., 2010; Lütkepohl et al., 2021] to name a few. More complete surveys of this

*Mailing address: Department of Economics, McGill University. e-mail: gabriel.rodriguezrondon@mail.mcgill.ca. Wed page: https://grodriguezrondon.com

†William Dow Professor of Economics, McGill University, Centre interuniversitaire de recherche en analyse des organisations (CIRANO), and Centre interuniversitaire de recherche en economie quantitative (CIREQ). Mailing address: Department of Economics, McGill University, Leacock Building, Room 414, 855 Sherbrooke Street West, Montreal, Quebec H3A 2T7, Canada. e-mail: jean-marie.dufour@mcgill.ca. Web page: http://www.jeanmariedufour.com

literature include Hamilton (2010), Hamilton (2016) and Ang and Timmermann (2012). Applications of MSMs outside of the macroeconomic and financial literature include: environmental and energy economics [Cevik et al., 2021; Charfeddine, 2017; Chevallier, 2011], industrial organization (Resende, 2008), health economics (Anser et al., 2021) and many others. An alternative but related model is the Hidden Markov Model (HMM). Like MSMs, HMMs are used to describe a process Y_t which depends on a latent Markov process S_t. However, HMMs depend only on S_t, which takes discrete values $\{1, \ldots, M\}$ where M is the number of regimes. In contrast, as described by An et al. (2013), when the process Y_t also depends on lags of Y_t (*e.g.*, $\{Y_{t-1}, \ldots, Y_{t-p}\}$), it is called a Hidden Markov-switching model, or simply a Markov-switching model. The dependence on past observations allows for more general interactions between Y_t and S_t, which can be used to model more complicated causal links between economic or financial variables of interest, so that MSMs are a generalization of the basic HMM. However, it is worth noting that HMMs have many applications including computational molecular biology [Baldi et al., 1994; Krogh et al., 1994], handwriting and speech recognition [Jelinek, 1997; Nag et al., 1986; L. Rabiner and Juang, 1986; L. R. Rabiner and Juang, 1993], computer vision and pattern recognition (Bunke and Caelli, 2001), and other machine learning applications.

An important issue with MSMs and HMMs is that the number of states or regimes must be determined *a priori*. Since the number of regimes is not always known, it is of interest to test the fit of a given model with a certain number of regimes (*e.g.*, M_0 regimes), against an alternative model with a different number of regimes (*e.g.*, $M_0 + m$ regimes). This highlights the importance of valid test procedures to determine the number of regimes in these types of models. However, standard hypothesis testing techniques are not easily applicable in this setting, because certain parameters of the model are unidentified under the null hypothesis, and usual regularity conditions needed to derive the asymptotic distribution of test statistics are not satisfied. The study of the asymptotic distribution of the likelihood ratio test for MSMs is a problem that has received a lot of attention [see Carter and Steigerwald, 2012; Cho and White, 2007; Garcia, 1998; Hansen, 1992; Kasahara and Shimotsu, 2018; Qu and Zhuo, 2021]. Most procedures focusing on the likelihood ratio test approach currently available are only able to deal with settings where the null hypothesis is that of a linear model (*i.e.*, $H_0 : M_0 = 1$) and the alternative hypothesis is a MSM with two regimes (*i.e.*, $H_1 : M_0 + m = 2$, where $M_0 = m = 1$). The exception is Kasahara and Shimotsu (2018) which study the asymptotic distribution of the likelihood ratio test statistic when the null hypothesis is of a model with M_0 regimes and the alternative hypothesis is that of a model with $M_0 + 1$ regimes where $M_0 \geq 1$ (and $m = 1$). Interestingly, in this setting, the authors establish the asymptotic validity of the parametric bootstrap procedure (see Proposition 21 of Kasahara and Shimotsu, 2018). Qu and Zhuo (2021) also show the asymptotic validity of the parametric bootstrap for specific data generating processes in the more simple setting where we would like to compare a linear model to a MSM with two regimes. At the same time, others have proposed alternative test procedures based on moments of least-squares residuals (see Dufour and Luger, 2017), parameter stability (see Carrasco et al., 2014), or other moment-matching conditions (see Antoine et al., 2022).

In Carrasco et al. (2014), the authors are interested in the case of testing a linear model against a MSM with only two regimes. This is because, as a test of parameter stability, the null hypothesis must always be that of a linear model and in this case the test only has good power against local alternatives. As a result, this test cannot be used to compare different general MSMs (with $M_0 > 1$). Other limitations of the test procedures mentioned so far (with the exception of those proposed in Dufour and Luger (2017)) is that they are aimed at establishing an asymptotic distribution of the test statistic, so they depend on assumptions needed to obtain asymptotic results which may be restrictive in many cases. For example, a common assumption is that the process studied is stationary with Gaussian errors. Within the likelihood ratio test literature, it is also common to assume a concentrated parameter space, in order to avoid the parameter boundary problem. On the other hand, Dufour and Luger (2017) propose a valid test for the null hypothesis of a linear model against an alternative of a MSM with two regimes, which relies on Monte Carlo test techniques. Specifically, the authors propose four test statistics based on the moments of the least-squares residuals, which are meant to capture different characteristics of a two-component mixture distribution. Approximate marginal p-values are computed for each moment specific test statistic and are combined using either the minimum or the product. This leads to a single test statistic that is free of nuisance parameters when there are no autoregressive lags in the model (*i.e.*, HMMs). When autoregressive lags are included in the model, the test procedure is no longer free of nuisance parameters. In this case, the authors use the Local Monte Carlo (LMC) and Maximized Monte Carlo (MMC) test procedure described in Dufour (2006).

