N° d'ordre : 76/2020-C/MT

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université des Sciences et de la Technologie « Houari Boumediene »
Faculté de Mathématiques
Département de Recherche Opérationnelle

THÈSE DE DOCTORAT

Présentée pour l'obtention du grade de Docteur

En : MATHÉMATIQUES

Spécialité : **Recherche Opérationnelle et Mathématiques Discrètes**

par

Abderaouf KHALFI

Titre

Étude de quelques modèles de séries chronologiques à seuil

Soutenue publiquement, le 26/12/2020 devant le jury composé de :

M. Mohamed BENTARZI,	Professeur à l'USTHB,	Président
M. Fayçal HAMDI,	Professeur à l'USTHB,	Directeur de thèse
Mme. Hafida GUERBYENNE,	Professeur à l'USTHB,	Examinatrice
M. Khaled KHALDI,	Professeur à l'UMB Boumerdès,	Examinateur
M. Tarek MEDKOUR,	Maître de Conférences /A à l'USTHB,	Examinateur
M. Saïd SOUAM,	Professeur à l'Université de Paris Nanterre,	Examinateur

Rssumé

Le but de cette thèse est de contribuer à l'analyse de la non linéarité en économétrie à travers l'étude de quelques modèles à seuil. Nos contributions concernent principalement deux volets de l'analyse des modèles économétriques. Dans un premier temps, nous nous intéressons au problème de la sélection du meilleur modèle dans la classe *Self-Exciting Threshold Autoregressive* ($SETAR$). Nous proposons une généralisation du critère de la densité prédictive (PDC), développé dans le cadre des modèles Autorégressifs (AR) linéaires par Djuric et Kay (1992), pour identifier conjointement le paramètre délai et les ordres Autorégressifs dans la classe $SETAR$. Dans un second temps, nous traitons le problème du changement de régime brutal imposé par la définition classique des modèles à seuil en adoptant le mécanisme de transition proposé par Li et al. (2015). Ce mécanisme est basé sur l'introduction d'intervalles au lieu de considérer des seuils ponctuels. Ces intervalles jouent le rôle de zones tampons permettant une transition lisse entre les régimes, i.e. une observation sera générée à partir de la même composante du mélange de celle qui la précède si sa valeur retardée se trouve dans la zone tampon. Nous nous intéressons dans notre travail à l'adoption de cette idée pour deux modélisations non linéaires. Tout d'abord, nous abordons le cas des modèles à volatilité stochastique à seuil en séries chronologiques (Breidt, 1996). Nous abordons ensuite, la modélisation des données de panel à seuil (Hansen, 1999a).

Mots clés : modèles à seuils, sélection des modèles, densité prédictive, asymétrie, données de panel, zone tampon, effet de levier, volatilité stochastique.

Abstract

The aim of this thesis is to contribute to the nonlinearity analysis in econometrics through the study of some threshold models. Our contributions mainly concern two aspects of the analysis of econometric models. First, we are interested in the problem of selecting the best model in the *Self-Exciting Threshold Autoregressive* ($SETAR$) class. We propose a generalization of the predictive density criterion (PDC), developed in the framework of linear Autoregressive (AR) models by Djuric and Kay (1992), to jointly identify the delay parameter and the Autoregressive orders in the $SETAR$ class. Secondly, we deal with the problem of sudden regime change imposed by the classical definition of threshold models by using the transition mechanism proposed by Li et al. (2015). This mechanism is based on the introduction of intervals instead of considering point thresholds. These intervals play the role of buffer zones allowing a smooth transition between regimes, i.e. an observation will be generated from the same component of the mixture of the one preceding it if its lagged value lies in the buffer zone. In our work, we adopted this idea for two nonlinear models dealing with two data structures. First, we extend the threshold stochastic volatility models (Breidt, 1996) in the case of time series. Secondly, we generalize the threshold panel data modelling (Hansen, 1999a).

Key words: threshlod models, model selection, predictive density, asymmetry, panel data, buffer zone, leverage effect, stochastic volatility.

Acknowledgments

In the name of Allah, the Most Gracious and the Most Merciful.

First and foremost, Alhamdulillah, praises and thanks to Allah, the Almighty, for his showers of blessings throughout my research work to complete the research successfully.

Although he was more than a thesis supervisor for me, and another manuscript like this will not be enough to express my gratitude to him, I would like to express my deepest thankfulness to Prof. Fayçal HAMDI for suggesting me this topic, without whose expert guidance, support, patience and encouragement none of this would have been possible. The extensive knowledge, vision, and creative thinking of Prof. Fayçal HAMDI have been the source of inspiration for me all along the way of this thesis. He has taught me the methodology to carry out the research and to present the research works as clearly as possible. It was a great privilege and honour to work and study under his guidance. I am extremely grateful for what he has offered me.

Profuse thanks go equally to the committee members: Prof. Mohamed BENTARZI for kindly accepting to be president of the jury, Prof. Hafida GUERBYENNE, Dr. Tarek MEDKOUR, Prof. Khaled KHALDI, Prof. Saïd SOUAM for kindly accepting to be examiners of the thesis. I would especially like to thank them for taking the time to evaluate this work and participate in my defense. Special thanks go to Prof. Saïd SOUAM for his efforts in developing part of the work presented in this thesis. Working alongside him was a wonderful experience for me that I hope that it will be repeated for future works.

To all those who have supported me from near or far. You will recognize yourselves. Thank you very much.

To *my mother, my* r_U*,*

To *my Mina, my* r_L*,*

You *are my thresholds.*

You *light my path in this nonlinear world,*

Your *existence is deterministic in this stochastic life.*

Contents

Introduction **1**

1 $SETAR$ model selection: A new Predictive Density based Criterion **9**

1.1 Introduction 9

1.2 Overview on the $SETAR$ models 12

1.2.1 Model definition 12

1.2.2 Stationnarity of the $SETAR$ models 13

1.2.3 Statistical inference for $SETAR$ models 14

1.2.3.1 Estimation 14

1.2.3.2 Hypothesis testing 15

1.3 Predictive Density Criterion for $SETAR$ model 16

1.3.1 Calculation of Predictive Density of $SETAR$ model 17

1.3.2 Predictive Density Criterion for $SETAR$ Model 22

1.4 Simulation study 26

1.5 Real Data examples 31

1.5.1 Annual Wolf's sunspot 31

1.5.2 GNP Growth 35

2 On Buffered Stochastic Volatility Model: Theory and Application **37**

2.1 Introduction 37

2.2 Buffered Threshold SV Models 41

2.3 Maximum Likelihood Estimation 42

2.4 Simulation Study 46

2.5 Empirical Study 47

3 Nonlinearity in Panel Data: A new Threshold Panel Data Model **54**

3.1 Introduction 54

3.2 Buffered Threshold Panel Data Model 56

3.3 Least Squares Estimation 58
3.4 Linearity test and determination of the number of regimes 62
3.5 Simulation study 65
3.6 Application 75

Conclusion **86**

Bibliography **89**

Introduction

"It's a nonlinear world". This is how Richard H. Enns chose to title his book (see Enns, 2010), which is a guide for mathematics students wanting to get into the nonlinearity mathematical treatment. Through this book, the author tried to illustrate how nonlinear dynamics plays a vital role in our world. This has been done by drawing examples from several disciplines such as pure mathematics, physics, chemistry, biology, engineering, economics, medicine, politics and sports. Indeed, nonlinearity presents a challenge both in terms of its form and the mathematical difficulty which it generates. Given that nonlinearity can have several sources leading to different mathematical formulations, this brings us to talk about several nonlinearities and not just one. In fact, while additivity and proportionality characterize linearity, nonlinearities appear as soon as one of these two characteristics is not verified. The major difficulty, when the dynamics of the studied variables, describing a system, breaks either the proportionality or additivity principles, is in the visualization of the interactions between the inputs of the system and its outputs. Thus, an irregularity in the dynamics can be observed, where small input changes of the studied system might lead to large output changes, and large input changes induces small output changes. Otherwise, we might not even be able to separate out parts of the input and output for analysis.

In statistics, the regression model is one of the most used tools in statistical modelling where dealing with the nonlinearity is a major concern. If the regression model describing the phenomenon under study cannot be adequately represented by a linear formulation, a nonlinear representation is therefore required. This opens a wide field of study in regression models and gives an opportunity to develop a range of models that allow a better understanding of the studied phenomena.

Our major interest through this thesis is to contribute, on different aspects, to the nonlinearity treatment in econometrics and regression analysis. Indeed, in recent years, nonlinearity, as well as nonstationarity, has been considered as one of the dominant properties in statistical analysis and modelling of economic data. In fact, many empirical studies have shown that the linear $ARMA$-type time-series modelling does not always make it possible to characterize the dynamics of some economic processes. The causes of this deficiency can be varied. However, two factors are mainly mentioned: the presence of asymmetry in the data and a temporal instability of the relationships induced by structural changes.

Many tracks have been explored for the nonlinearity modelling in this context. Among them, the regime change in the dynamics of the studied phenomena was often one of the features in which nonlinearity appears. Thus, one of those paths that has proven to be the most successful is that of regime switching models. This latter has the advantage of providing a realistic explanation of nonlinearity in observed economic data. The need to take into account nonlinearity and particularly, regime changes, tends to profoundly modify approaches to applied econometrics since its introduction. Quandt (1958) can be considered as the founder of the regime change representations. This seminal work has been completed by Goldfeld and Quandt (1973) and Baum and Petrie (1966) in which a complete formalism of such representation is provided.

Depending on the nature of the mechanism that governs the transition, the regime change can be deterministic or stochastic. Briefly, it is deterministic if it always occurs on a known date in advance. Otherwise, it is stochastic. Examples of this kind of models are those of Franses and Van Dijk (2000) and Beaudry and Koop (1993). The deterministic regime change is not addressed in this thesis and the stochastic regime change constitutes our general framework of study.

In stochastic regime change models, the main problem lies in the definition of the transition mechanism between the different regimes. This mechanism depends either on an unobservable or observable transition variable. This difference makes it possible to distinguish between the two main classes of models existing in econometrics and which first appeared in time series analysis framework. The first class contains Markov Switching (MS) models of Hamilton (1989) and the Mixture Autoregressive (MAR) models proposed by Wong and Li (2000). Note that in this class of models, the transition mechanism is based on an unobservable state variable which is assumed to follow a Markov chain for the MS model; while for the MAR, the transition is done according to an $i.i.d$ process. It should also be noted that for both models, at each period of time, the observation is generated by a component of the conditional distribution (by a regime) with some non-zero probability. The second class is that of Threshold Autoregressive (TAR) models introduced by Tong (1978), which historically constitute the first class of regime switching time series models. It is worth noting that the threshold idea was conceived in 1977 and it has been recorded in Tong's contributions to the discussion of a paper by Tony Lawrance in 1977 (Tong, 1977) and a *NATO ASI series* in 1978. Thereafter, the paper, *Threshold autoregression, limit cycles and cyclical data*, was accepted for reading to the Royal Statistical Society (RSS), on 19^{th} March 1980, where the most popular threshold model, namely the Self-Exciting Autoregressive ($SETAR$) model, has been presented.

The paper did not attain instant acclaim and according to Tong (2009), this is due to the fact that the idea was rather new, although its form was deceptively simple. In fact, there were still so many rough edges to smooth out (e.g., how to choose the threshold variables? Can the regime switching be continuous rather than discontinuous?), so many unresolved theoretical issues (e.g., what are the sampling properties of the parameters estimates? How to test for linearity within the context of Threshold models? How to obtain theoretical multi-step forecast formula?), and so many data-analytic techniques to develop.

Through his work, Howell Tong defined a new approach to model the regime change in time series and not only a restricted model. Indeed, there are many different, but conceptually equivalent, ways to express a threshold model, each having its advantages, depending on the context and purpose.

A general form of a TAR model is

$$Y_t = a_0^{(R_t)} + \sum_{i=1}^{p} a_i^{(R_t)} Y_{t-i} + b^{(R_t)} \varepsilon_t,$$

where ε_t's are $i.i.d$ $(0, \sigma^2)$ and R_t is an (indicator) time series taking values in the set $\{1, 2, \cdots, K\}$according to the values of the threshold variable. This indicator acts as the switching mechanism. Note that the TAR model can be easily extended to a Threshold Autoregressive Moving Average model ($TARMA$) by replacing $b^{(R_t)}\varepsilon_t$ by $\sum_{i=0}^{q} b_i^{(R_t)}\varepsilon_{t-j}$. Further extension to include some exogenous time series is obvious and may be indicated by the acronym $TARMAX$.

The basic idea of a threshold model is the piecewise linearization through the introduction of the indicator process, R_t, whose values depend on the thresholds. This idea has been called the *Threshold Principle* (TP) for easy reference, and this principle has been investigated in the econometrics literature. Indeed, the TP has been adapted for other data structures considered in the economic analysis such that the Panel Data.

Panel data analysis provides a better understanding of several economic phenomena (see e.g., Baltagi, 2008; Wooldridge, 2010; Hsiao, 2003, 2014). In fact, since the founding work of Balestra and Nerlove (1966), interests in the panel data econometrics has only increased over time. This obviously leads researchers to question the factors that caused this infatuation. Their double-dimensional structure provides richer information than that are usually available in cross-section or time series. More specifically, it becomes possible to study several individuals simultaneously over a period of time. The resulting gain is the ability to explain more complex individual behaviors and changing dynamics. From a technical point of view, using panel data has multiple advantages linked to the addition of an individual dimension, as Hsiao (2007) points out. Among the most important advantages is the increase in data variability by observing the variables over several individuals. The use of panel data also makes it possible to reduce the frequent problems in time series of colinearity between the explanatory variables thanks to the possibility of introducing inter-individual differences (see Pakes and Griliches; 1984). It is clear that these individual effects have also advantage of being able to highlight more, the heterogeneity between individuals.

Taking into account the individual and temporal dimensions makes it possible to bring out interesting characteristics which could not be distinguished by insufficient observations if only one dimension is considered. Moreover, by working simultaneously on several individuals instead of aggregating the risks of bias, it is possible to minimize these risks (see e.g., Fok et al., 2005; Lee, 2000).

All these advantages linked to the econometrics of panel data are not without compensation, which explains that the use of time series or cross-section models has not been abandoned. More precisely, the major difficulty lies in the modelling of heterogeneity between individuals. Indeed, the omission of individual differences can seriously bias the estimated coefficients as shown by Hsiao (2003).

Macroeconomic data processing has also generated an increase in the temporal dimension, which results in a transposition of the questions usually asked in time series, such as stationarity, nonlinearity or the temporal stability of relations. In this context, the use of regime change models, precisely, the TP in panel data makes it possible both to combine the advantages of working with panel data and, simultaneously, to deal with the problems of nonlinearity, heterogeneity and temporal instability of the relationship over time. More precisely, these models allow the existence of distinct individual dynamics which can moreover evolve over time while taking into account asymmetries. Therefore, threshold models are an interesting solution to the new challenges posed by the use of macroeconomic panel data. Being the first who thought about it, Hansen (1999*a*) deserves the merit for adapting the TP in panel data framework.

Another aspect of econometric modelling, where nonlinearity is an important issue, is the financial time series modelling. Indeed, the nonlinearity is also accepted, sometimes axiomatically, as a pattern and a challenge in the financial return time series analysis. In fact, several researchers have found that financial markets react nervously to political disorder and disagreement, economic crises, wars and natural disasters. Consequently, the financial series exhibit some statistical regularities, in addition to asymmetry, known as stylized facts such as the volatility clustering, the excess kurtosis, the leverage effect, the multimodality of the marginal distribution, and others. These stylized facts present many difficulties in the financial time series analysis. Indeed, the classical linear models of $ARMA$-type which assume a constant variance of errors, showed quickly their limits in the macroeconomic and financial time series modelling. Hence, these regularities, particularly, asymmetry in the volatility provide an ample room for employing and extending Tong's threshold model to this area. In fact, this development has been predicted by Tong (1990) who suggested the so-called *second-generation* models.

It should be noted that the TP was first applied in modelling the daily closing Hong Kong Hang Seng Index (HSI) by Li and Lam (1995), where a two-regime threshold autoregressive model has been proposed. An interesting finding was that during the study period, the autoregressive parameters are essentially positive when index value is positive while they are negative otherwise.

All the considerations presented above can constitute a satisfactory answer to the questions that followed Tong's discussion in front of the RSS. However, in practice, it was found that the TP, as defined by Tong, works usually well except around the boundaries between different regimes (Wu and Chen, 2007). It is partially due to the sudden jump of the theoretical structure with respect to the threshold variable, while a certain continuity may exist in the real examples. Although the smooth-transition threshold autoregressive ($STAR$) model (Chan and Tong, 1986; Van Dijk et al., 2002) can reduce the problem to some extent, it may not perform well for some complicated cases. This is because the

$STAR$ model is not always a piecewise linear. Moreover, Hamilton (1989), McCulloch and Tsay (1994) considered the discrete state MS autoregressive models for the financial and economic time series analysis. Its regime switching is completely controlled by a latent random variable. This latent indicator makes the regime switching mechanism more flexible meanwhile it also makes the interpretation more difficult in real application.

Wu and Chen (2007) proposed a Threshold variable Driven Switching Autoregressive ($TD-SAR$) model. The regime switching mechanism of the $TD-SAR$ model is jointly controlled by a latent indicator and some observable threshold variables. This latent indicator makes the regime switching mechanism more flexible. However, it is not always easy to find a physical interpretation of the fitted model.

Despite the existence of such useful models, there is still romm for formulating new classes of nonlinear econometric models, dealing with the problem discussed above. Following the successful foot steps of the classical TAR models, a new type of piecewise linear models that permit the delay of the regime switching has been introduced by Li et al. (2015). This new model is based on replacing the single threshold parameter in the classical formulation, by an interval which acts as a buffer zone for the regime switching, allowing a smooth transition.

From all what has been presented previously, the TP in the econometric modelling arouses our interest and opens several research perspectives for us to explore. The objective of this thesis is to contribute to the threshold models study in econometrics by focusing on two main aspects. First, we are interested in the identification problem in this class. This consists to find, among all candidate models, the best model able to adjust data. Saying the best requires the existence of a tool allowing to define an order relation between the models, which therefore allows to distinguish the best model from the others. Such a tool is commonly known as a *selection criterion*. Usually, model selection is based mainly on the estimation of the Kullback-Leibler divergence which has its roots in the information theory, based on the statistical entropy developed by the famous physicist Ludwig Boltzmann in thermodynamics field (see Tong, 1983). The merit of that goes to Akaike in his works (e.g., Akaike, 1969, 1973) who initiated the information criteria, which subsequently became the most widely used tools for model selection in different time series classes. An information criterion is basically founded on the log-likelihood which is a biased estimation of the Kullback-Leibler divergence and a penalty term which in turn constitutes the bias correction. The challenge was often in the improvement of this bias correction. Secondly, we study the problem of the abrupt regime change, imposed by the classical definition of threshold models. We, more precisely, we propose extensions of Li et al. (2015) buffering idea in order to benefit from it in different modelling situations, and to provide new econometric models allowing a better understanding of nonlinearities in the studied phenomena.

The outline of the thesis is as follow:

The first chapter deals with the problem of joint determination of the delay parameter and autoregressive orders of the $SETAR$ models. Specifically, we propose a variant of the Predictive Density Criterion (PDC) for the purpose of $SETAR$ model selection. The performance of this variant is evaluated by means of Monte Carlo experiments and two real data examples (the annual Wolf's sunspot number series and the GNP growth series of United States).

