WASHINGTON UNIVERSITY IN ST. LOUIS
Department of Economics

Dissertation Examination Committee:
Werner Ploberger, Chair
Gaetano Antinolfi
Ismael Mourifie
Guofu Zhou
Martin Xiumin
Fei Tan

Essays on Optimal Test for Markov Switching Parameters

by
Li Zhang

A dissertation presented to
The Graduate School
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

Fall 2023
St. Louis, Missouri

Table of Contents

Acknowledgments

First and foremost, I would like to express my sincere appreciation to my advisor, Professor Werner Ploberger, for his persistent support, patient instructions and insightful directions, which have significantly helped me to improve my research skills and prepared me for future challenges. Without his motivation and guidance, the dissertation would not have been possible. I would like to thank Professor Guofu Zhou for his encouragement, valuable comments and constant suppor. I am grateful to Professor Antinolfi Gaetano, Professor Martin Xiumin and Professor Fei Tan for serving on my dissertation committee and providing helpful comments. I would also like to thank Professor Wei Zhou for leading me to the road of economic research.

In addition I would like to thank my friends and fellow Ph.D. students at Washington University in St. Louis for their discussions, friendship and making this journal more fun. Many thanks to seminar participants at Midwest Theory Conference, North American Meeting of the Econometric Society, I acknowledge the financial support from Department of Economics and travel support through Center for Research in Economics and Strategy (CRES), in the Olin Business School, Washington University in St. Louis.

Last but not least, I would like to express my heartfelt gratitude to my parents for their continuing love, encouragement and support throughout my life.

Li Zhang

Washington University
Fall 2023

Dedicated to my parents.

ABSTRACT OF THE DISSERTATION

Essays on Optimal Test for Markov Switching Parameters

by

Li Zhang

Doctor of Philosophy in Economics

Washington University in St. Louis, 2023

Professor Werner Ploberger, Chair

We propose a likelihood ratio test for fixed unit root against time switching unit root models. Our test is different from the existing jump detection literature centered on the application of various forms of Augmented Dickey-Fuller tests. Our methodology involves the inclusion of random coefficients, which effectively capture both expansion and contraction behaviors. We show that the contiguous alternatives converge to the null hypothesis at the order of $T^{-3/4}$, where T is the sample size. Our test is asymptotically optimal in the sense that it maximizes a weighted power function. We derive the asymptotic distribution of our test under the null and local alternatives.

Chapter 1

Optimal test for stochastic unit root with i.i.d shocks

This chapter studies an optimal test in the i.i.d case. We start from a likelihood ratio in which the local alternatives are of the order $T^{-3/4}$. This is due to the fact that the regime change η_t is unknown, and we first assume it to be an i.i.d with normal distribution.

1.1 Introduction

In this paper I propose an likelihood ratio test for identifying bubble behavior in a stochastic unit root model. A bubble in asset prices can be characterized by an explosive surge in prices series that significantly surpasses the fundamental values. The finance profession has debated the possibility of bubbles from two main perspectives. Efficient market theory argues that when price levels deviate from the true economics values, no arbitrage condition will eventually force the prices back to the rational levels. On the other hand, behavior theory suggest that investors make their decisions with irrational impulses leading to deviations in prices. For instance, investors may base their decisions on past prices, assuming a continuous upward trend. Furthermore, the tendency to buy more intensifies when more people join the market. By empirically testing the evidence, we can gain a deeper understanding of whether there is evidence of bubble behavior and what characteristics it exhibits.

Using Bitcoin prices as an example, the cryptocurrency surged from 1,000 to its peak at $20,000 in Dec 2017. During pandemic period, Bitcoin's price burst into action once again. Bitcoin took less than a month in 2021 to smash its 2020 price record, surpassing $40,000 by Jan, 2021. By April, Bitcoin prices reached new all-time highs of over $60,000 as Coinbase, a cryptocurrency exchange, went public. What's the reasons for price to rise and fall. Is it a rational bubble or behavior ones. There still are a lot different arguments for that debate. Are those sudden jumps are results of a random walk or they do comes from

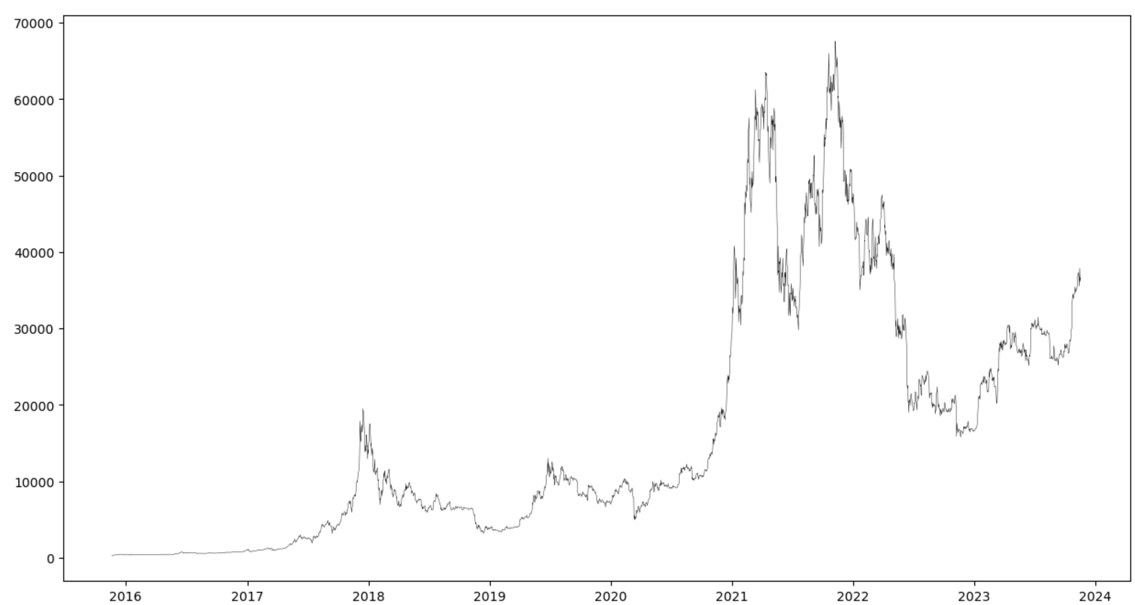

Figure 1.1: Time Series of Bitcoin from Nov 2015 to Nov 2023

explosive behavior. I consider a class of autoregressive models where the AR coefficient is around 1. Following a common model used in the literature, we assume the data $y_1, y_2, ..., y_T$, are generated as

$$y_t = \theta_t y_{t-1} + \epsilon_t$$

where θ_t can be written as $1 + \eta_t$ with η_t being the unobserved switching variable. η_t is time varing. We assume that it would follow a Markov chain. If η_t is larger than 0 in some regimes we allow for explosive behavior. If it is less than 0, that is the contraction. We want to test the null hypothesis $\eta_t = 1$ against the alternative $\theta_t = 1 + \eta_t$. The parameters driving the dynamic of the underlying Markov chain are not identified under the null hypothesis. As a result, the testing problem is nonstandard.

We believe that such a model is a better representation of the behavior of macroeconomic aggregates compared to the conventional fixed unit root model. For instance, models with fixed coefficients ignore the impacts that a succession of small and varied macroeconomic shocks may have on the structure of appropriate dynamic economic models, especially when applied to highly aggregated data series.Policy changes stand as a significant source of these shocks. For example, Lucas Jr (1976) suggested that using a simultaneous equations model for policy analysis and simulation is only valid when agent's expectation rules are embedded in its functional form since they are supposed to change with the changes in policy. And he used rational expectation equilibrium (REE) framework in which this property is satisfied by the expectation rules.

Likewise, the emergence of these shocks raises doubts about the time-invariant structures of unit roots. This issue has been explored in prior studies by Leybourne et al. (1996) and Granger and Swanson (1997), falling within the realm of stochastic unit root analysis. Itâs argued that stochastic unit root models provide a better description of the behavior of macroeconomic variables than fixed-unit-root models since they allow more general forms of nonstationarity.

As noted by Granger and Swanson (1997), augmented Dickey Fuller tests lack power to detect a variety of processes in the category of stochastic unit root. So specification tests are proposed. In Leybourne et al. (1996), they make the assumption that the AR coefficients are independent and identically distributed (i.i.d. thereafter) around mean 1 with some variance σ^2. They proposed to use a LM test for the null hypothesis of $\sigma^2 = 0$. This method was applied to several U.S. macroeconomic series widely believed to contain fixed unit roots. They found that for about half of the series the fixed-unit-root null is rejected.

Some people suggest to use a threshold model to capture the deviation of the series from unit root. For example, Caner and Hansen (2001) introduced a two-regime threshold autoregressive (TAR) model incorporating an autoregressive unit root. Inspired by the classic work of Hamilton (1989), our model is different to the existing literature on stochastic unit roots or threshold unit roots in that we assume the AR coefficients are driven by some unobserved Markov chain , allowing for a comprehensive exploration of correlations. Our model is different to the TAR model since we donât assume the regimes are observable. And there is no direct economic theory to justify the observable regimes. Moreover, our model deviates from prevailing literature on stochastic unit roots where assumptions either completely specify the dynamics of random coefficients or merely assume i.i.d. coefficients.

Our likelihood ratio test originate from Carrasco et al. (2014), (CHP thereafter), where fourth order Taylor expansion of the likelihood ratio are obtained. We modify the assumption that latent variable is stationary to be nonstationary.

