BANK OF CANADA
BANQUE DU CANADA

Staff working paper / Document de travail du personnel—2026-7

Last updated: March 4, 2026

MSTest: An R-Package for Testing Markov Switching Models

Gabriel Rodriguez Rondon
Canadian Economic Analysis Department
Bank of Canada
grodriguezrondon@bankofcanada.ca

Jean-Marie Dufour
McGill University
jean-marie.dufour@mcgill.ca

DOI: https://doi.org/10.34989/swp-2026-7 | ISSN 1701-9397

MSTest: An R-Package for Testing Markov Switching Models

Gabriel Rodriguez–Rondon* Jean–Marie Dufour†

Bank of Canada McGill University

December 16, 2025

Abstract

We present the R package MSTest, which implements hypothesis testing procedures to determine the number of regimes in Markov switching models. These models have wide-ranging applications in economics, finance, and many other fields. MSTest provides several testing frameworks, including Monte Carlo likelihood ratio tests (Rodriguez-Rondon and Dufour (2025)), moment-based tests (Dufour and Luger (2017)), parameter stability tests (Carrasco et al. (2014)), and classical likelihood ratio procedures (Hansen (1992)). In addition, the package offers tools for simulating and estimating univariate and multivariate Markov switching and hidden Markov models using either the expectation–maximization algorithm or maximum likelihood estimation. The functionality of the package is demonstrated through simulation-based examples.

Key words: Hypothesis testing, Monte Carlo tests, Likelihood ratio, Exact inference, Markov switching, Nonlinearity, Regimes, R software

JEL codes: C12, C15, C22, C52

*Mailing address: Bank of Canada, 234 Wellington St. W, Ottawa, ON K1A 0G9. e-mail: grodriguezrondon@bankofcanada.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://jeanmariedufour.com.

Résumé

Nous présentons le package R MSTest, qui implémente des procédures de tests d'hypothèses visant à déterminer le nombre de régimes dans les modèles à changements de régime de type Markovien. Ces modèles trouvent de nombreuses applications en économie, en finance et dans plusieurs autres domaines. MSTest inclus plusieurs cadres de test, notamment des tests du rapport de vraisemblance par Monte Carlo (Rodriguez-Rondon and Dufour (2025)), des tests fondés sur les moments (Dufour and Luger (2017)), des tests de stabilité des paramètres (Carrasco et al. (2014)) ainsi que des procédures classiques de rapport de vraisemblance (Hansen (1992)). De plus, le package fournit des outils permettant de simuler et d'estimer des modèles à changements de régime et des modèles de Markov cachés, univariés et multivariés, en utilisant soit l'algorithme d'espérance–maximisation, soit l'estimation par maximum de vraisemblance. Le fonctionnement du package est illustré à l'aide d'exemples fondés sur des simulations.

1 Introduction

Markov switching models were first introduced by Goldfeld and Quandt (1973), but were later popularized and became an active area of research in economics after Hamilton (1989) proposed modeling the first difference of U.S. GNP as a nonlinear stationary process rather than the linear stationary process that was typically used. The nonlinearity arises from discrete shifts in the process.

These models have since been applied in a wide range of macroeconomic and financial settings. For example, they have been used to identify business cycles and provide probabilistic statements about the state of the economy (Chauvet, 1998; Chauvet and Hamilton, 2006; Chauvet et al., 2002; Diebold and Rudebusch, 1996; Hamilton, 1989; Kim and Nelson, 1999; Qin and Qu, 2021), to model stock market volatility using Markov switching ARCH, GARCH, and stochastic volatility models (Hamilton, 1994; Gray, 1996; Klaassen, 2002; Haas et al., 2004; Pelletier, 2006; So et al., 1998), to model interest rate dynamics (Cai, 1994; Garcia and Perron, 1996), to study state-dependent impulse response functions (Sims and Zha, 2006; Caggiano et al., 2017), to identify structural shocks in SVAR models (Lanne et al., 2010; Herwartz and Lütkepohl, 2014; Lütkepohl et al., 2021), and more recently to improve measures of core inflation by allowing for multiple inflation regimes (Rodriguez-Rondon, 2024; Ahn and Luciani, 2024; Le Bihan et al., 2024). Hamilton (2016) provides a detailed survey of regime switching models in macroeconomics.

Outside of macroeconomic and financial applications, these models have also been applied in climate change research (Golosov et al., 2014; Dietz and Stern, 2015), environmental and energy economics (Cevik et al., 2021; Charfeddine, 2017), industrial organization(Aguirregabiria and Mira, 2007; Sweeting, 2013), and health economics (Hernández and Ochoa, 2016; Anser et al., 2021). Additionally, there is a related class of models–the Hidden Markov model–which has been widely used in computational molecular biology (Krogh et al., 1994; Baldi et al., 1994), handwriting and speech recognition (Rabiner and Juang, 1986; Nag et al., 1986; Rabiner and Juang, 1993; Jelinek, 1997), computer vision and pattern recognition (Bunke and Caelli, 2001), and other machine learning applications.

Given their empirical relevance, it is important to determine the number of regimes required to properly capture the nonlinearities present in the data, as this is not determined endogenously when estimating Markov switching models. However, the asymptotic results underlying conventional hypothesis testing procedures do not apply in this setting, because the regularity conditions required for such results are violated. Consequently, alternative hypothesis testing procedures have been proposed in the literature.

Notable contributions to testing the null hypothesis of a linear model against a model with two regimes include Hansen (1992), Hansen (1996a), Garcia (1998), Cho and White (2007), Marmer (2008), Carrasco et al. (2014), Kasahara et al. (2014), Dufour and Luger (2017), and Qu and Zhuo (2021). Testing the null of an M-regime model against the alternative of an $M+m$-regime model for $M \geq 1$ and $m = 1$ has been considered by Kasahara and Shimotsu (2018), where the authors show that the parametric bootstrap test can be asymptotically valid when imposing certain restrictions on the parameter space and focusing on univariate models with fixed or predetermined regressors. Qu and Zhuo (2021) present similar results regarding the parametric bootstrap for $M = m = 1$, but for a broader class of univariate models, albeit still under constrained parameter spaces. More recently, Rodriguez-Rondon and Dufour (2025) propose Monte Carlo likelihood ratio tests that accommodate cases with both $M \geq 1$ and $m \geq 1$ and also consider multivariate settings, which had not been addressed previously. Their testing procedures are the most general available and deal transparently with issues arising from violations of regularity conditions. Importantly, they do not require parametric restrictions, normality of errors, or stationarity of the underlying process, as the existence of an asymptotic distribution is unnecessary. This makes the tests applicable in a wider range of settings than the parametric bootstrap and allows their use in cases where the asymptotic validity of the parametric bootstrap has not been established. The maximized Monte Carlo version of their test even controls test size in finite samples, which is particularly relevant for many macroeconomic applications using quarterly data. In addition, this version of the test is robust to identification problems, which are common when working with Markov switching models.

Testing the number of regimes that a Markov switching model should include is an important step when determining the model's specification. However, carrying out these tests is not necessarily trivial, and conducting more than one test can quickly become tedious. As a result, we introduce **MSTest**, an R package designed to test the null hypothesis of M regimes against the alternative hypothesis of $M + m$ regimes for both univariate and multivariate models. The purpose of this R package is to enable users to determine the appropriate number of regimes for a given process by making hypothesis testing procedures readily available to a broad audience. It aims to facilitate the comparison of different testing procedures and the determination of the number of regimes in a model for economic research and policy relevant applications.

The package also allows users to simulate and estimate univariate and multivariate Markov switching models, as well as hidden Markov models. Estimation is provided through the use of the expectation–maximization (EM) algorithm or maximum likelihood estimation (MLE). **MSTest** utilizes **Rcpp** (Eddelbuettel and Balamuta (2018)), **RcppArmadillo** (Eddelbuettel et al. (2024)), and in some cases includes parallel computing options for computational efficiency. This is especially important given the computational burden associated with testing in the presence of nuisance parameters, as is the case here. The **MSTest** package includes the methodologies presented in Rodriguez-Rondon and Dufour (2025), Dufour and Luger (2017), Carrasco et al. (2014), and Hansen (1992). The parametric bootstrap discussed by Qu and Zhuo (2021) and Kasahara and Shimotsu (2018) is not explicitly implemented, but it can be performed using specific settings in combination with the local Monte Carlo likelihood ratio test of Rodriguez-Rondon and Dufour (2025).

The paper first describes the Markov switching and hidden Markov models for which the hypothesis tests are implemented. It then discusses the testing procedures included in the package and how they fit within the broader literature on hypothesis testing for Markov switching models, with particular attention to how they address violations of regularity conditions and identification challenges. This is followed by an overview of the main package functions and practical guidance intended to complement the documentation available on CRAN. The paper concludes with a brief summary.

2 Markov switching models

The **MSTest** package considers Markov switching models in which only the mean and variance are governed by the Markov process S_t, and we therefore describe such models here. We also consider a specific case of the Markov switching model–the hidden Markov model–in which no autoregressive coefficients are included as explanatory variables. In both cases, other exogenous explanatory variables may be included.

2.1 First-order Markov process

We begin by describing the first-order Markov process S_t that governs changes in the parameters of the Markov switching model. We assume that the process S_t is unobserved and evolves according to a first-order ergodic Markov chain with an $(M \times M)$ transition probability matrix given by

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

Here, $p_{ij} = P(S_t = j \mid S_{t-1} = i)$ denotes the probability that state i is followed by state j, and M is the total number of regimes. When considering M regimes, the process takes integer values $S_t \in \{1, \dots, M\}$. Additionally, the columns of the transition matrix $\mathbf{P}$ must sum to one in order for $\mathbf{P}$ to be a well-defined transition matrix (i.e., $\sum_{j=1}^{M} p_{ij} = 1$ for all i).

Considering the example in Hamilton (1994), where the Markov process has only two regimes, we require only a (2×2) transition matrix to summarize the transition probabilities $\mathbf{P}$ as follows:

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

We can also obtain the ergodic probabilities, $\pi = (\pi_1, \pi_2)'$, which are given by

$$\pi_1 = \frac{1 - p_{22}}{2 - p_{11} - p_{22}} \qquad \pi_2 = 1 - \pi_1$$

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

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

where $\mathbf{e}_{M+1}$ is the $(M+1)$th column of $\mathbf{I}_{M+1}$. These ergodic probabilities tell us on average, in the long-run, the proportion of time the process S_t spends in each regime.

2.2 Markov switching autoregressive models

A Markov switching model can be expressed as

$$y_t = \mu_{S_t} + \sum_{k=1}^{p} \phi_k (y_{t-1} - \mu_{S_{t-1}}) + Z_t \beta_z + \sigma_{S_t} \epsilon_t \tag{1}$$

where, in a univariate setting, y_t is a scalar, Z_t is a $(1 \times q_z)$ vector of exogenous variables whose coefficients do not depend on the latent Markov process S_t, and ϵ_t represents the error process, which, for example, may be distributed as a $\mathcal{N}(0,1)$. The error term is multiplied by the standard deviation σ_{S_t}, which may either depend on the Markov process or remain constant throughout (i.e., σ).

This Markov switching autoregressive model is labeled `MSARmdl` in **MSTest** when exogenous regressors Z_t are excluded, and `MSARXmdl` when they are included. These are the versions most commonly used in economic and financial applications, as well as in other time series-related settings. Other error distributions, such as the Student-t distribution, may be considered in future versions of the package. Currently, **MSTest** only considers Markov switching autoregressive models with normally distributed errors when simulating processes, and we therefore focus on this setup in the present paper.

Continuing with the example where a Markov switching model given by equation (1) has $M = 2$ regimes, such that $S_t = \{1, 2\}$, the sample log likelihood conditional on the first p observations of y_t is given by

$$L_T(\theta) = \log f(y_1^T | y_{-p+1}^0; \theta) = \sum_{t=1}^{T} \log f(y_t | \mathcal{Y}_{t-1}; \theta) \tag{2}$$

where $\mathcal{Y}_{t-1} = \sigma\text{-field}\{\ldots, Z_{t-1}, y_{t-2}, Z_t, y_{t-1}\}$ and $\theta = (\mu_1, \mu_2, \beta, \sigma_1, \sigma_2, vec(\mathbf{P}))$ and $vec(\cdot)$ is the vectorization operator which stacks the columns of a matrix to form a column vector. Here,

$$f(y_t | \mathcal{Y}_{t-1}; \theta) = \sum_{s_t=1}^{2} \sum_{s_{t-1}=1}^{2} \cdots \sum_{s_{t-p}=1}^{2} f(y_t, S_t = s_t, S_{t-1} = s_{t-1}, \ldots, S_{t-p} = s_{t-p} | \mathcal{Y}_{t-1}; \theta) \tag{3}$$

and more specifically

$$\begin{aligned} f(y_t, S_t = s_t, \ldots, S_{t-p} = s_{t-p} | \mathcal{Y}_{t-1}; \theta) = & \frac{\Pr(S_t^* = s_t^* | \mathcal{Y}_{t-1}; \theta)}{\sqrt{2\pi\sigma_{s_t}^2}} \\ & \times \exp\left\{ \frac{-[(y_t - \mu_{s_t}) - \sum_{k=1}^{p} \phi_k (y_{t-1} - \mu_{s_{t-1}}) - Z_t \beta_z]^2}{2\sigma_{s_t}^2} \right\} \end{aligned} \tag{4}$$

where we set

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

and $\Pr(S_t^* = s_t^* | \mathcal{Y}_{t-1}; \theta)$ is the probability that this occurs. Note that as in, Rodriguez-Rondon and Dufour (2025), here we denote the latent variable that determines the regimes at time t as S_t and let s_t denote the (observed) realization of S_t.