In the second chapter, we propose a new threshold stochastic volatility model called "Buffered Stochastic Volatility model". This model generalizes the threshold stochastic volatility model proposed by Breidt (1996), and which may explain more nonlinear characteristics observed in financial assets namely asymmetry and leverage effect. Keeping the piecewise linear structure in log-volatility as in Breidt (1996), we introduce a new flexible regime-switching mechanism, while avoiding a sudden jump in the log-volatility imposed by Breidt's formulation. We develop a Sequential Monte Carlo method to estimate the parameters of the model. Next, we analyze the performance of the proposed method through a simulation study. Finally, we apply our proposed model to fit the daily HON index series observed from June 11, 1990 to September 12, 2006.

In the last chapter, we introduce the "Buffered Threshold Panel Data model" which extends the classical Panel Threshold Regression model (PTR) proposed by Hansen (1999a). Our model provides also an alternative for the Panel Smooth Transition Regression model ($PSTR$) of Gonzalez et al. (2017) to take into account the problem of sudden jumps between regimes in Hansen's model. We first provide Ordinary Least Squares (OLS) estimation of the proposed model's parameters. Thereafter, we develop linearity tests and a bootstrap procedure for the likelihood ratio test. We show, through a simulation study, that the desirable consistency property of the OLS estimators is satisfied and that the proposed estimation and test procedures provide good results. We, finally, provide an empirical application in which we analyze the impact of natural resources dependence, and quality of institutions on economic growth for a panel of 19 countries, in the period $1996 - 2017$.

Finally, we comment globally our main findings by giving some concluding remarks regarding our contributions, and the research perspectives and challenges that may be considered for future works.

Chapter 1

$SETAR$ model selection: A new Predictive Density based Criterion

1.1 Introduction

Threshold models are one of the most popular and competitive formulations designed to take into account the nonlinearity in the statistical modelling of observed data. In fact, since its introduction by Tong (1978), the threshold-type models, in particular the Threshold Autoregressive Model (TAR), have been proven useful in many applications (e.g., Tong and Lim, 1980; Clements and Smith, 1997, 2000; Chen et al., 2011; Hansen, 2011). This class of models has shown a high ability to reflect many stylized facts of time series, as limit cycles, amplitude dependent frequencies and jump phenomena, facts that cannot be explained by linear models.

The genius of Tong's idea lies in the possibility of benefiting from the mathematical simplicity of linear models while taking into account the nonlinearity that exists. Indeed, Tong showed that a general nonlinear autoregressive model can be approximated through a piecewise linear structure and thresholds (e.g., Tong, 1983). This is done by assuming that data are generated from a mixture of linear models where the switching between model's components (regimes) is governed by the threshold variable. The basic formulation of threshold models is that studied by Tong and Lim (1980) where they introduce the Self-Exciting Threshold Autoregressive ($SETAR$) model and the process is "Self-Exciting" because the transition between regimes is stimulated by the process itself (i.e. a lagged (delayed) variable of the underlying process). Several papers investigated the probabilistic properties and statistical inference of this class of models (e.g., Tong and Lim, 1980; Chan and Tong, 1985; Petruccelli and Woolford, 1984; Chan et al., 1985; Tsay, 1989; Hili, 1992; Chan, 1993; Ling, 1999; Hansen, 1996, 1997, 1999*b*, 2000).

When the $SETAR$ model proves to be adequate to fit data, the main question is how to configure the model and how to find the best setting of its parameters. Indeed, to ensure the best use of the model, it is necessary to detect the exact number of regimes, know the thresholds values, fix the delay parameter and determine the order of each autoregressive component. The aim of this chapter is to contribute to this field of study. The $SETAR$ model parameters can be classified into two categories. The first one contains the number of regimes, the threshold values and the delay parameters and it can be called the "structural parameters category" which means that according to their values, observations will be assigned to the adequate regimes. The second category consists of autoregressive orders of the model's linear components. The selection of the first two elements in the structural parameters is not addressed in this chapter and in our opinion, this problem can be considered as open since there is no exact procedure that allows to get explicitly these parameters. Note that in some cases, their values are intimately linked to the specificity of the phenomenon, or left to the practitioner's skill. However, some suggestions for the estimation of the thresholds can be found in Tsay (1989) and Geweke and Terui (1993).

We focus in our contribution on joint identification of the autoregressive orders and the delay parameter of the $SETAR$ model. It can be pointed out that this problem did not receive much consideration in the literature. Intuitively, the piecewise linear structure of $SETAR$ model leads researchers to adopt procedures based on the choice of the best order for each linear component separately and take the sum over the number of regimes. In fact, Tong (1983) and Li (1988) suggest the use of the Akaike Information Criterion (AIC, Akaike, 1973) for the order selection without a theoretical justification. Unlike Tong's strategy, Wong and Li (1998) showed that their Corrected Akaike Information Criterion ($AICc$) is an unbiased estimate of the Kulback-Leibler divergence of the $SETAR$ model. This corrected version keeps the inconsistency property but it behaves better than AIC in the small sample situation. De Gooijer (2001) proposed three cross-validation criteria. By Monte Carlo study, he examined the performance of the suggested criteria compared to others and he found that the criterion inspired from the unbiased AIC (AIC_u, McQuarrie et al., 1997) can consistently outperform all other criteria when the sample size is moderate to large. Galeano and Peña (2007) extend the improved estimators of Kullback-Leibler divergence proposed by Hurvich et al. (1990) using the determinant term provided by Van der Leeuw (1994). They suggest some modifications to Wong and Li (1998) and De Gooijer (2001) procedures that improve the selection in $SETAR$ framework. More precisely, they modified the criteria AIC, $AICc$, BIC (Schwarz, 1978) and $AICu$ by adding a determinant term in each regime. Recently, some bootstrap-based model selection criteria have been proposed by adopting a Block bootstrap scheme (e.g., Hall,1985; Politis and Romano, 1994). Fenga and Politis (2015) studied the performance of a bootstrap version of the AIC. They compared the quality of the proposed criterion with its non-bootstrap counterpart. It should be noted that all the criteria we previously reported fall within the information theory framework and are structurally similar. They are based on the expected log-likelihood estimation of the candidate model and a penalty term. Therefore, these criteria necessarily require an estimation method for their evaluation.

Apart from the information theory based criteria, the predictive density function of each observation of the data has been proven useful to solve the order selection problem for some classes of time series models. Indeed, Djuric and Kay (1992) proposed a Bayesian criterion called the Predictive Density Criterion (PDC) for selecting the order of a linear autoregressive (AR) model. This criterion needs no estimation method for its evaluation. Bentarzi et al. (2008) extended the PDC to solve the order selection problem for periodic AR models. They also show, by a simulation study, the goodness of their criterion by comparison to some adapted information criteria for the periodic framework. This shows that the desirable consistency property of the PDC satisfied in the linear non-periodic case (see Djuric and Kay, 1992) is also met in their periodic case. Hence, the demonstrated effectiveness of this criterion as a model selection tool in these two settings motivates the exploration of an extended variant for non-linear modelling. For this purpose and once we get the expression of the predictive likelihood function, we propose an extended PDC version for the problem of joint determination of the delay parameter and autoregressive orders of $SETAR$ models.

The remainder of this chapter is organized as follows. In the next section, we give a brief overview on the $SETAR$ Model. We provide a precise definition and some probabilistic properties as well as the fundamentals of the inference theory on this class. More precisely, we give some elements of the hypothesis test relating to the detection of the absence of the threshold effects in data (linearity test) and unless, how many thresholds are controlling the dynamics (which means subsequently, the number of regimes which the phenomenon may follow). Our contribution will be presented in Section 1.3. Indeed, using an equivalent representation of the *SETAR* processes and a noninformative prior density, we determine an explicit expression of the predictive density function which allows us to derive our new criterion. The performance of the proposed criterion is discussed through a simulation study in as well as two real data examples in Section 1.4 and Section 1.5 respectively.

1.2 Overview on the *SETAR* models

1.2.1 Model definition

A *SETAR* model fitting a time series $\{y_t; t \in \mathbb{Z}\}$ can be described as a piecewise linear approximation to the general autoregressive model of order p such that

$$y_t = h\left(y_t, y_{t-1}, \cdots, y_{t-p}\right) + e_t,$$

where h denotes a nonlinear function and $\{e_t; t \in \mathbb{Z}\}$ is an innovation process. More specifically, the K-regimes *SETAR* model ($K \geq 2$) of orders p_1, p_2, $\cdots$, p_K, denoted $SETAR(K; p_1, p_2, \ldots, p_K)$, is given by

$$y_t = \sum_{k=1}^{K} \left(\phi_{k,0} + \sum_{j=1}^{p_k} \phi_{k,j} y_{t-j} + e_{k,t} \right) \mathbf{1}_{(r_{k-1} < y_{t-d} \leq r_k)}, \tag{1.1}$$

where $\{e_{k,t}; t \in \mathbb{Z}\}$ is a white noise process with a finite variance σ_k^2. $\mathbf{1}_A$ is the indicator function of the set A and $-\infty = r_0 < r_1 < \cdots < r_{K-1} < r_K = +\infty$ is a partition of the set of real numbers $\mathbb{R}$ in K parts which defines K regimes of the model. The number K of regimes and the orders p_1, p_2, $\cdots$, p_K of model (1.1) are positive integers. d is also a positive integer commonly referred to as the delay parameter. Note that each regime k ($k = 1, \ldots, K$) is a linear autoregressive of order p_k. This last may differ from a regime to another. This specific linear mechanism at any given point in time depends on the values taken by the switching variable y_{t-d}, where the observation y_t comes from the k-th regime if the threshold variable y_{t-d} falls within the k-th part of the partition.

We recall that the model is self-exciting because the transitions between regimes are a function of the past realizations of the process $\{y_t; t \in \mathbb{Z}\}$ itself. Here, the autoregressive parameters $\phi_{k,j}$ ($k = 1, \cdots, K$ and $j = 0, \cdots, p_k$) are supposed to be real and not necessarily the same across the regimes.

Let $e_{k,t} = \sigma_k \varepsilon_t$, where $\{\varepsilon_t; t \in \mathbb{Z}\}$ is a white noise process with zero mean and unit variance. Then, the $SETAR$ model (1.1) is expressed as follow

$$y_t = \sum_{k=1}^{K} \left(\phi_{k,0} + \sum_{j=1}^{p_k} \phi_{k,j} y_{t-j} + \sigma_k \varepsilon_t \right) \mathbf{1}_{(r_{k-1} < y_{t-d} \leq r_k)}.$$

1.2.2 Stationnarity of the $SETAR$ models

Many researchers have investigated some basic probabilistic properties of the $SETAR$ models. Petruccelli and Woolford (1984) have studied the ergodicity of the simplest two-regime TAR model. They obtained that model (1.1) , with $k = 2$, $p_1 = p_2 = d = 1$, $r_1 = 0$ and homogeneous $i.i.d.$ white noise with an everywhere positive probability density function, is ergodic if and only if $\phi_{1,1}$ and $\phi_{2,1}$ satisfy

$$\phi_{1,1} < 1, \ \phi_{2,1} < 1 \text{ and } \phi_{1,1}\phi_{2,1} < 1.$$

An extension of the previous result for the case of K-regimes can be found in Chan et al. (1985), Chan and Tong (1985), Hili (1992) and Ling (1999). The previous studies give only sufficient conditions for the ergodicity and/or stationarity. It remains difficult to find necessary and sufficient conditions for ergodicity and stationarity for general orders and/or number of regimes. For the K-regime $SETAR(K; p_1, p_2, ..., p_K)$ model, Chan and Tong (1985) show that the following condition,

$$\max_{k \in \{1,...,K\}} \sum_{i=1}^{p_k} |\phi_{k,i}| < 1,$$

is a sufficient condition for stationarity and ergodicity.

1.2.3 Statistical inference for *SETAR* models

1.2.3.1 Estimation

It is clear that the estimation of autoregressive coefficients in *SETAR* models requires the knowledge of the threshold, the delay parameter and the autoregressive orders. However, if the latters are determined, the Conditional Least Squares Estimation (*CLSE*) can be straightforwardly applied. Without loss of generality, we have opted to present a brief description of the estimation only for the $SETAR(2; p_1, p_2)$ with a single threshold r. Then, let $\phi_1 = (\phi_{1,0}, \phi_{1,1}, \cdots, \phi_{1,p_1})'$, $\phi_2 = (\phi_{2,0}, \phi_{2,1}, \cdots, \phi_{2,p_2})'$ where the prime denotes the transpose of the vector or matrix. $\widehat{\phi}_1(r), \widehat{\phi}_2(r), \widehat{\sigma}_1^2(r)$ and $\widehat{\sigma}_2^2(r)$ are the *CLSEs* based on r. The sum of squares error functions is defined as

$$S(\phi_1, \phi_2, r) = S_1(\phi_1, r) + S_2(\phi_2, r),$$

where

$$S_1(\phi_1, r) = \sum (y_t - \phi_{1,0} - \phi_{1,1} y_{t-1}, \cdots, \phi_{1,p_1} y_{t-p_1})^2 \mathbf{1}_{(y_{t-d} \leq r)},$$

and

$$S_2(\phi_2, r) = \sum (y_t - \phi_{2,0} - \phi_{2,1} y_{t-1}, \cdots, \phi_{2,p_1} y_{t-p_1})^2 \mathbf{1}_{(y_{t-d} > r)}.$$

All the summations are from $p+1$ to n with $p = \max(p_1, p_2, d)$ and n is the sample size. The *CLSEs* of $\widehat{\phi}_1(\widehat{r})$, $\widehat{\phi}_2(\widehat{r})$, $\widehat{\sigma}_1^2(\widehat{r})$, $\widehat{\sigma}_2^2(\widehat{r})$ and $\widehat{r}$ are the values which minimize $S(\phi_1, \phi_2, r)$ where $\widehat{\sigma}_k^2(\widehat{r}) = \frac{S_k(\widehat{\phi}_k(\widehat{r}), \widehat{r})}{n_k}$ for $k = 1, 2$, $n_1 = \sum \mathbf{1}_{(y_{t-d} \leq r)}$ and $n_2 = n - n_1$. Note that the asymptotic behavior of the *CLSE* in the *SETAR* framework was addressed with details by several authors (see e.g. Chan, 1993; Wong and Li, 1998; Hansen, 1997). In fact, Wong and LI(1998) derive in the following theorem some asymptotic properties of the *CLSE* for the *SETAR* model.

Theorem 1. *Under the four assumptions given in Chan (*1993*, pp.* $522 - 523$*)*

- $\widehat{r} = r + O_p(1/n)$*,*
- $\widehat{\phi}_1(\widehat{r})$*,* $\widehat{\phi}_2(\widehat{r})$*,*$\widehat{\sigma}_1^2(\widehat{r})$ *and* $\widehat{\sigma}_2^2(\widehat{r})$ *are asymptotically independent of* $\widehat{r}$ *and their asymptotic distributions are the same as those for the case in which* r *is known,*

- *Given* r,

$$\sqrt{n}\begin{pmatrix}\widehat{\phi}_1(r)\\ \widehat{\phi}_2(r)\end{pmatrix} \sim \mathcal{N}_{p_1+p_2+2}\left\{\begin{pmatrix}\phi_1(r)\\ \phi_2(r)\end{pmatrix}, \begin{pmatrix}\sigma_1^2V_1^{-1}(r) & 0\\ 0 & \sigma_2^2V_2^{-1}(r)\end{pmatrix}\right\}$$

  where $\mathcal{N}_{p_1+p_2+2}(.,.)$ *denotes the* $\{p_1+p_2+2\}$*-th dimensional normal distribution and* $V_1 = \mathbb{E}(U_t(r)U_t(r)')$, $V_2 = \mathbb{E}(W_t(r)W_t(r)')$ *such that* $U_t = \mathbf{1}_{(y_{t-d}\leq r)}(1, y_{t-1}, y_{t-2}, ..., y_{t-p_1})$ *and* $W_t = \mathbf{1}_{(y_{t-d}>r)}(1, y_{t-1}, y_{t-2}, ..., y_{t-p_2})$,

- *Given* r, $n_k\widehat{\sigma}_k^2(r) \sim \sigma_k^2(r)\chi^2_{n_k-(p_k+1)}$, $k = 1, 2$ *where* χ^2_h *denotes the* χ^2 *distribution with* h *degrees of freedom.*

1.2.3.2 Hypothesis testing

The question if the nonlinear specification is adequate rather than the linear one is quite central in the analysis of *SETAR* model. Since the *SETAR* model with a single regime is a linear *AR* model, this question can be compactly addressed in a more general context as the problem of determining the number of regimes that the model contains. Therefore, the development of statistical tests to provide a decision support tool for this question has proved essential.

The subject of testing the nonlinearity in the context of threshold models has been studied by Tsay (1989), Chan (1990, 1991), Chan and Tong (1990), Hansen (1996, 1997), and Caner and Hansen (1998). The testing problem is algebraically quite similar to the issue of testing for structural change of unknown timing, which dates back to Quandt (1960). Tests for single structural change in stationary models has been studied by Andrews (1993, 2003) and Andrews and Ploberger (1994), and in the context of non-stationary models by Hansen (1999*b*, 2002), Seo (1998), and Kuo (1998). Tests for multiple structural change have been studied by Bai (1997) and Bai and Perron (1998), and similar techniques have been applied to threshold models by Hansen (1999*b*). Both testing problems fall in the class of tests in the presence of unidentified nuisance parameters, which have been studied by Davies (1977, 1987), Andrews and Ploberger (1994), Hansen (1996), and Stinchcomb and White (1998).

Thus, testing for linearity (within the *SETAR* class of models) consists to a test of the null hypothesis of a linear *AR* model against the alternative of *SETAR* model with K regimes, for some $K > 1$. Similarly, it can be tested the null hypothesis of the two-regime

$SETAR$ model against the alternative of a K-regime $SETAR$, for some $K > 2$.

In more general cases, the natural Least Squares test of the null hypothesis of j-regime $SETAR$ against k-regime $SETAR$, with $k > j$, consists to reject the null for large values of

$$F_{jk} = n\left(\frac{S_j - S_k}{S_k}\right),$$

where S_j and S_k are the sum of squared residuals under the null and the alternative hypothesis respectively.

This is the Likelihood Ratio (LR) test when the errors are independent and normally distributed. It is also the conventional F (or Wald) test, and it is equivalent to the conventional Lagrange multiplier (or score) test. It has been pointed that this test is likely to have excellent power relative to alternative tests. The primary complication is that this testing problem is non-standard, due to the presence of a nuisance parameter which is only defined under the alternative, so the asymptotic distribution of the test statistics is also non-standard. However, simulations methods have been proven efficient to approximate the distribution of the test statistics. In fact, Hansen (1996, 1997, 1999a, 2000) proposed bootstrap approach to simulate the asymptotic distribution of the LR statistic.

1.3 Predictive Density Criterion for $SETAR$ model

It is interesting to exploit the piecewise structure of the $SETAR$ model to extend the Predictive Density Criterion (PDC) considered by Djuric and Kay (1992) for linear autoregressive model to this nonlinear model. However, the selection of the best $SETAR$ model using a predictive density based approach should be done with respect to its features and needs more careful treatment, since the model identification problem in addition to the order selection, consists on determining the delay parameter, thresholds values and number of regimes. To derive the PDC version for the selection of $SETAR$ model, we must explicitly determine the expression of the predictive density function of each observation of the data. Note that the number of regimes K and the threshold parameters $(r_1, ..., r_{K-1})$ are assumed to be known in this section and we will concentrate only on the delay parameter and autoregressive orders selection problem. We also assume that the $\{\varepsilon_t; t \in \mathbb{Z}\}$ are normally distributed and we take the initial conditions $(y_1, y_2, \ldots, y_{m-1})$

as given, where $m = \max\{\max\{p_1, ..., p_K\}, d\} + 1$.