The remainder of the paper is organized as follows. The model is described in Section 2. We also characterize likelihood ratio and the formula for our test statistics in this section. Limiting distribution is studied in Section 3. Asymmetric case in η_t is analyzed in section 4. We conclude in Section 5. All the proofs are relegated in the Appendix.

1.2 The Model

We build a basic time series model:

$$y_t = \theta_t y_{t-1} + \epsilon_t$$

where $\epsilon_t \sim N(0, \sigma^2)$ Here, we test the null with a fixed unit root against a stochastic root, namely

$$H_0 : \theta_t = \theta_0 = 1$$

against

$$H_1 : \theta_t = 1 + \eta_t$$

where η_t is the unobservable switching variable, and η_t is a homogeneous Markov chain; η_t could be greater than one, such that is shows explosive behavior on certain regimes. If it is less than one, there is contraction.

The conditional density of y_t is

$$f_t(\theta) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y_t - \theta_t y_{t-1})^2}{2\sigma^2}}.$$

Let Q_T^β denote the joint distribution of $(\eta_1, ..., \eta_T)$. β would be σ^2 in our symmetric case.

The ratio would be

$$\begin{aligned} l_t &= \int \prod_{t=1}^{T} f_t(\theta_t) dQ_T^\eta / \prod_{t=1}^{T} f_t(\theta_0) \\ &= E_\beta[\prod_{t=1}^{T} e^{\frac{-\eta_t^2 y_{t-1}^2}{2\sigma^2}} e^{\frac{(y_t - y_{t-1})\eta_t y_{t-1}}{\sigma^2}}]. \end{aligned}$$

If the null hypothesis is true, the ratio would be

$$E_\beta[\prod_{t=1}^{T} e^{-\frac{\eta_t^2 y_{t-1}^2}{2\sigma^2}} e^{\frac{\epsilon_t \eta_t y_{t-1}}{\sigma^2}}].$$

If we assume η_t has the following binomial distribution

$$\eta_t = \begin{cases} r, & p = 0.5 \\ -r, & q = 0.5 \end{cases}$$

Our Log likelihood would be

$$\ln l_t = -\frac{r^2}{2\sigma^2}\sum_{t=1}^{T} y_{t-1}^2 + T\ln(1/2) + \sum_{t=1}^{T}\ln(e^{\frac{r}{\sigma^2}\epsilon_t y_{t-1}} + e^{-\frac{r}{\sigma^2}\epsilon_t y_{t-1}}).$$

Considering the local alternative, η_t is around zero. Here the jump size is r, and we can detect how large it is. Put it another way, what is the size of r such that our likelihood ratio has a limit distribution?

Because r is small, we can apply a Taylor expansion to the third term in the equation above. Expand to the fourth moment(according to CHP,only fourth moments matter) as

$$\ln(e^x + e^{-x}) = \ln(2) + \frac{1}{2}x^2 - \frac{1}{12}x^4 + o(x^4)$$

where the likelihood ratio has the following form:

$$\ln l_t = \sum_{t=1}^{T}\left[\frac{r^2}{2\sigma^4} y_{t-1}^2(\epsilon_t^2 - \sigma^2) - \frac{r^4}{12\sigma^8}\epsilon_t^4 y_{t-1}^4\right]. \quad (1.1)$$

We would like to analyze the asymptotic behavior of the likelihood ratio l_t.

1.3 Limiting Distribution

In this section, we analyze the limiting distribution shown in (1.1). The main result is presented in the beginning of the section and then proven in steps. The main theorem is as follows:

Theorem 1. *Our log likelihood ratio has the following limiting distribution:*

$$\ln l_t \xrightarrow{d} c_1 \int_0^1 W_1^2(t) dW_2(t) + c_2 \int_0^1 W_1^4(t) dt$$

where $W_1(t)$ and $W_2(t)$ are two standard Brownian Motions generated from ϵ_t and $\epsilon_t^2 - \sigma^2$.

The proof of Theorem 1 is the main contribution of this paper, and the corresponding limiting distributions would be

$$\frac{\sum_{t=1}^{T} y_{t-1}^2 (\epsilon_t^2 - \sigma^2)}{\sqrt{2}\sigma^4 T^{3/2}} \xrightarrow{d} \int W_1(t)^2 dW_2(t)$$

$$\sum_{t=1}^{T} \frac{y_{t-1}^4}{T^3} \xrightarrow{d} \sigma^4 \int W_1^4(t) dt.$$

Under the null, we have

$$y_t = y_{t-1} + \epsilon_t = \sum_{s=1}^{t} \epsilon_s$$

where y_t is an integrated or partial sum process. Because ϵ_t is IID, it derives from the central limit theorem, such that

$$\frac{1}{\sqrt{T}} y_T \xrightarrow{d} N(0, \sigma^2). \tag{1.2}$$

In (1.1), we need to consider the limiting distribution for the terms $\sum y_t^2, \sum y_{t-1}\epsilon_t$. Both are non-normal, and here, functional CLT serves our purpose. Any fixed fraction of the sample has a corresponding limiting behavior. Therefore, let r denote any point of interval $[0, 1]$; then $[Tr]$ denotes the largest integer not exceeding Tr. Therefore, (1.2) can be generalized to

$$\frac{1}{\sqrt{T}} y_{[Tr]} \xrightarrow{d} N(0, r\sigma^2).$$

Observe that

$$Y_T(r) = \frac{1}{\sigma\sqrt{T}} y_{[Tr]},$$

which defines a discontinuous function.

Theorem (Functional Central Limit Theorem)**.** *Suppose that $\epsilon_t \sim iid(0, \sigma^2)$, and the stochastic process Y_t is defined by*

$$Y_T(r) = \frac{1}{\sigma\sqrt{T}} \sum_{t=1}^{[Tr]} \epsilon_t. \tag{1.3}$$

Then $Y_T \xrightarrow{d} W$, where W is a standard Wierner process.

For a multivariate vector case, the limit processes are the vector Wierner process $\mathbf{W}$

where

$$Y_T(t) = \frac{1}{\sigma\sqrt{T}} y_{[Tt]} + \big(Tt - [Tt]\big)\frac{1}{\sigma\sqrt{T}} \epsilon_{[Tt]+1}.$$

Because $\sigma^2 < \infty$, $X_T(t)$ weakly converges to a standard Brownian motion via the functional central limit theorem. The corresponding multivariate functional central limit theorem would be

$$\mathbf{Y}_T(t) = \Sigma^{-1/2}\Big(\frac{1}{\sqrt{T}}\mathbf{y}_{[Tt]} + \big(Tt - [Tt]\big)\frac{1}{\sqrt{T}}\boldsymbol{\epsilon}_{[Tt]+1}\Big) \xrightarrow{d} \mathbf{W}(\mathbf{t}),$$

where $\mathbf{W}(\mathbf{t})$ is a standard Brownian motion process on $\mathbb{R}^2$. In our case, let

$$\boldsymbol{\epsilon}_j = \begin{pmatrix} \epsilon_{1j} \\ \epsilon_{2j} \end{pmatrix} = \begin{pmatrix} \epsilon_j \\ \epsilon_j^2 - \sigma^2 \end{pmatrix}$$

with the variance matrix

$$\Sigma = \begin{pmatrix} \sigma^2 & 0 \\ 0 & 2\sigma^4 \end{pmatrix}$$

$$\mathbf{Y}_T(t) = \Sigma^{-1/2}\Big(\frac{1}{\sqrt{T}}\sum_{j=1}^{[Tt]}\boldsymbol{\epsilon}_j + \big(Tt - [Tt]\big)\frac{1}{\sqrt{T}}\boldsymbol{\epsilon}_{[Tt]+1}\Big) \xrightarrow{d} \mathbf{W}(\mathbf{t}),$$

namely,

$$\begin{pmatrix} \frac{1}{\sigma} & 0 \\ 0 & \frac{1}{\sqrt{2}\sigma^2} \end{pmatrix} \frac{1}{\sqrt{T}} \begin{pmatrix} \sum_{j=1}^{[Tt]}\epsilon_j + \big(Tt - [Tt]\big)\epsilon_{[Tt]+1} \\ \sum_{j=1}^{[Tt]}(\epsilon_j^2 - \sigma^2) + \big(Tt - [Tt]\big)(\epsilon_{[Tt]+1}^2 - \sigma^2) \end{pmatrix} \xrightarrow{d} \begin{pmatrix} W_1(t) \\ W_2(t) \end{pmatrix}.$$

Let

$$Y_{1T}(t) = \sum_{j=1}^{[Tt]}\epsilon_j + \big(Tt - [Tt]\big)\epsilon_{[Tt]+1}$$

$$Y_{2T}(t) = \sum_{j=1}^{[Tt]}(\epsilon_j^2 - \sigma^2) + \big(Tt - [Tt]\big)(\epsilon_{[Tt]+1}^2 - \sigma^2).$$

Therefore, we have

$$\begin{pmatrix} \frac{1}{\sigma\sqrt{T}} Y_{1T}(t) \\ \frac{1}{\sigma^2\sqrt{2T}} Y_{2T}(t) \end{pmatrix} \xrightarrow{d} \begin{pmatrix} W_1(t) \\ W_2(t) \end{pmatrix}.$$

By definition of $Y_{1T}(t)$ and $Y_{2T}(t)$, $y_{i-1} = Y_{1T}(\frac{i-1}{T})$, $Y_{2T}(\frac{i}{T}) - Y_{2T}(\frac{i-1}{T}) = \epsilon_i^2 - \sigma^2$

$$\sum_{t=1}^{T} y_{t-1}^2(\epsilon_t^2 - \sigma^2) = \sum_{i=1}^{T} Y_{1T}^2(\frac{i-1}{T})\big[Y_{2T}(\frac{i}{T}) - Y_{2T}(\frac{i-1}{T})\big].$$