2.3 Markov switching VAR model

Krolzig (1997) generalized the univariate Markov switching autoregressive model to the multivariate setting and thus introduced the Markov switching vector autoregressive (MS-VAR) model. The MS-VAR model considered in the **MSTest** package can be expressed as

$$\boldsymbol{y}_t = \boldsymbol{\mu}_{S_t} + \boldsymbol{\Phi}_1(\boldsymbol{y}_{t-1} - \boldsymbol{\mu}_{S_{t-1}}) + \cdots + \boldsymbol{\Phi}_p(\boldsymbol{y}_{t-p} - \boldsymbol{\mu}_{S_{t-p}}) + Z_t\boldsymbol{\beta} + \boldsymbol{\Sigma}_{S_t}^{1/2}\boldsymbol{\epsilon}_t \quad (5)$$

where $\boldsymbol{y}_t = [y_{1,t}, \ldots, y_{q,t}]'$, $\boldsymbol{\mu}_{S_t} = [\mu_{1,S_t}, \ldots, \mu_{q,S_t}]'$, $\boldsymbol{\epsilon}_t = [\epsilon_{1,t}, \ldots, \epsilon_{q,t}]'$, $\boldsymbol{\Phi}_k$ is a $(q \times q)$ matrix containing the autoregressive parameters at lag k, $\boldsymbol{\beta}$ is now a $(q_z \times q)$ matrix, and $\boldsymbol{\Sigma}_{S_t} = \boldsymbol{\Sigma}_{S_t}^{1/2}(\boldsymbol{\Sigma}_{S_t}^{1/2})'$ is the $(q \times q)$ regime dependent covariance matrix. As in the univariate setting, the **MSTest** package also includes a version without exogenous regressors Z_t, namely `MSVARmdl`, and a version that allows for the inclusion of exogenous regressors, `MSVARXmdl`. More sophisticated versions of the MS-VAR model, as well as their likelihood functions, are described in Krolzig (1997), and we direct the interested reader to that reference for further details on these models and their components.

2.4 Hidden Markov model

Hidden Markov models (HMMs) can be shown to be a special case of the more general Markov switching model defined above. In particular, hidden Markov models do not necessarily have to be applied to time series data, and for this reason they typically do not include lags of the endogenous variable y_t.

For example, we can recover a Hidden Markov model from (5) by simply excluding lags of y_t as explanatory variables, yielding

$$\boldsymbol{y}_t = \boldsymbol{\mu}_{S_t} + Z_t\boldsymbol{\beta} + \sigma_{S_t}\epsilon_t \quad (6)$$

When $q = 1$, we recover a univariate HMM from (1). This version and its multivariate counterpart are the HMMs considered in the package **MSTest** and are labeled as `HMmdl`.

As described by An et al. (2013), dependence on past observations allows for more general interactions between y_t and S_t, which can be used to model more complex causal relationships between economic or financial variables of interest. Including past observations is a very common practice in economic time series applications as a way to control for stochastic trends, which may help explain why Markov switching models are more popular than basic HMMs in this literature.

2.5 Model estimation

Typically, Markov switching models are estimated using the Expectation-Maximization (EM) algorithm (see Dempster et al. (1977)), Bayesian methods, or the Kalman filter when employing a state-space representation of the model. In very simple cases, Markov switching models can also be estimated using maximum likelihood estimation (MLE). However, since the Markov process S_t is latent and, more importantly, because the likelihood function may exhibit several modes of equal height along with other unusual features that complicate MLE, this approach is less commonly used.

The **MSTest** package enables estimation of the models described above via the EM algorithm by setting `control = list(method = "EM")` or via MLE by setting `control = list(method = "MLE")` in the estimation functions, which are described below. In practice, empirical estimates can sometimes be improved by using the EM algorithm results as initial values in a Newton-type optimization algorithm (i.e., MLE method). This two-step estimation procedure is used to obtain the results presented in the empirical section of Rodriguez-Rondon and Dufour (2025), as well as in some other related works.

We omit a detailed explanation of the EM algorithm and MLE, as our focus is on describing the **MSTest** package. For interested readers, the estimation of Markov switching models via the EM algorithm and MLE is described in detail in Hamilton (1990) and Hamilton (1994), and, for Markov switching VAR models, in Krolzig (1997).

3 Hypothesis testing for number of regimes

When estimating a Markov switching model, the number of regimes must be specified by the researcher, as it is not determined endogenously during the estimation process. However, when considering testing for the number of regimes in a Markov switching model, conventional hypothesis testing procedures are no longer valid as the regularity conditions needed for asymptotic validity are not met. To address these challenges, various studies have proposed alternative methods that yield valid testing procedures.

The general hypothesis of interest is

$$H_0 : M = M_0$$
$$H_1 : M = M_0 + m$$

where, in principal, both $M_0, m \geq 1$, though most classical procedures only handle the simplest case,

$$H_0 : M = 1, \qquad H_1 : M = 2,$$

In this section, we describe some key procedures, focusing on those included in **MSTest**. We will see that the Monte Carlo LR procedures introduced in Rodriguez-Rondon and Dufour (2025) and implemented in this package are the most general procedures, allowing for testing in much more general settings, including $M_0 \geq 1$, $m \geq 1$, non-Gaussian errors, non-stationary data, and multivariate models. Their generality may sometimes come at the cost of increased computational burden, making other methods attractive in the very simple cases for which they are well suited. Here, we explain how these procedures fit within the existing literature and briefly review how they address identification failures and violations of regularity conditions.

3.1 Likelihood ratio–type tests

Hansen (1992) was the first to propose a testing procedure for Markov switching models when $M_0 = m = 1$, and we therefore begin with a review of this procedure, which is available in **MSTest**. The author provides a thorough discussion of the issues that complicate the likelihood ratio approach for testing the number of regimes in a Markov switching model. First, it is typically assumed that the likelihood function is locally quadratic in the region containing the null hypothesis and the globally optimal parameter estimates. However, as noted by the author, because some parameters are not identified under the null, this region is likely to be flat with respect to those unidentified parameters rather than quadratic. Issues related to unidentified nuisance parameters under the null hypothesis have been studied in Davies (1977), Davies (1987), Andrews and Ploberger (1994), and Dufour (2006). Second, it is commonly assumed that the score is positive; however, as described, it can be identically 0 under the restricted maximum likelihood estimator of a linear model, corresponding to the null hypothesis. Third, some parameters, such as the transition probabilities, may take values of 0 or 1, leading to the parameter boundary problem discussed in Andrews (1999) and Andrews

(2001). In addition, the likelihood surface may exhibit multiple local optima, implying that the null hypothesis need not lie in the same region as the global optimum.

Hansen (1992) introduces a new approach that does not require the conventional assumptions associated with likelihood ratio tests. Instead, the author models the likelihood function as an empirical process of the unknown parameters and uses empirical process theory to establish a bound for the asymptotic distribution of the standardized likelihood ratio test. Hansen (1992) formulates the hypothesis as follows:

$$\begin{aligned} H_0 &: \alpha = \alpha_0 \\ H_1 &: \alpha \neq \alpha_0 \end{aligned} \tag{7}$$

where α_0 represents the parameter values under the null and α the parameter values under the alternative. They begin by decomposing the likelihood ratio as such:

$$\begin{aligned} LR_n(\alpha) &= L_n(\alpha) - L_n(\alpha_0) \\ &= \Sigma_{i=0}^{n}[l_i(\alpha) - l_i(\alpha_0)] \\ &= R_n(\alpha) + Q_n(\alpha) \end{aligned} \tag{8}$$

where $R_n(\alpha) = E[LR_n(\alpha)]$ is the expectation of the likelihood ratio function and $Q_n(\alpha) = \Sigma_{i=1}^{n} q_i(\alpha) = [l_i(\alpha) - l_i(\alpha_0)] - E[l_i(\alpha) - l_i(\alpha_0)]$ is the deviations from the mean. Fluctuations in Q play an important role in identifying an optimum as:

$$\begin{aligned} \frac{1}{\sqrt{n}} LR_n(\alpha) &= \frac{1}{\sqrt{n}} R_n(\alpha) + \frac{1}{\sqrt{n}} Q_n(\alpha) \\ &= \frac{1}{\sqrt{n}} R_n(\alpha) + Q(\alpha) + o_p(1) \end{aligned} \tag{9}$$

and by using the fact that $R_n(\alpha) \leq 0$ for all α when the null hypothesis is true, we can see that $\frac{1}{\sqrt{n}} LR_n(\alpha) \leq \frac{1}{\sqrt{n}} Q_n(\alpha)$ and so it follows that,

$$P[\frac{1}{\sqrt{n}} LR_n \geq x] \leq P[\sup_{\alpha} \frac{1}{\sqrt{n}} Q_n(\alpha) \geq x] \to P[\sup_{\alpha} Q(\alpha) \geq x] \tag{10}$$

Thus, the distribution of the empirical process Q can provide a bound for the asymptotic distribution of the LR statistic. The test statistic is further standardized:

$$LR_n^* = \sup_{\alpha} LR_n^*(\alpha) \tag{11}$$

where

$$LR_n^*(\alpha) = \frac{LR_n(\alpha)}{V_n(\alpha)^{1/2}} \tag{12}$$

As suggested by equation ((11)), the issue of nuisance parameters is addressed by evaluating the standardized likelihood ratio statistic for different values of the parameter vector α. Specifically, the test statistic is evaluated over a grid of parameter values and optimized with respect to these nuisance parameters. To be precise, Hansen (1992) defines $\alpha = (\mu_2, \sigma_2, p_{11}, p_{22})$ as the vector of nuisance parameters, which includes the parameters of the second regime as well as the transition probabilities. The first-regime parameters $\theta = (\mu_1, \sigma_1, \phi_1, \ldots, \phi_p)$ are fully identified. The vector α is further decomposed into $\beta = (\mu_2, \sigma_2)$ and $\gamma = (p_{11}, p_{22})$, where β is treated as a parameter of interest and is set equal to (μ_1, σ_1) under the null hypothesis.

However, the process Q may be serially correlated for some values of α. To address this issue, Hansen (1996a) proposes a correction that should be applied when computing the asymptotic distribution of the test statistic. This correction is also incorporated in the implementation of this testing procedure in **MSTest**.

There are two main drawbacks to consider when using this likelihood ratio procedure for testing Markov switching models. First, the test provides only a bound for the standardized likelihood ratio statistic, which can be conservative. As a result, it is important to note that the critical values reported by this test in the **MSTest** package are not those of the standardized likelihood ratio statistic itself, but rather those of the process Q, which provides a bound for the standardized likelihood ratio statistic. Second, the procedure requires optimizing over the nuisance parameters via a grid search. While this approach is manageable when analyzing models with only a few switching parameters—such as the regime-specific parameters and the transition probabilities p_{11} and p_{22}; it can quickly become computationally intensive as the number of parameters allowed to switch across regimes increases. Despite these drawbacks, we include this test in **MSTest** because its properties are well understood and it has frequently been used as a benchmark for comparison in the literature.

Several likelihood ratio–based procedures have been proposed for testing the number of regimes in Markov switching models following the work of Hansen (1992). Garcia (1998) extend the approach of Hansen (1992) by reducing the dimensionality of the nuisance parameter space, thereby lowering computational burden. However, their procedure relies on asymptotic assumptions that are known to be problematic in Markov switching settings (Hansen, 1996b; Andrews and Ploberger, 1994).

Cho and White (2007) develop a quasi-likelihood ratio test that explicitly accounts for parameters lying on the boundary of the parameter space. While their framework highlights the importance of boundary issues for the asymptotic distribution of the test statistic, subsequent work has shown that it may fail to properly account for time dependence induced by autoregressive dynamics when parameters switch across regimes (Carter and Steigerwald, 2012; Cho and White, 2011). For this reason, and given the importance of autoregressive structures in economic and financial applications, this test is also not included in **MSTest**.

More recently, Qu and Zhuo (2021) provide a refined asymptotic characterization of likelihood ratio–based tests by modeling the likelihood ratio as an empirical process over a restricted parameter space. Their results establish the asymptotic validity of parametric bootstrap procedures for a broad class of Markov switching models, while also clarifying how nuisance parameters affect the limiting distribution and explaining why certain bootstrap implementations and standard information criteria may be unreliable in specific settings. Although this procedure is not currently implemented in **MSTest**, it represents an important contribution and may be incorporated in future versions of the package.