1.3.1 Calculation of Predictive Density of *SETAR* model

Let be given a realization $y = (y_1, y_2, \ldots, y_n)$ generated from a stationary $SETAR(K; p_1, p_2, \ldots, p_K)$ defined by (1.1). For notational simplicity, let us write $\mathbf{p} = (p_1, ..., p_K)$, $\alpha = (\mathbf{p}, d)$ and $\mathbf{y}_{r,s} = (y_r, y_{r+1}, \cdots, y_s)'$, for any positive integers r and s such that $r < s$, where the prime denotes the transpose. Using these notations, the predictive density of y_t conditioning on the information available up to time $t-1$ can be expressed as follows

$$f(y_t|\mathbf{y}_{1,t-1}, \alpha) = \int_\sigma \int_\Phi f(y_t|\mathbf{y}_{1,t-1}, \alpha, \sigma, \Phi) f(\sigma, \Phi|\mathbf{y}_{1,t-1}, \alpha) d\Phi d\sigma, \tag{1.2}$$

where $\Phi = \left(\Phi'_{1,p_1}, \cdots, \Phi'_{K,p_K}\right)'$, $\Phi_{k,p_k} = (\phi_{k,0}, \phi_{k,1}, \cdots, \phi_{k,p_k})'$, for $k = 1, ..., K$, $\sigma = (\sigma_1, \cdots, \sigma_K)'$, and the function $f(\sigma, \Phi \mid y_{1,t-1}, \alpha)$ is the *a posteriori* density of the parameter vector $(\sigma', \Phi')'$ after observing $\mathbf{y}_{1,t-1}$. The essential problem in (1.2) is to find the form of $f(\sigma, \Phi|\mathbf{y}_{1,t-1}, \alpha)$. To do this, we can use the Bayes' rule which gives

$$f(\sigma, \Phi|\mathbf{y}_{1,t-1}, \alpha) \propto f(\mathbf{y}_{m,t-1}|\mathbf{y}_{1,m-1}, \alpha, \sigma, \Phi) f(\sigma, \Phi|\alpha),$$

where

$$\begin{aligned} &f(\mathbf{y}_{m,t-1}|\mathbf{y}_{1,m-1}, \alpha, \sigma, \Phi) \\ &\quad = \prod_{i=m}^{t-1} \left(\frac{1}{\sqrt{2\pi\sigma^2_{R_i}}} \exp\left\{ -\frac{1}{2\sigma^2_{R_i}} \left(y_i - \phi_{R_i,0} - \sum_{j=1}^{p_{R_i}} \phi_{R_i,j} y_{i-j} \right)^2 \right\} \right), \end{aligned}$$

and $\{R_i, i \in \mathbb{Z}\}$ is a sequence of discrete random variables taking values in the set $\{1, 2, \ldots, K\}$ such as $R_i = k$, if $r_{k-1} < y_{i-d} \leq r_k$.

Now, similarly to Djuric and Kay (1992) and Bentarzi et al. (2008), we assume that the prior distribution of the parameter vector $(\sigma', \Phi')'$ is given by the family of noninformative prior joint densities

$$f(\sigma, \Phi|\alpha) \propto \prod_{k=1}^{K} \frac{1}{\sigma_k}, \; \Phi_{k,p_k} \in \mathbb{R}^{p_k+1} \text{ and } \sigma_k \in \mathbb{R}^*_+, \text{ for } k = 1, ..., K.$$

Hence, one can rewrite (1.2) as follows

$$\begin{aligned} f\left(y_t \middle| \mathbf{y}_{1,t-1}, \alpha\right) &\propto \int_\sigma \int_\Phi f\left(y_t \middle| \mathbf{y}_{1,t-1}, \alpha, \sigma, \Phi\right) f\left(\mathbf{y}_{m,t-1} \middle| \mathbf{y}_{1,m-1}, \alpha, \sigma, \Phi\right) f\left(\sigma, \Phi \middle| \alpha\right) d\Phi d\sigma \\ &\propto \int_\sigma \int_\Phi f(\mathbf{y}_{m,t} \left|\mathbf{y}_{1,m-1}, \alpha, \sigma, \Phi\right.) f(\sigma, \Phi \left|\alpha\right.) d\Phi d\sigma \\ &\propto \int_\sigma \int_\Phi \left(\prod_{i=m}^{t} \frac{1}{\sqrt{2\pi\sigma_{R_i}^2}} \exp\left\{ - \frac{\left(y_i - \phi_{R_i,0} - \sum_{j=1}^{p_{R_i}} \phi_{R_i,j} y_{i-j}\right)^2}{2\sigma_{R_i}^2} \right\} \right) \left(\prod_{k=1}^{K} \frac{1}{\sigma_k} \right) d\Phi d\sigma. \end{aligned}$$

It may be pointed out that contrary to the linear case (see Djuric and Kay (1992) for *AR* model and Bentarzi et al. (2008) for periodic *AR* model), the previous integral is difficult to evaluate explicitly. In fact, since the switching between regimes is stochastic and irregular, we should specify the origin of the observation for each instant. To solve this problem, we use the arranged autoregression representation (Tsay, 1989). This representation provides an equivalent expression of the *SETAR* model (1.1). Indeed, Tsay (1989) suggests to arrange the equations in (1.1) for $t = m, ..., n$, such that the equations are sorted according to the threshold variable y_{t-d} which may take any value in $E_{n,d} = \{y_m, y_{m+1}, ..., y_{n-d}\}$. It may be noted that the predictive density given by (1.2) depends on the information available up to time $t-1$, therefore instead of using the set $E_{n,d}$, we must use the set $E_{t,d} = \{y_m, y_{m+1}, ..., y_{t-d}\}$.

Let $[t,i]$ be the time index in the original sample such that $y_{[t,i]}$ is the i-th smallest value in $E_{t,d}$ (see Figure 1.1). For example, if y_{18} is the smallest value in $E_{t,d}$, then $[t,1] = 18$; if y_6 is the second smallest value in $E_{t,d}$, then $[t,2] = 6$, etc. Let also $n_{t,k} = s_{t,k} - s_{t,k-1}$ denote the number of observations in $\{y_m, \ldots, y_t\}$ that follow the k-th component autoregressive of *SETAR* model.

Therefore, the model (1.1) can be written as a rearranged *SETAR* model

$$y_{[t,i]+d} = \phi_{k,0} + \sum_{j=1}^{p_k} \phi_{k,j} y_{[t,i]+d-j} + \sigma_k \varepsilon_{[t,i]+d}, \text{ if } s_{t,k-1} < i \leq s_{t,k}, \tag{1.3}$$

where $s_{t,0} = 0$ and $s_{t,k} = \min_{1 \leq i \leq t-m} \left\{i, \text{ such that } r_k < y_{[t,i+1]}\right\}$, for $k = 1, \cdots, K$. Hence, if the original time series is generated by a $SETAR\,(3; p_1, p_2, p_3)$ model and there are $s_{t,1}$ values in $E_{t,d}$ that are smaller than the threshold r_1, then the first $s_{t,1}$ equations in (1.3) correspond to the first regime. If there are also $n_{t,2} = s_{t,2} - s_{t,1}$ values in $E_{t,d}$ that lie between r_1 and r_2, then the $n_{t,2}$ following equations in (1.3) correspond to the second regime, and the remaining equations correspond to the third regime.

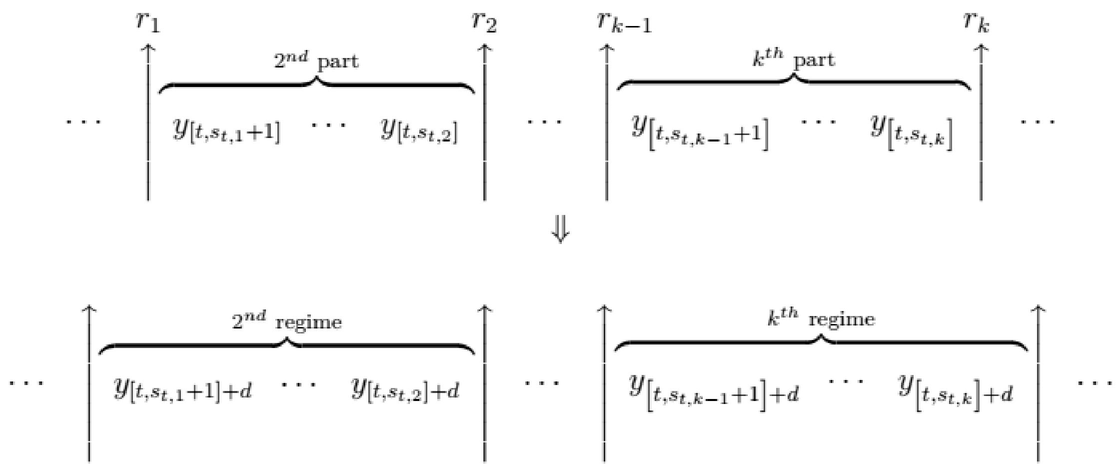

Figure 1.1: The arranged autoregression representation diagram

Using rearranged representation, the predictive density can be rewritten as follows

$$f\left(y_t \middle| \mathbf{y}_{1,t-1}, \alpha\right) = \frac{C}{(2\pi)^{\frac{t-m+1}{2}}} \times \int_\sigma \int_\Phi \prod_{k=1}^{K} \left(\frac{\exp\left\{ - \sum_{i=s_{t,k-1}+1}^{s_{t,k}} \left(y_{[t,i]+d} - \phi_{k,0} - \sum_{j=1}^{p_k} \phi_{k,j} y_{[t,i]+d-j} \right)^2 \Big/ 2\sigma_k^2 \right\}}{\sigma_k^{s_{t,k}-s_{t,k-1}+1}} \right) d\Phi d\sigma, \tag{1.4}$$

where C is the normalizing constant to be calculated later.

Now, considering the matrix $H_\alpha(l,k)$ and the vector $Y_\alpha(l,k)$ of dimensions $n_{l,k} \times (p_k+1)$ and $n_{l,k} \times 1$, respectively and defined by

$$H_\alpha(l,k) = \begin{bmatrix} 1 & y_{[l,s_{l,k-1}+1]+d-1} & y_{[l,s_{l,k-1}+1]+d-2} & \cdots & y_{[l,s_{l,k-1}+1]+d-p_k} \\ 1 & y_{[l,s_{l,k-1}+2]+d-1} & y_{[l,s_{l,k-1}+2]+d-2} & \cdots & y_{[t,s_{l,k-1}+2]+d-p_k} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & y_{[l,s_{l,k}]+d-1} & y_{[l,s_{l,k}]+d-2} & \cdots & y_{[l,s_{l,k}]+d-p_k} \end{bmatrix},$$

and

$$Y_\alpha(l,k) = \left(y_{[l,s_{l,k-1}+1]+d}, y_{[l,s_{l,k-1}+2]+d}, \ldots, y_{[l,s_{l,k}]+d}\right)'.$$

The following result gives closed-form expressions of the predictive density function of y_t conditioning on the information available up to time $t-1$. This result extends the one of Djuric and Kay (1992) for linear AR model to the $SETAR$ framework.

Proposition 2 (**Hamdi and Khalfi,** 2019)**.** *If* $n_{t-1,k} > p_k + 1$*, for all* $k \in \{1, ..., K\}$*, then the predictive density function of* $SETAR(K; p_1, \ldots, p_K)$ *defined in* (1.1) *is given by*

$$f\left(y_t \middle| \mathbf{y}_{1,t-1}, \alpha\right) = \frac{1}{\sqrt{2\pi}} \prod_{k=1}^{K} g_{p_k,d}\left(t, k\right), \tag{1.5}$$

where

$$g_{p_k,d}\left(t,k\right) = \frac{\left|\Sigma_\alpha\left(t-1,k\right)\right|^{\frac{1}{2}} \Gamma\left(\frac{n_{t,k}-p_k-1}{2}\right) \left(\frac{Y_\alpha'(t-1,k)P_\alpha(t-1,k)Y_\alpha(t-1,k)}{2}\right)^{\frac{n_{t-1,k}-p_k-1}{2}}}{\left|\Sigma_\alpha\left(t,k\right)\right|^{\frac{1}{2}} \Gamma\left(\frac{n_{t-1,k}-p_k-1}{2}\right) \left(\frac{Y_\alpha'(t,k)P_\alpha(t,k)Y_\alpha(t,k)}{2}\right)^{\frac{n_{t,k}-p_k-1}{2}}},$$

and

$$\begin{aligned} \Sigma_\alpha\left(t,k\right) &= H_\alpha'\left(t,k\right) H_\alpha\left(t,k\right), \\ P_\alpha\left(t,k\right) &= \mathbf{I}_{n_{t,k}} - H_\alpha\left(t,k\right) \Sigma_\alpha^{-1}\left(t,k\right) H_\alpha'\left(t,k\right). \end{aligned}$$

Here $\mathbf{I}_{n_{t,k}}$ *is the identity matrix of dimension* $n_{t,k}$ *and* $\Gamma\left(\cdot\right)$ *is a gamma function.*

Proof. As (1.3) is simply a linear regression, the expression may be rewritten more compactly in matrix notation by using the notation adopted in this section. Indeed, we can see that

$$\begin{aligned} &\exp\left\{-\frac{1}{2\sigma_k^2} \sum_{i=s_{t,k-1}+1}^{s_{t,k}} \left(y_{[t,i]+d} - \phi_{k,0} - \sum_{j=1}^{p_k} \phi_{k,j} y_{[t,i]+d-j}\right)^2\right\} \\ &\quad = \exp\left\{-\frac{Y_\alpha'(t,k)P_\alpha(t,k)Y_\alpha(t,k) + \left[\Phi_{k,p_k} - A_\alpha(t,k)\right]'\Sigma_\alpha(t,k)\left[\Phi_{k,p_k} - A_\alpha(t,k)\right]}{2\sigma_k^2}\right\}, \end{aligned}$$

where $A_\alpha\left(t,k\right) = \Sigma_\alpha^{-1}\left(t,k\right) H_\alpha'\left(t,k\right) Y_\alpha\left(t,k\right)$. Therefore, equation (1.4) may then be written as

$$\begin{aligned} f\left(y_t \middle| \mathbf{y}_{1,t-1}, \alpha\right) &= \frac{C}{(2\pi)^{\frac{t-m+1}{2}}} \\ &\times \prod_{k=1}^{K} \left[\int_{\sigma_k} \left(\frac{1}{\sigma_k^{n_{t,k}+1}} \exp\left\{-\frac{Y_\alpha'(t,k)P_\alpha(t,k)Y_\alpha(t,k)}{2\sigma_k^2}\right\}\right.\right. \\ &\times \left.\left. \int_{\Phi_{k,p_k}} \exp\left\{-\frac{\left[\Phi_{k,p_k} - A_\alpha(t,k)\right]'\Sigma_\alpha(t,k)\left[\Phi_{k,p_k} - A_\alpha(t,k)\right]}{2\sigma_k^2}\right\} d\Phi_{k,p_k}\right) d\sigma_k\right]. \end{aligned}$$

Now, by integrating out the vector parameter Φ_{k,p_k}, it can be easily shown that

$$\int_{\mathbb{R}^{p_k}} \exp\left\{-\frac{\left[\Phi_{k,p_k}-A_\alpha(t,k)\right]'\Sigma_\alpha(t,k)\left[\Phi_{k,p_k}-A_\alpha(t,k)\right]}{2\sigma_k^2}\right\} d\Phi_{k,p_k} = \frac{(2\pi)^{\frac{p_k+1}{2}}\sigma^{p_k+1}}{|\Sigma_\alpha(t,k)|^{\frac{1}{2}}},$$

hence

$$\begin{aligned} f\left(y_t|\,\mathbf{y}_{1,t-1},\alpha\right) &= \frac{C}{(2\pi)^{\frac{t-m+1}{2}}}\prod_{k=1}^{K}\left(\frac{(2\pi)^{\frac{p_k+1}{2}}}{|\Sigma_\alpha(t,k)|^{\frac{1}{2}}}\right) \\ &\quad\times\left(\int_0^{+\infty}\left(\sigma_k^{-n_{t,k}+p_k}\exp\left\{-\frac{Y_\alpha'(t,k)P_\alpha(t,k)Y_\alpha(t,k)}{2\sigma_k^2}\right\}\right)d\sigma_k\right) \\ &= \frac{C}{(2\pi)^{\frac{t-m+1}{2}}}\prod_{k=1}^{K}\left(\frac{(2\pi)^{\frac{p_k+1}{2}}\Gamma\left(\frac{n_{t,k}-p_k-1}{2}\right)}{2|\Sigma_\alpha(t,k)|^{\frac{1}{2}}\left(\frac{Y_\alpha'(t,k)P_\alpha(t,k)Y_\alpha(t,k)}{2}\right)^{\frac{n_{t,k}-p_k-1}{2}}}\right). \end{aligned} \tag{1.6}$$

Since the normalizing constant is given by

$$\begin{aligned} C^{-1} &= \int_\sigma\int_\Phi f(\mathbf{y}_{m,t-1}\,|\mathbf{y}_{1,m-1},\alpha,\sigma,\Phi\,)f(\sigma,\Phi\,|\alpha)d\Phi d\sigma \\ &= \frac{1}{(2\pi)^{\frac{t-m}{2}}} \\ &\times\int_\sigma\int_\Phi\prod_{k=1}^{K}\left[\frac{1}{\sigma_k^{n_{t-1,k}+1}}\exp\left\{\sum_{i=s_{t-1,k-1}+1}^{s_{t-1,k}}\frac{-\left(y_{[t-1,i]+d}-\phi_{k,0}-\sum_{j=1}^{p_k}\phi_{k,j}y_{[t-1,i]+d-j}\right)^2}{2\sigma_k^2}\right\}\right]d\Phi d\sigma. \end{aligned}$$

Finally, we can show that the density $f\left(y_t|\,\mathbf{y}_{1,t-1},\alpha\right)$ is well defined and the normalizing constant is equal to

$$C = (2\pi)^{\frac{t-m}{2}}\prod_{k=1}^{K}\frac{2\left|\Sigma_\alpha\left(t-1,k\right)\right|^{\frac{1}{2}}}{(2\pi)^{\frac{p_k+1}{2}}}\frac{\left(\frac{Y_\alpha'(t-1,k)P_\alpha(t-1,k)Y_\alpha(t-1,k)}{2}\right)^{\frac{n_{t-1,k}-p_k-1}{2}}}{\Gamma\left(\frac{n_{t-1,k}-p_k-1}{2}\right)}.$$

Replacing C by this last expression in (1.6), we obtain the predictive density given by (1.5). $\square$