Let

$$G_T = \sum_{i=1}^{T} Y_{1T}^2(\frac{i-1}{T})\big[Y_{2T}(\frac{i}{T}) - Y_{2T}(\frac{i-1}{T})\big].$$

Lemma 1. *Let $Y_T(r)$ be defined as in (1.3) and $X_T(r)$ defined similarily with respect to the partial sum $x_t = \sum_{s=1}^{t} w_s$ and $w_s \sim iid(0, \sigma_w^2)$. Here, (w_t, σ_t) are serially independent. If $(X_T, Y_T) \xrightarrow{d} (W_X, W_Y)$, then*

$$G_T^* = \sum_{j=0}^{k(T)-1} X_T(r_j)(Y_T(r_{j+1}) - Y_T(r_j))$$

where $r_0, ..., r_{k(T)}$ is a nested sequence of partition. Then, $k(T) \to \infty$ as $T \to \infty$ and

$$\min_{0<j<k(T)} |r_{j+1} - r_j| \to 0.$$

Then we have

$$G_T^* \xrightarrow{d} \int_0^1 W_X dW_Y.$$

Proof. The Skorokhod representation theorem states that a weakly convergent sequence of probability measures whose limit measure is sufficiently well-behaved can be represented as the distribution of a pointwise convergent sequence of random variables defined on a common probability space. Thus, there exists a sequence of random processes (X^T, Y^T) converging almost surely to (W_X, W_Y) in terms of Skorokhod topology. By definition of the Ito integral, if G^{*T} is constructed in the same way except in terms of (X^T, y^T), then

$$\left| G^{*T} - \int_0^1 W_X dW_Y \right| \xrightarrow{p} 0.$$

Convergence in probability implies convergence in distribution. G^{*T} and G_T^* have the same distribution by construction.Then, we are left to show that the differences between G_T and G_T^* can be neglectable. Define $T_j = [Tr_j]$

$$Y_T(r_j) - Y_T(r_{j-1}) = \frac{1}{\sigma_\epsilon \sqrt{T}} \sum_{t=T_{j-1}}^{T_j - 1} \epsilon_t$$

$$X_T(r_j) = \frac{1}{\sigma_w \sqrt{T}} \sum_{i=0}^{T_j - 1} w_{T_j - i}.$$

Thus,

$$\begin{aligned} G_T - G_T^* &= \frac{1}{\sigma_\epsilon \sigma_w T} \sum_{j=1}^{k(T)} \left(\sum_{t=T_{j-1}}^{T_j - 1} \sum_{i=0}^{T_j - 1} \epsilon_{t-i} w_{t+1} - \sum_{i=0}^{T_j - 1} \epsilon_{T_j - i} \sum_{t=T_{j-1}}^{T_j - 1} w_{t+1} \right) \\ &= \frac{1}{\sigma_\epsilon \sigma_w T} \sum_{j=1}^{k(T)} \sum_{t=T_{j-1}+1}^{T_j - 1} \left(\sum_{i=0}^{t - T_{j-1} - 1} \epsilon_{t-i} \right) w_{t+1}. \end{aligned}$$

If we recall that ϵ_t, w_t are serially independent, we can compute the variance

$$\begin{aligned} E(G_T - G_T^*)^2 &= \frac{1}{\sigma_\epsilon^2 \sigma_w^2 T^2} \sum_{j=1}^{k(T)} \sum_{t=T_{j-1}+1}^{T_j-1} E\left(\sum_{i=0}^{t-T_{j-1}-1} \epsilon_{t-i}\right)^2 E(w_{t+1})^2 \\ &= \frac{1}{T^2} \sum_{t=T_{j-1}+1}^{T_j-1} (t - T_{j-1}) \\ &\leq \frac{1}{T^2} \sum_{j=1}^{k(T)} (T_j - T_{j-1})^2 \\ &= O\left(\max_{1\leq j\leq k(T)} |r_j - r_{j-1}|\right) = o(1) \end{aligned}$$

where we obtain convergence in the probability. □

Lemma 2. *In our case,* ϵ_t, w_t *are* $\epsilon_t, \epsilon_t^2 - \sigma^2$*,* $plim(G_T - G_T^*) = 0$*.*

Proof. In our case, random processes (ϵ_t, w_t) are $(\epsilon_t, \epsilon_t^2 - \sigma^2)$,

$$plim(G_T - G_T^*) = \frac{1}{\sqrt{\omega_{\epsilon\epsilon}\omega_{ww}}} \sum_{i=1} E(\epsilon_{1-i} w_1) = 0.$$

□

Using lemmas 1 and 2 and the continuous mapping theorem, we obtain the following results:

Proposition 1.

$$\frac{\sum_{t=1}^{T} y_{t-1}^2(\epsilon_t^2 - \sigma^2)}{\sqrt{2}\sigma^4 T^{3/2}} \xrightarrow{d} \int W_1(t)^2 dW_2(t)$$

Proposition 2.

$$\sum_{t=1}^{T} \frac{y_{t-1}^4}{T^3} \xrightarrow{d} \sigma^4 \int W_1^4(t)dt$$

Proposition 3.

$$\frac{1}{\sqrt{96}\sigma^4 T^{5/2}} \sum_{t=1}^{T} y_{t-1}^4(\epsilon_t^4 - 3\sigma^4) \xrightarrow{d} \int W_1^4(t)dW_3(t)$$

where $W_1(t)$ and $W_3(t)$ are two independent standard Brownian motions.

Proof. Let

$$X_T^*(t) = \frac{1}{\sigma\sqrt{T}} y_{[Tr]};$$

FCLT imples that $X_T^*(t)$ converges weakly to Brownian motion $W_1(t)$. Considering that function $S_T(t)$ is given by

$$S_T(t) = \left[{X_T^*}^4(t)\right],$$

it follows that

$$\int_0^1 S_T(r)dr = \frac{1}{\sigma^4 T^3} \sum_{t=1}^{T} y_{t-1}^4.$$

By the continuous mapping theorem

$$\frac{1}{\sigma^4 T^3} \sum_{t=1}^{T} y_{t-1}^4 \xrightarrow{d} \int_0^1 W_1^4(t)dt.$$

Similarly,

$$\frac{1}{\sqrt{96}\sigma^4 T^{5/2}} \sum_{t=1}^{T} y_{t-1}^4(\epsilon_t^4 - 3\sigma^4) \xrightarrow{d} \int W_1^4(t)dW_3(t).$$

□

For the second term of the loglikelihood ratio,

$$\begin{aligned}\sum_{t=1}^{T} \epsilon_t^4 y_{t-1}^4 &= \sum_{t=1}^{T} (\epsilon_t^4 - E\epsilon_t^4 + E\epsilon_t^4) y_{t-1}^4 \\ &= \sum_{t=1}^{T} (\epsilon_t^4 - 3\sigma^4) y_{t-1}^4 + 3\sigma^4 \sum_{t=1}^{T} y_{t-1}^4.\end{aligned}$$

If r has the order of $\frac{1}{T^{3/4}}$, we can detect the jump, and the limit distribution would be

$$\ln l_t \xrightarrow{d} \frac{\sqrt{2}}{2} \int_0^1 W_1^2(t) dW_2(t) - \frac{1}{4} \int_0^1 W_1^4(t) dt.$$

Here, we complete the proof of the main theorem.

Our test can tell us whether y_t has an explosive root, which can help describe certain series. The estimation of r and how large of a jump we can detect is our main concern. These questions can be tackled with the Bayesian reference.

1.4 Asymmetric Case

In this section, we analyze the asymmetric case in which the jump size is different for the explosion and contraction. For an asymmetric case, we assume that η_t has the following binomial distribution:

$$\eta_t = \begin{cases} a, & p \\ b, & q \end{cases}$$

where $ab + pq = 0, p + q = 1$

$$\ln l_t = \sum_{1}^{T} \ln \left(p e^{\frac{2a\epsilon_t y_{t-1} - a^2 y_{t-1}^2}{2\sigma^2}} + q e^{\frac{2b\epsilon_t y_{t-1} - b^2 y_{t-1}^2}{2\sigma^2}} \right).$$

Let $x = e^{\frac{a^2 y_{t-1}^2 - 2a\epsilon_t y_{t-1}}{2\sigma^2}}$, $y = e^{\frac{b^2 y_{t-1}^2 - 2b\epsilon_t y_{t-1}}{2\sigma^2}}$, $\ln l_t = \sum \ln(pe^x + qe^y)$.