Kasahara and Shimotsu (2018) propose a likelihood ratio testing framework based on higher-order expansions and difference-in-quadratic-mean approximations (seeLiu and Shao (2003)), allowing for tests of $M_0 \geq 1$ regimes against $M_0 + 1$ regimes. They establish asymptotic validity of the parametric bootstrap under fixed or predetermined regressors. While the parametric bootstrap is not implemented directly in **MSTest**, it can be replicated using appropriate settings within the local Monte Carlo likelihood ratio test of Rodriguez-Rondon and Dufour (2025), which is included and applies in even more general settings.

In Rodriguez-Rondon and Dufour (2025), the authors propose the maximized Monte Carlo likelihood ratio test (MMC-LRT) and the local Monte Carlo likelihood ratio test (LMC-LRT), which can be used to compare very general Markov switching models. These procedures represent the most general class of testing methods currently available, as they are applicable in settings when both $M_0 \geq 1$ and $m \geq 1$. Further, as described in Rodriguez-Rondon and Dufour (2025), these Monte Carlo likelihood ratio tests remain valid when the process y_t is non-stationary, when the model exhibits non-Gaussian errors, when parameters lie on the boundary of the parameter space, and in multivariate settings, including the Markov switching VAR and multivariate hidden Markov models discussed above. More precisely, the MMC-LRT and LMC-LRT are the only test procedures available in the literature–and available in **MSTest**–that can be used to test multivariate Markov switching models. In addition, the MMC-LRT controls test size in finite samples and is robust to identification issues, properties that are particularly relevant for macroeconomic applications of Markov switching models, as emphasized in Rodriguez-Rondon and Dufour (2025).

Here, we provide a brief overview of the MMC-LRT and LMC-LRT procedures. Readers interested in further details are referred to the formal treatment in Rodriguez-Rondon and Dufour (2025) and to Dufour (2006) for additional discussion of the Monte Carlo techniques underlying these procedures. For simplicity of exposition, we consider the case of a null hypothesis corresponding to a linear model (i.e., $M_0 = 1$) against an alternative hypothesis with $M_0 + m = 2$ regimes, and focus on an autoregressive model in which only the mean and variance are allowed to switch across regimes.

Since these are likelihood ratio–based procedures, the log-likelihood values under both the null and alternative hypotheses are required. The log-likelihood for the model under the alternative hypothesis (and under the null hypothesis when $M_0 > 1$) is given by (2)–(4):

$$L_T(\theta_1) = \log f(y_1^T \,|\, y_{-p+1}^0;\, \theta_1) = \sum_{t=1}^{T} \log f(y_t \,|\, y_{-p+1}^{t-1};\, \theta_1) \tag{13}$$

where

$$\theta_1 = (\mu_1,\, \mu_2,\, \sigma_1,\, \sigma_2,\, \phi_1,\, \dots\,,\, \phi_p,\, p_{11},\, p_{22})' \in \Omega\,. \tag{14}$$

Here, 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

$$L_T^0(\theta_0) = \log f(y_1^T \,|\, y_{-p+1}^0;\, \theta_0) = \sum_{t=1}^{T} \log f(y_t \,|\, y_{-p+1}^{t-1};\, \theta_0) \tag{15}$$

where

$$f(y_t \,|\, y_{-p+1}^{t-1};\, \theta_0) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ \frac{-[y_t - \mu - \sum_{k=1}^{p} \phi_k (y_{t-k} - \mu)]^2}{2\sigma^2} \right\}, \tag{16}$$

$$\theta_0 = (\mu,\, \sigma^2, \phi_1,\, \dots\,,\, \phi_p)' \in \bar{\Omega}_0. \tag{17}$$

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

$$H_0 : \delta_1 = \delta_2 = \delta \quad \text{for some unknown} \delta = (\mu, \sigma)\,, \tag{18}$$

%

$$H_1 : (\delta_1,\, \delta_2) = (\delta_1^*,\, \delta_2^*) \;\; \text{for some unknown} \delta_1^* \neq \delta_2^*, \tag{19}$$

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

$$L_T^0(\theta_0) = L_T(\theta_1)\,, \quad \theta_1 \in \Omega_0, \tag{20}$$

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 vector of nuisance parameters: 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, LR_T, is in fact completely determined by $\theta_0 \in \bar{\Omega}_0$. As in Garcia (1998) and the parametric bootstrap procedure describe in Qu and Zhuo (2021) and Kasahara and Shimotsu (2018), it is assumed that the null hypothesis depends only on the mean, variance, and autoregressive coefficients. The likelihood ratio statistic for testing H_0 against H_1 can then written as

$$LR_T = 2[\bar{L}_T(H_1) - \bar{L}_T(H_0)] \tag{21}$$

where

$$\bar{L}_T(H_1) = \sup\{L_T(\theta_1) : \theta_1 \in \Omega\}\,, \tag{22}$$

$$\bar{L}_T(H_0) = \sup\{L_T^0(\theta_0) : \theta_0 \in \bar{\Omega}_0\} = \sup\{L_T(\theta_1) : \theta_1 \in \Omega_0\}\,. \tag{23}$$

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

$$LR(N,\, \theta_0) := [LR_T^{(1)}(\theta_0),\, \ldots\,, LR_T^{(N)}(\theta_0)]', \qquad \theta_0 \in \bar{\Omega}_0\,. \tag{24}$$

As discussed in Rodriguez-Rondon and Dufour (2025), the main assumptions required are that the random variables $LR_T^{(0)},\, LR_T^{(1)}(\theta_0),\, \ldots\,, LR_T^{(N)}(\theta_0)$ are exchangeable for some $\theta_0 \in \bar{\Omega}_0$ each with distribution function $F[x \,|\, \theta_0]$ (i.e., they are *i.i.d*). From here, we can compute the Monte Carlo p-value which is given by

$$\hat{p}_N[x \,|\, \theta_0] = \frac{N + 1 - R_{LR}[LR_T^{(0)};\, N]}{N + 1} \tag{25}$$

where

$$R_{LR}[LR_T^{(0)};\, N] = \sum_{i=1}^{N} I[LR_T^0 \geq LR_T^i(\theta_0)] \tag{26}$$

and $I(C) := 1$ if condition C holds, and $I(C) = 0$ otherwise. As can be seen from (26), $R_{LR}[LR_T^{(0)};$ $N]$ simply computes the rank of the test statistic from the observed data within the generated series $LR(N, \theta_0)$. Then, as shown in Rodriguez-Rondon and Dufour (2025), a critical region for this test statistic with level α is then given by

$$\sup_{\theta_0 \in \bar{\Omega}_0} \hat{p}_N[LR_T^{(0)} \,|\, \theta_0] \leq \alpha \tag{27}$$

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

$$\mathbb{P}\left[\sup\{\hat{p}_N[LR_T^{(0)} \,|\, \theta_0] : \theta_0 \in \bar{\Omega}_0\} \leq \alpha\right] \leq \alpha \tag{28}$$

under the null hypothesis and so it is a valid test with level α for H_0 and this result does not depend on the sample size T and so it is also valid in finite samples.

This is the maximized Monte Carlo likelihood ratio test, which requires searching for the maximum Monte Carlo p-value over the nuisance parameter space $\bar{\Omega}_0$. Since this space can be very large–and grows rapidly with the number of autoregressive components and the number of regimes–the authors propose a more efficient alternative based on searching over a consistent set C_T, as originally proposed in Dufour (2006). A consistent set can be constructed using a consistent point estimate. For example, let $\hat{\theta}_0$ denote a consistent point estimate of θ_0. Then, we can define

$$C_T = \{\theta_0 \in \bar{\Omega}_0 : \parallel \hat{\theta}_0 - \theta_0 \parallel < c\} \tag{29}$$

where c is a fixed positive constant that does not depend on T and $\|\cdot\|$ is the Euclidean norm in $\mathbb{R}^k$. A consistent set of interest may also be $C_T^* = C_T^{CI} \cup C_T^{\epsilon}$ where

$$C_T^{CI} = \{\theta_0 \in \bar{\Omega}_0 : \parallel \hat{\theta}_0 - \theta_0 \parallel < 2 \times S.E.(\hat{\theta}_0)\} \tag{30}$$

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

Hence, C_T^{CI} is defined by a confidence interval based on consistent point estimates, while C_T^{ϵ} is determined using a fixed constant ϵ that is independent of T. The union of these two sets allows for values that may lie outside the confidence interval for some parameters and within it for others, depending on the choice of ϵ. **MSTest** enables users to define the consistent set C_T by specifying only a fixed positive constant ϵ, only the confidence interval, or the union of both.

As discussed in Dufour (2006) and Rodriguez-Rondon and Dufour (2025), the solution to this optimization problem need not be unique, meaning that the maximum Monte Carlo p-value may correspond to multiple parameter vectors. For this reason, derivative-free numerical optimization methods are recommended to locate the maximum Monte Carlo p-value over the nuisance parameter space. **MSTest** allows users to employ the generalized simulated annealing algorithm (**GenSA**), the genetic algorithm (**GA**), and particle swarm optimization (**pso**); see Yang Xiang et al. (2013), Zambrano-Bigiarini et al. (2013), Scrucca et al. (2013), Dufour (2006), and Dufour and Neves (2019).

Finally, as described in Rodriguez-Rondon and Dufour (2025), the consistent set C_T can be chosen as the singleton $C_T = \{\hat{\theta}_0\}$, which yields the local Monte Carlo likelihood ratio test (LMC-LRT). In this case, the consistent set contains only the consistent point estimate $\hat{\theta}_0$, and so the Monte Carlo p-value depends solely on this estimate. The LMC-LRT can be interpreted as a finite-sample analogue of the parametric bootstrap. Here, asymptotic validity refers to the convergence of $\hat{\theta}_0$ to the true parameter value θ_0 as the sample size T increases, rather than to the existence of an asymptotic distribution for the test statistic or its critical values, as emphasized in studies such as Hansen (1992), Garcia (1998), Cho and White (2007), Qu and Zhuo (2021), and Kasahara and Shimotsu (2018).

Specifically, as with the parametric bootstrap, the LMC-LRT is valid only asymptotically as $T \to \infty$. However, unlike the parametric bootstrap, it does not require a large number of Monte Carlo replications (i.e., $N \to \infty$), since the procedure does not aim to approximate asymptotic critical values or rely on convergence of the test statistic to a limiting distribution. Instead, it uses critical values obtained directly from the simulated sample distribution. This design leads to substantial computational gains, as it avoids the need for extensive simulations to recover asymptotically valid critical values. Moreover, as shown in Rodriguez-Rondon and Dufour (2025), the LMC-LRT remains valid even in settings where an asymptotic distribution does not exist. This property further highlights why both the MMC-LRT and LMC-LRT procedures are more general than the parametric bootstrap.

That said, the parametric bootstrap procedures discussed in Qu and Zhuo (2021) and Kasahara and Shimotsu (2018) can be implemented by appropriately constraining the parameter space of the transition probabilities and (or) variance, in line with the assumptions of those studies. In such cases, using a sufficiently large number of simulations allows one to approximate asymptotic critical values when prior work has established the validity of the parametric bootstrap. Nevertheless, as emphasized in Rodriguez-Rondon and Dufour (2025), the MMC-LRT and LMC-LRT remain the most general likelihood ratio–based testing procedures currently available for Markov switching models.

3.2 Moment-based tests

Dufour and Luger (2017) propose an alternative approach to testing Markov switching models that avoids the statistical issues associated with likelihood ratio–type tests discussed above. In addition, their procedure is computationally less demanding than the likelihood ratio–based tests described previously, the parameter stability test discussed next, and it also allows the econometrician to perfectly control the test level through the Monte Carlo testing methods developed in Dufour (2006). However, their method is restricted to the case where $M_0 = m = 1$.