In this paper, we also use the Monte Carlo procedures described in Dufour (2006) to deal with the issues plaguing conventional testing procedures discussed so far, but in a likelihood ratio test setting. This also allows us to deal with nuisance parameters in the distribution of the likelihood ratio test statistic. Specifically, we propose the Local Monte Carlo Likelihood Ratio Test (LMC-LRT) and the Maximized Monte Carlo Likelihood Ratio Test (MMC-LRT), which can be used to identify the number of regimes in both MSMs and HMMs in the more general case where we would like to compare models with M_0 regimes under the null hypothesis against models with $M_0 + m$ regimes under the alternative, where here both $M_0 \geq 1$ and $m \geq 1$. Since MSMs are more general than HMMs in the sense that we can recover a HMM by simply setting the number of autoregressive lags to zero, we focus on MSMs throughout this study but the results of the tests proposed here are also applicable to HMMs. An important contribution of the MMC-LRT proposed here is that it is an exact test for determining the number of regimes in a MSM and is valid both in finite samples and asymptotically. Further, since we are not working with the asymptotic distribution of the test statistic, we can also relax some of the assumptions typically required to obtain asymptotic results. There are four main advantages of using such a framework. The first is that the violation of the regularity conditions needed to drive an asymptotic distribution are no longer problematic. This means that we can consider the full nuisance parameter space rather than a concentrated parameter space as in Qu and Zhuo (2021) and Kasahara and Shimotsu (2018). The second advantage is that this allows us to determine the appropriate number of regimes even when dealing with a non-stationary process Y_t. The third advantage is that we can deal with cases where the asymptotic distribution is more complicated to obtain or even infeasible, such as specific cases where the errors are non-Gaussian. Finally, the fourth advantage is that LMC-LRT and MMC-LRT can be applied to multivariate settings (*e.g.*, Markov-switching VAR models or multivariate HMM). It is also worth noting that non-stationary processes, non-Gaussian errors and multivariate settings have not received a lot of attention in the literature on hypothesis testing for the number of regimes in MSMs. Simulation results indicate that both the LMC-LRT and MMC-LRT procedures presented here are able to control the probability of a type I error as suggested by the theory proposed in Dufour (2006), and have better power than other test proposed in the literature and considered here for comparison. Another noteworthy contributions of this paper includes tabulating results of the test proposed by Dufour and Luger (2017) when the process Y_t is non-stationary. Further simulation results with more complicated DGPs and an empirical application using U.S. GNP growth can also be found in Rodriguez Rondon and Dufour (2022a). All tests results presented in this paper and in Rodriguez Rondon and Dufour (2022a) are obtained using the R package **MSTest** described in a companion paper

Rodriguez Rondon and Dufour (2022b).

The next sections are structured as follows. Section 2 reviews the Markov-switching autoregressive model we are interested in and briefly discusses estimation procedures. In section 3, we introduce and discuss the MMC-LRT and LMC-LRT proposed in this paper. Section 4 provides simulation results for the size and power of the proposed testing procedures and compares them to those of other tests proposed in the literature. Finally, section 5 provides concluding remarks.

2. Markov switching

In general, a MSM can be expressed as

$$y_t = x_t\beta + z_t\delta_{s_t} + \sigma_{s_t}\epsilon_t\,. \tag{1}$$

In a univariate setting, y_t is a scalar, x_t is a fixed (or predetermined) $1\times n$ vector of variables whose coefficients do not depend on the latent Markov process S_t, z_t is an $1 \times \nu$ vector of variables whose coefficient depend on the Markov process S_t, and ϵ_t is an error process, which for example may follow a $\mathcal{N}(0,1)$ distribution, and σ_{s_t} a standard deviation which may also depend on the Markov process S_t or remain constant throughout (*i.e.*, σ). For our testing procedures, other distributions on the error processes may be considered, but for simplicity we assume a normal distribution. Similarly, in a multivariate setting, we could allow y_t to be a $1 \times q$ vector [*i.e.*, $y_t = (y_{1,t},\, \dots\,,\, y_{q,t})$] and ϵ_t also a $1 \times q$ vector, which may be distributed as $\mathcal{N}(\mathbf{0},\Sigma_{s_t})$ or $\mathcal{N}(\mathbf{0},\Sigma)$ [if the variance-covariance matrix does not depend on the latent Markov process S_t].

In order to have an autoregressive MSM, as described above, lags of y_t are included in either x_t or z_t depending on whether we want to allow the autoregressive coefficients to depend on the regimes. This general setting also allows one to consider a trend function within x_t or z_t. A HMM can also be recovered by considering only a constant term in z_t and excluding x_t. For the sake of exposition, in the following we consider a MSM where only the mean and the variance may be subject to change and autoregressive coefficient remain constant. That is, $x_t = (y_{t-1},\, \dots\,,\, y_{t-p})$ and $z_t = 1$. We also reformulate the model to make the dependence on the mean μ_{s_t} more explicit. This leads to the following model:

$$y_t = \mu_{s_t} + \sum_{k=1}^{p}\phi_k(y_{t-k} - \mu_{s_{t-k}}) + \sigma_{s_t}\epsilon_t \tag{2}$$

where we can see that the mean and variance of the observed process y_t are governed by the latent Markov chain process S_t. As described in Hamilton (1994), for a model with M regimes, the one-step transition probabilities can be gathered into a transition matrix such as

$$\mathbf{P} = \begin{bmatrix} p_{11} & \dots & p_{M1} \\ \vdots & \ddots & \vdots \\ p_{1M} & \dots & p_{MM} \end{bmatrix} \tag{3}$$

where for example $p_{ij} = P(S_t = j\,|\,S_{t-1} = i)$ is the probability that state i switches to state j. For example, if we consider a model with only two regimes, we only need a 2×2 transition matrix to summarize the transition probabilities $\mathbf{P}$:

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{21} \\ p_{12} & p_{22} \end{bmatrix} \tag{4}$$

In either case, the columns of the transition matrix must sum to one in order to have a well defined transition matrix (*i.e.*, $\sum_{j=1}^{M} = p_{ij} = 1$). We can also obtain the ergodic probabilities, $\pi = (\pi_1, \pi_2)'$, which are given by

$$\pi_1 = \frac{1 - p_{22}}{2 - p_{11} - p_{22}}\,, \quad \pi_2 = 1 - \pi_1\,, \tag{5}$$

in a setting with two regimes or, more generally, for any number of M regimes we could use

$$\boldsymbol{\pi} = (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\mathbf{e}_{N+1}\,, \quad \mathbf{A} = \begin{bmatrix} \mathbf{I}_M - \mathbf{P} \\ \mathbf{1}' \end{bmatrix}\,, \tag{6}$$

where $\mathbf{e}_{M+1}$ is the $(M+1)$-th column of $\mathbf{I}_{M+1}$.