Take the Taylor expansion to the fourth moment when x, y are relatively small. We then obtain

$$\begin{aligned}\ln(pe^{x_1} + qe^{x_2}) \approx\ & qy + \frac{pq}{2}y^2 + \frac{pq(p-q)}{6}y^3 + \frac{pq(p^2-4pq+q^2)}{24}y^4 \\ & + px + \frac{pq}{2}x^2 - \frac{pq(p-q)}{6}x^3 + \frac{pq(p^2-4pq+q^2)}{24}y^4 \\ & - pqxy + \frac{pq(p-q)}{2}xy^2 - \frac{pq(p^2-4qp+q^2)}{6}xy^3 \\ & + \frac{pq(p-q)}{2}x^2y + \frac{pq(p^2-4qp+q^2)}{4}x^2y^2 - \frac{pq(p^2-4pq+q^2)}{6}x^3y.\end{aligned}$$

That is

$$\begin{aligned}\ln l_t \approx \sum_{t=1}^{T} \Bigg[& (\frac{ap+bq}{\sigma^2})\epsilon_t y_{t-1} - (\frac{a^2p+b^2q}{2\sigma^2})y_{t-1}^2 + \frac{pq}{8\sigma^4}(a^2-b^2)^2 y_{t-1}^4 \\ & - \frac{pq}{2\sigma^4}(a^3 - a^2b - ab^2 + b^3)\epsilon_t y_{t-1}^3 + \frac{pq}{2\sigma^4}(a-b)^2\epsilon_t^2 y_{t-1}^2 \\ & + \frac{pq(p-q)}{4\sigma^6}(a^2-b^2)(a-b)^2\epsilon_t^2 y_{t-1}^4 + \frac{pq(p^2-4pq+q^2)}{24\sigma^8}(a-b)^4\epsilon_t^4 y_{t-1}^4 \Bigg].\end{aligned}$$

We can rewrite it as

$$\begin{aligned}\ln l_t \approx \sum_{t=1}^{T} \Bigg[& (\frac{ap+bq}{\sigma^2})\epsilon_t y_{t-1} - \frac{(ap+bq)^2}{2\sigma^2}y_{t-1}^2 + \frac{pq}{2\sigma^4}(a-b)^2(\epsilon_t^2 - \sigma^2)y_{t-1}^2 \\ & + \big[\frac{pq}{8^4}(a^2-b^2)^2 + (a-b)^4\frac{pq(p^2-4pq+q^2)}{8\sigma^4} + \frac{pq(p-q)}{4\sigma^4}(a^2-b^2)(a-b)^2\big]y_{t-1}^4 \\ & + \frac{pq(p^2-4pq+q^2)}{24\sigma^8}(a-b)^4(\epsilon_t^4 - 3\sigma^4)y_{t-1}^4 + \frac{pq(p-q)}{4\sigma^6}(a^2-b^2)(a-b)^2(\epsilon_t^2-\sigma^2)y_{t-1}^4 \\ & - \frac{pq}{2\sigma^4}(a^3 - a^2b - ab^2 + b^3)\epsilon_t y_{t-1}^3 \Bigg].\end{aligned}$$

Because $ap + bq = 0$, the first and second terms vanish. We can derive the limit distributions for the third and fourth terms if a and b have the order $\frac{1}{T^{3/4}}$ as proposition 1 and 2.

The fifth, sixth and seventh terms vanish if a and b have that order by FCLT. To be more precise,

$$\frac{1}{\sqrt{2}\sigma^6 T^{5/2}} \sum_{t=1}^{T} (\epsilon_t^2 - \sigma^2) y_{t-1}^4 \to \sigma^4 \int_0^1 W_1^4(t) dt$$

$$\frac{1}{\sigma^4} T^2 \sum_{t=1}^{T} \epsilon_t y_{t-1}^3 \to \int_0^1 W_1^3(r) dW_1(r).$$

The limiting distribution for $\ln l_t$ should be

$$\ln l_t \to c_1 \int_0^1 W_1^2(t) dW_2(t) + c_2 \int_0^1 W_1^4(t) dt.$$

1.5 Concluding Remarks

We propose a new likelihood ratio test against Markov switching unit root models. This test applies to a wide range of models that are popular in macroeconomics and finance. It is simple to implement, as it requires only the estimation of the parameters under the null hypothesis of constant parameters. The testing problem is challenging due to nonstationarity. We derive the limiting distribution with i.i.d case.

Chapter 2

Optimal test for Markov dynamics

In recent years, the field of econometrics has seen the development of various tools designed to test for the existence of financial bubbles. These tests carry significant practical implications, because their detection can impact a wide range of stakeholders, including investors, portfolio managers, regulators, and policymakers. Identifying the presence of financial bubbles is crucial for taking appropriate countermeasures. With Markov dynamics in our jump parameter η, our model should better capture the market changes. The parameters driving the dynamic of the underlying Markov chain are not identified under the null hypothesis. As a result, the testing problem is nonstandard. We believe that such a model is a better representation of the behavior of prices series than the usual fixed unit root model.

2.1 Introduction

In this paper we concentrate on the likelihood ratio test with Markov switching. We assume the data are generated as

$$y_t = \alpha_t y_{t-1} + \epsilon_t$$

where α_t could be written as $1 + \eta_t$ with η_t being the unobserved switching variable. η_t is assumed to be Markov. We want to test the null hypothesis $\alpha_t = 1$ against $\alpha_t = 1 + \eta_t$. η_t could be greater than zero, that is, we allow for explosive behavior in some regimes. If η_t is lower than zero, contraction is obtained. We believe that such a model is a better representation of the behavior of prices series than the usual fixed unit root model.For instance, models with fixed coefficients ignore the effects that a succession of small and varied macroeconomic shocks may have on the structure of appropriate dynamic economic models.A main source of these shocks is policy change.And people have noticed this problem long ago.For example,in his famous 1976 paper,Lucas suggested that using a simultaneous equations model for policy analysis and simulation is only valid when agent's expectation rules are embedded in its functional form since they are supposed to change with the changes in policy. And he used rational expectation equilibrium framework in which this property is satisfied by the expectation rules.

Few research papers have proposed tests for Markov switching. Garcia (1998) examined the asymptotic distribution of a likelihood ratio test of the sup-type. Similarly, (Hansen, 1992) approached the problem by considering likelihood as an empirical process with parameters as indices and based his test on the supremum of the likelihood ratio over nuisance parameters. Both of these approaches entail estimating the model under alternative scenarios, which can be quite labor-intensive. Surprisingly, neither of these studies investigates local statistical power.

Gong et al. (1997) took a different approach by reparameterizing their linear model in the frequency domain. They devised a test that relied on differences in the spectrum between null and alternative hypotheses. On a Bayesian front, Kim and Nelson (2001) proposed a model selection procedure for Markov switching.

The remainder of the paper is organized as follows. The model is described in Section 2. We also characterize likelihood ratio and the formula for our test statistics in this section.

Taylor expansion of likelihood is studied in Section 3. We derive the limit distribution in section 4. We conclude in Section 5. All the proofs are relegated in the Appendix.

2.2 Model depscription

For the Markov case we need to integrate with all the states. We assume that η_t is a three-states Markov chain with a 3×3 transition matrix as shown below:

$$P = \begin{pmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33}. \end{pmatrix}$$

The three states are:

$$\eta_t = \begin{cases} a, & State1 \\ 0, & State2 \\ b, & State3. \end{cases} \cdot$$

The sample is split into blocks, where a B_N block has length $B(T)$, and i is the index for block $i = 1, ..., B_N$. We decompose the sum as follows:

$$\sum_{t=1}^{T} = \sum_{i=1}^{B_N} \sum_{t=(i-1)B(T)+1}^{iB(T)}.$$

Our analysis relies on the derivatives of the logarithm of the likelihood function. We denote the conditional parametric densities as $f_t(\theta_t)$, which are functions of the parameter vector described in the Markov process. As in Chapter 1, we obtain the likelihood ratio as

$$
\begin{aligned}
L_t^Q &= \int \prod_{t=1}^{T} f_t(\theta_t) dQ_T^\eta / \prod_{t=1}^{T} f_t(\theta_0) \\
&= \int e^{-\frac{1}{2\sigma^2}\sum_{t=1}^{T}[y_t-(1+\eta_t)y_{t-1}]^2} dQ_T^\eta / e^{-\frac{1}{2\sigma^2}\sum_{t=1}^{T}(y_t-y_{t-1})^2} \\
&= E_\eta(\prod_{t=1}^{T} e^{\frac{1}{\sigma^2}\Delta y_t \eta_t y_{t-1} - \frac{1}{2\sigma^2}\eta_t^2 y_{t-1}^2}).
\end{aligned}
$$

For now we assume that $\sigma = 1$.

The main difference with CHP is that here the local alternatives are of order $T^{-3/4}$ Let

$$
G_i = \sum_{t=(i-1)B(T)+1}^{iB(T)} \frac{1}{\sqrt[4]{T}} \Delta y_t \frac{y_{t-1}}{\sqrt{T}} \eta_t - \frac{1}{2\sqrt{T}} \frac{y_{t-1}^2}{T} \eta_t^2.
$$

Our local likelihood would be

$$
L = E_\eta \left[exp \sum_{i=1}^{B_N} G_i \right].
$$

Let

$$
c_i = -log E_\eta [exp G_i | \mathcal{F}_{i-1}].
$$

Thus

$$
\begin{aligned}
E\left[\prod_{i=1}^{T}(exp(G_i)exp(c_i))\right] &= E\left[E\left(\prod_{i=1}^{T} exp(G_i)exp(c_i)\middle|\mathcal{F}_{T-1}\right)\right] \\
&= E\left[\prod_{i=1}^{T-1} exp(G_i)exp(c_i)E\left(exp(G_T)exp(c_T)|\mathcal{F}_{T-1}\right)\right] \\
&= E\left[\prod_{i=1}^{T-1} exp(G_i)exp(c_i)\right] \\
&\vdots \\
&= 1
\end{aligned}
\tag{2.1}
$$

First equation is the result of law of iterated expectation and the second one is due to the construction of c_i.