Remark 3. *When* $\phi_{k,0}=0$, *for all* $k\in\{1,...,K\}$, *we can consider* $\Phi_{k,p_k}=(\phi_{k,1},\cdots,\phi_{k,p_k})'$

as just a p_k-vector. Then, if $n_{t-1,k} > p_k$ the expression of $g_{p_k,d}\left(t,k\right)$ is replaced by

$$g_{p_k,d}\left(t,k\right)=\begin{cases}\dfrac{\Gamma\left(\frac{n_{t,k}}{2}\right)}{\Gamma\left(\frac{n_{t-1,k}}{2}\right)}\dfrac{\left(\frac{Y_{\alpha}'(t-1,k)Y_{\alpha}(t-1,k)}{2}\right)^{\frac{n_{t-1,k}-p_k}{2}}}{\left(\frac{Y_{\alpha}'(t,k)Y_{\alpha}(t,k)}{2}\right)^{\frac{n_{t,k}-p_k}{2}}}, & \textit{if } p_k=0,\\ \dfrac{\left|\Sigma_{\alpha}(t-1,k)\right|^{\frac{1}{2}}}{\left|\Sigma_{\alpha}(t,k)\right|^{\frac{1}{2}}}\dfrac{\Gamma\left(\frac{n_{t,k}-p_k}{2}\right)}{\Gamma\left(\frac{n_{t-1,k}-p_k}{2}\right)}\dfrac{\left(\frac{Y_{\alpha}'(t-1,k)P_{\alpha}(t-1,k)Y_{\alpha}(t-1,k)}{2}\right)^{\frac{n_{t-1,k}-p_k}{2}}}{\left(\frac{Y_{\alpha}'(t,k)P_{\alpha}(t,k)Y_{\alpha}(t,k)}{2}\right)^{\frac{n_{t,k}-p_k}{2}}}, & \textit{otherwise,}\end{cases}$$

where

$$H_{\alpha}\left(l,k\right)=\begin{bmatrix} y_{\left[l,s_{l,k-1}+1\right]+d-1} & y_{\left[l,s_{l,k-1}+1\right]+d-2} & \cdots & y_{\left[l,s_{l,k-1}+1\right]+d-p_k}\\ y_{\left[l,s_{l,k-1}+2\right]+d-1} & y_{\left[l,s_{l,k-1}+2\right]+d-2} & \cdots & y_{\left[t,s_{l,k-1}+2\right]+d-p_k}\\ \vdots & \vdots & & \vdots\\ y_{\left[l,s_{l,k}\right]+d-1} & y_{\left[l,s_{l,k}\right]+d-2} & \cdots & y_{\left[l,s_{l,k}\right]+d-p_k}\end{bmatrix},$$

and $\Sigma_{\alpha}\left(t,k\right)$, $A_{\alpha}\left(t,k\right)$ and $P_{\alpha}\left(t,k\right)$ are defined as in Proposition 1*.*

*When $K=1$, this predictive density function coincides with the one obtained by Djuric and Kay (*1992*, p.* 1830*) for the case of a linear AR model.*

1.3.2 Predictive Density Criterion for *SETAR* Model

We turn now to the *PDC* criterion which was first proposed by Djuric and Kay (1992) for the linear *AR* optimal order selection problem. Optimality was defined according to the overall probability of selection error. This error is minimized when the selection is carried out according to the predictive densities models. Recall that the form of its criterion is given as follows

$$PDC(p)=-\sum_{t=2}^{n}\log f(\left.y_t\right|\mathbf{y}_{1,t-1},p),\ p\in\left\{0,...,P\right\},$$

where $f(\left.y_t\right|y_{1,t-1},p)$ is the predictive density of y_t conditioning on the information available up to time $t-1$ and the autoregressive order p of the candidate model. Here P is a large enough positive integer. Note that in the *AR* framework, we can write the predictive density of y_t only if there are enough past observations, i.e. if $t>2p+1$. Considering the different possible values of p, it is clear that if we just consider the times t for which the previous constraint is satisfied, we cannot compare the candidate models by exploiting the same amount of data samples. However, in order to use the same number of observations for comparing different models, Djuric and Kay (1992) have designed an initialization strategy for $t \leq 2p+1$ to allow the derivation of the predictive densities of y_t without violating the constraint $t > 2p+1$. They have replaced the quantities $PDC(p)$ by

$$PDC^*(p) = -\sum_{i=1}^{p} \sum_{t=2(i-1)+2}^{2i+1} \log f(y_t | \mathbf{y}_{1,t-1}, i-1) - \sum_{t=2p+2}^{n} \log f(y_t | \mathbf{y}_{1,t-1}, p),$$

which are all computed on the basis of the same number of observations $n-1$.

The same problem is encountered by Bentarzi et al (2008) where they propose a similar procedure to Djuric and Kay (1992) strategy to derive the Predictive Density Criterion for a univariate Periodic autoregressive model (PAR) with period S. Note that in Bentarzi et al. (2008), the order selection consists in finding the non negative integer p such as $p = \max_{i \in \{1,2,\cdots,S\}} \{p_i\}$ which minimize

$$BGH_p = -\sum_{r=0}^{p-1} \sum_{t=2(r+1)S}^{2(r+2)S-1} \log f(y_t | \mathbf{y}_{1,t-1}, r) - \sum_{t=2(p+1)S}^{n} \log f(y_t | \mathbf{y}_{1,t-1}, p).$$

These last quantities are all computed on the basis of the same number of observations, $(N-2S+1)$. This allows an adequate comparison between all candidate models and the optimal order is p^* such as

$$p^* = \arg\left\{ \min_{p \in \{0,2,\cdots,P\}} BGH_p \right\},$$

with P is also the maximum lag considered in the selection procedure.

In our $SETAR$ framework, the expression of the PDC criterion defined from the predictive density (1.5) is given by

$$\begin{aligned} PDC(\alpha) &= -\sum_{t=t_0}^{n} \log f(y_t | \mathbf{y}_{1,t-1}, \alpha) \\ &= \tfrac{1}{2}\log(2\pi) - \sum_{t=t_0}^{n} \sum_{k=1}^{K} \log\left(g_{p_k,d}(t,k)\right), \\ &\qquad \text{for } \mathbf{p} \in \{0,...,P\}^K \text{ and } d \in \{1,...,D\}, \end{aligned} \tag{1.7}$$

where $t_0 = \max_{k=1,\ldots,K} t_{k,0}$ and $t_{k,0}$ is the smallest t such that $n_{t-1,k} > 1$, i.e. t_0 is the first time from which we can evaluate (1.5) for $\mathbf{p} = (0,...,0)$. The constant D is a positive integer.

Since α is a $K+1$-vector, the question that arises is how to approximate the predictive functions whose evaluation is not possible, while using the same number of observations? In what follows, we try to use the same reasoning given by Djuric and Kay (1992) in the computation of the PDC^{*}. However, the approach developed in Djuric and Kay's paper does not trivially extend to the $SETAR$ case and the approximation of the predictive densities of the first observations is more complex in the $SETAR$ framework since we do not have a single constraint to satisfy. Indeed, if we assume, for example, that $\mathbf{p} = (0, 0, ..., 0)$, then (1.7) can be written as follows

$$PDC^{*}(0,0,...,0,d) = \frac{1}{2}\log(2\pi) - \sum_{t=t_0}^{n}\sum_{k=1}^{K}\log\left(g_{0,d}(t,k)\right),$$

where the number of observations used for the calculation of the PDC^{*}, in this $SETAR(K; 0, 0, ..., 0)$ case, is equal to $n - t_0 + 1$. Now, if $\mathbf{p} = (p_1, 0, ..., 0)$, and in order to use the same number of observations $n - t_0 + 1$, we can use the following initialization procedure

$$\begin{aligned} PDC^{*}(p_1,0,...,0,d) &= \frac{1}{2}\log(2\pi) \\ &- \sum_{i=1}^{p_1}\sum_{t=t_{(i-1,0,...,0,d)}}^{t_{(i,0,...,0,d)}-1}\left(\log\left(g_{i,d}(t,1)\right) - \sum_{k=2}^{K}\log\left(g_{0,d}(t,k)\right)\right) \\ &- \sum_{t=t_{(p_1,0,...,0,d)}}^{n}\left(\log\left(g_{p_k,d}(t,1)\right) - \sum_{k=2}^{K}\log\left(g_{0,d}(t,k)\right)\right), \end{aligned}$$

where $t_{(i,0,...,0,d)}$ is the first $t \geq t_0$ satisfying the constraint $n_{t-1,1} > i+1$ and the convention that $t_0 = t_{(0,0,...,0,d)}$. In the general case, we propose the following initialization strategy

$$\begin{aligned} PDC^{*}(\alpha) &= \frac{1}{2}\log(2\pi) - \sum_{t=t_0}^{t_\alpha-1}\sum_{k=1}^{K}\log\left(g_{\min\left\{n_{t-1,k}-2,p_k\right\},d}(t,k)\right) \\ &- \sum_{t=t_\alpha}^{n}\sum_{k=1}^{K}\log\left(g_{p_k,d}(t,k)\right), \end{aligned} \tag{1.8}$$

where t_α is the first $t \geq t_0$ satisfying the constraint $n_{t-1,k} > p_k + 1$, for all $k \in \{1, ..., K\}$. Thus, the delay parameter and autoregressive orders estimation are given by the value of the vector α which minimizes (1.8).

We now give an illustrative example allowing us to obtain a more precise idea of our proposed initialization procedure and how to evaluate the PDC^{*}.

Example 4. *Consider the case of a SETAR model with two regimes. For a given realization $y = (y_1, y_2, \ldots, y_n)$ of a stationary SETAR model in which the true value of α is $\alpha_0 = (\mathbf{p}_0, d_0)$, the delay parameter and autoregressive orders estimations are given by the value of the vector $\alpha = (\mathbf{p}, d) \in \{0, ..., P\}^2 \times \{1, ..., D\}$ which minimizes (1.8). We will require that the set of all candidate models under consideration includes α_0 (i.e. $\max \mathbf{p}_0 < P$ and $d_0 < D$).*

When we set the values of P and D, we can illustrate how to evaluate the PDCs of the different candidate models based on our proposed initialization procedure. For instance, if we put $P \leq 5$ and we set $D = 1$, Table 1.1 shows, for different candidate models, the values of $n_{t,k}$, $k = 1, 2$ obtained from a simulation of size 50 of the considered SETAR model.

1. *For the case $P = 5$, let us first note that when $\mathbf{p} = (0, 0)$, we can evaluate the predictive density only from $t = 6$, but we have to wait until $t = 21$ for $\mathbf{p} = (5, 5)$ (see the fourth and last columns of Table 1.1). Note also that for this last case $\mathbf{p} = (5, 5)$, we cannot proceed to the replacement of the predictive density by that of the model $\alpha = (0, 0, 1)$ only from $t_0 = 10$. Thus, the PDC^* takes the following form*

$$PDC^*(\alpha) = \tfrac{1}{2}\log(2\pi) - \sum_{t=10}^{t_\alpha - 1}\sum_{k=1}^{2}\log\left(g_{\min\left\{n_{t-1,k}-2, p_k\right\},1}(t,k)\right) - \sum_{t=t_\alpha}^{50}\sum_{k=1}^{2}\log\left(g_{p_k,1}(t,k)\right), \ \mathbf{p} \in \{1, 2, ..., 5\}^2.$$

On the other hand, it is clear, from Table 1.1, that when $\mathbf{p} = (0, 3)$ the first time from which we can evaluate the predictive density is $t = 16$. Moreover, for $t = 10, 11, 12, 13, 14$ and 15, the predictive density of y_t can be replaced respectively by that of $\mathbf{p} = (0, 0), (0, 1), (0, 1), (0, 1), (0, 2)$ and $(0, 2)$. Therefore, we get

$$\begin{aligned} PDC^*(0,3,1) = {} & \tfrac{1}{2}\log(2\pi) - \sum_{k=1}^{2}\log\left(g_{0,1}(10,k)\right) \\ & - \sum_{t=11}^{13}\left\{\log\left(g_{0,1}(t,1)\right) + \log\left(g_{1,1}(t,2)\right)\right\} \\ & - \sum_{t=14}^{15}\left\{\log\left(g_{0,1}(t,1)\right) + \log\left(g_{2,1}(t,2)\right)\right\} - \sum_{t=16}^{50}\sum_{k=1}^{2}\log\left(g_{p_k,1}(t,k)\right). \end{aligned}$$

2. *For the case $P = 2$, we use, unlike in the previous case, only the two blocks ($m = 1$ and $m = 2$) of Table 1.1 to evaluate the PDC^*. The findings indicate that we cannot proceed to the replacement of the predictive density by that of the model $\alpha = (0, 0, 1)$*

only from $t_0 = 9$*. When* $\mathbf{p} = (0, 2)$*, it is clear that the first time from which we can evaluate* $f(y_t | \mathbf{y}_{1,t-1}, \alpha)$ *is* $t = 11$*. Furthermore, for* $t = 9$ *and* 10*, the predictive density of* y_t *can be replaced respectively by that of* $\mathbf{p} = (0, 0)$ *and* $(0, 1)$*. We finally obtain*

$$\begin{aligned} PDC^*(0,2,1) = \tfrac{1}{2}\log(2\pi) - \sum_{k=1}^{2} \log\left(g_{0,1}(9,k)\right) - \log\left(g_{0,1}(10,1)\right) \\ - \log\left(g_{1,1}(10,2)\right) - \sum_{t=11}^{50}\sum_{k=1}^{2} \log\left(g_{p_k,1}(t,k)\right). \end{aligned}$$

1.4 Simulation study

In order to investigate the performance of the proposed PDC criterion for $SETAR$ orders identification problem, we carried out a simulation study based on five stationary $SETAR$ models. We assume that K is known and equal to 2 or 3 in this study, i.e. one or two thresholds. We simulated 1000 data samples with different lengths (n = 50, 100, 200 and 400). The corresponding parameter values are chosen to satisfy the stationary and ergodic sufficient condition. For each of the realizations, candidate models of orders $\mathbf{p} \in \{0, 1, ..., 7\}^K$ and delay parameters $d \in \{1, 2, ..., 5\}$ (for $K = 2$ and $K = 3$, we have respectively 320 and 512 models in total) are fitted to the data and the criteria AIC, $AICc$, BIC and PDC^* are evaluated. We recall that the definitions of AIC, $AICc$ and BIC used for $SETAR$ models are given by

$$\begin{aligned} AIC(\alpha) &= \sum_{k=1}^{K} n_k \ln \widehat{\sigma}_{k,d}^2 + 2(p_k + 1), \\ AICc(\alpha) &= \sum_{k=1}^{K} n_k \ln \widehat{\sigma}_{k,d}^2 + \frac{n_k(n_k + p_k + 1)}{n_k - p_k - 3}, \end{aligned}$$

and

$$BIC(\alpha) = \sum_{k=1}^{K} n_k \ln \widehat{\sigma}_{k,d}^2 + (p_k + 1)\ln n_k,$$

where

$$n_k = \sum_{t=m}^{n} \mathbf{1}_{(r_{k-1} < y_{t-d} \le r_k)},$$

Table 1.

	$m=2$					$m=3$					$m=4$					$m=5$					$m=6$				
			Candidate Models					Candidate Models					Candidate Models					Candidate Models					Candidate Models		
t	$n_{t,1}$	$n_{t,2}$	$(0,0)$	$\cdots$	$(1,1)$	$n_{t,1}$	$n_{t,2}$	$(0,2)$	$\cdots$	$(2,2)$	$n_{t,1}$	$n_{t,2}$	$(0,3)$	$\cdots$	$(3,3)$	$n_{t,1}$	$n_{t,2}$	$(0,4)$	$\cdots$	$(4,4)$	$n_{t,1}$	$n_{t,2}$	$(0,5)$	$\cdots$	$(5,5)$
1	0	0	–	–	–	0	0	–	–	–	0	0	–	–	–	0	0	–	–	–	0	0	–	–	–
2	0	1	–	–	–	0	0	–	–	–	0	0	–	–	–	0	0	–	–	–	0	0	–	–	–
3	0	2	–	–	–	0	1	–	–	–	0	0	–	–	–	0	0	–	–	–	0	0	–	–	–
4	1	2	–	–	–	1	1	–	–	–	1	0	–	–	–	0	0	–	–	–	0	0	–	–	–
5	2	2	–	–	–	2	1	–	–	–	2	0	–	–	–	1	0	–	–	–	0	0	–	–	–
6	3	2	$(0,0)$		$(0,0)$	3	1	–	–	–	3	0	–	–	–	2	0	–	–	–	1	0	–	–	–
7	4	2	$(0,0)$		$(1,0)$	4	1	–	–	–	4	0	–	–	–	3	0	–	–	–	2	0	–	–	–
8	4	3	$(0,0)$		$(1,0)$	4	2	–	–	–	4	1	–	–	–	3	1	–	–	–	2	1	–	–	–
9	4	4	$(0,0)$		$(1,1)$	4	3	$(0,0)$		$(2,0)$	4	2	–	–	–	3	2	–	–	–	2	2	–	–	–
10	4	5	$(0,0)$		$(1,1)$	4	4	$(0,1)$		$(2,1)$	4	3	$(0,0)$		$(2,0)$	3	3	$(0,0)$		$(1,0)$	2	3	$(0,0)$	–	$(0,0)$
11	5	5	$\vdots$		$\vdots$	5	4	$(0,2)$		$(2,2)$	5	3	$(0,1)$		$(2,1)$	4	3	$(0,1)$		$(1,1)$	3	3	$(0,1)$		$(0,1)$
12	6	5				6	4	$(0,2)$	$\vdots$	$(2,2)$	6	3	$(0,1)$		$(3,1)$	5	3	$(0,1)$		$(2,1)$	4	3	$(0,1)$		$(1,1)$
13	6	6				6	5	$\vdots$		$\vdots$	6	4	$(0,1)$		$(3,1)$	5	4	$(0,1)$		$(3,1)$	4	4	$(0,1)$		$(2,1)$
14	7	6				7	5				7	4	$(0,2)$		$(3,2)$	6	4	$(0,2)$		$(3,2)$	5	4	$(0,2)$		$(2,2)$
15	7	7				7	6				7	5	$(0,2)$		$(3,2)$	6	5	$(0,2)$		$(4,2)$	5	5	$(0,2)$		$(3,2)$
16	7	8				7	7				7	6	$(0,3)$		$(3,3)$	6	6	$(0,3)$		$(4,3)$	5	6	$(0,3)$		$(3,3)$
17	7	9				7	8				7	7	$(0,3)$		$(3,3)$	6	7	$(0,4)$		$(4,4)$	5	7	$(0,4)$		$(3,4)$
18	7	10				7	9				7	8	$\vdots$		$\vdots$	6	8	$(0,4)$		$(4,4)$	5	8	$(0,5)$		$(3,5)$
19	8	10				8	9				8	8				7	8	$\vdots$		$\vdots$	6	8	$(0,5)$		$(3,5)$
20	9	10				9	9				9	8				8	8				7	8	$\vdots$		$(4,5)$
21	9	11				9	10				9	9				8	9				7	9			$(5,5)$
$\vdots$	$\vdots$	$\vdots$				$\vdots$	$\vdots$				$\vdots$	$\vdots$				$\vdots$	$\vdots$				$\vdots$	$\vdots$			$\vdots$
50			$(0,0)$	$\cdots$	$(1,1)$			$(0,2)$	$\cdots$	$(2,2)$			$(0,3)$	$\cdots$	$(3,3)$			$(0,4)$	$\cdots$	$(4,4)$			$(0,5)$	$\cdots$	$(5,5)$

is the number of observations in regime k and $\widehat{\sigma}^2_{k,d}$ is defined as

$$\widehat{\sigma}^2_{k,d} = \frac{1}{n_k} \sum_{t=m}^{n} \left(y_t - \phi_{k,0} - \sum_{j=1}^{p_k} \phi_{k,j} y_{t-j} \right)^2 \mathbf{1}_{(r_{k-1} < y_{t-d} \leq r_k)}.$$

First, we focus our attention on the case in which $\phi_{k,0} = 0$, for $k = 1, 2$, and $d = 1$. We carried out two simulation experiments to compare the performance of AIC, $AICc$, BIC and PDC^*. The $SETAR\,(2; p_1, p_2)$ models simulated are of the form

$$y_t = \begin{cases} \sum_{j=1}^{p_1} \phi_{1,j} y_{t-j} + \sigma_1 \varepsilon_t & y_{t-1} < r_1, \\ \sum_{j=1}^{p_2} \phi_{2,j} y_{t-j} + \sigma_2 \varepsilon_t & y_{t-1} \geq r_1. \end{cases}$$

For Model I, $r_1 = 0$, $\sigma = (1,1)'$, $\Phi = (0, 0.5, 0, -0.5)'$ (this model has been previously studied by Wong and Li, 1998) and Model II, $r_1 = 0.5$, $\sigma = (1,2)'$ and $\Phi = (0, -0.5, -0.2, 0, 0.3, 0.4)'$. For the case $d = 2$, we also considered a model without constant (Model III), i.e, we specify $\phi_{k,0} = 0$, for $k = 1, 2$. The autoregressive parameter of regime 1 is $\phi_{1,1} = -0.8$ and $\sigma_1 = 1$, while those of regime 2 are set to $\phi_{1,2} = -0.2$ and $\sigma_2 = 2$. The threshold parameter $r_1 = 0$. We next performed a simulation study on two models with constant (Models IV and V) to examine how well the criteria work for model selection in this case. The true parameters for Model IV are specified as $d = 1$, $r_1 = 0.5$, $\sigma = (1,1)'$ and $\Phi = (.3, -0.5, -0.5, 0)'$. For Model V, they are given by $d = 1$, $r_1 = -0.5$, $r_2 = 0.5$, $\sigma = (1,1,1)'$ and $\Phi = (-1.5, -.7, 0.5, 0, -1, 0)'$.