Continuing with the example of a MSM such as the one given by (2) with $S_t = \{1, 2\}$ [*i.e.*, $M = 2$ regimes], the log-likelihood conditional on the first p observations of y_t is given by

$$L_T(\theta) = \log f(y_1^T \,|\, y_{-p+1}^0;\, \theta) = \sum_{t=1}^{T} \log f(y_t \,|\, y_{-p+1}^{t-1};\, \theta) \tag{7}$$

where $\theta = (\mu_1, \mu_2, \sigma_1, \sigma_2, \phi_1, \ldots, \phi_p, p_{11}, p_{22})'$ and

$$f(y_t \,|\, y_{-p+1}^{t-1};\, \theta) = \sum_{s_t=1}^{2} \sum_{s_{t-1}=1}^{2} \cdots \sum_{s_{t-p}=1}^{2} f(y_t,\, S_t = s_t,\, S_{t-1} = s_{t-1},\, \ldots,\, S_{t-p} = s_{t-p} \,|\, y_{-p+1}^{t-1};\, \theta)\,. \tag{8}$$

Under Gaussianity, we have:

$$f(y_t,\, S_t = s_t,\, \ldots,\, S_{t-p} = s_{t-p} \,|\, y_{-p+1}^{t-1};\, \theta) = \frac{\mathbb{P}(S_t^* = s_t^* \,|\, y_{-p+1}^{t-1};\, \theta)}{\sqrt{2\pi\sigma_{s_t}^2}} \times \exp\left\{ \frac{-[y_t - \mu_{s_t} - \sum_{k=1}^{p} \phi_k (y_{t-k} - \mu_{s_{t-k}})]^2}{2\sigma_{s_t}^2} \right\} \tag{9}$$

where

$$S_t^* = s_t^* \;\text{ if }\; S_t = s_t,\, S_{t-1} = s_{t-1},\, \ldots,\, S_{t-p} = s_{t-p} \tag{10}$$

and $\mathbb{P}(S_t^* = s_t^* \,|\, y_{-p+1}^{t-1}; \theta)$ is the probability that this occurs.

Typically, MSMs are estimated using the Expectation Maximization (EM) algorithm (see Dempster et al., 1977), Bayesian methods or through the use of the Kalman filter (using the state-space representation of the model). In very simple cases, MSMs can be estimated using Maximum Likelihood Estimation (MLE). However, since the Markov process S_t is latent and more importantly the likelihood function can have several modes of equal height in addition to other unusual features that can complicate estimation by MLE this is not often used. In this study, we use the EM algorithm when estimating MSMs. It is worth noting that, in practice, empirical estimates can sometimes be improved by using the results of the EM algorithm as initial values in a Newton-type of optimization algorithm. This two-step estimation procedure is used to obtain results presented in the empirical section of this paper. We omit a detailed explanation of the EM algorithm as our focus is on the hypothesis testing procedures proposed here. For the interested reader, the estimation of a Markov switching model via the EM algorithm is describe in detail in Hamilton (1990) and Krolzig (1997).

3. Monte Carlo likelihood ratio tests

In this section, we introduce the Maximized Monte Carlo likelihood ratio test (MMC-LRT) and the Local Monte Carlo likelihood ratio test (LMC-LRT) obtained assuming normally distributed errors and a model of the form (2). For simplicity, we use an example where we are interested by a null hypothesis of linear model (*i.e.*, only $M_0 = 1$ regime) and an alternative hypothesis of $M_0 + m = 2$ regimes. However, it is easy to see this methodology can be extended to more general cases with $M_0 \geq 1$ and $m \geq 1$. As in Garcia (1998) and the parametric bootstrap procedure describe in Qu and Zhuo (2021) and Kasahara and Shimotsu (2018), we assume that the null hypothesis depends only on the mean, variance and autoregressive coefficients.

The LRT approach requires that we estimate the model both under the null and alternative hypothesis, so that we can obtain the log-likelihoods for each model. The log-likelihood for the model under the alternative (and under the null hypothesis if $M_0 > 1$) is given by (7) - (9):

$$L_T(\theta_1) = \log f(y_1^T \,|\, y_{-p+1}^0;\, \theta_1) = \sum_{t=1}^{T} \log f(y_t \,|\, y_{-p+1}^{t-1};\, \theta_1) \tag{11}$$

where

$$\theta_1 = (\mu_1,\, \mu_2,\, \sigma_1,\, \sigma_2,\, \phi_1,\, \dots\, ,\, \phi_p,\, p_{11},\, p_{22})' \in \Omega\,. \tag{12}$$

The subscript of 1 underscores the fact that θ_1 is the parameter vector under the alternative hypothesis. The set Ω satisfies any theoretical restrictions we may wish to impose on θ_1 [such as $\sigma_1 > 0$ and $\sigma_2 > 0$]. On the other hand, the log-likelihood under the null hypothesis ($M_0 = 1$) is given by

$$L_T^0(\theta_0) = \log f(y_1^T \,|\, y_{-p+1}^0;\, \theta_0) = \sum_{t=1}^{T} \log f(y_t \,|\, y_{-p+1}^{t-1};\, \theta_0) \tag{13}$$

where

$$f(y_t \,|\, y_{-p+1}^{t-1};\, \theta_0) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ \frac{-[y_t - \mu - \sum_{k=1}^{p} \phi_k (y_{t-k} - \mu)]^2}{2\sigma^2} \right\}, \tag{14}$$

$$\theta_0 = (\mu,\, \sigma^2, \phi_1,\, \dots\, ,\, \phi_p)' \in \bar{\Omega}_0. \tag{15}$$