Proposition 1. *There exists c_i' which is irrelevant to η such that*

$$\sum_i c_i' / \sum_i c_i \xrightarrow{p} 1.$$

and $\sum_i c_i'$ coverge to constant in probability on some set A_T such that $\lim\limits_{T\to\infty} P(A_T) = 1$ where A_T is $\mathcal{F}_T$ measurable and independent of η

The following lemma in CHP is used in the proof of Proposition 1

Lemma 1. *Assume that for any $\epsilon > 0$, we can find $1 - \epsilon \leq \frac{f_T}{f_{T^*}} \leq 1 + \epsilon$ on some set A_T^ϵ so that $\lim\limits_{T\to\infty} P(A_T^\epsilon) = 1$ where A_T^ϵ is $\mathcal{F}_T$ measurable and independent of η. Then $\frac{E(f_T|\mathcal{F}_T)}{E(f_T^*|\mathcal{F}_T)} \xrightarrow{P} 1$*

Along the lines of CHP, we can use the taylor expansion to approximate likelihood ratio on some set with probability one.

2.3 Taylor Expansion of likelihood

For each c_i, taking taylor expansion at the fourth order, we obtain

$$E_\eta(expG_i|\mathcal{F}_{i-1}) = E_\eta\left[1 + G_i + \frac{1}{2}G_i^2 + \frac{1}{6}G_i^3 + \frac{1}{24}G_i^4 + \frac{1}{5!}G_i^{*5}\middle|\mathcal{F}_{i-1}\right]. \tag{2.2}$$

Let

$$a = \sum_{t=(i-1)B(T)+1}^{iB(T)} \frac{1}{\sqrt[4]{T}}\Delta y_t \frac{y_{t-1}}{\sqrt{T}}\eta_t$$

$$b = \sum_{t=(i-1)B(T)+1}^{iB(T)} \frac{-1}{2\sqrt{T}}\frac{y_{t-1}^2}{T}\eta_t^2$$

According to the definition of G_i,

$$\begin{aligned}1 + G_i + \frac{1}{2}G_i^2 + \frac{1}{6}G_i^3 + \frac{1}{24}G_i^4 &= 1 + (a+b) + \frac{1}{2}(a+b)^2 + \frac{1}{6}(a+b)^3 + \frac{1}{24}(a+b)^4 \\ &= 1 + a + b + \frac{1}{2}a^2 + ab + \frac{1}{2}b^2 + \frac{1}{6}a^3 + \frac{1}{2}a^2b + \frac{1}{2}ab^2 + \frac{1}{6}b^3 \\ &+ \frac{1}{24}a^4 + \frac{1}{6}a^3b + \frac{1}{4}a^2b^2 + \frac{1}{6}ab^3 + \frac{1}{24}b^4.\end{aligned}$$

We obtain 14 terms in the form of a and b. The next step is to evaluate that each term a and b are clear as above. Let the second order

$$a^2 = \sum_{B_i}\frac{1}{\sqrt{T}}(\Delta y_t)^2\frac{y_{t-1}^2}{T}\eta_t^2 + \sum_{k\neq j}\frac{1}{\sqrt{T}}\Delta y_k\frac{y_{k-1}}{\sqrt{T}}\Delta y_j\frac{y_{j-1}}{\sqrt{T}}\eta_k\eta_j.$$

Taking conditional expectation, Term 3

$$E_\eta \frac{1}{2}a^2 = \sum_{B_i} \frac{1}{2\sqrt{T}}(\Delta y_t)^2 \frac{y_{t-1}^2}{T} E(\eta_t^2|\mathcal{F}_{i-1}) + \sum_{k\neq j} \frac{1}{2\sqrt{T}} \Delta y_k \frac{y_{k-1}}{\sqrt{T}} \Delta y_j \frac{y_{j-1}}{\sqrt{T}} E(\eta_k \eta_j|\mathcal{F}_{i-1})$$

We know η is exponentially decayed in our Markov case. If we replace the conditional expectation, given that $\mathcal{F}_{i-1}$ with an unconditional expectation, the differences between them would be neglectable. Term 3 can be written as

$$E\frac{1}{2}a^2 = \sum_{B_i} \frac{1}{2\sqrt{T}}(\Delta y_t)^2 \frac{y_{t-1}^2}{T} E\eta_t^2 + \sum_{k} \frac{1}{\sqrt{T}} \Delta y_k \frac{y_{k-1}}{\sqrt{T}} \sum_{j<k} \Delta y_j \frac{y_{j-1}}{\sqrt{T}} E(\eta_k \eta_j).$$

We can do the same thing for the remaining terms. The results are listed as follows.

Term 4 (ab)

$$Eab = \sum_{B_i} \frac{1}{2\sqrt{T}\sqrt[4]{T}} \Delta y_t \frac{y_{t-1}}{\sqrt{T}} \frac{y_{t-1}^2}{T} E\eta_t^3 + \sum_{k\neq j} \frac{-1}{2\sqrt{T}\sqrt[4]{T}} \Delta y_k \frac{y_{k-1}}{\sqrt{T}} \frac{y_{j-1}^2}{T} E(\eta_k \eta_j^2)$$

Following Lemma 1 shows that third moment would vanish:

Lemma 2. *Under our assumptions for* η_t*,*

$$\sum_i \sum_{B_i} \frac{1}{2\sqrt{T}\sqrt[4]{T}} \Delta y_t \frac{y_{t-1}}{\sqrt{T}} \frac{y_{t-1}^2}{T} E\eta_t^3 \xrightarrow{p} 0.$$

Proof: We can see that Δy_t, $\frac{y_{t-1}}{\sqrt{T}}$, $\frac{y_{t-1}^2}{T}$ all have the order of 1. The magnitude of variance for the sum would be $O(T^{-1.5} \times Blocksize \times constant)$. In the logL we sum all the blocks the maginitude would be $O(T^{-0.5} \times constant)$ Convergence in mean square implies convergence in probability.

Lemma 3.

$$\sum_i \sum_{k\, in\, B_i} \frac{1}{2\sqrt[3/4]{T}} \Delta y_k \frac{y_{k-1}}{\sqrt{T}} \sum_{j<k} \frac{y_{j-1}^2}{T} \eta_k \eta_j^2 \xrightarrow{p} 0$$

Proof: As η is exponentially decayed, we have

$$|E\eta_k \eta_j^2| < |\lambda^{k-j} \eta_j^3|.$$

so $\sum_{j<k} \frac{y_{j-1}^2}{T} \eta_k \eta_j^2$ exists. Hence, the sums are (as a product of a martingale difference sequence with terms determined in the past)are m.d.s. Again, we can calculate the variance of each term like we did in Lemma 1, where the magnitude would be $O(T^{-0.5} \times constant.)$ Therefore, it converges to zero in probability. For the case $j > k$, we can switch index. Hence, Term 4 vanish, and we obtain the following proposition:

Proposition 2.

$$\sum_{B_i} \frac{1}{2\sqrt{T}\sqrt[4]{T}} \Delta y_t \frac{y_{t-1}}{\sqrt{T}} \frac{y_{t-1}^2}{T} E\eta_t^3 + \sum_{k\neq j} \frac{-1}{2\sqrt{T}\sqrt[4]{T}} \Delta y_k \frac{y_{k-1}}{\sqrt{T}} \frac{y_{j-1}^2}{T} E(\eta_k \eta_j^2) \xrightarrow{p} 0$$

Term 5 ($\frac{1}{2}b^2$)

$$E\frac{1}{2}b^2 = \frac{1}{2} \sum \frac{1}{4T} \frac{y_{t-1}^4}{T^2} E\eta_t^4 + \sum_k \frac{1}{4T} \frac{y_{k-1}^2}{T} \sum_{j<k} \frac{y_{j-1}^2}{T} E\eta_k^2 \eta_j^2$$

Term 6 ($\frac{1}{6}a^3$)

$$
\begin{aligned}
E\frac{1}{6}a^3 = & \frac{1}{6}\sum_t \frac{1}{T^{3/4}}(\Delta y_t)^3 \left(\frac{y_{t-1}}{\sqrt{T}}\right)^3 E\eta_t^3 \\
& + \frac{1}{2}\sum_k \frac{1}{\sqrt[4]{T}}\Delta y_k \frac{y_{k-1}}{\sqrt{T}} \sum_{j<k} \frac{1}{\sqrt{T}}(\Delta y_j)^2 \frac{y_{j-1}^2}{T} E\eta_j^2\eta_k \\
& + \frac{1}{2}\sum_k \frac{1}{\sqrt[4]{T}}\Delta y_k \frac{y_{k-1}}{\sqrt{T}} \sum_{j>k} \frac{1}{\sqrt{T}}(\Delta y_j)^2 \frac{y_{j-1}^2}{T} E\eta_j^2\eta_k \\
& + \sum_{l<j<k} \left(\frac{1}{\sqrt[4]{T}}\right)^3 \Delta y_l \Delta y_j \Delta y_k \frac{y_{l-1}}{\sqrt{T}} \frac{y_{j-1}}{\sqrt{T}} \frac{y_{k-1}}{\sqrt{T}} E\eta_l\eta_j\eta_k
\end{aligned}
$$

Following Lemmas 1 and 2, the first three lines converge to zero. Looking at the last term, we obtain

Lemma 4.

$$
\sum_i \sum_{l<j<k} \left(\frac{1}{\sqrt[4]{T}}\right)^3 \Delta y_l \Delta y_j \Delta y_k \frac{y_{l-1}}{\sqrt{T}} \frac{y_{j-1}}{\sqrt{T}} \frac{y_{k-1}}{\sqrt{T}} E\eta_l\eta_j\eta_k \xrightarrow{p} 0
$$