The results of the 1000 sampling experiments are summarized using the following measures of performance: the frequency $f_{\mathbf{p}}$ of selecting the correct order, the frequency f_d of selecting the correct delay d and the frequency f_α of selecting the correct order $\mathbf{p}$ and the correct delay d jointly. These measures are reported in Tables 1.2-1.6.

Table 1.2. Frequencies of selection of the Model I. The parameter values are $d = 1$, $r_1 = 0$, $\sigma = (1,1)'$ and $\Phi = (0, 0.5, 0, -0.5)'$.

	$n = 50$			$n = 100$			$n = 200$			$n = 400$		
Criterion	$f_{\mathbf{p}}$	f_d	f_α	$f_{\mathbf{p}}$	f_d	f_α	$f_{\mathbf{p}}$	f_d	f_α	$f_{\mathbf{p}}$	f_d	f_α
PDC^*	299	711	293	586	914	585	836	994	836	942	999	942
AIC	152	651	149	361	904	361	448	992	448	434	998	434
$AICc$	188	642	177	257	893	256	224	992	224	186	998	186
BIC	220	681	215	583	956	583	815	997	815	909	1000	909

Table 1.3. Frequencies of selection of the Model II. The parameter values are $d = 1$, $r_1 = 0.5$, $\sigma = (1, 2)'$ and $\Phi = (0, -0.5, -0.2, 0, 0.3, 0.4)'$.

	$n = 50$			$n = 100$			$n = 200$			$n = 400$		
Criterion	$f_{\mathbf{p}}$	f_d	f_α	$f_{\mathbf{p}}$	f_d	f_α	$f_{\mathbf{p}}$	f_d	f_α	$f_{\mathbf{p}}$	f_d	f_α
PDC^*	131	959	131	165	991	165	464	1000	464	821	1000	821
AIC	88	978	88	138	994	137	223	1000	223	252	1000	252
$AICc$	66	978	66	84	993	84	89	1000	89	67	1000	67
BIC	128	977	127	161	996	161	460	1000	460	788	1000	788

Table 1.4. Frequencies of selection of the Model III. The parameter values are $d = 2$, $r_1 = 0$, $\sigma = (1, 4)'$ and $\Phi = (0, -0.8, 0, -0.2)'$.

	$n = 50$			$n = 100$			$n = 200$			$n = 400$		
Criterion	$f_{\mathbf{p}}$	f_d	f_α	$f_{\mathbf{p}}$	f_d	f_α	$f_{\mathbf{p}}$	f_d	f_α	$f_{\mathbf{p}}$	f_d	f_α
PDC^*	167	876	117	238	993	233	360	999	359	553	1000	553
AIC	62	807	55	119	986	118	231	999	231	284	1000	284
$AICc$	123	848	96	103	985	102	119	999	119	117	1000	117
BIC	111	861	88	213	992	208	355	1000	355	548	1000	548

Table 1.5. Frequencies of selection of the Model IV. The parameter values are $d = 1$, $r_1 = 0.5$, $\sigma = (1, 1)'$ and $\Phi = (.3, -0.5, -0.5, 0)'$.

	$n = 50$			$n = 100$			$n = 200$			$n = 400$		
Criterion	$f_{\mathbf{p}}$	f_d	f_α	$f_{\mathbf{p}}$	f_d	f_α	$f_{\mathbf{p}}$	f_d	f_α	$f_{\mathbf{p}}$	f_d	f_α
PDC^*	401	580	261	607	777	539	803	901	799	922	983	922
AIC	107	406	73	248	623	234	372	799	372	440	956	440
$AICc$	272	382	113	177	550	155	186	719	186	193	935	193
BIC	235	510	155	527	749	474	776	894	775	907	985	907

Table 1.6. Frequencies of selection of the Model V. The parameter values are $d = 1$, $r_1 = -0.5, r_2 = 0.5$, $\sigma = (1, 1, 1)'$ and $\Phi = (-1.5, -.7, 0.5, 0, -1, 0)'$.

	$n = 50$			$n = 100$			$n = 200$			$n = 400$		
Criterion	$f_{\mathbf{p}}$	f_d	f_α	$f_{\mathbf{p}}$	f_d	f_α	$f_{\mathbf{p}}$	f_d	f_α	$f_{\mathbf{p}}$	f_d	f_α
PDC^*	292	865	276	578	987	578	781	1000	781	857	1000	857
AIC	20	514	20	204	991	204	287	1000	287	270	1000	270
$AICc$	142	581	135	243	996	243	180	1000	180	126	1000	126
BIC	54	524	54	495	995	495	763	1000	763	856	1000	856

It can be observed that for each instance, the performance of PDC^* is much better, than the other criteria, in terms of higher frequencies of jointly selecting the correct order and delay. It can also be observed that the PDC^* obtains the most correct order selections, followed by BIC. Note that for Model I (respectively Model II, III, IV and V) and $n = 50$ the AIC successfully jointly selected the correct order and delay only 14.9% (respectively 8.8%, 5.5%, 7.3% and 0.2%) of the time, whereas for $n = 400$ the frequency is 43.4% (respectively 25.2%, 28.4%, 44% and 27%) which is a relatively low rate. Although $AICc$ and BIC are designed, for the purpose of small-sample model selection, to overcome the inconsistency of AIC, their performance is not as good as PDC^*. Indeed, for $n = 50$, the PDC^* rates of correct identification of the vector α are higher (about respectively 29.3%, 13.1%, 11.7%, 26.1% and 27.1% for Models I, II, III, IV and V) than for BIC (about respectively 21.5%, 12.7%, 8.8%, 15.5% and 0.5% for Models I, II, III, IV and V). Furthermore, one of the results which prominently stands out most from Tables 1.2-1.6 is that the true value of the delay parameter is chosen more often for all the criteria. The frequency of selecting the correct delay varies from 64.2% to 97.8% (respectively 89.3% to 99.6%, 99.2% to 100%, 99.8% to 100%) when $n = 50$ (respectively $n = 100$, 200 and 400). While for the Models IV and V, the rate of selecting the true value of the delay by the three criteria AIC, $AICc$ and BIC, is less than 60% when the sample size is 50. However, if we increase the sample size ($n = 200$ or 400), this low rate increases up to 98% and 100% for the BIC criterion, as we can remark from Tables 5-6. Also, the PDC^* criterion obtains the most correct delay parameter selections when the sample size is small. Finally, when the sample size increases, the frequencies of selection of the true p and/or d performed by the BIC criterion and our PDC^* criterion become approximately equal.

It may be noted that all identification methods discussed in our paper (AIC, $AICc$, BIC and PDC^*) have the advantage of allowing automatic determination of the delay parameter and autoregressive orders of a $SETAR$ process. It may also be noted that the computational efficiency of the different criteria depends on several factors: most importantly, the sample size n, the dimension of the largest model in the candidate class, i.e. the value of P, and the number of regimes K. We note here that the delay parameter has no effect on the evaluation cost of the four studied criteria.

In contrast to AIC, $AICc$ and BIC criteria, our PDC^* procedure does not need to estimate all candidate models. Despite this, the PDC^* is usually computationally exorbitant compared to the other three criteria. However, advances in computing now allow complex methods to be used for identifying the generating model, instead of classical methods such that AIC, $AICc$ and BIC criteria as long as this type of computational burden is not very costly and we will get a better performance.

It is important to note that all our simulations were carried out on an Intel(R) Core(TM) i5, 2.50 GHz processor of a PC desktop with 8.00 Go of RAM, using Matlab version $R2016$. For example, in the Model IV case, the procedure PDC^* (for all the 320 possible $SETAR$ models) requires around 11 seconds for $n = 100$ and around 148 seconds for $n = 400$. While for the Model V, the same procedure (for $K = 3$, there are 512 candidate models) requires as much as ten times more. It is around 112 seconds for $n = 100$ and around 1297 seconds for $n = 400$.

1.5 Real Data examples

1.5.1 Annual Wolf's sunspot

The sequence of Wolf's sunspot numbers is very famous in time series analysis. It may be said that the study of this series is the basis for the analysis of time series as we know it nowadays. In fact, the era of linear time series modelling began with the linear autoregressive models of Yule (1927), first introduced in the study of this series. Various linear and nonlinear models have been applied to analyse the dynamics of this series. We can cite the two-regime models proposed in Tong (1990), and the three-regime threshold models in Tsay (1989). Recently, Li et al. (2015) validate a three-regime threshold using the AIC and BIC.

We consider the annual mean Wolf's sunspot numbers from 1700 to 2013. This provides a total of 314 observations and in total, there are 314 observations. The dataset was downloaded from the website of Sunspot Index and Longterm Solar Observations. Figure 1.2 displays the Annual Wolf's sunspot series plot.

We apply our PDC for the joint determination of the delay parameter and the autoregressive orders of a three-regime $SETAR$ model. The AIC and BIC criteria were used for the model selection in Li et al. (2015) with the maximum order $P = 13$ and a maximum delay $D = 6$. The performance of the linear AR was also examined. Both information criteria select the same autoregressive model but two different threshold autoregressive

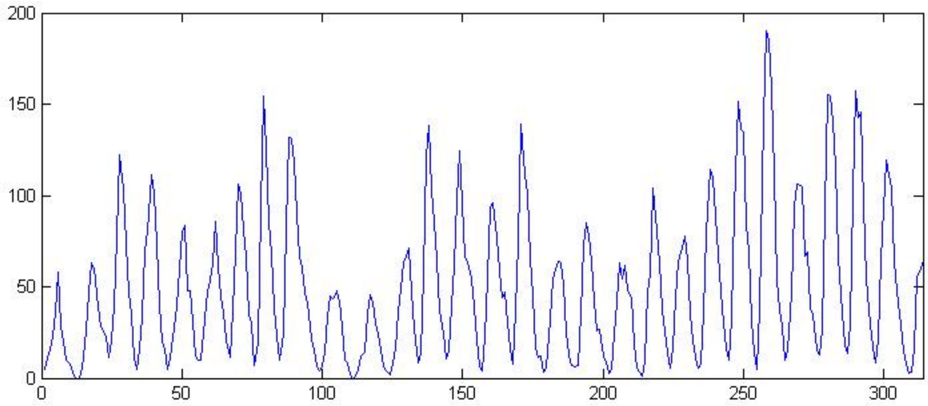

Figure 1.2: Annual Wolf's sunspot series plot from 1700 to 2013.

models. For a comparison purpose, we consider three-regime threshold autoregressive models validated by Li et al. (2015) and we opt for the same choices of P and D in the application of our PDC by considering the resulting AIC's and BIC's thresholds estimations.

The resulting fitted delay parameter is $\widehat{d} = 2$ for both PDC and AIC, and this meets the results of Tsay (1989). Figure 1.3 and Figure 1.4 provide the sample autocorrelation functions, ACFs, of the residuals from the fitted models, and we may conclude that they are all adequate.

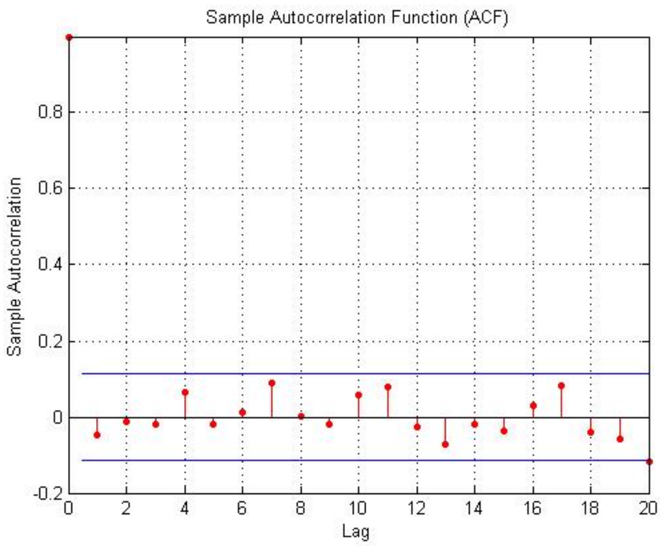

Figure 1.3: The sample ACF of the residuals from the fitted three-regime $SETAR$ model using AIC.

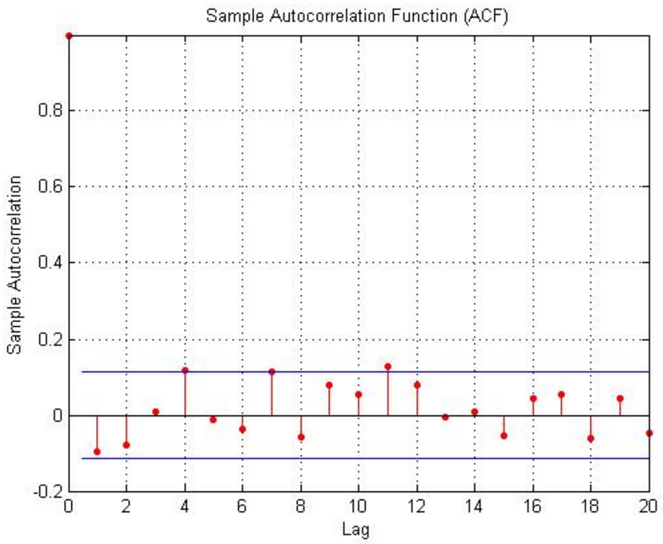

Figure 1.4: The sample ACF of the residuals from the fitted three-regime $SETAR$ model using PDC and AIC's thresholds estimates.

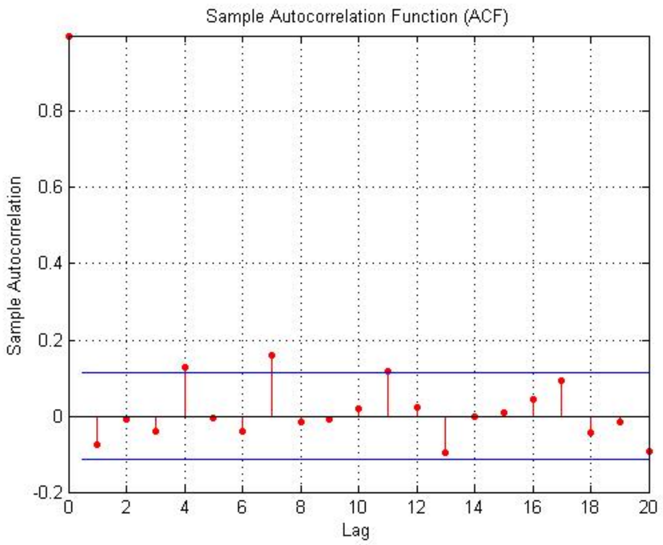

Figure 1.5: The sample ACF of the residuals from the fitted three-regime $SETAR$ model using BIC.

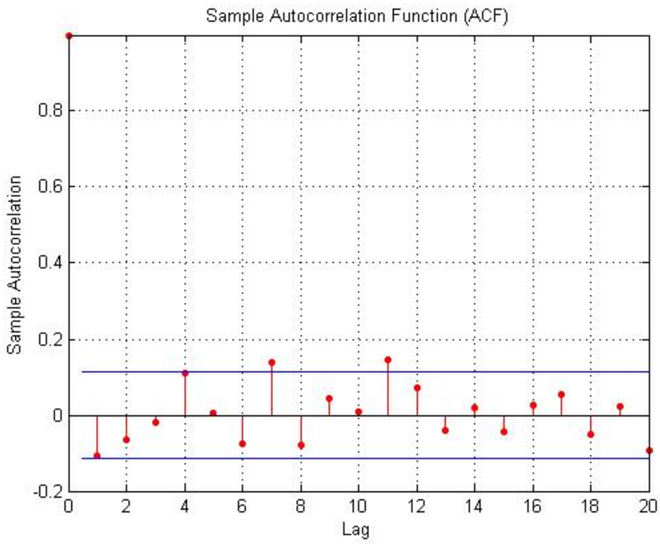

Figure 1.6: The sample ACF of the residuals from the fitted three-regime $SETAR$ model using PDC and BIC's thresholds estimates.

The other estimated parameters of these fitted models are also reported in Table 1.7 where the variances of the autoregressive parameters are in brackets.

Table 1.7. Estimated models for the Annual Wolf's sunspot number series.

	$SETAR(3;p_1,p_2,p_3)$											
	PDC			AIC			PDC			BIC		
Lag	Lower	Middle	Upper	Lower	Middle	Upper	Lower	Middle	Upper	Lower	Middle	Upper
0	−2.64 (18.49)	18.18 (18.26)	10.22 (8.98)	−1.68 (18.44)	8.71 (23.69)	11.59 (10.15)	−2.64 (18.49)	14.86 (14.59)	11.07 (13.35)	−0.02 (17.62)	15.57 (14.38)	13.32 (15.11)
1	2.33 (0.02)	1.59 (0.00)	0.88 (0.00)	2.17 (0.03)	1.60 (0.00)	0.68 (0.00)	2.33 (0.02)	1.53 (0.00)	0.86 (0.00)	2.27 (0.03)	1.54 (0.00)	0.65 (0.00)
2	−3.2 (0.66)	−0.89 (0.04)	0.09 (0.00)	−2.95 (0.71)	−0.75 (0.04)	0.06 (0.00)	−3.2 (0.66)	−0.70 (0.03)	0.10 (0.00)	−2.91 (0.78)	−0.70 (0.03)	0.06 (0.00)
3	0.53 (0.26)	−0.53 (0.03)	−0.30 (0.00)	0.71 (0.26)	−0.28 (0.04)	−0.21 (0.00)	0.53 (0.26)	−0.42 (0.02)	−0.29 (0.00)	0.76 (0.27)	−0.45 (0.03)	−0.15 (0.00)
4	−0.70 (0.14)	0.30 (0.01)	–	−0.86 (0.15)	0.34 (0.02)	−0.00 (0.00)	−0.70 (0.14)	0.25 (0.02)	–	−0.97 (0.17)	0.28 (0.02)	-0.02 (0.00)
5	0.58 (0.06)	–	–	0.82 (0.07)	−0.26 (0.02)	−0.09 (0.00)	0.58 (0.06)	−0.27 (0.01)	–	0.98 (0.05)	0.27 (0.01)	−0.09 (0.00)
6	0.28 (0.01)	–	–	−0.21 (0.07)	0.17 (0.01)	0.01 (0.01)	0.28 (0.01)	0.21 (0.00)	–	–	0.20 (0.00)	0.02 (0.01)
7	–	–	–	0.26 (0.02)	0.09 (0.01)	0.14 (0.01)	–	–	–	–	–	0.14 (0.01)
8	–	–	–	–	−0.20 (0.07)	−0.18 (0.01)	–	–	–	–	–	−0.14 (0.02)
9	–	–	–	–	0.35 (0.01)	0.14 (0.01)	–	–	–	–	–	0.32 (0.00)
10	–	–	–	–	−0.33 (0.02)	0.05 (0.00)	–	–	–	–	–	–
11	–	–	–	–	0.20 (0.02)	0.12 (0.00)	–	–	–	–	–	–
12	–	–	–	–	0.13 (0.02)	–	–	–	–	–	–	–
13	–	–	–	–	−0.29 (0.01)	–	–	–	–	–	–	–
	Other estimates											
Nobs	52	111	145	49	108	144	52	137	119	49	134	118
$\widehat{d}$		2			2			2			2	
$\widehat{\sigma}^2$	137.46	193.85	147.6	129.96	144.35	111.68	137.46	180.77	139.47.6	142.70	171.15	107.04
(r_1, r_2)			$(10.7, 43.9)$						$(10.7, 54.1)$			

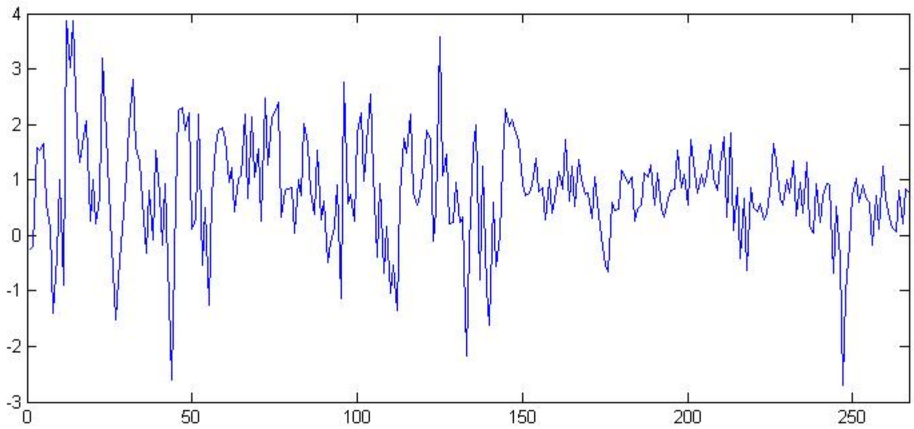

Figure 1.7: The plot of GNP growth rates

1.5.2 GNP Growth

Nonlinear time-series models, more particularly threshold models, have shown a high power in the business cycles modelling. Indeed, the traditional threshold autoregressive models have been applied to model the series of the Gross National Product (GNP) and the Gross Domestic Product (GDP), as aggregated outputs of a country. Potter (1995) suggested to distinguish between the expansion periods and the recession ones of the economy using a two-regime threshold autoregressive model for GNP growth rates, where each regime corresponds to one of the two states.