Note that $\bar{\Omega}_0$ has lower dimension than Ω. The null and alternative hypotheses can be written as:

$$H_0 : \delta_1 = \delta_2 = \delta \quad \text{for some unknown } \delta = (\mu,\, \sigma)\,, \tag{16}$$

$$H_1 : (\delta_1,\, \delta_2) = (\delta_1^*,\, \delta_2^*) \;\; \text{for some unknown } \delta_1^* \neq \delta_2^*\,, \tag{17}$$

where $\delta_1 = (\mu_1, \sigma_1)$ and $\delta_2 = (\mu_2, \sigma_2)$. Clearly, H_0 is a restricted version of H_1: for each $\theta_0 \in \bar{\Omega}_0$, we can find θ_1 such that

$$L_T^0(\theta_0) = L_T(\theta_1)\,, \quad \theta_1 \in \Omega_0, \tag{18}$$

where Ω_0 is the subset of vectors $\theta_1 \in \Omega$ such that θ_1 satisfies H_0. Under H_0, the vector $\theta_0 \in \bar{\Omega}_0$ is a nuisance parameter: the null distribution of any test statistic for H_0 depends on $\theta_0 \in \bar{\Omega}_0$. In this problem, the null distribution of the test statistic is in fact completely determined by θ_0.

The likelihood ratio statistic for testing H_0 against H_1 can then written as

$$LR_T = 2[\bar{L}_T(H_1) - \bar{L}_T(H_0)] \tag{19}$$

where

$$\bar{L}_T(H_1) = \sup\{L_T(\theta_1) : \theta_1 \in \Omega\}\,, \tag{20}$$

$$\bar{L}_T(H_0) = \sup\{L_T^0(\theta_0) : \theta_0 \in \bar{\Omega}_0\} = \sup\{L_T(\theta_1) : \theta_1 \in \Omega_0\}\,. \tag{21}$$

The null distribution of LR_T depends on the parameter $\theta_0 \in \bar{\Omega}_0$. Since the model is parametric, we can generate a vector N i.i.d replications of LR_T for any given value of $\theta_0 \in \bar{\Omega}_0$:

$$LR(N,\, \theta_0) := [LR_T^{(1)}(\theta_0),\, \dots\, , LR_T^{(N)}(\theta_0)]', \qquad \theta_0 \in \bar{\Omega}_0\,. \tag{22}$$

Let us denote $LR_T^{(0)} := LR_T$ the test statistic based in the observed data. Given the model considered, we can assume [as in (4.10) of Dufour (2006)] that:

$$\text{the random variables } \; LR_T^{(0)},\, LR_T^{(1)}(\theta_0),\, \dots\, , LR_T^{(N)}(\theta_0) \; \text{ are exchangeable for some } \theta_0 \in \bar{\Omega}_0, \; \text{ each with distribution function } \; F[x \,|\, \theta_0]\,. \tag{23}$$

Set

$$\hat{F}_N[x \,|\, \theta_0] := \hat{F}_N[x;\, LR(N,\, \theta_0)] = \frac{1}{N}\sum_{i=1}^{N} I[LR_T^{(i)}(\theta_0) \le x] \tag{24}$$

$$\hat{G}_N[x \,|\, \theta_0] := \hat{G}_N[x;\, LR(N,\, \theta_0)] = 1 - \hat{F}_N[x;\, LR(N,\, \theta_0)] \tag{25}$$

where $I(C) := 1$ if condition C holds, and $I(C) = 0$ otherwise. $\hat{F}_N[x \,|\, \theta_0]$ is the sample distribution of the simulated statistics, and $\hat{G}_N[x \,|\, \theta_0]$ is the corresponding survival function. Then, the Monte Carlo p-value is given by

$$\hat{p}_N[x \,|\, \theta_0] = \frac{N\hat{G}_N[x \,|\, \theta_0] + 1}{N+1}\,. \tag{26}$$

Alternatively, using the relationship

$$\begin{aligned} R_{LR}[LR_T^{(0)};\, N] &= N\hat{F}_N[x;\, LR(N,\, \theta_0)] \\ &= \sum_{i=1}^{N} I[LR_T^0 \ge LR_T^i(\theta_0)] \end{aligned} \tag{27}$$

we can define a Monte Carlo p-value as

$$\hat{p}_N[x \,|\, \theta_0] = \frac{N + 1 - R_{LR}[LR_T^{(0)};\, N]}{N+1} \tag{28}$$

where, as can be seen from (27), $R_{LR}[LR_T^{(0)}; N]$ simply computes the rank of the test statistic using the observed data within the generated series $LR(N, \theta_0)$. As discussed in Dufour (2006), a critical region with level α is then given by

$$\sup_{\theta_0 \in \bar{\Omega}_0} \hat{p}_N[LR_T^{(0)} \,|\, \theta_0] \le \alpha \tag{29}$$

More precisely, if $(N+1)\alpha$ is an integer, we have

$$\mathbb{P}\left[\sup\{\hat{p}_N[LR_T^{(0)} \,|\, \theta_0] : \theta_0 \in \bar{\Omega}_0\} \le \alpha\right] \le \alpha \tag{30}$$

under the null hypothesis: we get a valid test with level α for H_0; see Proposition 4.1 in Dufour (2006). In the present case, we call this procedure the Maximized Monte Carlo Likelihood Ratio Test (MMC-LRT).

The parameter space, however, can be very large. Specifically, it grows as the number of autoregressive components increases and as the number of regimes increases. Additionally, the solution may not be unique in the sense that the maximum p-value may be obtained by more than one parameter vector. For this reason, numerical optimization methods that do not depend on the use of derivatives are recommended to find the maximum Monte Carlo p-value within the nuisance parameter space. Such algorithms include: Generalized Simulated Annealing, Genetic Algorithms, and Particle Swarm [see Dufour, 2006; Dufour and Neves, 2019].