Proof: the sum can be written as

$$
\sum_k \left(\frac{1}{\sqrt[4]{T}}\right)^3 \Delta y_k \frac{y_{k-1}}{\sqrt{T}} \sum_{j<k} \Delta y_j \frac{y_{j-1}}{\sqrt{T}} \sum_{l<j} \Delta y_l \frac{y_{l-1}}{\sqrt{T}}.
$$

Again, the variance is bounded, and $\sum_{l<j}, \sum_{j<k}$ are terms in the "past" sums are m.d.s Hence, our assumption for exponential mixing guarantees convergence to zero in mean square. Thus, we have the following proposition:

Proposition 3. *If η is exponentially mixing, all the third order term in the log likelihood Taylor expansion converge to zero in probability.*

Term 7 ($\frac{1}{2}a^2b$)

$$E\frac{1}{2}a^2b = \frac{1}{2}\sum \frac{-1}{2T}(\Delta y_t)^2\frac{y_{t-1}^4}{T^2}E\eta_t^4 + \sum_{k\neq j}\frac{-1}{2T}(\Delta y_k)^2\frac{y_{k-1}^2}{T}\frac{y_{j-1}^2}{T}E\eta_k^2\eta_j^2$$
$$+ \frac{1}{2}\sum_{j\neq k,l}\frac{-1}{2T}\Delta y_j\frac{y_{j-1}}{\sqrt{T}}\Delta y_k\frac{y_{k-1}}{\sqrt{T}}\frac{y_{l-1}^2}{T}E\eta_j\eta_k\eta_l^2$$

Term 8 ($\frac{1}{2}ab^2$)

$$E\frac{1}{2}ab^2 = \frac{1}{2}\sum_{l,k}\frac{1}{4T\sqrt[4]{T}}\Delta y_l\frac{y_{l-1}}{\sqrt{T}}\frac{y_{k-1}^4}{T^2}E\eta_l\eta_k^4$$
$$+ \frac{1}{2}\sum_{k\neq j,l}\frac{1}{\sqrt[4]{T}}\Delta y_l\frac{y_{l-1}}{\sqrt{T}}\frac{1}{4T}\frac{y_{k-1}^2}{T}\frac{y_{j-1}^2}{T}E\eta_k^2\eta_j^2\eta_l$$

Lemma 5. *Term 8 converges to zero in probability.*

Proof: $\Delta y_l\frac{y_{l-1}}{\sqrt{T}}\frac{y_{k-1}^4}{T^2}$ are all of order 1, and $E\eta_l\eta_k^4$ is exponentially decayed. Following Lemma 1, it would vanish. So does the the second sum. We have the following proposition:

Proposition 4. *If η is exponentially decayed, all the fifth order terms in the log likelihood Taylor expansion converge to zero in probability.*

Term 9 ($\frac{1}{6}b^3$)

$$E\frac{1}{6}b^3 = \frac{1}{6}\sum_{B_i}\frac{-1}{8T\sqrt{T}}\frac{y_{t-1}^6}{T^3}E\eta_t^6$$
$$+ \frac{1}{2}\sum_{j\neq k}\frac{y_{j-1}^4}{T^2}\frac{-1}{8T\sqrt{T}}\frac{y_{k-1}^2}{T}\eta_j^4\eta_k^2$$
$$+ \sum_{l<j<k}\frac{-1}{8T\sqrt{T}}\frac{y_{l-1}^2y_{j-1}^2y_{k-1}^2}{T^3}E\eta_l^2\eta_j^2\eta_k^2$$

Proposition 5. *If η is exponentially decayed, all the sixth and higher order terms in the log likelihood Taylor expansion converge to zero in probability.*

Proof: $\frac{y_{t-1}^6}{T^3}, \frac{y_{k-1}^2}{T}, \frac{y_{l-1}^2 y_{j-1}^2 y_{k-1}^2}{T^3}$ are all of order 1. Mean square of term 9 would converge to zero in probability. Note that there is no exponential decay in sixth order.

Term 10 ($\frac{1}{24}a^4$)

$$
\begin{aligned}
E\frac{1}{24}a^4 = &\frac{1}{24}\sum_{B_i}\frac{1}{T}(\Delta y_t)^4\frac{y_{t-1}^4}{T^2}E\eta_t^4 \\
&+\frac{1}{6}\sum_{j\neq k}\frac{1}{T}(\Delta y_j)^3(\frac{y_{j-1}}{\sqrt{T}})^3\Delta y_k\frac{y_{k-1}}{\sqrt{T}}E\eta_j^3\eta_k \\
&+\frac{1}{4}\sum_{j<k}\frac{1}{T}\left(\Delta y_j\frac{y_{j-1}}{\sqrt{T}}\right)^2\left(\Delta y_k\frac{y_{k-1}}{\sqrt{T}}\right)^2E\eta_j^2\eta_k^2 \\
&+\frac{1}{2}\sum_{l<j<k}\frac{1}{T}\left(\Delta y_l\frac{y_{l-1}}{\sqrt{T}}\right)^2\left(\Delta y_j\frac{y_{j-1}}{\sqrt{T}}\right)\left(\Delta y_k\frac{y_{k-1}}{\sqrt{T}}\right)E\eta_l^2\eta_j\eta_k \\
&+\frac{1}{2}\sum_{l<j<k}\frac{1}{T}\left(\Delta y_l\frac{y_{l-1}}{\sqrt{T}}\right)\left(\Delta y_j\frac{y_{j-1}}{\sqrt{T}}\right)^2\left(\Delta y_k\frac{y_{k-1}}{\sqrt{T}}\right)E\eta_l\eta_j^2\eta_k \\
&+\frac{1}{2}\sum_{l<j<k}\frac{1}{T}\left(\Delta y_l\frac{y_{l-1}}{\sqrt{T}}\right)\left(\Delta y_j\frac{y_{j-1}}{\sqrt{T}}\right)\left(\Delta y_k\frac{y_{k-1}}{\sqrt{T}}\right)^2E\eta_l\eta_j\eta_k^2 \\
&+\sum_{l<j<k<m}\frac{1}{T}\left(\Delta y_l\frac{y_{l-1}}{\sqrt{T}}\right)\left(\Delta y_j\frac{y_{j-1}}{\sqrt{T}}\right)\left(\Delta y_k\frac{y_{k-1}}{\sqrt{T}}\right)\left(\Delta y_m\frac{y_{m-1}}{\sqrt{T}}\right)E\eta_l\eta_j\eta_k\eta_m
\end{aligned}
$$

Term 11 ($\frac{1}{6}a^3b$)

$$
\begin{aligned}
E\frac{1}{6}a^3b &= E\frac{1}{6}\sum_{l,j,k,m}\left(\frac{1}{\sqrt[4]{T}}\Delta y_l\frac{y_{l-1}}{\sqrt{T}}\eta_l\right)\left(\frac{1}{\sqrt[4]{T}}\Delta y_j\frac{y_{j-1}}{\sqrt{T}}\eta_j\right)\left(\frac{1}{\sqrt[4]{T}}\Delta y_k\frac{y_{k-1}}{\sqrt{T}}\eta_k\right)\left(\frac{-1}{2\sqrt{T}}\frac{y_{m-1}}{T}\eta_m^2\right) \\
&= -\frac{1}{6}\sum_{l,j,k,m}\frac{1}{T^{3/4}}\Delta y_l\frac{y_{l-1}}{\sqrt{T}}\Delta y_j\frac{y_{j-1}}{\sqrt{T}}\Delta y_k\frac{y_{k-1}}{\sqrt{T}}\frac{y_{m-1}^2}{2T\sqrt{T}}E\eta_l\eta_j\eta_k\eta_m^2
\end{aligned}
$$

Proposition 6. *If η is exponentially decayed, Term 11 in the log likelihood Taylor expansion converge to zero in probability.*

Term 12 $(\frac{1}{4}a^2b^2)$

$Term\ \ 12 = \frac{1}{4}(Term\ \ 4)^2$. According to Proposition 1 Term 12 converge to zero in probability.

Term 13 $(\frac{1}{6}ab^3)$

$$E\frac{1}{6}ab^3 = E\frac{1}{6}\sum_{l,j,k,m}\left(\frac{1}{\sqrt[4]{T}}\Delta y_l \frac{y_{l-1}}{\sqrt{T}}\eta_l\right)\left(\frac{-1}{2\sqrt{T}}\frac{y_{j-1}^2}{T}\eta_j^2\right)\left(\frac{-1}{2\sqrt{T}}\frac{y_{k-1}^2}{T}\eta_k^2\right)\left(\frac{-1}{2\sqrt{T}}\frac{y_{m-1}^2}{T}\eta_m^2\right)$$
$$= -\frac{1}{48}\sum_{l,j,k,m}\frac{1}{T^{7/4}}\Delta y_l \frac{y_{l-1}}{\sqrt{T}}\frac{y_{j-1}^2}{T}\frac{y_{k-1}^2}{T}\frac{y_{m-1}^2}{T}E\eta_l\eta_j^2\eta_k^2\eta_m^2$$

Proposition 7. *If η is exponentially decayed, Term 13 in the log likelihood Taylor expansion converge to zero in probability.*

Term 14 $(\frac{1}{24}b^4)$

$$E\frac{1}{24}b^4 = E\frac{1}{24}\sum_{l,j,k,m}\left(\frac{-1}{2\sqrt{T}}\frac{y_{l-1}^2}{T}\eta_l^2\right)\left(\frac{-1}{2\sqrt{T}}\frac{y_{j-1}^2}{T}\eta_j^2\right)\left(\frac{-1}{2\sqrt{T}}\frac{y_{k-1}^2}{T}\eta_k^2\right)\left(\frac{-1}{2\sqrt{T}}\frac{y_{m-1}^2}{T}\eta_m^2\right)$$
$$= \frac{1}{384T^2}\sum_{l,j,k,m}\frac{y_{l-1}^2}{T}\frac{y_{j-1}^2}{T}\frac{y_{k-1}^2}{T}\frac{y_{m-1}^2}{T}E\eta_l^2\eta_j^2\eta_k^2\eta_m^2$$