Pesaran and Potter (1997) tried to push the analysis and they argued that three regimes may be more suitable to respectively encompass bad, good, and normal periods. Tiao and Tsay (1994) considered a four-regime threshold autoregressive model. Recently, Li et al. (2015) suggest to use AIC and BIC to analyze a three regime threshold-based dynamics of this series. This motivates us to consider our PDC in the model selection step to fit the GNP growth rates of US economy.

We consider the seasonally adjusted sequence Y_t of quarterly US real GNP from the first quarter of 1947 to the last quarter of 2013, which provides a total of 268 observations. We focus on its growth rates as percentages in this application, which is defined as $y_t = 100\{\log(Y_t) - \log(Y_{t-})\}$. This transformation induces a loss of one observation. Therefore, the sample size is 267. The plot of this series is given in Figure 1.7. This GNP growth series has a sample mean of 0.79 and a sample variance of 0.94.

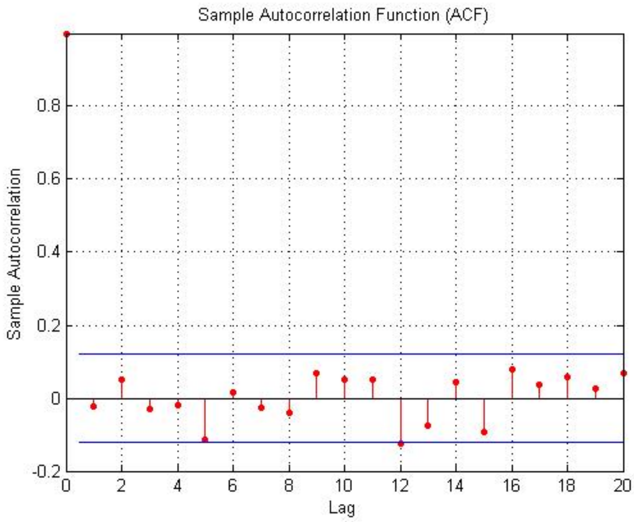

Figure 1.8: The sample ACF of the residuals from the fitted thres regime $SETAR$ model using PDC

As in Li et al. (2015), we fit a three-regime $SETAR$ model to this series where the PDC is used to select the delay parameter and the autoregressive orders with $D = 6$ and $P = 4$. The considered threshold parameters are the same as those estimated by Li et al. (2015). The fitted model is given by

$$y_t = \begin{cases} 0.57_{(0.00)} + 0.37_{(0.01)} y_{t-1} + 0.98\varepsilon_t & \text{if } y_{t-1} \leq 0.78, \\ -0.36_{(0.20)} + 0.92_{(0.18)} y_{t-1} + 0.44_{(0.01)} y_{t-2} & \\ \qquad\qquad -0.24_{(0.00)} y_{t-3} + 0.44\varepsilon_t & \text{if } 0.78 < y_{t-1} \leq 1.53, \\ 0.38_{(0.19)} + 0.44_{(0.04)} y_{t-1} + 0.71\varepsilon_t & \text{if } 1.53 < y_{t-1}, \end{cases}$$

where the variances of the autoregressive parameters are in brackets. It is worth mentioning that the PDC gives the same autoregressive orders that those identified by Li and al. (2015). However, the difference between our PDC identification and their selection lies in the delay parameter which is $\widehat{d} = 1$ for our PDC and $\widehat{d} = 2$ for AIC selection. This leads to the estimation of two different models since the considered delay parameter impacts the assignment of observations to regimes. In fact, our PDC identification begets the allocation of 129 observations to the first regime, 77 in the second and 58 in the third regime, contrary to the AIC selection where 130 observations was in the first regime, 78 in the second and 53 observations in the third. Only 261 observations have been considered unlike our PDC where we use 264 observations. The ACF of the residuals obtained from our model is given in Figure 1.8 We may conclude from this last figure that our fitted model is adequate.

Chapter 2

On Buffered Stochastic Volatility Model: Theory and Application

2.1 Introduction

It is widely accepted that the volatility (instantaneous variation) of financial returns is not constant over time and that the variance of returns shows a high serial autocorrelation, often called volatility clustering. The study of the instantaneous change and volatility clustering plays a considerable role in the financial time series analysis. An important tool for modelling conditional volatility is the so-called Stochastic Volatility (SV) model introduced by Taylor (1986) and which is considered as a satisfactory alternative to the reference class commonly known as the $ARCH/GARCH$ family. These two main families of econometric models were proposed to analyze time-dependent variances and they stand out from a construction point of view. Thus, in the $ARCH/GARCH$ models, time-dependent variances modelling is mainly characterized by specifying the volatility as a function of powers of previous observations and/or past volatility and, consequently, the volatility can be observed one-step ahead. Instead of assuming this deterministic feature, the SV modelling is designed to capture the stochastic nature of volatility by allowing volatility to evolve following a stochastic process. From an empirical point of view, the introduction of an innovation term into the latent volatility process makes the SV formulation more flexible than the $ARCH/GARCH$ specification (see e.g., Shephard, 1996; Carnero et al., 2004; Malmsten and Teräsvirta, 2004).

Furthermore, important and well documented empirical studies find evidence about the existence of an asymmetric response of volatility to positive and negative past returns. This property, known as leverage effect, was first described by Black (1976). In fact, financial markets become more volatile in response to “bad news” (or negative shocks) than to “good news” (or positive shocks) of the equivalent magnitude. Despite the success of the two previously presented models in their basic versions, however, neither $GARCH$ nor SV models incorporate a leverage effect and asymmetry, and this may present a limitation of basic formulations for both approaches.

By considering a deterministic behavior of the volatility, Nelson (1991) suggests analyzing this type of stylized facts by introducing the Exponential $GARCH$ model ($EGARCH$), which takes into account the asymmetric responses to positive and negative returns. Rodriguez and Ruiz (2012) also investigated various asymmetric GARCH specifications. In the stochastic case of the volatility, one way of incorporating leverage effect and asymmetry into the SV model is to allow a correlation between the noise of the volatility process and the noise of the observations series (Harvey and Shephard, 1996; Jacquier et al., 2004).

However, allowing a correlation between returns and volatility is not the only way to construct an asymmetric stochastic volatility model. In fact, threshold modelling introduced by Tong (1978), has been adopted to reproduce asymmetric behavior and to capture the leverage effect exhibited by financial returns. The use of threshold models to explain asymmetric volatility has already been known in the deterministic volatility framework. Indeed, Zakoïan (1991, 1994) introduce the so-called threshold $GARCH$ ($TGARCH$). This class of models has been enlarged by Rabemananjara and Zakoian (1993). In the same context, Li and Li (1996) introduce the double-threshold autoregressive conditional heteroscedastic ($DTARCH$) where a piecewise linear representation has been considered for both conditional mean and conditional variance. With this in mind, Liu et al. (1997) further extend this model to the Double Threshold $GARCH$ model. Li (2009) gives an extensive survey of the threshold approaches for volatility changes modelling with $ARCH$-type models.

Seeing the contribution of threshold models in the volatility modelling for the deterministic context and from an intuitive point of view, it seems natural to extend the threshold idea to the SV case. Indeed, this approach has been introduced by Breidt (1996) to the stochastic volatility analysis area by the definition of the Threshold Stochastic Volatility model (TSV). Based on Tong’s pioneering work, the proposed model assumes that, according to the nature of the information good or bad, the volatility dynamics follows a two-regime threshold model, where the log-volatility in each regime is represented by a first order autoregressive model and the transition between regimes is determined according to the lagged stock returns signs. A similar approach has been proposed by So et al. (2002) where the threshold SV model has been constructed to capture simultaneously the mean and variance asymmetries. Chen et al. (2008) generalized the TSV model and incorporates a heavy-tailed error distribution. TSV models are quite a popular choice for representing the volatility of financial returns (see Mao et al. (2017) and references therein).

In addition, empirical studies have shown that asset prices have a particular volatility reaction for good and bad news (Bekaert and Wu, 2000). This reaction can be described in some way as hysterical. Assuming the threshold effect in the volatility and without loss of generality, for a two-regime situation, there are some cases where the shift of regime may not happen at the same threshold value. In fact, when the return on the price of an asset exceeds a particular positive threshold r_U, the market can affirm the advent of good news, while the bad news is only confirmed when return crosses another negative threshold, r_L. The interval $(r_L; r_U]$ serves therefore as a buffer zone, no information comes in when the return falls within this zone and volatility structure is also assumed to remain unchanged. It is clear that this finding cannot be captured by the classical threshold model since it considers only a single threshold parameter. Although threshold models have been hugely successful, they have been observed to have poor performance around the boundaries between the different regimes. This is due to the sudden change in the probabilistic structure, which may not be the case in the real world (see Wu and Chen, 2007). To deal with this kind of situations, we consider in this chapter the buffering idea firstly introduced by Li et al. (2015). We recall that this idea has been discussed in the introductory paper untitled "Hysteresis autoregressive time series model" where a new more flexible regime switching mechanism is defined. This mechanism is based on the replacement of the single threshold parameter by a zone which is called "Hysteresis zone". This approach provides a rigorous definition of the regime indicator. To illustrate and highlight this point, let us limit ourselves, without loss of generality, to the case of a two-regime model. Instead of taking a single threshold parameter, we consider an interval consisting of a lower and an upper thresholds and which acts as a buffer zone. If the threshold variable is less than the lower boundary of the zone $z_t \leq r_L$, then the observation is from the first regime. Conversely, the observation comes from the second regime when the threshold variable is greater than the upper one $z_t > r_U$. When the threshold variable z_t falls within the buffer zone $(r_L, r_U]$, the regime indicator keeps the value of its most recent past.

We emphasize the fact that this idea has been proven to be useful in the study of the quarterly dataset on U.S. GNP considered in Zhu et al. (2014) and Li et al. (2015). Indeed, their empirical results showed that the Buffered Autoregressive (BAR) model has a better fitting than the TAR model for this series. Based on the fitted BAR model, they found that the regime of GNP series does not shift unless we have experienced a big 'contraction' or 'expansion' two years before. This finding is clearly of practical interest and cannot be detected by the TAR model.

This idea was adopted in Lo et al. (2016) where they proposed a new threshold conditional heteroscedastic model called the buffered threshold $GARCH$. The combination of the BAR specification for the conditional mean and the buffer $GARCH$ specification for the conditional variance is an important extension which has many appealing features and includes a better specification of the forecast variance. From this basis, two similar models were independently introduced. The first one is the Double Hysteretic Heteroskedastic Model proposed by Chen and Truong (2016) to analyze the relationship between the underlying stock markets and index futures markets. In parallel, Zhu et al. (2017) introduced the so called "Buffered Autoregressive model with Generalized Autoregressive Conditional Heteroscedasticity (BAR-$GARCH$)" to deal with volatility in exchange rates.

In this chapter, we introduce this idea in the context of SV modelling. By considering the same new flexible regime-switching mechanism, we define in the next section a new threshold stochastic volatility model that we call Buffered Stochastic Volatility (BSV) model. This model generalizes Breidt′s threshold stochastic volatility model and attempts to highlight the nonlinearity in assets returns dynamics and provides a better explanation of the asymmetry and leverage effect in financial time series. Section 2.3 is dedicated to the estimation problem of BSV. By considering our formulation as a nonlinear state space model, we employ a Sequential Monte Carlo method to provide a Maximum Likelihood estimate of the proposed model. The performance of the proposed estimation method is discussed through a simulation study in Section 2.4. Finally, an application of the model to financial markets is provided in the last section.

2.2 Buffered Threshold SV Models

The realization of the univariate process $\{\varepsilon_t; t \in \mathbb{Z}\}$ is said to be generating from a two-regime Buffered Threshold SV (BSV) model if it satisfy

$$\begin{cases} \varepsilon_t = \eta_t \exp\left\{\frac{1}{2}x_t\right\}, \ t \in \mathbb{Z}, \\ x_t = \beta_0^{(k)} + \beta_1^{(k)} x_{t-1} + Q^{(k)} e_t, \ \text{if } R_t = k, \end{cases} \tag{2.1}$$

where ε_t is an observable time series of return, x_t is known as the log-volatility of the process ε_t, and

$$R_t = \begin{cases} 1 & \text{if } \varepsilon_{t-1} \leq r_L, \\ R_{t-1} & \text{if } r_L < \varepsilon_{t-1} \leq r_U, \\ 0 & \text{if } \varepsilon_{t-1} > r_U, \end{cases} \tag{2.2}$$

The two processes (η_t) and (e_t) are two independent sequences of independent and identically distributed $i.i.d.$ random variables with zero mean and unit variance. It is worth noting that the asymmetry in the volatility is introduced by assuming that the log-volatility switches between two different first-order AR dynamics. Here, the regime specification is determined according to whether the previous return value is less or greater than the boundary (or threshold) parameters, r_L and r_U, respectively, and when the previous return value falls within the buffer zone $(r_L, r_U]$ the regime indicator stays unchanged. The main originality in this formulation is in this flexibility in the transition mechanism which ensures that volatility does not move from one regime to another suddenly (brutally) and therefore, it allows to better highlight the leverage effect and the asymmetry in the financial return.

For a notational simplicity, (2.1) can be summarized compactly in (2.3) as follows,

$$\begin{cases} \varepsilon_t = \eta_t \exp\left\{\frac{1}{2}x_t\right\}, \\ x_t = \beta_0^{(R_t)} + \beta_1^{(R_t)} x_{t-1} + Q^{(R_t)} e_t, \end{cases} \quad t \in \mathbb{Z}. \tag{2.3}$$

Note that when $r_L = r_U = 0$, model (2.3) is reduced to the traditional threshold SV model (Breidt, 1996). This symmetrical nature of the reaction with respect to 0 can be a strong restriction in practice where the threshold value may be unknown. However, it is worth mentioning that the formulation (2.3) is more general and differs from that of Breidt (1996) in which the transition is brutal since the log-volatility can switch from one regime to another if the status of the lagged value of ε_t crosses up or down 0. In other words, the regime indicator R_t only depends on the sign of the lagged value of ε_t. In our proposed model, and thanks to the relaxation of the symmetry assumption of the reaction, the formulation becomes more realistic and the regime indicator may depend on the past of the regime indicators infinitely far away. As in Li et al. (2015), it can be iteratively calculated in the almost surely sense as follows

$$\begin{aligned} R_t &= \mathbf{1}_{(\varepsilon_{t-1} \leq r_L)} + R_{t-1}\mathbf{1}_{(r_L < \varepsilon_{t-1} \leq r_U)} \\ &= \mathbf{1}_{(\varepsilon_{t-1} \leq r_l)} + \sum_{j=0}^{\infty} \prod_{i=0}^{j} \mathbf{1}_{(r_L < \varepsilon_{t-1-i} \leq r_U)} \mathbf{1}_{(\varepsilon_{t-j-2} \leq r_L)}. \end{aligned}$$

2.3 Maximum Likelihood Estimation

Consider the representation (2.2) and (2.3) of the BSV model, which we assume to be stationary. Let $\underline{\varepsilon} = (x_0,\, x_1, ..., x_n, \varepsilon_1, \varepsilon_2, ..., \varepsilon_n)$ be the vector of the complete data. For a given realization $\varepsilon = (\varepsilon_1, \varepsilon_2, ..., \varepsilon_n)$ of the model (2.3), and under the assumptions that $x_0 \sim \mathcal{N}(\mu_0, {\delta_0}^2)$, the normality of $\{e_t\}$ and $\{\eta_t\}$, the complete likelihood function of the vector parameters $\theta = \left(\beta_0^{(1)}, \beta_0^{(0)}, \beta_1^{(1)}, \beta_1^{(0)}, Q^{(1)}, Q^{(0)}\right)'$ can be expressed as follows

$$\begin{aligned} \mathbf{L}(\theta; \underline{\varepsilon}) &= f(x_0) \prod_{t=1}^{n} f(x_t | x_{t-1}) \prod_{t=1}^{n} f(\varepsilon_t | x_t) \\ &= \frac{1}{\sqrt{2\pi}\delta_0} \exp\left(\frac{-\left(x_0 - \mu_0\right)^2}{2\delta_0^2}\right) \prod_{t=1}^{n} \frac{1}{\sqrt{2\pi}Q^{(R_t)}} \exp\left(\frac{-\left(x_t - \beta_0^{(R_t)} - \beta_1^{(R_t)} x_{t-1}\right)^2}{2\left(Q^{(R_t)}\right)^2}\right) \\ &\quad \times \prod_{t=1}^{n} \frac{1}{\sqrt{2\pi}Q^{(R_t)}} \exp\left(\frac{\varepsilon_t^2 \exp\left(-x_t\right)}{2}\right) \exp\left(\frac{-x_t}{2}\right). \end{aligned}$$

Thus, the complete log-likelihood function is given by

$$
\begin{aligned}
-2\log \mathbf{L}(\theta;\underline{\varepsilon}) = {} & C + \log \delta_0^2 + \frac{(x_0-\mu_0)^2}{\delta_0^2} \\
& + \sum_{k=0}^{1}\sum_{t=1}^{n}\left(2\log Q^{(k)} + \left(\frac{x_t-\beta_0^{(k)}-\beta_1^{(k)}x_{t-1}}{Q^{(k)}}\right)^2\right)\mathbf{1}_{(R_t=k)} \\
& + \sum_{t=1}^{n}\left(\varepsilon_t^2\exp\{-x_t\}+x_t\right)\mathbf{1}_{(R_t=k)},
\end{aligned}
$$

where C is a constant independent of θ.