In order to facilitate optimization [as described in Dufour (2006)], it is also possible to search within a smaller consistent set of the parameter space C_T. A consistent set can be defined using the consistent point estimate. For example, let $\hat{\theta}_0$ be the consistent point estimate of θ_0. Then, we can define

$$C_T = \{\theta_0 \in \bar{\Omega}_0 : \parallel \hat{\theta}_0 - \theta_0 \parallel < c\} \tag{31}$$

where c is a fixed positive constant that does not depend on T and $\|\cdot\|$ is the Euclidean norm in $\mathbb{R}^k$. An empirically interesting consistent set that we consider in this study is $C_T^* = C_T^{CI} \cup C_T^{\epsilon}$ where

$$C_T^{CI} = \{\theta_0 \in \bar{\Omega}_0 : \parallel \hat{\theta}_0 - \theta_0 \parallel < 2 \times S.E.(\hat{\theta}_0)\} \tag{32}$$

$$C_T^{\epsilon} = \{\theta_0 \in \bar{\Omega}_0 : \parallel \hat{\theta}_0 - \theta_0 \parallel < \epsilon\} \tag{33}$$

C_T^{CI} is defined by a 95% confidence interval of the consistent point estimates, while C_T^{ϵ} is defined using a fixed constant ϵ that does not depend on T. The union of these two sets allows us to consider values that may be outside the confidence interval of the autoregressive parameters and the transition probabilities, depending on the choice of ϵ, while also constraining the values considered for the mean and the variances to a reasonable region.

Finally, we can also define C_T to be the singleton set $C_T = \{\hat{\theta}_0\}$, which gives us the Local Monte Carlo Likelihood Ratio Test (LMC-LRT). Here, the consistent set includes only the consistent point estimate $\hat{\theta}_0$. Generic conditions for the asymptotic validity of such a test are discussed in section 5 of Dufour (2006), but these are more restrictive than those for the MMC-LRT procedure. The LMC test can be interpreted as the finite-sample analogue of the parametric bootstrap. To reflect this, we replace $\hat{F}_N[x \mid \theta_0]$ with $\hat{F}_{TN}[x \mid \theta_0] = \hat{F}_N[x; LR_T(N, \theta_0)]$ and $\hat{G}_N[x \mid \theta_0]$ with $\hat{G}_{TN}[x \mid \theta_0] = \hat{G}_N[x; LR_T(N, \theta_0)]$ where the subscript T is meant to allow the test statistics and functions to change based on increasing sample sizes. As a result, the Monte Carlo p-value is given by

$$\hat{p}_{TN}[x \mid \theta_0] = \frac{N\hat{G}_{TN}[x \mid \theta_0] + 1}{N + 1} \tag{34}$$

The asymptotic validity in this case refers to the estimate $\hat{\theta}_0$ converging asymptotically to the true parameters in θ_0 as the sample size increases. This is not related to the asymptotic validity of the critical values as desired in Hansen (1992), Garcia (1998), Cho and White (2007), Qu and Zhuo (2021) and Kasahara and Shimotsu (2018). Specifically, like the parametric bootstrap, the LMC procedure is only valid asymptotically as $T \to \infty$ but, unlike the parametric bootstrap, we do not need a large number of simulations (*i.e.*, $N \to \infty$), since we do not try to approximate the asymptotic critical values nor assume that the distribution of the test statistic converges asymptotically but rather work with the critical values from the sample distribution $\hat{F}[x \mid \theta_0]$. This allows the procedure to be computationally efficient in the sense that we will not need to perform a large number of simulations with the aim of obtaining asymptotically valid critical values. In fact, as can be seen from equations

Table 1: Empirical size of test when $H_0 : M_0 = 1$ and $H_1 : M_0 + m = 2$

	$\phi = 0.1$		$\phi = 0.9$		$\phi = 1.0$	
Tests	$T = 100$	$T = 200$	$T = 100$	$T = 200$	$T = 100$	$T = 200$
LMC-LRT	5.8	5.3	4.9	4.8	4.9	4.9
MMC-LRT	0.4	0.5	0.9	0.9	0.9	0.6
LMC_{min}	4.4	5.9	4.7	4.7	4.2	4.8
LMC_{prod}	5.3	5.2	4.9	5.1	4.9	4.3
MMC_{min}	0.2	0.2	0.2	0.7	0.2	0.0
MMC_{prod}	0.1	0.2	0.4	0.8	0.2	0.5
supTS	4.8	5.1	6.0	4.5	-	-
expTS	6.8	6.2	5.4	6.9	-	-

(28) and (34), the number of replications N is taken into account in the calculation of the p-value both in the numerator and the denominator so that it essentially remains fixed as N increases. As discussed in Dufour (2006), building a test with level $\alpha = 0.05$ requires as few as 19 replications but using more replications can increase the power of the test. For this reason, in our simulations results we use $N = 99$ for our Monte Carlo procedure as in Dufour and Khalaf (2001) and Dufour and Luger (2017) though it is also possible to use the procedure described in Davidson and MacKinnon (2000) to determine the optimal number of simulations to minimize experimental randomness and loss of power.

4. Simulation evidence

Here, we present tables summarizing the empirical size and power (in percentage) of the two tests proposed in this paper, the LMC-LRT and MMC-LRT. We also present results for the moment-based test of Dufour and Luger (2017) and the parameter stability test of Carrasco et al. (2014) for comparison. In what follows, the nominal level is set to be $\alpha = 0.05$ and results are based on 1000 replications of the data generating process (DGP). Throughout, we will consider a simple AR(1) model given by

$$y_t = \mu_{s_t} + \phi_1(y_{t-1} - \mu_{s_{t-1}}) + \sigma_{s_t}\epsilon_t \tag{35}$$

where $\epsilon_t \sim \mathcal{N}(0, 1)$, such that only the mean and variance are governed by the Markov process S_t. It is understood that the LMC-LRT, supTS, and expTS procedures should perform better in large sample sizes since they are asymptotic tests. However, many economic applications using quarterly observations are limited to as few as 100 to 200 observations so in the following we consider these sample sizes to get an idea of the finite sample performance of these tests. Also, whenever considering a Maximized Monte Carlo test, we use the set $C_T^* = C_T^{CI} \cup C_T^{0.1}$. All results presented here can be obtained using the R-package **MSTest** described in the companion paper Rodriguez Rondon and Dufour (2022b).