Proposition 8. *If η is exponentially decayed, Term 14 in the log likelihood Taylor expansion converge to zero in probability.*

2.4 Limit Distribution

The remaining parts in $logL$ would be second order and fourth order. Fourth order should compensate. The Taylor series expansion for log(1+x) informs that the fourth order are exactly $-\frac{1}{2}$ square of the second order. Then, we have

Proposition 9. *The fourth order compensates for the second order in the Taylor expansion* *of log likelihood*

To be more specific, we plug in σ,where the second order in $LogL$ would be

$$\sum_{i=1}^{B_N}\left(\sum_{B_i}\frac{1}{\sigma^4}\frac{1}{2\sqrt{T}}\frac{y_{t-1}^2}{T}[(\Delta y_t)^2-\sigma^2]E\eta_t^2+\sum_k\frac{1}{\sigma^4\sqrt{T}}\Delta y_k\frac{y_{k-1}}{\sqrt{T}}\sum_{j<k}\Delta y_j\frac{y_{j-1}}{\sqrt{T}}E\eta_k\eta_j\right) \quad (2.3)$$

$$=\sum_t\frac{1}{\sigma^4}\frac{1}{2\sqrt{T}}\frac{y_{t-1}^2}{T}[(\Delta y_t)^2-\sigma^2]E\eta_t^2+\sum_t\frac{1}{\sigma^4\sqrt{T}}\Delta y_t\frac{y_{t-1}}{\sqrt{T}}\sum_{j<t}\Delta y_j\frac{y_{j-1}}{\sqrt{T}}E\eta_t\eta_j. \quad (2.4)$$

Our main aim would be the asymptotic behavior of $LogL$ in this Markov situation.

Theorem 1. *Under H_0, the limit distribution for the second order would be*

$$\sqrt{2}\int W_1^2(t)dW_2(t)E\eta_t^2+L\int W_1(t)^2dW_3(t)E\eta_t\eta_{t-s}.$$

Here L is finite.

Theorem 1 is our main contribution to Markov dynamics. The following proposition shows the asymptotic behavior for the first sum in (2.2):

Under H_0, we have

$$\sum_t\frac{1}{\sigma^4}\frac{1}{2\sqrt{T}}\frac{y_{t-1}^2}{T}[\epsilon_t^2-\sigma^2]E\eta_t^2\to\sqrt{2}\int W_1^2(t)dW_2(t)E\eta_t^2 \quad (2.5)$$

which is exactly proposition 1 in Chapter 1.

Lemma 6.

$$plim\left|\sum_t\frac{1}{\sigma^4\sqrt{T}}\Delta y_t\frac{y_{t-1}}{\sqrt{T}}\sum_{j<t}\Delta y_j\frac{y_{j-1}}{\sqrt{T}}E\eta_t\eta_j-\sum_t\frac{1}{\sigma^4\sqrt{T}}\Delta y_t\frac{y_{t-1}}{\sqrt{T}}\sum_{s=1}^{U}\Delta y_{t-s+1}\frac{y_{t-s}}{\sqrt{T}}E\eta_t\eta_j\right|=0$$

Proof When j is very different from t, $|E\eta_t\eta_j|<\lambda^{(t-j)}E\eta_j^2$. Then, only finite j matters, and s is finite from 1 to U.

For each s, we have

$$\sum_t \frac{1}{\sigma^4\sqrt{T}}\Delta y_t \frac{y_{t-1}}{\sqrt{T}}\Delta y_{t-s}\frac{y_{t-s-1}}{\sqrt{T}}E\eta_t\eta_{t-s}. \tag{2.6}$$

Replace y_{t-1} with $y_{t-1} = \epsilon_{t-1} + \epsilon_{t-2} + ... + \epsilon_{t-s} + y_{t-s-1}$, and we have

$$\sum_t \frac{1}{\sigma^4\sqrt{T}}\Delta y_t \frac{y_{t-1}}{\sqrt{T}}\Delta y_{t-s}\frac{y_{t-s-1}}{\sqrt{T}}E\eta_t\eta_{t-s} \tag{2.7}$$

$$= \sum_t \frac{1}{\sigma^4\sqrt{T}}\Delta y_t \frac{(\epsilon_{t-1} + ... + \epsilon_{t-s})}{\sqrt{T}}\Delta y_{t-s}\frac{y_{t-s-1}}{\sqrt{T}}E\eta_t\eta_{t-s} \tag{2.8}$$

$$+ \sum_t \frac{1}{\sigma^4\sqrt{T}}\Delta y_t \frac{y_{t-s-1}}{\sqrt{T}}\Delta y_{t-s}\frac{y_{t-s-1}}{\sqrt{T}}E\eta_t\eta_{t-s}. \tag{2.9}$$

Lemma 7. *Under H_0,*

$$\sum_t \frac{1}{\sigma^4\sqrt{T}}\Delta y_t \frac{(\epsilon_{t-1} + ... + \epsilon_{t-s})}{\sqrt{T}}\Delta y_{t-s}\frac{y_{t-s-1}}{\sqrt{T}}E\eta_t\eta_{t-s} \xrightarrow{p} 0.$$

Proof: $\Delta y_t = \epsilon_t$,where ϵ_i is uncorrelated with ϵ_j when $i \neq j$. Again, y_{t-s-1} is a term in the "past"; thus, m.d.s. The terms in (2.6) are all of order one. Mean square would converge to zero, which completes the proof of Lemma 5.

Rewrite (2.7) as

$$\sum_t \frac{1}{\sigma^4\sqrt{T}}\Delta y_t \frac{y_{t-1}}{\sqrt{T}}\Delta y_{t-s}\frac{y_{t-s-1}}{\sqrt{T}}E\eta_t\eta_{t-s} \tag{2.10}$$

$$= \sum_t \frac{1}{\sigma^4\sqrt{T}}\epsilon_t\epsilon_{t-s}\frac{y^2_{t-s-1}}{T}E\eta_t\eta_{t-s}. \tag{2.11}$$

We follow the same methodology as in Chapter 1 to obtain the limiting distribution. Let

$$Y_T(t) = \frac{1}{\sqrt{T}\sigma}y_{[Tt]} + \big(Tt - [Tt]\big)\frac{1}{\sqrt{T}\sigma}\epsilon_{[Tt]+1}.$$

Because $\sigma^2 < \infty$, $Y_T(t)$ weakly converges to a standard Brownian motion by FCLT. The corresponding multivariate FCLT would be

$$\mathbf{Y}_T(t) = \Sigma^{-1/2}\Big(\frac{1}{\sqrt{T}}\mathbf{y}_{[Tt]} + \big(Tt - [Tt]\big)\frac{1}{\sqrt{T}}\boldsymbol{\epsilon}_{[Tt]+1}\Big) \xrightarrow{d} \mathbf{W}(\mathbf{t}).$$

where $\mathbf{W}(\mathbf{t})$ is a standard Brownian motion process on $\mathbb{R}^2$. In Markov dynamics,

$$\boldsymbol{\epsilon}_j = \begin{pmatrix} \epsilon_{1j} \\ \epsilon_{2j} \end{pmatrix} = \begin{pmatrix} \epsilon_j \\ \epsilon_j \epsilon_k \end{pmatrix}$$

with the variance matrix

$$\Sigma = \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^4 \end{pmatrix}$$

$$\mathbf{Y}_T(t) = \Sigma^{-1/2}\Big(\frac{1}{\sqrt{T}}\sum_{j=1}^{[Tt]}\boldsymbol{\epsilon}_j + \big(Tt - [Tt]\big)\frac{1}{\sqrt{T}}\boldsymbol{\epsilon}_{[Tt]+1}\Big) \xrightarrow{d} \mathbf{W}(\mathbf{t})$$

namely,

$$\begin{pmatrix} \frac{1}{\sigma} & 0 \\ 0 & \frac{1}{\sigma^2} \end{pmatrix} \frac{1}{\sqrt{T}} \begin{pmatrix} \sum_{j=1}^{[Tt]}\epsilon_j + \big(Tt - [Tt]\big)\epsilon_{[Tt]+1} \\ \sum_{j=1}^{[Tt]}(\epsilon_j\epsilon_{j+s}) + \big(Tt - [Tt]\big)(\epsilon_{[Tt]+1}\epsilon_{[Tt]+1+s}) \end{pmatrix} \xrightarrow{d} \begin{pmatrix} W_1(t) \\ W_3(t) \end{pmatrix}.$$

Let

$$Y_{1T}(t) = \sum_{j=1}^{[Tt]}\epsilon_j + \big(Tt - [Tt]\big)\epsilon_{[Tt]+1}$$

$$Y_{3T}(t) = \sum_{j=1}^{[Tt]}(\epsilon_j\epsilon_{j+s}) + \big(Tt - [Tt]\big)(\epsilon_{[Tt]+1}\epsilon_{[Tt]+1+s})$$

Thus, we have

$$\begin{pmatrix} \frac{1}{\sqrt{T}\sigma}Y_{1T}(t) \\ \frac{1}{\sqrt{T}\sigma^2}Y_{3T}(t) \end{pmatrix} \xrightarrow{d} \begin{pmatrix} W_1(t) \\ W_3(t) \end{pmatrix}$$