The moment-based test of Dufour and Luger (2017) is based on computing moments of the least squares residuals from autoregressive models estimated under the null hypothesis of $M_0 = 1$. More specifically, they focus on the mean, variance, skewness, and excess kurtosis of the least squares residuals. These moments are calculated as

$$M(\hat{\epsilon}) = \frac{|m_2 - m_1|}{\sqrt{s_1^2 + s_2^2}} \tag{32}$$

where, $m_1 = \frac{\Sigma_{t=1}^{T}\hat{\epsilon}_t\mathbf{1}[\hat{\epsilon}_t<0]}{\Sigma_{t=1}^{T}\mathbf{1}[\hat{\epsilon}_t<0]}$, $m_2 = \frac{\Sigma_{t=1}^{T}\hat{\epsilon}_t\mathbf{1}[\hat{\epsilon}_t>0]}{\Sigma_{t=1}^{T}\mathbf{1}[\hat{\epsilon}_t>0]}$, $s_1^2 = \frac{\Sigma_{t=1}^{T}(\hat{\epsilon}_t-m_1)^2\mathbf{1}[\hat{\epsilon}_t<0]}{\Sigma_{t=1}^{T}\mathbf{1}[\hat{\epsilon}_t<0]}$ and $s_2^2 = \frac{\Sigma_{t=1}^{T}(\hat{\epsilon}_t-m_2)^2\mathbf{1}[\hat{\epsilon}_t>0]}{\Sigma_{t=1}^{T}\mathbf{1}[\hat{\epsilon}_t>0]}$

$$V(\hat{\epsilon}) = \frac{\vartheta_2(\hat{\epsilon})}{\vartheta_1(\hat{\epsilon})} \tag{33}$$

where, $\vartheta_1 = \frac{\Sigma_{t=1}^{T}\hat{\epsilon}_t^2\mathbf{1}[\hat{\epsilon}_t^2<\hat{\sigma}^2]}{\Sigma_{t=1}^{T}\mathbf{1}[\hat{\epsilon}_t^2<\hat{\sigma}^2]}$, $\vartheta_2 = \frac{\Sigma_{t=1}^{T}\hat{\epsilon}_t^2\mathbf{1}[\hat{\epsilon}_t^2>\hat{\sigma}^2]}{\Sigma_{t=1}^{T}\mathbf{1}[\hat{\epsilon_t}^2>\hat{\sigma}^2]}$ and $\hat{\sigma}^2 = T^{-1}\Sigma_{t=1}^{T}\hat{\epsilon}_t^2$

$$S(\hat{\epsilon}) = |\frac{\Sigma_{t=1}^{T}\hat{\epsilon}_t^3}{T(\hat{\sigma}^2)^{3/2}}| \tag{34}$$

and

$$K(\hat{\epsilon}) = |\frac{\Sigma_{t=1}^{T}\hat{\epsilon}_t^4}{T(\hat{\sigma}^2)^2} - 3| \tag{35}$$

The testing procedure involves computing a test statistic for each moment, obtaining the corresponding individual p-values, and then combining these statistics using one of two different methods for aggregating independent tests. The first method is based on the minimum of the p-values and was originally proposed by Tippett et al. (1931) and Wilkinson (1951). In this case, the test statistic is given by

$$F_{min}(\hat{\epsilon}) = 1 - min\{\hat{G}_M[M(\hat{\epsilon})], \hat{G}_V[V(\hat{\epsilon})], \hat{G}_S[S(\hat{\epsilon})], \hat{G}_K[K(\hat{\epsilon})]\} \tag{36}$$

where for example, $\hat{G_M}[M(\hat{\epsilon})] = 1 - \hat{F_M}[M(\hat{\epsilon})]$ is the Monte Carlo p-value of $M(\hat{\epsilon})$. The second method of combining the test statistics involves taking the product of them. This method of combining test statistics was suggested by Fisher (1932) and Pearson (1933). In this case the the test statistic becomes,

$$F_{\times}(\hat{\epsilon}) = 1 - \{\hat{G}_M[M(\hat{\epsilon})] \times \hat{G}_V[V(\hat{\epsilon})] \times \hat{G}_S[S(\hat{\epsilon})] \times \hat{G}_K[K(\hat{\epsilon})]\} \tag{37}$$

Interested readers are referred to Dufour et al. (2004) and Dufour et al. (2014) for further discussion of these methods for combining test statistics.

Finally, the Monte Carlo p-value of the combined test statistic is given by

$$G_{F_{min}}[F_{min}(\hat{\epsilon}); N] = \frac{N + 1 - R_{F_{min}}[F_{min}(\hat{\epsilon}); N]}{N} \tag{38}$$

and

$$G_{F_{\times}}[F_{\times}(\hat{\epsilon}); N] = \frac{N + 1 - R_{F_{\times}}[F_{\times}(\hat{\epsilon}); N]}{N} \tag{39}$$

where $R_{F_{min}}$ and $R_{F_{\times}}$ are the ranks of $F_{min}(\hat{\epsilon})$ and $F_{\times}(\hat{\epsilon})$ in $F_{min}(\hat{\eta}_1)$, ..., $F_{min}(\hat{\eta}_{N-1})$ and $F_{\times}(\hat{\eta}_1)$, ...,$F_{\times}(\hat{\eta}_{N-1})$ respectively, when ordered. Also, $\hat{\eta} = \eta - \bar{\eta}$ and $\eta \sim N(0, I_T)$.

The computational efficiency of this test makes it easily extendable to a maximized Monte Carlo framework when nuisance parameters are present. Moreover, it is not subject to the same degree of statistical difficulties–such as unidentified parameters under the null hypothesis–that arise in Hansen (1992), Garcia (1998), and Carrasco et al. (2014). This is because the transition probabilities, as well as the mean and variance, do not need to be treated as nuisance parameters. Instead, only the parameters associated with explanatory variables may be unidentified under the null and are therefore treated as nuisance parameters.

While Garcia (1998) also reduce the nuisance parameter space by treating only the transition probabilities p_{11} and p_{22} as nuisance parameters, the reduction achieved by the moment-based test is even more substantial, making it particularly tractable for autoregressive models with multiple lags. Although this moment-based test is limited to comparing linear models against two-regime Markov switching alternatives, it is the least computationally intensive procedure available and can be computed in a matter of seconds, even when using the maximized Monte Carlo version of the test.

3.3 Optimal test for regime switching

Carrasco et al. (2014) propose a test for parameter constancy in random coefficient and Markov switching models that can be viewed as an extension of the information matrix test of White (1982). A key advantage of this procedure is that it requires estimation only under the null hypothesis, a feature it shares with the moment-based test of Dufour and Luger (2017). This is particularly attractive in Markov switching settings, where estimating models under the alternative can be computationally demanding due to nonlinearity and the presence of multiple local optima. In contrast, likelihood ratio–based procedures such as those of Hansen (1992), Garcia (1998), Cho and White (2007), Qu and Zhuo (2021), Kasahara and Shimotsu (2018), and Rodriguez-Rondon and Dufour (2025) require estimation under both the null and alternative hypotheses.

Using the Neyman–Pearson lemma, Carrasco et al. (2014) show that their test is asymptotically locally equivalent to the likelihood ratio test. However, simulation evidence in Rodriguez-Rondon and Dufour (2025) and discussion in Qu and Zhuo (2021) suggest that likelihood ratio–based approaches may exhibit higher power in certain settings, such as when only the mean switches and regime persistence is high. Moreover, the procedure of Carrasco et al. (2014) involves bootstrap methods and optimization over nuisance parameters, which can make it more computationally intensive than simpler alternatives such as the moment-based test of Dufour and Luger (2017) and even the MMC-LRT and LMC-LRT of Rodriguez-Rondon and Dufour (2025) (e.g., when variance is allowed to switch).

The authors formulate the hypothesis in the following way:

$$\begin{aligned} H_0 &: \theta_t = \theta_0 \\ H_1 &: \theta_t = \theta_0 + \eta_t \end{aligned} \tag{40}$$

where the switching variable η_t is unobservable, stationary, and may depend on nuisance parameters β. Their test makes use of the second derivatives of the log-likelihood and the outer products of the scores, as in the information matrix test, with the addition of an extra term that captures the serial dependence of the time-varying coefficients. This means that the form of the test depends on the latent process η_t only through its second-order properties. Additionally, the distribution of η_t is assumed to exist even under the null, but it does not play a role in the distribution of the data $(y_T, y_{T-1}, y_{T-2}, \ldots, y_1)$ under the null. That is, under the null hypothesis, they are mutually exclusive.

The authors first propose a Sup-type test as in Davies (1987) to combat the presence of nuisance parameters. They set $\eta_t = chS_t$, where c is a scalar specifying the amplitude of the change, h a vector specifying the direction of the alternative and S_t is a Markov-chain, which follows an autoregressive process such as $S_t = \rho S_{t-1} + e_t$, where e_t is i.i.d. U[-1,1] and $-1 < \rho < 1$ so that S_t is bounded by support $(-1/(1-|\rho|), 1/(1-|\rho|)$ and has zero mean. Letting $\beta = (c^2, h', \rho')$ be the vector of nuisance parameters, we can write

$$\mu_{2,t}(\beta,\theta) = \frac{1}{2}c^2 h'[(\frac{\partial l_t}{\partial\theta\partial\theta'} + (\frac{\partial l_t}{\partial\theta})(\frac{\partial l_t}{\partial\theta})') + 2\sum_{s<t}\rho^{(t-s)}(\frac{\partial l_t}{\partial\theta})(\frac{\partial l_t}{\partial\theta})']h \tag{41}$$

which allows us to get the expression

$$\text{supTS} = \sup_{\{h,p:||h||=1,\underline{\rho}<\rho<\bar{\rho}\}} = \frac{1}{2}(max(0, \frac{\Gamma_T^*}{\sqrt{\hat{\epsilon^{*}}'\hat{\epsilon^*}}}))^2 \tag{42}$$

as in Carrasco et al. (2014), where $\mu^*_{2,t}(\beta,\theta) = \mu_{2,t}(\beta,\theta)/c^2$, $\Gamma^*_T = \Gamma^*_T(\beta,\theta) = \sum_t \mu^*_{2,t}(\beta,\theta)/\sqrt{T}$ and $\hat{\epsilon^*}$ are the residuals from regressing $\mu^*_{2,t}(\beta,\theta)$ on $l_t^{(1)}(\hat{\theta})$ so that Γ^*_T and $\hat{\epsilon^*}$ are not dependent on c^2. As previously mentioned, this methodology involves bootstrapping over the distributions of the nuisance parameters. As a result, one must choose a prior distribution for the nuisance parameters. The most commonly used distribution in this case is the uniform distribution, and this is what is implemented in the **MSTest** package. However, since the parameter c^2 is not necessarily bounded from above, a uniform distribution may not always be appropriate. As a result, Carrasco et al. (2014) also suggest using an Exponential-type test as in Andrews and Ploberger (1994). They propose the following statistic:

$$expTS = \int_{\{\underline{\rho}\leq\rho\leq\bar{\rho},||h||<1\}} \Psi(h,\rho)d\rho dh \tag{43}$$

where

$$\Psi(h,\rho) = \begin{cases} \sqrt{2\pi} exp[\frac{1}{2}(\frac{\Gamma_T^*}{\sqrt{\hat{\epsilon^*}'\hat{\epsilon^*}}}-1)^2]\Phi(\frac{\Gamma_T^*}{\sqrt{\hat{\epsilon^*}'\hat{\epsilon^*}}}-1) & \text{if } \hat{\epsilon^*}'\hat{\epsilon^*} \neq 0. \\ 1 & \text{otherwise.} \end{cases} \tag{44}$$

The tests proposed by Carrasco et al. (2014) have been widely used in empirical applications, including Hamilton (2005), Warne and Vredin (2006), Kahn and Rich (2007), Morley and Piger (2012), and Dufrénot et al. (2011), in testing MS-GARCH models by Hu and Shin (2008), and as benchmark procedures in Dufour and Luger (2017), Qu and Zhuo (2021), and Rodriguez-Rondon and Dufour (2025). Owing to their broad use and optimality properties, these tests are also included in the **MSTest** package.

4 The R package MSTest

The **MSTest** package is designed to conduct hypothesis tests for the number of regimes in Markov switching models. Because many of the testing procedures implemented in the package require estimation of restricted and unrestricted models, as well as simulation under the null hypothesis, **MSTest** also provides tools for simulating and estimating Markov switching models. These features, however, are intended primarily to support the hypothesis testing procedures and are not the main focus of the package. In particular, they are implemented to ensure full compatibility with the testing framework.

For these reasons, **MSTest** distinguishes itself through its implementation of testing procedures that remain valid under violations of standard regularity conditions. In addition, the use of **Rcpp** and parallel computing options enhances computational efficiency, making it feasible to handle nuisance parameters and Monte Carlo–based procedures in practice. While alternative software may exist for simulating and estimating Markov switching models, **MSTest** offers a unique combination of generality, computational efficiency, and access to modern hypothesis testing procedures, which sets it apart from existing tools.

4.1 Data sets

The **MSTest** package includes three samples of U.S. real GNP and one sample of U.S. real GDP, all of which are readily accessible once the package is loaded. Specifically, it provides the original sample used in Hamilton (1989), the sample ending in 2010 first considered in Carrasco et al. (2014), and an extended sample spanning from the second quarter of 1947 to the second quarter

Label	Description
`hamilton84GNP`	Sample originally considered in Hamilton (1989), spanning 1951Q2 to 1984Q4.
`chp10GNP`	U.S. real GNP sample used in Carrasco et al. (2014) and Dufour and Luger (2017), spanning 1951Q2 to 2010Q4.
`USGNP`	U.S. real GNP sample used in Rodriguez-Rondon and Dufour (2025), spanning 1947Q2 to 2024Q2.
`USRGDP`	U.S. real GDP sample used in Rodriguez-Rondon and Dufour (2025), spanning 1947Q2 to 2024Q2.