Table 1 reports the empirical size of the tests when $\phi = 0.1$ in the first two columns, when $\phi = 0.9$ in the next two columns, and when $\phi = 1.0$ in the last two columns. For each value of ϕ we consider a sample size of $T = 100$ or $T = 200$. As suggested by the theory laid out in Dufour (2006), the maximized Monte Carlo tests has empirical rejection frequencies $\leq 5\%$ under the null hypothesis. The LMC-LRT, LMC_{min}, and LMC_{prod}, also appear perform very well with a rejection frequency of approximately 5% even in finite samples. The simulation results for the supTS and expTS show that these tests also appear

Table 2: Empirical power of test when $H_0 : M_0 = 1$ and $H_1 : M_0 + m = 2$ and $(p_{11}, p_{22}) = (0.9, 0.9)$

	$\phi = 0.1$		$\phi = 0.9$		$\phi = 1.0$	
Tests	$T = 100$	$T = 200$	$T = 100$	$T = 200$	$T = 100$	$T = 200$
$\Delta\mu = 2$, $\Delta\sigma = 0$						
LMC-LRT	48.7	81.2	19.5	33.8	12.5	15.8
MMC-LRT	30.6	51.0	4.6	8.4	4.6	2.4
LMC_{min}	3.8	5.8	14.6	21.3	18.8	28.5
LMC_{prod}	4.1	5.9	15.2	24.2	19.6	30.8
MMC_{min}	0.0	0.1	2.8	4.1	1.9	5.0
MMC_{prod}	0.1	0.1	1.9	4.0	2.3	5.2
supTS	24.3	49.9	8.4	12.5	-	-
expTS	15.6	25.4	21.7	32.6	-	-
$\Delta\mu = 0$, $\Delta\sigma = 1$						
LMC-LRT	66.6	93.4	67.0	94.5	67.2	94.6
MMC-LRT	39.4	69.6	35.6	73.0	31.7	64.7
LMC_{min}	36.5	64.7	42.5	64.9	39.6	64.4
LMC_{prod}	40.6	66.8	43.3	69.1	42.5	65.3
MMC_{min}	9.0	30.2	11.2	27.8	10.4	16.3
MMC_{prod}	10.7	31.4	10.8	31.0	7.9	18.0
supTS	32.4	58.0	32.2	67.4	-	-
expTS	40.1	62.6	54.1	84.7	-	-
$\Delta\mu = 2$, $\Delta\sigma = 1$						
LMC-LRT	83.7	99.4	45.3	77.2	29.5	43.9
MMC-LRT	60.3	90.1	24.0	52.0	14.0	29.5
LMC_{min}	51.9	81.6	39.9	62.3	35.4	57.7
LMC_{prod}	45.9	74.2	42.5	65.1	38.1	60.9
MMC_{min}	10.5	39.0	10.2	24.0	8.8	13.7
MMC_{prod}	14.3	37.9	11.8	29.6	9.0	18.1
supTS	72.7	96.2	34.6	62.9	-	-
expTS	75.6	97.0	53.9	77.9	-	-

to control the empirical size well but less so than the Monte Carlo based tests as they have a small degree of over-rejection in some cases. This however, should be expected given the small sample sizes.

Table 2 and 3 report the empirical power of the tests when the underlying DGP is a MSM with two regimes. We consider the same values for ϕ and T as in Table 1 but in Table 2 the transition probabilities are $(p_{11}, p_{22}) = (0.9, 0.9)$ resulting in $\boldsymbol{\pi} = (0.5, 0.5)$ while in Table 3 the transition probabilities are $(p_{11}, p_{22}) = (0.90, 0.50)$ so that $\boldsymbol{\pi} = (0.83, 0.17)$. The first case corresponds to having spent, on average in the long run, the same amount of time in both regimes whereas the latter case suggests, on average in the long run, more time is spent in the first regime over the entire sample. The first panel in each table corresponds to a DGP where only the mean changes, the second panel a DGP where only the variance changes and the third bottom panel a DGP where both the mean and the variance are different across regimes. As can be seen from these tables, power is lowest when only the mean is subject to change for all tests. The LMC_{min}, LMC_{prod}, MMC_{min},

Table 3: Empirical power of test when $H_0 : M_0 = 1$ and $H_1 : M_0 + m = 2$ and $(p_{11}, p_{22}) = (0.9, 0.5)$

	$\phi = 0.1$		$\phi = 0.9$		$\phi = 1.0$	
Tests	$T = 100$	$T = 200$	$T = 100$	$T = 200$	$T = 100$	$T = 200$
$\Delta\mu = 2$, $\Delta\sigma = 0$						
LMC-LRT	34.0	72.3	27.4	48.3	20.8	24.5
MMC-LRT	20.3	41.0	7.1	16.9	4.4	8.1
LMC_{min}	14.8	29.9	13.5	21.8	16.3	28.4
LMC_{prod}	11.9	22.4	15.0	23.4	16.9	29.0
MMC_{min}	1.0	3.7	1.4	2.6	1.6	2.1
MMC_{prod}	1.8	3.3	1.2	3.8	1.6	3.6
supTS	23.8	47.0	11.9	18.2	-	-
expTS	24.6	47.1	22.1	33.5	-	-
$\Delta\mu = 0$, $\Delta\sigma = 1$						
LMC-LRT	55.7	88.7	61.6	88.4	58.4	89.9
MMC-LRT	36.4	69.4	30.7	60.8	29.2	54.5
LMC_{min}	48.7	69.4	49.4	72.0	45.6	74.4
LMC_{prod}	48.8	70.7	49.5	71.6	47.8	73.2
MMC_{min}	20.5	44.4	19.7	49.2	16.6	34.3
MMC_{prod}	20.3	42.6	18.8	44.5	16.2	31.2
supTS	29.9	46.4	30.0	50.3	-	-
expTS	43.9	68.3	52.8	78.6	-	-
$\Delta\mu = 2$, $\Delta\sigma = 1$						
LMC-LRT	83.4	99.4	60.2	88.1	53.1	67.4
MMC-LRT	66.2	91.6	41.5	74.5	29.7	53.7
LMC_{min}	84.1	99.0	65.9	89.4	63.7	88.2
LMC_{prod}	83.7	99.2	68.5	91.6	63.9	89.9
MMC_{min}	47.3	50.3	28.1	49.8	23.4	46.7
MMC_{prod}	48.3	44.6	34.9	43.6	28.3	42.1
supTS	80.8	96.9	53.2	79.8	-	-
expTS	86.6	99.4	75.1	94.7	-	-