By definition of $Y_{1T}(t)$ and $Y_{3T}(t)$, $y_{i-1} = X_{1T}(\epsilon_i\epsilon_{i+s} = \frac{i-1}{T})$, $Y_{3T}(\frac{i}{T}) - Y_{3T}(\frac{i-1}{T})$

$$\sum_{t=s+1}^{T} y_{t-s-1}^2(\epsilon_t\epsilon_{t-s}) = \sum_{i=1}^{T} X_{1T}^2(\frac{i-1}{T})\big[X_{3T}(\frac{i}{T}) - X_{3T}(\frac{i-1}{T})\big].$$

Let

$$G'_T = \sum_{i=1}^{T} X_{1T}^2(\frac{i-1}{T})\big[X_{3T}(\frac{i}{T}) - X_{3T}(\frac{i-1}{T})\big].$$

Lemma 8.

$$G'_T \xrightarrow{d} \int [\sigma\sqrt{T}W_1(t)]^2 d(\sqrt{T}\sigma^2)W_3(t)$$

Proof: *The Skorokhod representation theorem* plays the same role here for Y_{1T}, Y_{3T} as it did in Chapter 1. There exist sequences of random processes (Y^{1T}, Y^{3T}) converge almost surely to limit processes that are jointly distributed as $\big(\sqrt{T}\sigma W_1(t), \sqrt{T}\sigma^2 W_3(t)\big)$. If a random variable G'^T is constructed like G'_T in terms of the Skorokhod process $\big(Y^{1T}, Y^{3T}\big)$, then by the definition of the Ito integral

$$\left|G'^T - \int [\sigma\sqrt{T}W_1(t)]^2 d(\sqrt{T}\sigma^2)W_3(t)\right| \xrightarrow{p} 0.$$

Convergence in probability implies convergence in distribution, and G'_T and G'^T have the same distribution by construction. Thus,

Proposition 10.

$$\frac{1}{\sigma^4 T\sqrt{T}} \sum_{t=s+1} \epsilon_t\epsilon_{t-s} y_{t-s-1}^2 \xrightarrow{d} \int W_1(t)^2 dW_3(t)$$

(2.9) would be

$$\frac{1}{\sigma^4 T\sqrt{T}}\sum_{t=s+1}\epsilon_t\epsilon_{t-s}y_{t-s-1}^2 E\eta_t\eta_{t-s}\xrightarrow{d}\int W_1(t)^2 dW_3(t)E\eta_t\eta_{t-s}$$

The Limiting distribution for the second order

$$\sqrt{2}\int W_1^2(t)dW_2(t)E\eta_t^2 + L\int W_1(t)^2 dW_3(t)E\eta_t\eta_{t-s}.$$

2.5 Discussions and Concluding Remarks

This paper present a likelihood ratio test against Markov switching alternatives. This test is applicable across a broad spectrum of models commonly utilized in the fields of macroeconomics and finance. Its implementation is straightforward. We fully exploit the Markov property of the random coefficients. We derive the limiting distribution with Markov case.

2.6 Appendix

Proof of Proposition 1

Proof. Replace $E_\eta(expG_i|\mathcal{F}_{i-1})$ in c_i with its first 4 Taylor expansion in (2.2), we obtain c_i'. Because of proposition 4-8, all the orders higher than 4 converge to zero in probability. Thus we have

$$\sum_i c_i'/\sum_i c_i \xrightarrow{p} 1.$$

Rewrite (2.1) as $E\left[\prod_{i=1}^T(exp(G_i)\frac{exp(ci)}{exp(c_i')}exp(c_i'))\right] = 1$. Along the lines in the appendix of CHP, $\frac{exp(ci)}{exp(c_i')}$ can be taken out of the expectation on the some set with probability one.

We can approximate likelihood ratio with $\frac{1}{\prod_{i=1}^{T} exp(c_i')}$. In section (2.4) we show the limiting distribution of the 2nd order in Taylor expansion. $\sum c_i'$ will converge to constant in probability □

Bibliography

AIT-SAHALIA, Y. (2004): "Disentangling diffusion from jumps," *Journal of Financial Economics*, 74, 487–528.

ANDERSEN, T. G., T. BOLLERSLEV, F. X. DIEBOLD, AND P. LABYS (2003): "Modeling and forecasting realized volatility," *Econometrica*, 71, 579–625.

ANDREWS, D. W. (1993): "Tests for parameter instability and structural change with unknown change point," *Econometrica: Journal of the Econometric Society*, 821–856.

ANDREWS, D. W. AND W. PLOBERGER (1994): "Optimal tests when a nuisance parameter is present only under the alternative," *Econometrica: Journal of the Econometric Society*, 1383–1414.

ARSHANAPALLI, B. AND W. B. NELSON (2016): "Testing for stock price bubbles: a review of econometric tools," *The International Journal of Business and Finance Research*, 10, 29–42.

BEC, F., M. BEN SALEM, AND M. CARRASCO (2004): "Tests for unit-root versus threshold specification with an application to the purchasing power parity relationship," *Journal of Business & Economic Statistics*, 22, 382–395.

BILLINGSLEY, P. (2013): *Convergence of probability measures*, John Wiley & Sons.

BREEN, J. D. AND L. HU (2021): "The predictive content of oil price and volatility: New evidence on exchange rate forecasting," *Journal of International Financial Markets, Institutions and Money*, 75, 101454.

Caner, M. and B. E. Hansen (2001): "Threshold autoregression with a unit root," *Econometrica*, 69, 1555–1596.

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

Choi, I. (2001): "Unit root tests for panel data," *Journal of international money and Finance*, 20, 249–272.

Davidson, J. (1994): *Stochastic limit theory: An introduction for econometricians*, OUP Oxford.

——— (2006): "Asymptotic methods and functional central limit theorems," Palgrave.

Davidson, J. and R. M. De Jong (2000): "The functional central limit theorem and weak convergence to stochastic integrals II: fractionally integrated processes," *Econometric Theory*, 16, 643–666.

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

Dufour, J.-M. and M. L. King (1991): "Optimal invariant tests for the autocorrelation coefficient in linear regressions with stationary or nonstationary AR (1) errors," *Journal of Econometrics*, 47, 115–143.

Elliott, G., T. J. Rothenberg, and J. H. Stock (1992): "Efficient tests for an autoregressive unit root," .

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

Gong, F., R. S. Mariano, et al. (1997): *Testing under non-standard conditions in frequency domain: with applications to Markov regime switching models of exchange rates and the federal funds rate*, 23, Federal Reserve Bank of New York.

Granger, C. W. and N. R. Swanson (1997): "An introduction to stochastic unit-root processes," *Journal of Econometrics*, 80, 35–62.

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*, 357–384.

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

Heckman, J. J. and E. Leamer (2007): *Handbook of econometrics*, Elsevier.

Herrera, A. M., L. Hu, and D. Pastor (2018): "Forecasting crude oil price volatility," *International Journal of Forecasting*, 34, 622–635.

Hu, L. (2006): "Optimal Test for Stochastic Unit Root with Markov Switching," Working paper.

Hu, L. and W. Ploberger (2004): "A Complete Class of Tests When Nuisance Parameter is Present Only Under the Alternative," .

Hu, L. and Y. Shin (2008): "Optimal test for Markov switching GARCH models," *Studies in Nonlinear Dynamics & Econometrics*, 12.

——— (2014): "Testing for cointegration in markov switching error correction models," in *Essays in Honor of Peter CB Phillips*, Emerald Group Publishing Limited, 123–150.

Iglehart, D. L. (1974): "Functional central limit theorems for random walks conditioned to stay positive," *The Annals of Probability*, 2, 608–619.

Karatzas, I. and S. Shreve (2012): *Brownian motion and stochastic calculus*, vol. 113, Springer Science & Business Media.

Kim, C.-J. and C. R. Nelson (2001): "A Bayesian approach to testing for Markov-switching in univariate and dynamic factor models," *International Economic Review*, 42, 989–1013.

Lehmann, E. L., J. P. Romano, and G. Casella (1986): *Testing statistical hypotheses*, vol. 3, Springer.

Leybourne, S. J., B. P. McCabe, and A. R. Tremayne (1996): "Can economic time series be differenced to stationarity?" *Journal of Business & Economic Statistics*, 14, 435–446.

Lucas Jr, R. E. (1976): "Econometric policy evaluation: A critique," in *Carnegie-Rochester conference series on public policy*, North-Holland, vol. 1, 19–46.

Mills, T. C. and K. Patterson (2009): *Palgrave Handbook of Econometrics: Volume 2: Applied Econometrics*, vol. 2, Springer.

Perron, P. (1988): "Trends and random walks in macroeconomic time series: Further evidence from a new approach," *Journal of economic dynamics and control*, 12, 297–332.

Phillips, P. C., Y. Wu, and J. Yu (2011): "Explosive behavior in the 1990s Nasdaq: When did exuberance escalate asset values?" *International economic review*, 52, 201–226.

Phillips, P. C. and J. Yu (2011): "Dating the timeline of financial bubbles during the subprime crisis," *Quantitative Economics*, 2, 455–491.

Pollard, D. (2012): *Convergence of stochastic processes*, Springer Science & Business Media.

Shao, J. (2003): *Mathematical statistics*, Springer Science & Business Media.

Skorokhod, A. V. (1956): "Limit theorems for stochastic processes," *Theory of Probability & Its Applications*, 1, 261–290.