Table 1: U.S. real GNP and GDP data sets included in **MSTest**

of 2024. The U.S. real GDP series also covers this extended period. These data sets have been used in Hansen (1992), Carrasco et al. (2014), Dufour and Luger (2017), and Rodriguez-Rondon and Dufour (2025), among others, to test for the number of regimes in Markov switching models and to illustrate the performance of various testing procedures. Table 1 reports the labels used to identify each data set in **MSTest** and summarizes their sample spans.

These data sets can be loaded using the following commands once **MSTest** has been attached:

```
data("hamilton84GNP", package = "MSTest")
data("chp10GNP", package = "MSTest")
data("USGNP", package = "MSTest")
data("USRGDP", package = "MSTest")
```

Each data set contains three columns: (1) `Date`, (2) `GNP` or `RGDP`, and (3) `GNP_gr` or `RGDP_gr`. The `Date` column is of class `Date`, constructed using `as.Date()`. The second column reports the level of U.S. real GNP or GDP, while the third column contains the corresponding growth rate.

4.2 Simulation

This section describes the simulation functions available in **MSTest**. Simulation plays a central role in **MSTest**, as many of the testing procedures rely on Monte Carlo methods to obtain valid critical values in the presence of nuisance parameters and identification failures. These functions are primarily used internally by the hypothesis testing procedures; particularly those that rely on simulation to obtain sample or asymptotic null distributions, such as `LMCLRTest`, `MMCLRTest`, `DLMCTest`, `DLMMCTest`, and `CHPTest`. They may also be useful to users developing new estimation or testing procedures for Markov switching models who wish to assess performance through controlled experiments.

Function	Description
simuNorm	Simulates a normally distributed process, with optional exogenous regressors.
simuAR	Simulates an autoregressive process with p lags (AR(p)).
simuARX	Simulates an AR(p) process with exogenous regressors.
simuVAR	Simulates a vector autoregressive process with p lags (VAR(p)).
simuVARX	Simulates a VAR(p) process with exogenous regressors.
simuMSAR	Simulates a Markov switching AR(p) process.
simuMSARX	Simulates a Markov switching AR(p) process with exogenous regressors.
simuMSVAR	Simulates a Markov switching VAR(p) process.
simuMSVARX	Simulates a Markov switching VAR(p) process with exogenous regressors.
simuHMM	Simulates a hidden Markov model (HMM), with optional exogenous regressors.

Table 2: Simulation functions available in **MSTest**

Table 2 lists the simulation functions provided by **MSTest** and the classes of processes they generate. Each function takes as input a `List` specifying the data-generating process (DGP). While we present illustrative examples below, a complete description of each function's arguments is available in the package documentation on CRAN.

As an illustration, consider the `simuNorm` function, which generates a univariate or multivariate Gaussian process. The required input is a `List` containing: the sample size (`n`), the number of series (`q`) (where `q=1` indicates a univariate process and `q>1` indicates a multivariate process), a $(q \times 1)$ vector of means, and a $(q \times q)$ covariance matrix. Users may optionally specify a `burnin` period to discard initial observations; for `simuNorm`, the default is zero, whereas autoregressive processes use a larger default burn-in to reduce sensitivity to initialization. Alternatively, users may supply a custom matrix of innovations via the `eps` argument.

The code below illustrates how to simulate several linear and regime-switching processes using the functions in **MSTest**, including multivariate normal, AR, VAR, Markov switching AR and VAR, and Hidden Markov models.

```
mdl_norm <- list(n     = 500,
                 q     = 2,
                 mu    = c(5, -2),
                 sigma = rbind(c(5.0, 1.5),
                               c(1.5, 1.0)))
simu_norm <- simuNorm(mdl_norm)
```

```
mdl_ar <- list(n     = 500,
               mu    = 5,
               sigma = 1,
               phi   = c(0.75))
simu_ar <- simuAR(mdl_ar)

mdl_var <- list(n     = 500,
                p     = 1,
                q     = 2,
                mu    = c(5, -2),
                sigma = rbind(c(5.0, 1.5),
                              c(1.5, 1.0)),
                phi   = rbind(c(0.50, 0.30),
                              c(0.20, 0.70)))
simu_var <- simuVAR(mdl_var)

mdl_hmm <- list(n     = 500,
                q     = 2,
                mu    = rbind(c(5, -2),
                              c(10, 2)),
                sigma = list(rbind(c(5.0, 1.5),
                                   c(1.5, 1.0)),
                             rbind(c(7.0, 3.0),
                                   c(3.0, 2.0))),
                k     = 2,
                P     = rbind(c(0.95, 0.10),
                              c(0.05, 0.90)))
simu_hmm <- simuHMM(mdl_hmm)

mdl_ms <- list(n     = 500,
               mu    = c(5,10),
               sigma = c(1,1),
```

```
               phi   = c(0.75),
               k     = 2,
               P     = rbind(c(0.95, 0.10),
                             c(0.05, 0.90)))
simu_msar <- simuMSAR(mdl_ms)


mdl_msvar <- list(n     = 500,
                  p     = 1,
                  q     = 2,
                  mu    = rbind(c(5, -2),
                                c(10, 2)),
                  sigma = list(rbind(c(5.0, 1.5),
                                     c(1.5, 1.0)),
                               rbind(c(7.0, 3.0),
                                     c(3.0, 2.0))),
                  phi   = rbind(c(0.50, 0.30),
                                c(0.20, 0.70)),
                  k     = 2,
                  P     = rbind(c(0.95, 0.10),
                                c(0.05, 0.90)))
simu_msvar <- simuMSVAR(mdl_msvar)
```

The simulated processes are shown in Figure 1. Regime switching is clearly visible in the Markov switching specifications, particularly when comparing the autoregressive process (middle left) with the Markov switching autoregressive process (middle right). Similar differences arise in the multivariate setting when comparing the VAR and MS-VAR processes.

4.3 Model estimation

Here, we briefly describe the functions available in **MSTest** for estimating linear and Markov switching models. In total, ten model classes can be estimated, which are listed in Table @ref(tab:models} along with their labels in **MSTest** and corresponding model specifications. Although `Nmdl` and `HMmdl` are written using multivariate notation, the functions automatically detect a univariate setting

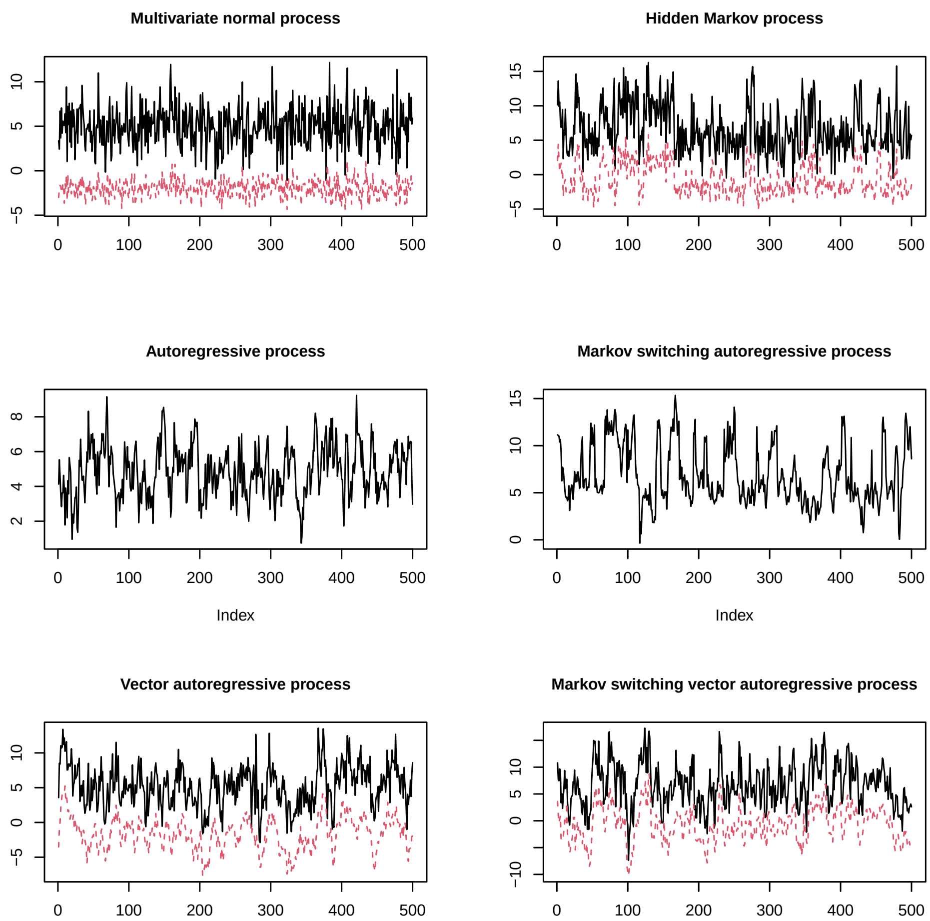

Figure 1: Simulated linear (left column) and Markov switching (right column) processes..

Model	Label	Equation
$N(\boldsymbol{\mu}+\boldsymbol{x}_t\boldsymbol{\beta},\boldsymbol{\Sigma})$	`Nmdl`	$\boldsymbol{y}_t=\boldsymbol{\mu}+\boldsymbol{x}_t\boldsymbol{\beta}+\boldsymbol{\Sigma}^{1/2}\boldsymbol{\epsilon}_t$
AR(p)	`ARmdl`	$y_t=\mu+\sum_{k=1}^{p}\phi_k(y_{t-k}-\mu)+\sigma\epsilon_t$
ARX(p)	`ARXmdl`	$y_t=\mu+\sum_{k=1}^{p}\phi_k(y_{t-k}-\mu)+\boldsymbol{x}_t\boldsymbol{\beta}+\sigma\epsilon_t$
VAR(p)	`VARmdl`	$\boldsymbol{y}_t=\boldsymbol{\mu}+\sum_{k=1}^{p}(\boldsymbol{y}_{t-k}-\boldsymbol{\mu})\boldsymbol{\Phi}_k+\boldsymbol{\Sigma}^{1/2}\boldsymbol{\epsilon}_t$
VARX(p)	`VARXmdl`	$\boldsymbol{y}_t=\boldsymbol{\mu}+\sum_{k=1}^{p}(\boldsymbol{y}_{t-k}-\boldsymbol{\mu})\boldsymbol{\Phi}_k+\boldsymbol{x}_t\boldsymbol{\beta}+\boldsymbol{\Sigma}^{1/2}\boldsymbol{\epsilon}_t$
MS-AR(p)	`MSARmdl`	$y_t=\mu_{s_t}+\sum_{k=1}^{p}\phi_k(y_{t-k}-\mu_{s_t})+\sigma_{s_t}\epsilon_t$
MS-ARX(p)	`MSARXmdl`	$y_t=\mu_{s_t}+\sum_{k=1}^{p}\phi_k(y_{t-k}-\mu_{s_t})+\boldsymbol{x}_t\boldsymbol{\beta}+\sigma_{s_t}\epsilon_t$
MS-VAR(p)	`MSVARmdl`	$\boldsymbol{y}_t=\boldsymbol{\mu}_{s_t}+\sum_{k=1}^{p}\boldsymbol{\Phi}_k(\boldsymbol{y}_{t-k}-\boldsymbol{\mu}_{s_{t-k}})+\boldsymbol{\Sigma}_{s_t}^{1/2}\boldsymbol{\epsilon}_t$
MS-VARX(p)	`MSVARXmdl`	$\boldsymbol{y}_t=\boldsymbol{\mu}_{s_t}+\sum_{k=1}^{p}\boldsymbol{\Phi}_k(\boldsymbol{y}_{t-k}-\boldsymbol{\mu}_{s_{t-k}})+\boldsymbol{x}_t\boldsymbol{\beta}+\boldsymbol{\Sigma}_{s_t}^{1/2}\boldsymbol{\epsilon}_t$
HMM	`HMmdl`	$\boldsymbol{y}_t=\boldsymbol{\mu}_{s_t}+\boldsymbol{x}_t\boldsymbol{\beta}+\boldsymbol{\Sigma}_{s_t}^{1/2}\boldsymbol{\epsilon}_t$

Table 3: Models and their specifications available in **MSTest**.

when a $(T \times 1)$ vector is provided as input. Models whose labels include an “X’’ allow for the inclusion of exogenous regressors. While `Nmdl` and `HMmdl` do not follow this naming convention, exogenous regressors can always be included in these models as well.