Since the volatility is unobservable, the estimation procedure requires missing data treatment techniques. We call therefore here the EM algorithm combined with particle filters and smoothers (for more details see Kim and Stoffer, 2008; Boussaha and Hamdi, 2017; Boussaha et al. 2018). The expected conditional complete data likelihood defined in E-step is given by $\mathbf{Q}$ as follows,

$$
\begin{aligned}
\mathbf{Q}\left(\theta,\widehat{\theta}^{(i-1)}\right) &= \mathbb{E}\left[-2\log \mathbf{L}\left(\theta;\underline{\varepsilon}\right)\Big|\varepsilon,,\widehat{\theta}^{(i-1)}\right] \\
&= C + 2\sum_{k=0}^{1}\sum_{t=1}^{n}\log Q^{(k)}\mathbf{1}_{(R_t=k)} \\
&\quad + \sum_{k=0}^{1}\sum_{t=1}^{n}\left\{\frac{\left(x_t^{(n)}-\beta_0^{(k)}-\beta_1^{(k)}x_{t-1}^{(n)}\right)^2+P_t^{(n)}+\left(\beta_1^{(k)}\right)^2P_{t-1}^{(n)}-2\beta_1^{(k)}P_{t,t-1}^{(n)}}{\left(Q^{(k)}\right)^2}\right\}\mathbf{1}_{(R_t=k)} \\
&\quad + \sum_{k=0}^{1}\sum_{t=1}^{n}\mathbb{E}\left[\exp\{-x_t\}\Big|\varepsilon,\widehat{\theta}^{(i-1)}\right]\mathbf{1}_{(R_t=k)} + \sum_{t=1}^{n}x_t^{(n)}, \qquad (2.4)
\end{aligned}
$$

where

$$
\begin{aligned}
x_t^{(n)} &= \mathbb{E}\left(x_t\Big|\varepsilon,,\widehat{\theta}^{(i-1)}\right), \\
P_t^{(n)} &= \mathbb{E}\left(\left(x_t-x_t^{(n)}\right)^2\Big|\varepsilon,,\widehat{\theta}^{(i-1)}\right),
\end{aligned}
$$

and

$$
P_{t,t-1}^{(n)} = \mathbb{E}\left(\left(x_t-x_t^{(n)}\right)\left(x_{t-1}-x_{t-1}^{(n)}\right)\Big|\varepsilon,,\widehat{\theta}^{(i-1)}\right).
$$

Before we go to M-step, we have to evaluate the quantities $x_t^{(n)}$, $P_t^{(n)}$, $P_{t,t-1}^{(n)}$ and $\mathbb{E}\left[\exp(-x_t)|\,\varepsilon,\theta^{(i-1)}\right]$. To deal with this, we use the following Particle Filter Algorithm:

Algorithm 1 Particle filter algorithm

1. Initialisation: sample $f_0^{(j)} \sim p_0(x_0)$ with $w_0^{(j)} = 1/M$, $j = \overline{1, M}$, where M is the number of the particles.
2. At time $t \geq 1$
   (a) for $j = \overline{1, M}$
      i. Sample $e_t^{(j)} \sim \mathcal{N}(0, 1)$.
      ii. Calculate $p_t^{(j)} = \beta_0^{(R_t)} + \beta_1^{(R_t)} f_{t-1}^{(j)} + Q^{(R_t)} e_t^{(j)}$.
      iii. Compute the importance weights:
      $w_t^{(j)} \propto w_{t-1}^{(j)} \exp\left(-\frac{\exp\left(-p_t^{(j)}\right)\varepsilon_t^2}{2}\right) \exp\left(-p_t^{(j)}/2\right)$.
   (b) Normalise the importance weights.
   (c) Compute the measure of degeneracy $\left(n^{eff} = 1 \Big/ \sum_{j=1}^{M} \widetilde{w}_t^{(j)2}\right)$.
   If $n^{eff} \leq n^T$(typically $n^T = M/2$), resample with replacement M equally weighted particles $\left\{f_t^{(j)}\right\}_{j=1}^{j=M}$ from the set $\left\{p_t^{(j)}\right\}_{j=1}^{j=M}$ according to the normalised weights.
   Else $f_t^{(j)} = p_t^{(j)}, j = \overline{1, M}$.
3. Assign the particle $\left\{f_t^{(j)}\right\}_{j=1}^{j=M}$ a weight $\left\{\widetilde{w}_t^{(j)}\right\}_{j=1}^{j=M}$.

The following particle smoothing algorithm, from which, we take into account the information available after time t, gives approximations of the quantities $x_t^{(n)}$, $P_t^{(n)}$, $P_{t,t-1}^{(n)}$ and $\mathbb{E}\left[\exp\left(-x_t\right) \middle| \varepsilon, \widehat{\theta}^{(i-1)}\right]$, which are necessary in E-step.

Algorithm 2 Particle smoothing algorithm

1. For $j = \overline{1, M}$, choose $s_n^{(j)} = f_n^{(i)}$, with probability $\widetilde{w}_n^{(i)}$ and set $W_n^{(j)} = 1/M$.
2. For $j = \overline{1, M}$, for $t = n, ..., 0$,
   (a) Calculate the smoothed weights for $i = \overline{1, M}$
   $$W_{t-1|t}^{(i)} = \widetilde{w}_{t-1}^{(i)} \exp\left(\frac{-\left(s_t^{(j)} - \beta_0^{(R_t)} - \beta_1^{(R_t)} f_{t-1}^{(i)}\right)^2}{2\left(Q^{(R_t)}\right)^2} \right),$$
   and normalise them.
   (b) Choose $s_{t-1}^{(j)} = f_{t-1}^{(i)}$, with the normalised smoothed weights $W_{t-1|t}^{(i)}$.
3. Finally, we can obtain the EM estimates by using the sample averages of the functions of particle smoothers and which are needed for step M. For example,
   $$\widehat{P}_{t,t-1}^{(n)} = \frac{\sum_{j=1}^{M} \left(s_t^{(j)} - \widehat{x}_t^{(n)}\right)\left(s_{t-1}^{(j)} - \widehat{x}_{t-1}^{(n)}\right)}{M}, \; t = \overline{1, n}.$$

After replacing $x_t^{(n)}$, $P_t^{(n)}$ and $P_{t,t-1}^{(n)}$ by their approximations, we then go to M-step , where $\widehat{\theta}^{(i)}$ is obtained by
$$\widehat{\theta}^{(i)} = \arg\max_{\theta} Q\left(\theta, \widehat{\theta}^{(i-1)}\right).$$

The first-order derivatives of equation (2.4), with respect to the unknown parameters, for $k = 0, 1$, are given by

$$\begin{aligned}
\frac{\partial \mathbf{Q}\left(\theta,\widehat{\theta}^{(i-1)}\right)}{\partial \beta_1^{(k)}} &= \frac{-2}{\left(Q^{(k)}\right)^2} \sum_{t=1}^{n} \left[x_{t-1}^{(n)}\left(x_t^{(n)} - \beta_0^{(k)}\right) + P_{t,t-1}^{(n)} - \beta_1^{(k)}\left(\left(x_{t-1}^{(n)}\right)^2 + P_{t-1}^{(n)}\right)\right] \mathbf{1}_{(R_t=k)}, \\
\frac{\partial \mathbf{Q}\left(\theta,\widehat{\theta}^{(i-1)}\right)}{\partial \beta_0^{(k)}} &= \frac{-2}{\left(Q^{(k)}\right)^2} \sum_{t=1}^{n} \left(x_t^{(n)} - \beta_0^{(k)} - \beta_1^{(k)} x_{t-1}^{(n)}\right) \mathbf{1}_{(R_t=k)}, \\
\frac{\partial \mathbf{Q}\left(\theta,\widehat{\theta}^{(i-1)}\right)}{\partial Q^{(k)}} &= \frac{2S_k}{Q^{(k)}} - \frac{2\sum_{t=1}^{n}\left(x_t^{(n)} - \beta_0^{(k)} - \beta_1^{(k)} x_{t-1}^{(n)}\right)^2 \mathbf{1}_{(R_t=k)}}{\left(Q^{(k)}\right)^3} \\
&\quad - \frac{2\sum_{t=1}^{n}\left\{P_t^{(n)} + \left(\beta_1^{(k)}\right)^2 P_{t-1}^{(n)} - 2\beta_1^{(k)} P_{t,t-1}^{(n)}\right\}\mathbf{1}_{(R_t=k)}}{\left(Q^{(k)}\right)^3},
\end{aligned}$$

where $S_k = \sum_{t=1}^{n} \mathbf{1}_{(R_t=k)}$.

Therefore, solution of the M-step exists in a closed form, so at the i^{th} iteration we have

$$\begin{cases} \widehat{\beta}_1^{(k,i)} = \dfrac{\left(\sum\limits_{t=1}^{n} x_t^{(n)}\mathbf{1}_{(R_t=k)}\right)\left(\sum\limits_{t=1}^{n} x_{t-1}^{(n)}\mathbf{1}_{(R_t=k)}\right) \sum\limits_{t=1,r_t=k}^{n}\left(x_t^{(n)}x_{t-1}^{(n)}+P_{t,t-1}^{(n)}\right)}{\sum\limits_{t=1}^{n}\left(\left(x_{t-1}^{(n)}\right)^2+P_{t-1}^{(n)}\right)\mathbf{1}_{(R_t=k)}-\frac{1}{S_k}\left(\sum\limits_{t=1}^{n} x_{t-1}^{(n)}\mathbf{1}_{(R_t=k)}\right)^2}, \\ \widehat{\beta}_0^{(k,i)} = \frac{1}{S_k}\left[\sum\limits_{t=1}^{n} x_t^{(n)}\mathbf{1}_{(R_t=k)} - \widehat{\beta}_1^{(k,i)}\sum\limits_{t=1}^{n} x_{t-1}^{(n)}\mathbf{1}_{(R_t=k)}\right], \\ \widehat{Q}^{(k,i)} = \left[\frac{1}{S_k}\sum\limits_{t=1}^{n}\left\{\left(x_t^{(n)} - \widehat{\beta}_0^{(k,i)} - \widehat{\beta}_1^{(k,i)}x_{t-1}^{(n)}\right)^2 + P_t^{(n)} + \left(\widehat{\beta}_1^{(k,i)}\right)^2 P_{t-1}^{(n)} - 2\widehat{\beta}_1^{(k,i)}P_{t,t-1}^{(n)}\right\}\mathbf{1}_{(R_t=k)}\right]^{1/2}. \end{cases}$$

2.4 Simulation Study

In order to assess the quality of our proposed estimation procedure, a simulation study is conducted in this section. The simulations are replicated 1000 times in the first two experiments with different sample sizes $n = 250, 500, 750$ and 1000. Without loss of generality, the corresponding parameter values were chosen so that $|\beta_1| < 1$ and $|\beta_2| < 1$ to avoid a nonstationary and an explosive autoregressive regimes. In the last experiment, to start the EM algorithm, we opted to use a basic SV modelling. Once we obtain convergent estimators for the basic SV model, we duplicate them for the two regimes to initialize the EM algorithm for BSV modelling. We have carried out a lot of simulation experience in this direction, they give similar findings, so we have chosen to report only the one presented in Table 2.3. the simulation is replicated 100 times with sample sizes $n = 250$. We note here that in each case we used a sequence of $M = 200$ particles.

Table 2.1. Results of a simulation study for BSV with $r_L = -0.3$ and $r_U = 0.3$

		$n = 250$		$n = 500$		$n = 750$		$n = 1000$	
	TV	MLE	Sd-MLE	MLE	Sd-MLE	MLE	Sd-MLE	MLE	Sd-MLE
$\beta_0^{(1)}$	1	1.0822	0.2815	1.0518	0.1874	1.0255	0.1464	1.0156	0.1296
$\beta_0^{(0)}$	0.5	0.5682	0.4108	0.5418	0.2739	0.5208	0.2293	0, 5156	0.2099
$\beta_1^{(1)}$	0.9	0.8836	0.0500	0.8908	0.0340	0.8956	0.0265	0.8958	0.0213
$\beta_1^{(0)}$	0.6	0.5763	0.0809	0.5907	0.0549	0.5926	0.0444	0.5934	0.0370
$Q^{(1)}$	1	1.0669	0.2228	1.0483	0.1508	1.0360	0.1182	1.0140	0.0914
$Q^{(0)}$	4	3.9676	0.3281	3.9652	0.2367	3.9478	0.1854	3.9573	0.1284

Table 2.2. Results of a simulation study for BSV with $r_L = 0.15$ and $r_U = 0.35$

		$n = 250$		$n = 500$		$n = 750$		$n = 1000$	
	TV	MLE	Sd-MLE	MLE	Sd-MLE	MLE	Sd-MLE	MLE	Sd-MLE
$\beta_0^{(1)}$	$1,2$	$1,3143$	$0,3650$	$1,2832$	$0,2642$	$1,2640$	$0,1792$	$1,2781$	$0,1557$
$\beta_0^{(0)}$	$0,8$	$0,9035$	$0,3627$	$0,8364$	$0,2231$	$0,8173$	$0,2140$	$0,7981$	$0,1605$
$\beta_1^{(1)}$	$0,8$	$0,7667$	$0,1085$	$0,7734$	$0,0794$	$0,7775$	$0,0576$	$0,7773$	$0,0479$
$\beta_1^{(0)}$	$0,5$	$0,4605$	$0,1080$	$0,4809$	$0,0717$	$0,4927$	$0,0700$	$0,4965$	$0,0514$
$Q^{(1)}$	1	$1,0459$	$0,1724$	$0,9987$	$0,1145$	$0,9924$	$0,0900$	$1,0041$	$0,0896$
$Q^{(0)}$	1	$0,9798$	$0,1486$	$0,9827$	$0,0987$	$0,9922$	$0,8192$	$0,9924$	$0,0850$

Table 2.3. Results of a simulation study for BSV with $r_L = -0.3$ and $r_U = 0.3$

	$\beta_0^{(1)}$	$\beta_0^{(0)}$	$\beta_1^{(1)}$	$\beta_1^{(0)}$	$Q^{(1)}$	$Q^{(0)}$
TV	1	0.5	0.9	0.6	1	4
MLE	$1,0920$	$0,5913$	$0,8820$	$0,5749$	$1,0532$	$3,9814$
Sd-MLE	$0,3055$	$0,4500$	$0,0575$	$0,0905$	$0,3071$	$0,3697$

From Tables 2.1-2.3, it can be seen, in general, that the means of the estimates are very close to the true parameter values and the bias is of small magnitude even for small sample sizes. This indicates that the proposed estimation procedure can provide stable estimators around true values and once the size increases, the bias in turn decreases which ensures the empirical convergence towards the true parameters values.

2.5 Empirical Study

To showcase the contribution of BSV model in practice to the financial time series analysis, we fitted the BSV model to the daily prices, adjusted for dividends and splits, of Honeywell International Inc (HON). This series represents one of the thirty components of the second oldest U.S. market index Dow Jones Industrial Average (DJIA) index. The period of analysis is taken from June 11, 1990 to September 12, 2006, and we have therefore 4100 observations in our studied sample. This time series has been studied by Lo et al. (2016) where the $BGARCH$ model has been proposed and that is what motivated basically our choice of this series. Note that this series has been also studied by Caiado and Crato (2010) where they showed by the Ljung-Box test to have no significant serial correlation effect. Here, the analysis is based on the sequence of log-returns in percentage defined as $r_t = 100 (\ln P_t - \ln P_{t-1})$ where P_t is the closing price on day t. The P_t and r_t are shown in Figure 2.1. The data are collected from Yahoo-Finance.

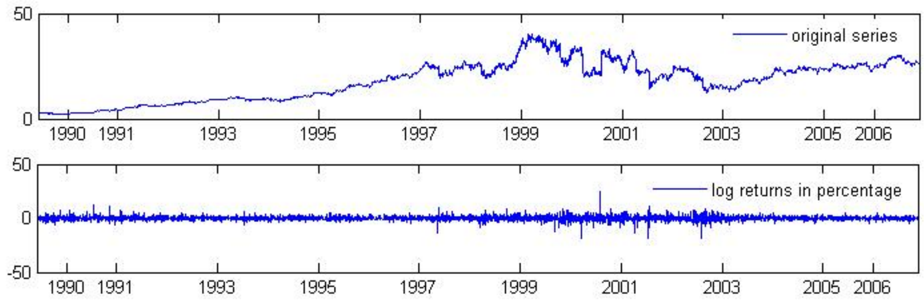

Figure 2.1: HON series and sequence of log-returns in percentage

Table 2.4 presents the estimation results corresponding to the estimation procedure discussed in the previous section. For comparison issues, this section considers also a basic SV, TSV and $BGARCH$ modelling. We also report the results of estimation of these three different modelling in the same table. For the estimated parameters of the $BGARCH$ model, we take those obtained by Lo et al. (2016) , and they are indicated by * in the Table.

Table 2.4. Parameter estimates ($M = 200$ and $nbre\_it\_EM = 200$)

	SV	TSV	BSV		$BGARCH$
		$\gamma = 0$	$r_L = -1.3209$ $r_U = 1.3212$		$r_L = -1.3209$ $r_U = 1.3212$
$\beta_0^{(1)}$	0.1940	0.1050	0.1164	ω_1^*	0.2840
$\beta_0^{(0)}$	–	0.0728	0.0552	ω_2^*	0.1384
$\beta_1^{(1)}$	0.7996	0.9673	0.9473	α_1^*	0.1647
$\beta_1^{(0)}$	–	0.8637	0.8919	α_2^*	0.0772
$Q^{(1)}$	0.5864	0.3002	0.3466	β_1^*	0.8353
$Q^{(0)}$	–	0.3802	0.3123	β_2^*	0.8366

The threshold value γ in TSV model was chosen to be zero assuming that there exists an asymmetric behavior in the volatility process characterizing the distinctive effects of news whether they are good or bad. This point is considered in quite a few papers (So et al., 2002; So et al., 2009). Other works suppose that the threshold value is not equal to zero, which implies its estimation (Xu, 2012).

As previously mentioned, in many empirical works, we cannot deduce the advent of good news only if the price is positive, and if the price is negative we deduce the advent of bad news; there is no reason to impose such strong assumption on all the financial data. These works allow to the threshold value to be flexible with respect of the nature of data and the specificity of the financial series under study. As already argued, this also reinforces the introduction of the buffer zone that we take it containing the zero threshold value, and which is not necessarily symmetrical by reference to zero. The buffering process will encompass both zero-threshold and flexible-threshold cases and which will allow, as described above, a smooth transition between regimes.

Concerning the boundary parameters, r_L and r_U, of the buffer zone, we have taken the one obtained by Lo et al. (2016) to allow a possible comparison between BSV and $BGARCH$. They choose the values of boundary parameters from the 15^{th} to the 85^{th} empirical percentiles of observations and use the BIC criterion to identify the best setting of all model parameters. According to this choice of r_L and r_U, and despite this zone constitutes only $\frac{r_U - r_L}{\max\limits_{t>1}(r_{t-1}) - \min\limits_{t>1}(r_{t-1})} = 0.0601$ of the amplitude of the interval to which the threshold variables belong, we have more than half (2492) of the observations in the buffer zone.

Although the estimated parameters are close together for TSV and BSV both, the interpretation of the two resulting fitted models differ from a philosophical point of view. We recall that the two-regime TSV model for financial return, attempts to capture the advent of good or bad news with respect to zero threshold value. In addition, the BSV model is designed to explain more asymmetric features which requires relaxation of the constraint of symmetry with respect to zero. The BSV does not only react according to the nature of the news. Then, it is clear that the introduction of the buffer zone in the BSV model induces a change in the assignment of observations to regimes for threshold-based SV models. Indeed, 1003 observations from the buffer zone were assigned to the first regimes under the fitted BSV, which can be seen as a reaction to the good news. These same observations correspond to the advent of bad news according to the TSV, since these observations were identified in the second regime under the fitted TSV. We find only 54 observations in the buffer zone which are in the first regime according to both models.