and MMC_{prod} procedures have the lowest power when only the mean is subject to change. The LMC-LRT procedure proposed here on the other hand has comparable power to the supTS and expTS tests and in many cases even performs better. When the variance is subject to change, all tests have higher power. The LMC-LRT test proposed here appears to have the highest power in most cases when the variance is subject to change, with the supTS and expTS tests having comparable results in some cases. Given that the MMC-LRT procedure considers a wider set of nuisance parameter values in comparison to the LMC-LRT procedure, the power of the MMC-LRT is lower than that of the LMC-LRT in all cases. The same is true when comparing the MMC_{min}, and MMC_{prod} power results to those of LMC_{min} and LMC_{prod} and this should be expected. However, we find that for the set C_T^*, the power of the MMC-LRT procedure proposed here is quite good and in some cases even outperforms the LMC_{min} and LMC_{prod} procedures when only the mean is subject to change.

In Rodriguez Rondon and Dufour (2022a) we further demonstrate the use and performance of the LMC-LRT and MMC-LRT proposed here by including simulation results of two other cases of interest in the univariate setting. Specifically, we consider an alternative of $M_0 + m = 3$ and cases where we compare $M_0 = 2$ vs. $M_0 + m = 3$ regimes. In addition, multivariate settings (*e.g.*, MS-VAR models) are also considered. Results presented in Rodriguez Rondon and Dufour (2022a) suggest that the LMC-LRT and the parametric Bootstrap procedure discussed in Qu and Zhuo (2021) and Kasahara and Shimotsu (2018) are outperformed by the MMC-LRT proposed here for controlling the size of the test in finite samples when considering more complicated DGPs (*e.g.*, MSM with two or more regimes under the null hypothesis). This further highlighting the value of having a procedure such as the MMC-LRT which is identification robust and valid even in finite samples.

5. Conclusion

We have shown how to use the Monte Carlo procedures described in Dufour (2006) for the setting of a likelihood ratio test for MSMs. In doing so, we propose the Maximized Monte Carlo Likelihood Ratio Test (MMC-LRT) and the Local Monte Carlo Likelihood Ratio Test (LMC-LRT) that can be used to determine the number of regimes in MSMs and in HMMs. Specifically, the tests proposed here are general enough where they can deal with settings where we are interested in comparing models with M_0 regimes under the null hypothesis against models with $M_0 + m$ regimes under the alternative, where here both $M_0, m \geq 1$. Further, they can also be applied to settings where we have a non-stationary process, a process with non-Gaussian errors, and multivariate settings. To the best of our knowledge, we are the first to consider hypothesis testing of multivariate MSMs. Although we work we the sample distribution of the test statistic, asymptotic results have not been provided for LRT in a multivariate setting which brings forward an interesting direction for future research. The simulation results suggest that both versions of the Monte Carlo likelihood ratio test are able to control the level of the test very well. An important contribution is the MMC-LRT, which perform well in that it maintains a rejection frequencies $\leq \alpha$ in all cases as suggested by the theory proposed in Dufour (2006) and is an identification-robust procedure that is valid in finite samples and asymptotically. Further, simulation results also suggest both test have good power. Specifically, the LMC-LRT has comparable or better power than the supTS and expTS tests while both the LMC-LRT and MMC-LRT outperform their moment-based counterparts.

Acknowledgements

We are grateful to participants of the 56th Canadian Economics Association conference bootstrap and simulation session, the 17th CIREQ Ph.D. Students Conference, and the Joint Statistical Meetings business and economic statistics session for their valuable comments. Financial support from the Fonds de Recherche du Québec sur la Société et la Culture (FRQSC) Doctoral Research Scholarships (B2Z) is also gratefully acknowledged.

References

An, Y., Hu, Y., & Shum, M. (2013). *Identifiability and inference of hidden Markov models* (tech. rep.).

Ang, A., & Timmermann, A. (2012). Regime changes and financial markets. *The Annual Review of Financial Economics*, *4*(1), 313–337.

Anser, M. K., Godil, D. I., Khan, M. A., Nassani, A. A., Zaman, K., & Abro, M. M. Q. (2021). The impact of coal combustion, nitrous oxide emissions, and traffic emissions on covid-19 cases: A Markov-switching approach. *Environmental Science and Pollution Research*, *28*(45), 64882–64891.

Antoine, B., Khalaf, L., Kichian, M., & Lin, Z. (2022). Identification-robust inference with simulation-based pseudo-matching. *Journal of Business & Economic Statistics*, 1–18.

Augustyniak, M. (2014). Maximum likelihood estimation of the Markov-switching GARCH model. *Computational Statistics & Data Analysis*, *76*, 61–75.

Baldi, P., Chauvin, Y., Hunkapiller, T., & McClure, M. A. (1994). Hidden Markov models of biological primary sequence information. *Proceedings of the National Academy of Sciences*, *91*(3), 1059–1063.

Bunke, H., & Caelli, T. M. (2001). *Hidden Markov models: Applications in computer vision* (Vol. 45). World Scientific.

Carrasco, M., Hu, L., & Ploberger, W. (2014). Optimal test for Markov switching parameters. *Econometrica*, *82*(2), 765–784.

Carter, A. V., & Steigerwald, D. G. (2012). Testing for regime switching: A comment. *Econometrica*, *80*(4), 1809–1812.

Cevik, E. I., Yıldırım, D. Ç., & Dibooglu, S. (2021). Renewable and non-renewable energy consumption and economic growth in the us: A Markov-switching var analysis. *Energy & Environment*, *32*(3), 519–541.

Charfeddine, L. (2017). The impact of energy consumption and economic development on ecological footprint and co2 emissions: Evidence from a Markov switching equilibrium correction model. *Energy Economics*, *65*, 355–374.

Chauvet, M. (1998). An econometric characterization of business cycle dynamics with factor structure and regime switching. *International economic review*, 969–996.

Chauvet, M., & Hamilton, J. D. (2006). Dating business cycle turning points. *Contributions to Economic Analysis*, *276*, 1–54.