The simulation functions introduced in the previous subsection were used to generate a multivariate hidden Markov process with $q = 2$ series, a Markov switching autoregressive process with $p = 1$ lag, and a Markov switching vector autoregressive process with $p = 1$ lag and $q = 2$ series. The output of each simulation function is a list containing the simulated data, the true latent state variable S_t, and other elements of the data-generating process (DGP). These simulated series are now used as inputs to the corresponding estimation functions, as illustrated below for the hidden Markov model and the two Markov switching models.

```
control <- list(msmu   = TRUE,
                msvar  = TRUE,
                method = "EM",
                use_diff_init = 30)
mdl_est_hmm <- HMmdl(simu_hmm[["y"]], k = 2, control = control)
summary(mdl_est_hmm)
```

```
#>
#> Hidden Markov Model
#>               coef     s.e.
#> mu_1,1    9.927900 0.194200
#> mu_2,1    1.910400 0.108350
```

```
#> mu_1,2    5.065500 0.132740
#> mu_2,2   -1.971900 0.062588
#> sig_11,1  7.194500 0.738910
#> sig_12,1  3.083600 0.369440
#> sig_22,1  2.110100 0.231420
#> sig_11,2  5.028700 0.421850
#> sig_12,2  1.625900 0.170000
#> sig_22,2  1.059200 0.094782
#> p_11      0.918020 0.069019
#> p_12      0.081976 0.021094
#> p_21      0.056677 0.014690
#> p_22      0.943320 0.057771
#>
#> log-likelihood =  -1850.738
#> AIC =  3729.476
#> BIC =  3788.481
#>
#> Residuals:
#>        Min       1Q    Median      3Q    Max
#> Y1 -7.0846 -1.53990 -0.048561 1.59840 6.4115
#> Y2 -2.9423 -0.85197  0.010227 0.73751 3.8617
```

```
control <- list(msmu   = TRUE,
                msvar  = FALSE,
                method = "EM",
                use_diff_init = 30)
mdl_est_msar <- MSARmdl(simu_msar[["y"]], p = 1, k = 2, control = control)
summary(mdl_est_msar)
```

```
#>
#> Markov Switching Autoregressive Model
#>             coef      s.e.
#> mu_1  10.075000 0.2365200
```

```
#> mu_2   4.902400 0.1774800
#> phi_1  0.746200 0.0296370
#> sig    0.952880 0.0605680
#> p_11   0.881820 0.0933220
#> p_12   0.118180 0.0341480
#> p_21   0.029999 0.0086569
#> p_22   0.970000 0.0489730
#>
#> log-likelihood =  -786.1795
#> AIC =  1588.359
#> BIC =  1622.06
#>
#> Residuals:
#>        Min       1Q     Median      3Q  Max
#> Y1 -2.5877 -0.66312 0.00085905 0.70483 2.43
```

```
control <- list(msmu   = TRUE,
                msvar  = TRUE,
                method = "EM",
                use_diff_init = 30)
mdl_est_msvar <- MSVARmdl(simu_msvar[["y"]], p = 1, k = 2, control = control)
summary(mdl_est_msvar)
```

```
#>
#> Markov Switching Vector Autoregressive Model
#>              coef     s.e.
#> mu_1,1   10.713000 0.407430
#> mu_2,1    2.753400 0.381970
#> mu_1,2    4.486100 0.434060
#> mu_2,2   -2.367900 0.395930
#> phi_1,11  0.460710 0.058102
#> phi_1,12  0.225140 0.084482
#> phi_1,21  0.170420 0.030609
```

```
#> phi_1,22  0.657350 0.041134
#> sig_11,1  8.246400 0.764890
#> sig_12,1  3.714500 0.396390
#> sig_22,1  2.465400 0.249870
#> sig_11,2  4.907800 0.459810
#> sig_12,2  1.417900 0.190820
#> sig_22,2  1.075300 0.109260
#> p_11      0.960050 0.080117
#> p_12      0.039947 0.015131
#> p_21      0.040533 0.015296
#> p_22      0.959470 0.077868
#>
#> log-likelihood =  -1878.007
#> AIC =  3792.014
#> BIC =  3867.841
#>
#> Residuals:
#>        Min      1Q    Median      3Q    Max
#> Y1 -7.8858 -1.7021 -0.040940 1.70520 7.0405
#> Y2 -3.9873 -0.7955  0.047291 0.83282 3.6043
```

In all three examples, we set `method = "EM"` in the `control` list to estimate the models using the expectation–maximization (EM) algorithm. As discussed previously and detailed in the package documentation, users may alternatively specify `method = "MLE"` to employ direct maximum likelihood estimation. The arguments `msmu` and `msvar` determine whether the mean and variance are allowed to switch across regimes. Setting either option to `FALSE` imposes regime invariance on the corresponding parameter.

The option `use_diff_init = 30` instructs the estimation routine to run the optimizer thirty times using different initial values. The solution associated with the highest log-likelihood is retained as the primary output, while results from all optimization runs are stored in the `trace` element of the output list. All estimation functions return `S3` objects, for which dedicated `print()` and `summary()` methods are provided. The `summary()` method reports parameter estimates, log-likelihood values, information criteria (AIC and BIC), and diagnostic measures such as residual quantiles. In this

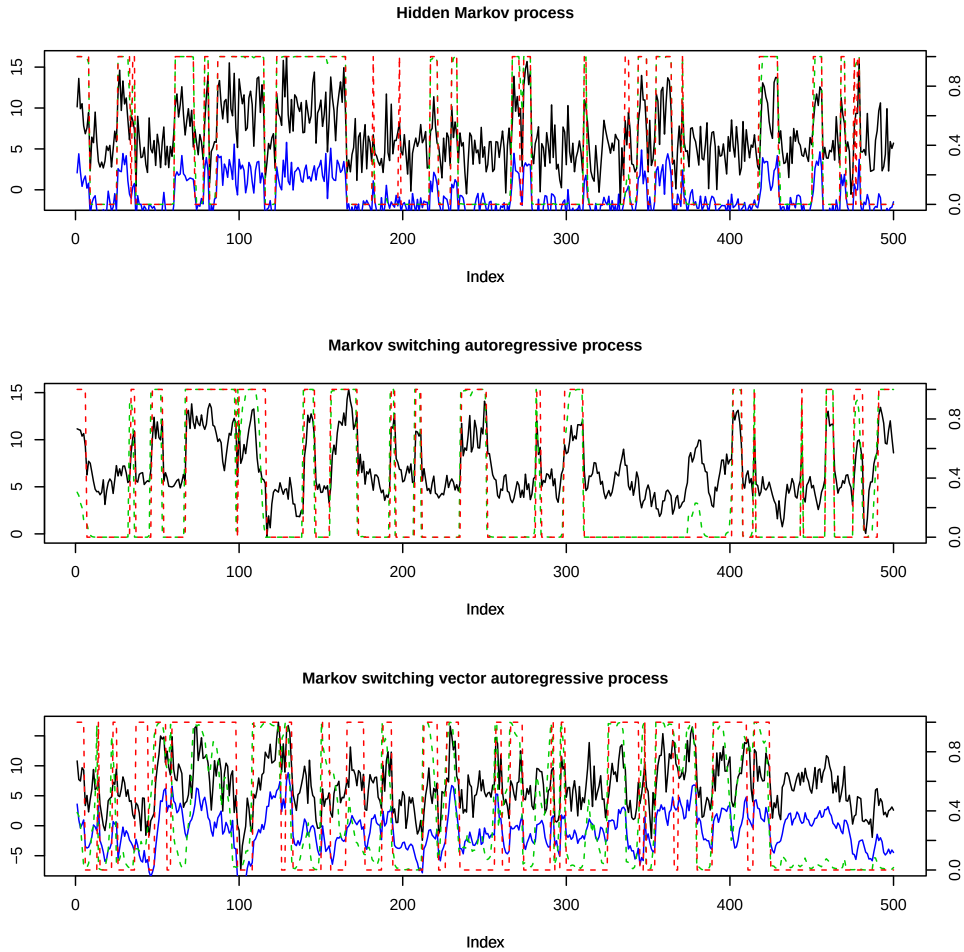

Figure 2: Simulated processes (black and blue), true regime states (red - dashed), and model estimated smoothed probabilities (green - dashed).

controlled setting, the estimated parameters can also be directly compared with their true DGP values.

Figure 2 displays the simulated series (black and blue), the true latent regime states S_t used in the data-generating process (red, dashed), and the estimated smoothed regime probabilities obtained from the fitted models (green, dashed). Overall, the estimated smoothed probabilities track the true regime changes closely across all three models.

For the hidden Markov and Markov switching autoregressive models, regime transitions are captured particularly well. In the Markov switching VAR case, regime classification is somewhat less precise over certain intervals, reflecting the increased complexity of multivariate models. This behavior is expected and can often be mitigated by increasing the sample size or by using a larger number of initial values in the estimation procedure.

Function	Description
`LMCLRTest`	Local Monte Carlo likelihood ratio test proposed in Rodriguez-Rondon and Dufour (2025).
`MMCLRTest`	Maximized Monte Carlo likelihood ratio test proposed in Rodriguez-Rondon and Dufour (2025).
`DLMCTest`	Local Monte Carlo moment-based test for Markov switching autoregressive models proposed in Dufour and Luger (2017).
`DLMMCTest`	Maximized Monte Carlo moment-based test for Markov switching autoregressive models proposed in Dufour and Luger (2017).
`CHPTest`	Optimal test for parameter constancy in Markov switching models proposed in Carrasco et al. (2014).
`HLRTest`	Likelihood ratio–based test of Hansen (1992), using empirical process theory to obtain the asymptotic distribution of a bound for the LR statistic.

Table 4: Hypothesis tests available in **MSTest**.

4.4 Hypothesis testing

This section describes the hypothesis testing functions available in **MSTest**. The package is designed with ease of use in mind and, in most cases, requires the user to provide only the series y_t (and Z_t when applicable) along with the number of regimes to be tested.

Table 4 summarizes the hypothesis testing functions currently implemented in **MSTest**. All testing functions return the value of the test statistic, the associated p-value, and, when applicable, critical values and parameter estimates under the null and alternative hypotheses.

It is important to clarify the interpretation of the reported critical values. In the case of `HLRTest()`, the values labeled as critical values correspond to those of the process Q introduced in Hansen (1992), which provide an upper bound for the likelihood ratio statistic rather than critical values for the LR statistic itself. Similarly, the values returned by `LMCLRTest()`, `MMCLRTest()`, `DLMCTest()`, and `DLMMCTest` and referred to as critical values correspond to empirical quantiles of the simulated null distribution. Throughout the package, the term "critical values' ' is therefore used in a generic sense to denote rejection thresholds derived from the relevant reference distribution.

4.4.1 Monte Carlo likelihood ratio test

We begin by reviewing the local Monte Carlo likelihood ratio test (LMC-LRT). As indicated in Table 4, this test is implemented in **MSTest** via the function `LMCLRTest()`. Because the LMC-LRT requires estimating both the restricted and unrestricted models in order to simulate the null distribution, the same estimation options that are passed to the model estimation functions can also be supplied to the testing procedure through the arguments `mdl_h0_control` and `mdl_h1_control`.

The same structure applies to the maximized Monte Carlo likelihood ratio test discussed below.

Through these control lists, users can specify whether the mean and variance are allowed to switch across regimes, the estimation method, the number of initial values used in optimization, and related options. The `LMCLRTest()` function also allows users to specify a different number of initial values for model estimation during null distribution simulation via the `use_diff_init_sim` argument. By default, this option is set equal to the value used when estimating the model on the observed data, as specified in `mdl_h0_control` and `mdl_h1_control`, which is generally recommended.

In addition to estimation controls, the `LMCLRTest()` function requires the user to specify the number of lags `p`, the number of regimes under the null hypothesis `k0`, and the number of regimes under the alternative hypothesis `k1`. Below, we illustrate the LMC-LRT using the Markov switching vector autoregressive model with $p = 1$ lag simulated in the previous subsection. We test the null hypothesis $H_0 : M = 1$ against the alternative $H_1 : M = 2$. Since the data were generated from a two-regime model, rejection of the null hypothesis is expected.

```
lmc_control = list(N = 19,
                   mdl_h0_control = list(const  = TRUE,
                                         getSE  = FALSE),
                   mdl_h1_control = list(msmu   = TRUE,
                                         msvar  = TRUE,
                                         getSE  = FALSE,
                                         method = "EM",
                                         use_diff_init = 1))
lmclrt <- LMCLRTest(simu_msvar[["y"]], p = 1, k0 = 1 , k1 = 2, control = lmc_control)
summary(lmclrt)
```

```
#>
#> Restricted Model
#>              coef
#> mu_1     7.77650
#> mu_2     0.33925
#> phi_1,11 0.52276
#> phi_1,12 0.33784
```

```
#> phi_1,21 0.19297
#> phi_1,22 0.67143
#> sig_11   8.03740
#> sig_12   3.48390
#> sig_22   2.44520
#>
#> log-likelihood =  -1919.314
#> AIC =  3856.629
#> BIC =  3894.542
#>
#> Unrestricted Model
#>                coef
#> mu_1,1     4.486100
#> mu_2,1    -2.367900
#> mu_1,2    10.713000
#> mu_2,2     2.753400
#> phi_1,11   0.460710
#> phi_1,12   0.225140
#> phi_1,21   0.170420
#> phi_1,22   0.657350
#> sig_11,1   4.907800
#> sig_12,1   1.417900
#> sig_22,1   1.075300
#> sig_11,2   8.246400
#> sig_12,2   3.714500
#> sig_22,2   2.465400
#> p_11       0.959470
#> p_12       0.040533
#> p_21       0.039947
#> p_22       0.960050
#>
#> log-likelihood =  -1878.007
```

```
#> AIC =  3792.014
#> BIC =  3867.841
#>
#> Rodriguez-Rondon & Dufour (2025) Local Monte Carlo Likelihood Ratio Test
#>          LRT_0  0.90% 0.95%  0.99% p-value
#> LMC_LRT 82.615 11.682 11.75 15.763    0.05
```

The test output reports the likelihood ratio statistic, Monte Carlo critical values, and the corresponding Monte Carlo p-value. In this example, the null hypothesis of a single regime is rejected. It is important to note that the Monte Carlo p-value is computed using a finite number of simulations, set here to `N = 19`. As a result, the Monte Carlo p-value is discrete and can only take values on the grid $\{1/(N+1), 2/(N+1), \ldots, N/(N+1)\}$. Consequently, with $N = 19$, the smallest attainable p-value is $1/(19+1) = 0.05$. The reported p-value of 0.05 therefore corresponds to the strongest possible evidence against the null hypothesis given this simulation size. As discussed in (Dufour, 2006; Rodriguez-Rondon and Dufour, 2025), larger values of N may be desirable to improve test power, but very large values are not required, since these methods do not attempt to approximate asymptotic critical values ($N = 99$ is typically recommended). Here, we use $N = 19$ purely for illustrative purposes.

As discussed previously, the `LMCLRTest()` function can also be used to replicate the parametric bootstrap procedures studied in Qu and Zhuo (2021) and Kasahara and Shimotsu (2018). To do so, the user can set `mdl_h1_control = list(method = "MLE")`, which allows for direct maximum likelihood estimation and facilitates the imposition of parameter constraints. Such constraints can be specified through the arguments `mle_theta_low` and `mle_theta_upp` in `mdl_h1_control`. In addition to restricting the transition probabilities, Kasahara and Shimotsu (2018) also impose constraints on the variance parameters, which can be implemented within the same framework.

When using the LMC-LRT as a parametric bootstrap, it is typically desirable to increase the number of Monte Carlo simulations to better approximate asymptotic critical values. For instance, Qu and Zhuo (2021) use $N = 199$, while Kasahara and Shimotsu (2018) use $N = 299$. In practice, lower values of N may still be reasonable given the computational cost of repeatedly estimating Markov switching models.

The maximized Monte Carlo likelihood ratio test (MMC-LRT) is implemented via the `MMCLRTest()` function. This procedure shares the same estimation controls as the LMC-LRT for fitting the restricted and unrestricted models and similarly requires the user to specify `p`, `k0`, and k1. In addition, the MMC-LRT includes several options specific to the maximization step over the nuisance parameter space.

In particular, users may specify the constant eps used to define the consistent set over which the maximization is performed. Setting CI_union = TRUE constructs the set $C^*_T = C^{CI}_T \cup C^{\epsilon}_T$, while setting eps = 0 and CI_union = TRUE restricts the search to C^{CI}_T. The numerical optimization algorithm used in this search can be selected via the type argument. By default, the algorithm stops when a Monte Carlo p-value of 1 is reached, corresponding to the maximum possible value. Alternatively, the search can be terminated once the test fails to reject, that is, when the p-value exceeds the chosen significance level α, by setting threshold_stop.

Below, we illustrate the MMC-LRT by testing a linear univariate model with $M = 1$ regime, simulated using a simple normal distribution in the previous subsection. We set eps = 0.3 and CI_union = FALSE and use the simulated annealing algorithm via type = "GenSA". To avoid searching for long, we set threshold_stop = 0.05 + 1e-6, so that the algorithm stops once the null fails to be rejected. Though not shown here, parallel computation can be enabled by setting workers = 8, which distributes the null distribution simulation across multiple cores. To use this option, the user must first register a parallel backend and close the cluster after completion.

```
mmc_control = list(N = 19,
                   eps = 0.3,
                   threshold_stop = 0.05 + 1e-6,
                   type = "GenSA",
                   CI_union = FALSE,
                   mdl_h0_control = list(const  = TRUE,
                                         getSE  = FALSE),
                   mdl_h1_control = list(msmu   = TRUE,
                                         msvar  = TRUE,
                                         getSE  = FALSE,
                                         method = "EM"),
                   maxit  = 100)
mmclrt <- MMCLRTest(simu_norm[["y"]], p = 0, k0 = 1 , k1 = 2, control = mmc_control)
```

```
#> Have got accurate energy -0.25 <= -0.050001 in smooth search
#> Emini is: -0.25
```

```
#> xmini are:
#> 4.998107919 -2.046390673 4.370604258 1.204649505 0.8673397858
#> Totally it used 0.000118 secs
#> No. of function call is: 1
```

```
summary(mmclrt)
```

```
#>
#> Restricted Model
#>            coef
#> mu_1    4.99810
#> mu_2   -2.04640
#> sig_11  4.37060
#> sig_12  1.20460
#> sig_22  0.86734
#>
#> log-likelihood =  -1631.436
#> AIC =  3272.871
#> BIC =  3293.944
#>
#> Unrestricted Model
#>                 coef
#> mu_1,1    4.5675e+00
#> mu_2,1   -2.1688e+00
#> mu_1,2    5.8802e+00
#> mu_2,2   -1.7956e+00
#> sig_11,1  4.5928e+00
#> sig_12,1  1.0072e+00
#> sig_22,1  7.2573e-01
#> sig_11,2  2.7574e+00
#> sig_12,2  1.2798e+00
#> sig_22,2  1.0638e+00
#> p_11      5.1200e-01
```

```
#> p_12      4.8800e-01
#> p_21      1.0000e+00
#> p_22      2.1791e-08
#>
#> log-likelihood =  -1625.303
#> AIC =  3278.605
#> BIC =  3337.61
#>
#> Rodriguez-Rondon & Dufour (2025) Maximized Monte Carlo Likelihood Ratio Test
#>          LRT_0 p-value
#> MMC_LRT 12.266    0.25
```

In this example, the MMC-LRT fails to reject the null hypothesis, as indicated by the reported Monte Carlo p-value. As with the LMC-LRT, the Monte Carlo nature of the procedure implies that the p-value resolution depends on the number of simulations N. Nevertheless, the MMC-LRT remains the most general likelihood ratio–based testing procedure currently available for Markov switching models, as it allows for maximization over nuisance parameters while retaining finite-sample validity.

4.4.2 Moment-based tests

The Monte Carlo moment-based test proposed by Dufour and Luger (2017) is implemented in **MSTest** via the functions `DLMCTest()` (local version) and `DLMMCTest()` (maximized version). As with the other testing procedures, the user must provide the series y_t and the number of autoregressive lags p. The parameter N determines the number of Monte Carlo replications used to compute the test. The default value is $N = 99$, as Dufour et al. (2004) show that using more than 100 replications (including the observed statistic) has little effect on test power.

In addition, this procedure requires approximating the distribution of the p-value associated with each moment-based statistic through an auxiliary simulation. The number of replications used for this approximation is controlled by `N2`, which defaults to 10,000.

Below, we illustrate the moment-based local Monte Carlo test using the previously simulated Markov switching autoregressive process. The test is computationally inexpensive and completes almost instantaneously. The output of the `summary()` method reports the nuisance parameter estimate (here, ϕ_1), the test statistics associated with each of the four moments (mean, variance, skewness, and excess kurtosis), the combined test statistic based on $F(\varepsilon)$, the corresponding critical values (quantiles of simulated distribution), and the Monte Carlo p-values. As expected, the null hypothesis of linearity is clearly rejected, since the data were generated from a Markov switching process with $M = 2$ regimes.

```
lmc_control = list(N = 99,
                   simdist_N = 10000,
                   getSE = TRUE)
lmcmoment <- DLMCTest(simu_msar[["y"]], p = 1, control = lmc_control)

summary(lmcmoment)

#>
#> Restricted Model
#>          coef     s.e.
#> mu    5.93390 0.374020
#> phi_1 0.82887 0.025029
#> sig   2.04380 0.129390
#>
#> log-likelihood =  -886.3994
#> AIC =  1778.799
#> BIC =  1791.437
#>
#> Dufour & Luger (2017) Moment-Based Local Monte Carlo Test
#>            phi_1 M(eps) V(eps)  S(eps) K(eps) F(eps)   0.90%   0.95%   0.99%
#> LMC_min  0.82887 1.3948  14.86 0.55248 4.2263      1 0.96417 0.97974 0.98997
#> LMC_prod 0.82887 1.3948  14.86 0.55248 4.2263      1 0.99870 0.99918 0.99978
#>          p-value
#> LMC_min     0.01
#> LMC_prod    0.01
```

The computational efficiency of this procedure extends naturally to the maximized version of the moment-based test, implemented via `DLMMCTest()`. Despite the additional step of maximizing the Monte Carlo p-value over the nuisance parameter space, the test remains very fast to compute in practice.

In this case, additional options are available to control the numerical optimization. The argument `optim_type` determines the optimization algorithm used in the search. As with the MMC-LRT procedure, the arguments `eps` and `CI_union` allow the user to define the consistent set over which the maximization is performed. Many of the optimization-related options available for `MMCLRTest()` are also available here.

```
mmc_control <- list(N = 99,
                    getSE = TRUE,
                    eps = 0,
                    CI_union = TRUE,
                    optim_type = "GenSA",
                    threshold_stop = 0.05 + 1e-6,
                    maxit = 100)
mmcmoment <- DLMMCTest(simu_msar[["y"]], p = 1, control = mmc_control)
```

```
#> Stop. Nb function call=100 max function call=100.
#> Emini is: -0.01
#> xmini are:
#> 0.8288716498
#> Totally it used 0.00119 secs
#> No. of function call is: 100
#> Stop. Nb function call=100 max function call=100.
#> Emini is: -0.01
#> xmini are:
#> 0.8288716498
#> Totally it used 0.001093 secs
#> No. of function call is: 100
```

```
summary(mmcmoment)
```

```
#>
#> Restricted Model
#>          coef     s.e.
#> mu    5.93390 0.374020
```

```
#> phi_1 0.82887 0.025029
#> sig   2.04380 0.129390
#>
#> log-likelihood =  -886.3994
#> AIC =  1778.799
#> BIC =  1791.437
#>
#> Dufour & Luger (2017) Moment-Based Maximized Monte Carlo Test
#>          M(eps) V(eps)  S(eps) K(eps) F(eps) p-value
#> MMC_min  1.3948  14.86 0.55248 4.2263      1    0.01
#> MMC_prod 1.3948  14.86 0.55248 4.2263      1    0.01
```

In the example above, we again apply the test to the simulated Markov switching autoregressive process. We set the stopping threshold to 0.05 + 1e-6, so that the optimization terminates once the null hypothesis fails to be rejected, although this does not occur in this case. We also set `eps = 0`, restricting the search to the confidence interval–based consistent set, as in Dufour and Luger (2017). Once again, the null hypothesis of a linear model is rejected, in line with the results obtained from the other testing procedures.

4.4.3 Parameter stability test

The parameter stability test proposed by Carrasco et al. (2014) is implemented in **MSTest** via the function `CHPTest()`. As with the other testing procedures, the user must provide the series y_t and the number of autoregressive lags p. In this case, the parameter N specifies the number of bootstrap replications and is set to 3000 by default, following Carrasco et al. (2014). The parameter ρ_b controls the bound of one of the nuisance parameters and is set to 0.7 by default, again as in their study. Setting $\rho_b = 0.7$ implies that the grid search is performed over the parameter space $\rho \in [-0.7, 0.7]$.

```
chp_control = list(N = 1000,
                   rho_b = 0.7,
                   msvar = FALSE)
pstabilitytest <- CHPTest(simu_ar[["y"]], p = 1, control = chp_control)
summary(pstabilitytest)
```

```
#>
#> Restricted Model
#>          coef     s.e.
#> mu    5.11640 0.156620
#> phi_1 0.70483 0.031123
#> sig   1.06580 0.067474
#>
#> log-likelihood =  -723.9473
#> AIC =  1453.895
#> BIC =  1466.532
#>
#> Carrasco, Hu, & Ploberger (2014) Parameter Stability Test
#>
#> - Switch in Mean only
#>       test-stat  0.90%  0.95%  0.99% p-value
#> supTS   0.55996 2.0123 2.8253 4.3814   0.537
#> expTS   0.83337 1.3740 1.8493 5.0661   0.451
```

Above, we provide an example of the parameter stability test applied to the linear autoregressive process simulated earlier. In this example, we set `N = 1000` and `msvar = FALSE`, so that the test focuses on potential instability in the mean only. The `summary()` method reports the parameter estimates of the restricted model, which is the only model that needs to be estimated for this test, along with the results for both the supTS and expTS versions of the test statistic.