Chauvet, M., Juhn, C., & Potter, S. (2002). Markov switching in disaggregate unemployment rates. *Empirical Economics*, *27*(2), 205–232.

Chevallier, J. (2011). Evaluating the carbon-macroeconomy relationship: Evidence from threshold vector error-correction and Markov-switching var models. *Economic Modelling*, *28*(6), 2634–2656.

Cho, J.-S., & White, H. (2007). Testing for regime switching. *Econometrica*, *75*(6), 1671–1720.

Davidson, R., & MacKinnon, J. G. (2000). Bootstrap tests: How many bootstraps? *Econometric Reviews*, *19*(1), 55–68.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society. Series B*, *39*(1), 1–38.

Diebold, F. X., & Rudebusch, G. D. (1996). Measuring business cycles: A modern perspective. *The Review of Economics and Statistics*, *78*.

Dufour, J.-M. (2006). Monte Carlo tests with nuisance parameters: A general approach to finite-sample inference and nonstandard asymptotics. *Journal of Econometrics*, *133*(2), 443–477.

Dufour, J.-M., & Khalaf, L. (2001). Monte Carlo test methods in econometrics. *Companion to Theoretical Econometrics', Blackwell Companions to Contemporary Economics, Basil Blackwell, Oxford, UK*, 494–519.

Dufour, J.-M., & Luger, R. (2017). Identification-robust moment-based tests for Markov switching in autoregressive models. *Econometric Reviews*, *36*(6-9), 713–727.

Dufour, J.-M., & Neves, J. (2019). Finite-sample inference and nonstandard asymptotics with Monte Carlo tests and r. In *Handbook of statistics* (pp. 3–31, Vol. 41). Elsevier.

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

Garcia, R., & Perron, P. (1996). An analysis of the real interest rate under regime shifts. *Review of Economics and Statistics*, *78*, 111–125.

Goldfeld, S. M., & Quandt, R. E. (1973). A Markov model for switching regressions. *Journal of Econometrics*, *1*(1), 3–15.

Gray, S. F. (1996). Modeling the conditional distribution of interest rates as a regime-switching process. *Journal of Financial Economics*, *42*(1), 27–62.

Haas, M., Mittnik, S., & Paolella, M. S. (2004). A new approach to Markov-switching GARCH models. *Journal of Financial Econometrics*, *2*(4), 493–530.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, *57*(2), 357–384.

Hamilton, J. D. (1990). Analysis of time series subject to changes in regime. *Journal of Econometrics*, *45*(1-2), 39–70.

Hamilton, J. D. (2010). Regime switching models. In *Macroeconometrics and time series analysis* (pp. 202–209). Palgrave Macmillan, London.

Hamilton, J. D. (2016). Macroeconomic regimes and regime shifts. *Handbook of macroeconomics*, *2*, 163–201.

Hamilton, J. D., & Susmel, R. (1994). Autoregressive conditional heteroskedasticity and changes in regime. *Journal of Econometrics*, *64*(1-2), 307–333.

Hamilton, J. D. (1994). *Time series analysis*. Princeton university press.

Hansen, B. E. (1992). The likelihood ratio test under nonstandard conditions: Testing the Markov switching model of GNP. *Journal of applied Econometrics*, *7*(S1), S61–S82.

Herwartz, H., & Lütkepohl, H. (2014). Structural vector autoregressions with Markov switching: Combining conventional with statistical identification of shocks. *Journal of Econometrics*, *183*(1), 104–116.

Jelinek, F. (1997). *Statistical methods for speech recognition*. MIT press.

Kasahara, H., & Shimotsu, K. (2018). Testing the number of regimes in Markov regime switching models. *arXiv preprint arXiv:1801.06862*.

Kim, C.-J., & Nelson, C. R. (1999). Has the us economy become more stable? a bayesian approach based on a Markov-switching model of the business cycle. *Review of Economics and Statistics*, *81*(4), 608–616.

Klaassen, F. (2002). Improving GARCH volatility forecasts with regime-switching GARCH. In *Advances in Markov-switching models* (pp. 223–254). Springer.

Krogh, A., Mian, I. S., & Haussler, D. (1994). A hidden Markov model that finds genes in e. coli dna. *Nucleic acids research*, *22*(22), 4768–4778.

Krolzig, H.-M. (1997). *Markov-switching vector autoregressions*. Springer.

Lanne, M., Lütkepohl, H., & Maciejowska, K. (2010). Structural vector autoregressions with Markov switching. *Journal of Economic Dynamics and Control*, *34*(2), 121–131.

Lütkepohl, H., Meitz, M., Netšunajev, A., & Saikkonen, P. (2021). Testing identification via heteroskedasticity in structural vector autoregressive models. *The Econometrics Journal*, *24*(1), 1–22.

Marcucci, J. (2005). Forecasting stock market volatility with regime-switching GARCH models. *Studies in Nonlinear Dynamics & Econometrics*, *9*(4).

Nag, R., Wong, K., & Fallside, F. (1986). Script recognition using hidden Markov models. *ICASSP'86. IEEE International Conference on Acoustics, Speech, and Signal Processing*, *11*, 2071–2074.

Pelletier, D. (2006). Regime switching for dynamic correlations. *Journal of Econometrics*, *131*(1-2), 445–473.

Qin, A., & Qu, Z. (2021). Modeling regime switching in high-dimensional data with applications to U.S. business cycles. *Working paper*.

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

Rabiner, L., & Juang, B. (1986). An introduction to hidden Markov models. *IEEE ASSP Magazine*, *3*(1), 4–16.

Rabiner, L. R., & Juang, B. H. (1993). *Fundamentals of speech recognition*. Prentice Hall.

Resende, M. (2008). Mergers and acquisitions waves in the uk: A Markov-switching approach. *Applied Financial Economics*, *18*(13), 1067–1074.

Rodriguez Rondon, G., & Dufour, J.-M. (2022a). Monte Carlo likelihood ratio tests for Markov switching models. *Working paper*.

Rodriguez Rondon, G., & Dufour, J.-M. (2022b). MSTest: An R-package for testing Markov switching models. *Working paper*.