As expected, the test fails to reject the null hypothesis of parameter stability, which is consistent with the fact that the data were generated from a linear autoregressive model without regime switching.

4.4.4 Stochastic likelihood ratio test

The stochastic likelihood ratio test proposed by Hansen (1992) is implemented in **MSTest** via the function `HLRTest()`. As before, the user must provide the series y_t and specify the number of autoregressive lags p, as the test is designed to assess the null hypothesis of linearity in autoregressive models.

This procedure performs a grid search over nuisance parameters, and most of the available options are therefore related to the construction of this grid. In particular, users can specify the grid size via `gridsize`, as well as the starting values and step sizes for the mean and variance grids through `mugrid_from`, `mugrid_by`, and their variance counterparts. These grid values pertain to the unrestricted model. Conditional on these fixed values, the test optimizes the parameters of the restricted model; accordingly, the arguments `theta_null_low` and `theta_null_upp` can be used to define the optimization region for the null hypothesis.

While it is possible to allow for regime-dependent variances by setting `msvar = TRUE`, doing so increases the dimensionality of the nuisance parameter space and can substantially raise the computational cost of the procedure.

```
hlrt_control  <- list(msvar          = FALSE,
                      gridsize       = 20,
                      mugrid_from    = 0,
                      mugrid_by      = 1,
                      theta_null_low = c(0,-0.99,0.01),
                      theta_null_upp = c(20,0.99,20))
hlrt <- HLRTest(simu_msar[["y"]], p = 1, control = hlrt_control)
summary(hlrt)
```

```
#>
#> Restricted Model
#>          coef     s.e.
#> mu    5.93390 0.374020
#> phi_1 0.82887 0.025029
#> sig   2.04380 0.129390
#>
#> log-likelihood =  -886.3994
#> AIC =  1778.799
#> BIC =  1791.437
#>
#> Hansen (1992) Likelihood Ratio Bound Test -  Switch in Mean only
#>       test-stat 0.90 % 0.95 % 0.99 % p-value
```

```
#> M = 0    8.8412 2.0605 2.3458 2.8217       0
#> M = 1    8.8412 2.1715 2.4783 3.1255       0
#> M = 2    8.8412 2.1784 2.5079 3.1990       0
#> M = 3    8.8412 2.1743 2.4807 3.0618       0
#> M = 4    8.8412 2.1578 2.4666 3.0455       0
```

Above, we illustrate the stochastic likelihood ratio test using the previously simulated Markov switching autoregressive process. In this example, the `summary()` method reports parameter estimates for the restricted model, which is the only model that needs to be estimated in this framework. The output also presents test results for the different standardizations discussed in Hansen (1996a).

Across all standardizations, the null hypothesis of linearity is strongly rejected, with Monte Carlo p-values effectively equal to zero. This outcome is consistent with the fact that the data were generated from a Markov switching process with two regimes.

Taken together, the likelihood ratio–based, moment-based, and parameter stability tests implemented in **MSTest** provide complementary approaches to testing for regime switching. Their differing assumptions, computational requirements, and robustness properties allow users to select procedures that are well suited to their specific empirical setting.

5 Conclusion

The importance of testing the number of regimes in Markov switching models has motivated a substantial body of research, reflecting the statistical and computational challenges inherent in this problem. Seminal contributions such as Hansen (1992), Carrasco et al. (2014), and Dufour and Luger (2017) focus on testing the null hypothesis of a single regime (i.e., a linear model) against the alternative of two regimes. More recently, Rodriguez-Rondon and Dufour (2025) introduced a class of Monte Carlo likelihood ratio tests that allow one to test a null hypothesis with M_0 regimes against an alternative with $M_0 + m$ regimes, for any $M_0 \geq 1$ and $m \geq 1$, for a a broad class of models, including multivariate and non-stationary ones, thereby substantially expanding the scope of feasible hypothesis testing in this setting.

The **MSTest** package makes these testing procedures readily available to applied researchers by implementing the methods proposed in these four studies within a unified and user-friendly framework. In addition to hypothesis testing, the package provides tools for simulating and estimating a wide range of Markov switching and hidden Markov models, primarily to support the testing procedures themselves. This paper has reviewed the underlying methodology and illustrated how **MSTest** can be used in practice to conduct inference on the number of regimes.

Overall, **MSTest** aims to facilitate empirical work involving Markov switching models by lowering the computational and implementation barriers associated with regime testing. By offering multiple complementary testing procedures with differing assumptions and robustness properties, the package allows users to select methods that are well suited to their specific empirical setting and to make informed decisions about model specification.

References

Aguirregabiria, V. and Mira, P. (2007). Sequential estimation of dynamic discrete games. *Econometrica*, 75(1):1–53.

Ahn, H. J. and Luciani, M. (2024). Common and idiosyncratic inflation. *Finance and Economics Discussion Series 2020-024r1. Washington: Board of Governors of the Federal Reserve System.*

An, Y., Hu, Y., Hopkins, J., and Shum, M. (2013). Identifiability and inference of hidden markov models. Technical report, Technical report.

Andrews, D. W. (1999). Estimation when a parameter is on a boundary. *Econometrica*, 67(6):1341–1383.

Andrews, D. W. (2001). Testing when a parameter is on the boundary of the maintained hypothesis. *Econometrica*, 69(3):683–734.

Andrews, D. W. and Ploberger, W. (1994). Optimal tests when a nuisance parameter is present only under the alternative. *Econometrica: Journal of the Econometric Society*, pages 1383–1414.

Anser, M. K., Godil, D. I., Khan, M. A., Nassani, A. A., Zaman, K., and 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.

Baldi, P., Chauvin, Y., Hunkapiller, T., and 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. and Caelli, T. M. (2001). *Hidden Markov models: applications in computer vision*, volume 45. World Scientific.

Caggiano, G., Castelnuovo, E., and Figueres, J. M. (2017). Economic policy uncertainty and unemployment in the united states: A nonlinear approach. *Journal of Applied Econometrics*, 32(2):281–298.

Cai, J. (1994). A markov model of unconditional variance in arch. *Journal of Business and Economic Statistics*, 12(3):309–316.

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

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

Cevik, E. I., Yıldırım, D. Ç., and 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*, pages 969–996.

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

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

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

Cho, J.-S. and White, H. (2011). Testing for regime switching: Rejoinder. *Unpublished Manuscrpt, University of California, San Diego.*

Davies, R. B. (1977). Hypothesis testing when a nuisance parameter is present only under the alternative. *Biometrika*, 64(2):247–254.

Davies, R. B. (1987). Hypothesis testing when a nuisance parameter is present only under the alternative. *Biometrika*, 74(1):33–43.

Dempster, A. P., Laird, N. M., and 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. and Rudebusch, G. D. (1996). Measuring business cycles: A modern perspective. *The Review of Economics and Statistics*, 78.

Dietz, S. and Stern, N. (2015). Endogenous growth, convexity of damages and climate risk: How nordhaus' framework supports deep cuts in carbon emissions. *The Economic Journal*, 125(583):574–620.

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., Bernard, J.-T., and Genest, I. (2004). Simulation-based finite-sample tests for heteroskedasticity and arch effects. *Journal of Econometrics*, 122(2):317–347.

Dufour, J.-M., Khalaf, L., and Voia, M. (2014). Finite-sample resampling-based combined hypothesis tests, with applications to serial correlation and predictability. *Communications in Statistics-Simulation and Computation*, 44(9):2329–2347.

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

Dufour, J.-M. and Neves, J. (2019). Finite-sample inference and nonstandard asymptotics with monte carlo tests and r. In *Handbook of statistics*, volume 41, pages 3–31. Elsevier.

Dufrénot, G., Mignon, V., and Péguin-Feissolle, A. (2011). The effects of the subprime crisis on the latin american financial markets: An empirical assessment. *Economic Modelling*, 28(5):2342–2357.

Eddelbuettel, D. and Balamuta, J. J. (2018). Extending R with C++: A Brief Introduction to Rcpp. *The American Statistician*, 72(1):28–36.

Eddelbuettel, D., Francois, R., Bates, D., Ni, B., and Sanderson, C. (2024). *RcppArmadillo: 'Rcpp' Integration for the 'Armadillo' Templated Linear Algebra Library.* R package version 14.0.2-1.

Fisher, R. A. (1932). *Statistical methods for research workers.* Genesis Publishing Pvt Ltd.

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

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

Goldfeld, S. M. and Quandt, R. E. (1973). A markov model for switching regressions. *Journal of econometrics*, 1(1):3–15.

Golosov, M., Hassler, J., Krusell, P., and Tsyvinski, A. (2014). Optimal taxes on fossil fuel in general equilibrium. *Econometrica*, 82(1):41–88.

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., and 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: Journal of the Econometric Society*, pages 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. (1994). *Time series analysis.* Princeton university press.

Hamilton, J. D. (2005). What's real about the business cycle? Technical report, National Bureau of Economic Research.

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

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.

Hansen, B. E. (1996a). Erratum: The likelihood ratio test under nonstandard conditions: testing the markov switching model of gnp. *Journal of Applied econometrics*, 11(2):195–198.

Hansen, B. E. (1996b). Inference when a nuisance parameter is not identified under the null hypothesis. *Econometrica: Journal of the econometric society*, pages 413–430.

Hernández, P. and Ochoa, C. (2016). Switching models for the analysis of health care costs: An application to the spanish health care system. *Health Economics*, 25(6):830–843.

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

Hu, L. and Shin, Y. (2008). Optimal test for markov switching garch models. *Studies in Nonlinear Dynamics & Econometrics*, 12(3).

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

Kahn, J. A. and Rich, R. W. (2007). Tracking the new economy: Using growth theory to detect changes in trend productivity. *Journal of Monetary Economics*, 54(6):1670–1701.

Kasahara, H., Okimoto, T., and Shimotsu, K. (2014). Modified quasi-likelihood ratio test for regime switching. *The Japanese Economic Review*, 65(1):25–41.

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

Kim, C.-J. and 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*, pages 223–254. Springer.

Krogh, A., Mian, I. S., and 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). The markov-switching vector autoregressive model. In *Markov-Switching Vector Autoregressions*, pages 6–28. Springer.

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

Le Bihan, H., Leiva-León, D., and Pacce, M. (2024). Underlying inflation and asymmetric risks. *Review of Economics and Statistics*, pages 1–45.

Liu, X. and Shao, Y. (2003). Asymptotics for likelihood ratio tests under loss of identifiability. *The Annals of Statistics*, 31(3):807–832.

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

Marmer, V. (2008). Testing the null hypothesis of no regime switching with an application to gdp growth rates. *Empirical Economics*, 35(1):101–122.

Morley, J. and Piger, J. (2012). The asymmetric business cycle. *Review of Economics and Statistics*, 94(1):208–221.

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

Pearson, K. (1933). On a method of determining whether a sample of size n supposed to have been drawn from a parent population having a known probability integral has probably been drawn at random. *Biometrika*, 25(3/4):379–410.

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

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

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

Rabiner, L. and Juang, B. (1986). An introduction to hidden markov models. *ieee assp magazine*, 3(1):4–16.

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

Rodriguez-Rondon, G. (2024). Underlying core inflation with multiple regimes. *Discussion paper, McGill University (Economics Department).*

Rodriguez-Rondon, G. and Dufour, J.-M. (2025). Monte Carlo likelihood ratio tests for Markov switching models. *Discussion paper, McGill University (Economics Department).*

Scrucca, L. et al. (2013). Ga: a package for genetic algorithms in r. *Journal of Statistical Software*, 53(4):1–37.

Sims, C. A. and Zha, T. (2006). Were there regime switches in us monetary policy? *American Economic Review*, 96(1):54–81.

So, M. E. P., Lam, K., and Li, W. K. (1998). A stochastic volatility model with markov switching. *Journal of Business & Economic Statistics*, 16(2):244–253.

Sweeting, A. (2013). Dynamic product positioning in differentiated product markets: The effect of fees for musical performance rights on the commercial radio industry. *Econometrica*, 81(5):1763–1803.

Tippett, L. H. C. et al. (1931). The methods of statistics. *The Methods of Statistics.*

Warne, A. and Vredin, A. (2006). Unemployment and inflation regimes. *Studies in Nonlinear Dynamics & Econometrics*, 10(2).

White, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica: Journal of the Econometric Society*, pages 1–25.

Wilkinson, B. (1951). A statistical consideration in psychological research. *Psychological bulletin*, 48(2):156.

Yang Xiang, Gubian, S., Suomela, B., and Hoeng, J. (2013). Generalized simulated annealing for efficient global optimization: the GenSA package for R. *The R Journal Volume 5/1, June 2013.*

Zambrano-Bigiarini, M., Clerc, M., and Rojas, R. (2013). Standard particle swarm optimisation 2011 at cec-2013: A baseline for future pso improvements. In *2013 IEEE Congress on Evolutionary Computation*, pages 2337–2344. IEEE.