It is worth mentioning that from Table 2.4, classical variants of SV model check the sufficient condition of strictly stationarity. For BSV, each autoregressive component is also stationary. In the case of TSV model, we have $\left|\widehat{\beta}_1^{(1)}\right|^{\widehat{q}} \left|\widehat{\beta}_1^{(0)}\right|^{1-\widehat{q}} = 0.9111$ where $\widehat{q} = \frac{1935}{4099} = 0.4721$ represents, empirically, the probability of belonging to the first regime under the fitted TSV model. Note that this stationarity condition becomes sufficient and necessary in the basic SV modelling framework where $\left|\widehat{\beta}_1^{(1)}\right| = 0.7996$.

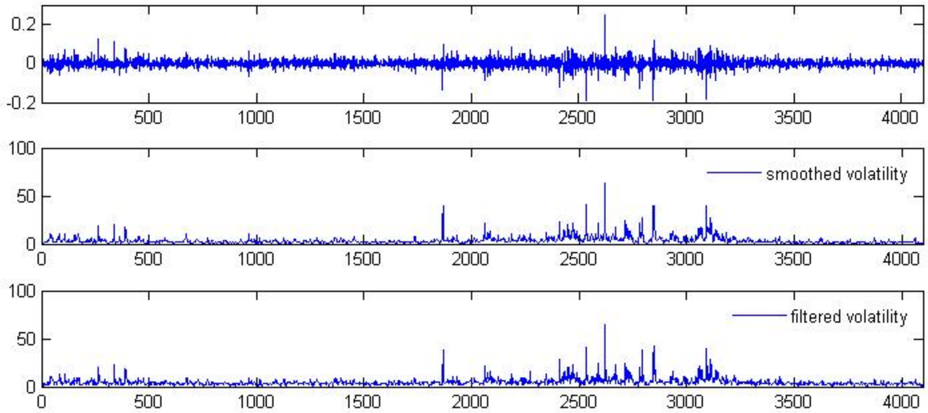

Figure 2.2: Estimated volatility from fitted SV model.

It should also be indicated that estimates of the persistence coefficients $\widehat{\beta}_1^{(1)}$ and $\widehat{\beta}_1^{(0)}$ in the volatility equation for BSV and TSV models are greater than these calculated for the basic SV model. Even more interesting is to observe that the corresponding $BGARCH$ estimates do not satisfy the strict stationarity constraint; the estimated sum $\max\{\alpha_1^*, \alpha_2^*\} + \max\{\beta_1^*, \beta_2^*\} = 1.0013$; exceeds slightly unity. This finding for the $GARCH$-family modelling has already been reported in other works (Teräsvirta, 2009). It may be appropriate to prove that even if the generating solution is in the nonstationary region, it is not yet explosive (see Regnard and Zakoïan, 2010). It is important to point out that the persistence of volatility implied by fitted $BGARCH(1,1)$ model is greater than that implied by the different variants of SV model. It should be underlined that this result is in line with similar findings in the case of SV versus $GARCH(1,1)$ and $EGARCH$ (see e.g., Carnero et al., 2004). Indeed, the persistence is one of the three elements that Carnero et al. (2004) used to analyze the relationship between the models SV and $GARCH$ where they conclude that the basic SV model is more flexible than the $GARCH(1,1)$ model.

Note that point estimate of the volatility of the three variants of the SV model can be constructed by using the filter and smoother algorithms. The volatility obtained is called filtered volatility and smoothed volatility respectively. Figures 2.3-2.7 plot the estimated volatilities for all models. The estimated volatilities displayed in Figures 2.2-2.5 show a prima facie the differences among them are smaller. Indeed, the estimated volatilities resemble the variation of the log-returns, in particular, the volatility computed from the fitted BSV and TSV (see Figures 2.6-2.7). Moreover, Figure 2.7 reveals that a large number of log-returns are outside the interval $[\overline{r} - 2\widehat{\sigma}_t, \overline{r} + 2\widehat{\sigma}_t]$ as part of the fit with the

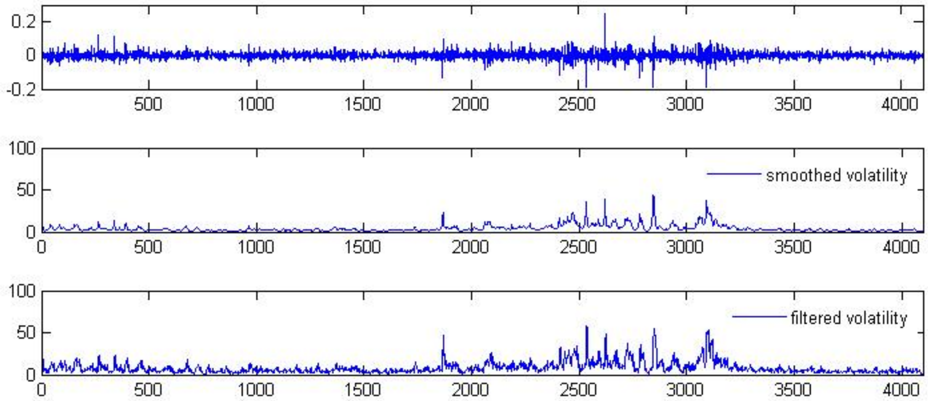

Figure 2.3: Estimated volatility from fitted TSV model.

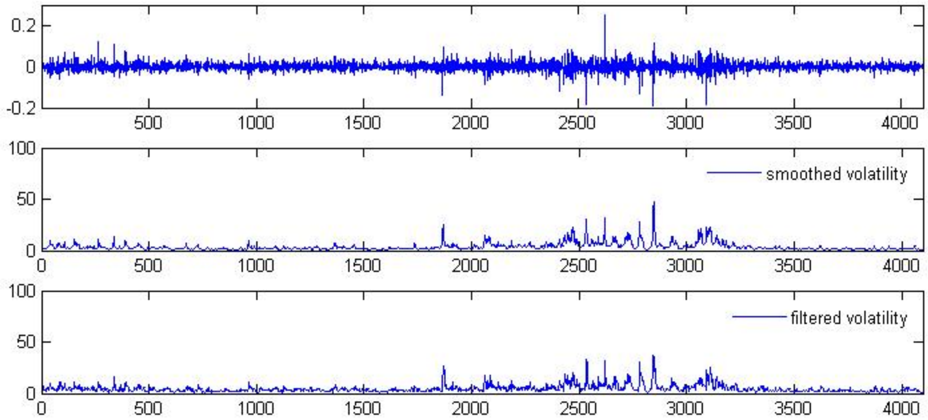

Figure 2.4: Estimated volatility from fitted BSV model.

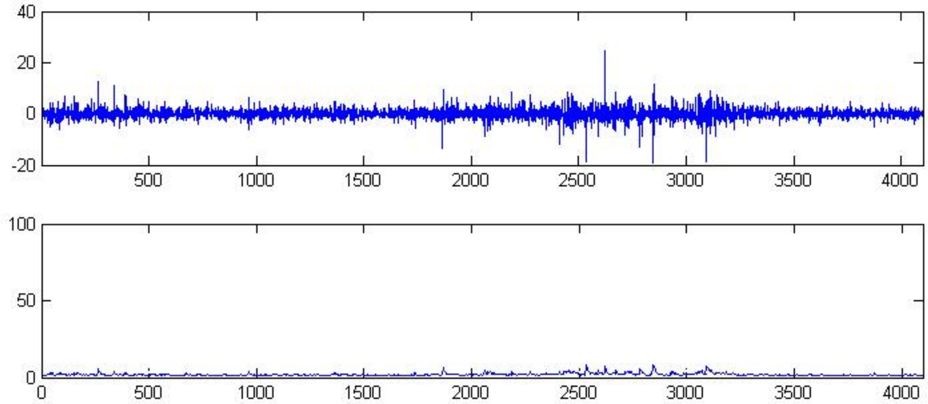

Figure 2.5: Estimated volatility from fitted $BGARCH$ model.

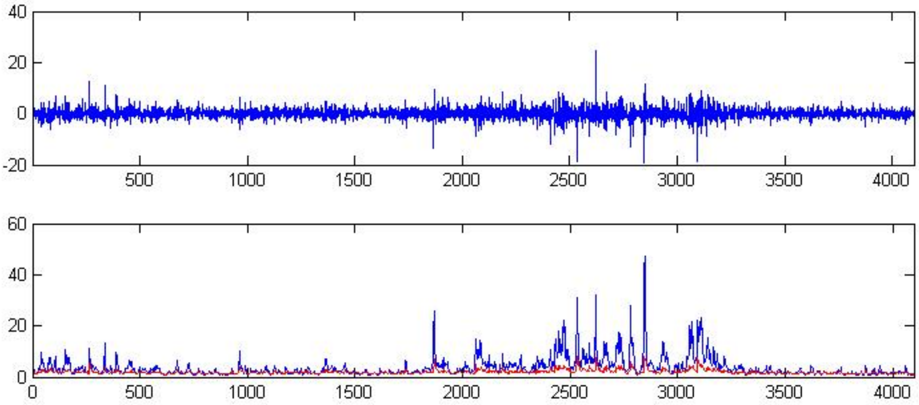

Figure 2.6: Estimated volatilities from fitted $BGARCH$ (red) and BSV (blue).

$BGARCH$ model.

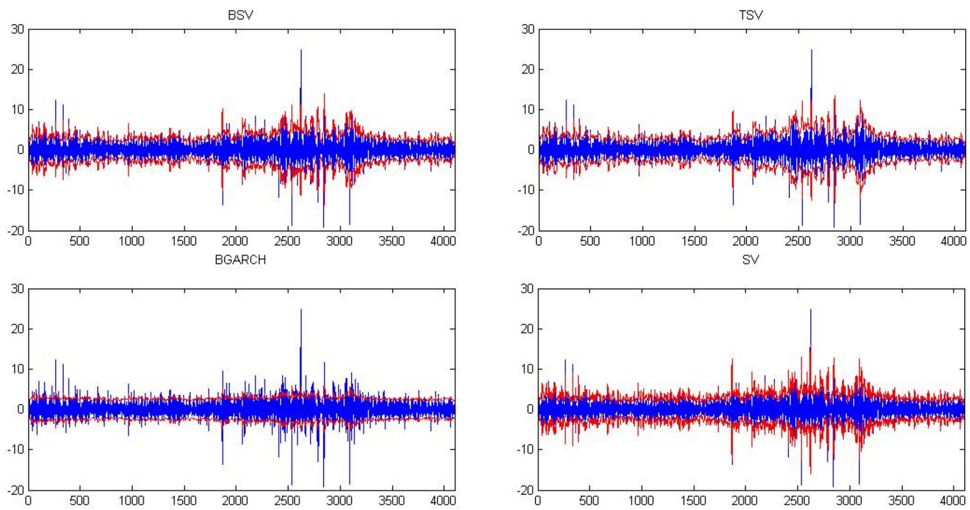

Figure 2.7: Returns r_t and confidence intervals $\overline{r} \mp 2\widehat{h}_t$ ($\overline{r} \mp 2\widehat{\sigma}_t$) where $\overline{r}$ is the empirical mean of the returns, $\widehat{h}_t$ is the smoothed volatility of SV approaches and $\widehat{\sigma}_t$ is the estimated volatility in the $BGARCH(1,1)$.

Table 2.5 confirms the results of Figure 2.7, where the empirical coverages of the prediction intervals based on BSV modelling are closer to the nominal coverages than obtained from other models.

Table 2.5. Empirical coverages of the $(1-\alpha)100\%$ one-step-ahead prediction intervals using smoothed volatility.

	50%	60%	70%	80%	90%	95%	99%
SV	48.22	59.32	**70.96**	83.14	94.12	98.63	99.83
TSV	52.46	62.81	73.11	83.19	93.07	**96.90**	99.63
BSV	**50.20**	**60.86**	71.11	**81.45**	**91.61**	97.00	**99.56**
$BGARCH$	43.22	52.44	61.91	71.16	81.97	88.29	95.14

Globally, from the analysis presented above, the BSV model gives a better fit to the dataset than the SV, TSV and $BGARCH(1,1)$. This is mainly due to the flexibility of the transition mechanism which allows a better highlighting of the non-linear characteristics exhibited in the HON series.

Chapter 3

Nonlinearity in Panel Data: A new Threshold Panel Data Model

3.1 Introduction

As previously stated, threshold models are recognized as a useful tool for time series analysis since their introduction by Tong (1978). Indeed, their structure has been used by several researchers, especially by econometricians. This approach gives a sophistical way to take into account the nonlinearity exhibited by several financial and macroeconomic phenomena. This is done by providing a simplified formulation to mimic these nonlinear stylized facts and more precisely the regime changes in their dynamics.

Many extensions and mathematical developments for threshold models have been adopted for the analysis of other data structures, in particular the panel data treatment framework. Indeed, Hansen (1999*a*) proposed a Threshold Regression model for the non-dynamic panels case. The main contribution of Hansen's model lies in the possibility of allowing the individuals constituting the panel to be in different regimes for a given period. This enables a better capture of the heterogeneity in the panel and visualizing of the nonlinearity in the interaction between the dependent variable and the explanatory variables for each panel's component. From a technical point of view, Hansen (1999*a*) has proposed a Least Squares Estimation (LSE) of the parameters and a non-standard asymptotic theory for inference. Moreover, he applied his model to study whether financial constraints affect investment decisions in a 15-year sample of 565 American firms.

However, the sudden change of regime characterizing the Hansen's formulation may be problematic in some situations in which the transition is smooth. In order to capture the absence of sudden jumps, Gonzalez et al. (2017) develop a non-dynamic Panel Smooth Transition Regression ($PSTR$) model with individual fixed effects. The parameters are allowed to change smoothly as a function of a threshold variable. The performance of this model may depend on the choice of the transition function according to the studied phenomenon. This modelling turns out to be useful when the number of regimes is sufficiently high and from a theoretical point of view, there exist an infinity of regimes in this modelling lying between extreme regimes.

In some circumstances, an interesting phenomenon happens when a past temporary change of a relevant forcing variable lead to a change in the economic behavior of the analyzed variable but a removal to the initial value of the forcing variable does not induce a return to the initial behavior (i.e. there is dependence of the state of a system on its history). The so called hysteresis phenomenon, originally stemming from physics, has been widely used in labor theory and foreign trade to explain the persistence effects of temporary stimulus (see e.g., Göcke, 2002). For example, in foreign trade, temporary exchange rate shocks could induce persistent consequences on quantities and prices due to sunk market-entry costs. Indeed, in order to sell in a foreign market a firm incurs some entry costs that cannot be recovered after exit (e.g., distribution and service networks). If the domestic currency temporarily depreciates, entering this foreign market becomes profitable for some domestic firms. But even if the exchange rate regains its initial level, it is still profitable to sell in the foreign market when the variable costs are recovered. This simple microeconomic hysteresis could thus be aggregated and gives rise to a continuous macro loop of overall exports (Borgersen and Göcke, 2007). This effect has been widely documented in the empirical literature (see for example Belke and Kronen, 2019 for a recent study of the four main EU exporting countries).

This smooth switching between different equilibria (finite-configurations or finite-states) of the studied variable (system) could thus be usefully mobilized in order to analyze the dynamics of its evolution. This is why we do propose, in this chapter, an alternative model, based on this idea of hysteresis, by defining a new smooth and flexible regime switching mechanism. Our definition is inspired from the buffered process, discussed in the previous chapter, and proposed by Li et al. (2015). This approach makes the dynamics of transition more smooth and flexible compared to the classical Threshold Regression model. Even though this idea is still in its infancy, it provides a new way to understand and explain the nonlinearity observed in data. As we saw in the previous chapter, it is worth noting that all applications of the buffered process have been done for time series. It seems quite natural to extend this approach to situations in which there are sudden jumps in the dynamics of a phenomenon besides an individual dimension. This is the purpose of our contribution through this chapter.

The remainder of the chapter is organized as follows. In Section 3.2, we provide a precise definition of our Buffered Threshold Panel Data ($BTPD$) model. In Section 3.3, a Least Squares Estimation method is established for the slope coefficients and its steps are detailed. The inference theory for our proposed model is developed in Section 3.4. It turns out that for a better use of our model, it is necessary to develop adequation tests in order to highlight the nonlinear patterns in the studied data and to measure the statistical significance of the buffer zones effects. More precisely, we develop a likelihood ratio test for testing linearity against the $BTPD$ model and the number of regimes. Since the distribution of this test is non standard, a bootstrap procedure is proposed to simulate the likelihood ratio test distribution. Section 3.5 proposes a simulation study of the finite sample properties of the proposed procedures (estimation and test) and confirms their consistency and adequacy. Section 3.6 is devoted to an empirical application dealing with the combined interaction effects of natural resource dependence and the quality of institutions on economic growth in rentier countries. We more precisely show that our $BTPD$ modelling gives better results than PTR and $PSTR$ modellings.

3.2 Buffered Threshold Panel Data Model

We consider the following balanced[1] panel $\{Y_{i,t}, q_{i,t}, X_{i,t} : 1 \leq i \leq N, 1 \leq t \leq T\}$ where i and t respectively denote the individual and temporal indices. We do consider a univariate context in which the dependent variable $Y_{i,t}$ and the threshold variable $q_{i,t}$ are scalars. $X_{i,t} = \left(X_{i,t}^{(1)}, X_{i,t}^{(2)}, ..., X_{i,t}^{(m)}\right)$ is a $(1 \times m)$-vector of exogenous explanatory variables. We first present the case of two regimes. The observed data are generated from a non-dynamic two-regime $BTPD$ model with fixed effects if they satisfy the following regression model

$$Y_{i,t} = \mu_i + X_{i,t}\beta_1 R_{i,t} + X_{i,t}\beta_2(1 - R_{i,t}) + \varepsilon_{i,t}, \tag{3.1}$$

[1]The panel is balanced when all the individuals constituting it have the same time dimension.

where μ_i is the fixed effect of individual i and $\varepsilon_{i,t}$ is the error term which is assumed to be independent and identically distributed with zero mean and a positive finite variance σ^2. β_k, $k = 1, 2$, represent the slope coefficients. $R_{i,t}$ is the regime indicator defined as follows,

$$R_{i,t} = \begin{cases} 1 & \text{if } q_{i,t} \leq r_L, \\ R_{i,t-1} & \text{if } r_L < q_{i,t} \leq r_U, \\ 0 & \text{if } r_U < q_{i,t}, \end{cases}$$

where r_L and r_U ($r_L \leq r_U$) are the boundary parameters which constitute the buffer or hysteresis zone. In model (3.1), the transition mechanism is modeled in the same way as for the buffered threshold of time series (see Li et al., 2015). The originality of model (3.1) lies in the representation of a panel by several distinct regimes. Each regime is characterized by a linear dynamics, in which the transition between a finite number of different regimes is gradual rather than abrupt. This proposal is justified by the fact that in practice the regime $R_{i,t}$ may not shift immediately, and there could be a buffering region in which the regime of $Y_{i,t}$ depends on the regime of $Y_{i,t-1}$. In addition, the contribution of the individual dimension in the smooth transition mechanism is to group in the same regime, and at a given date, the individuals following the same linear process. However, the individuals composing this regime are likely to evolve smoothly over time, because the transition variable also depends on the temporal dimension.

Note that from formulation (3.1), we can obtain the classical panel threshold regression model proposed by Hansen (1999a) as a special case by setting $r_L = r_U = r$. However, it is worth mentioning that the above formulation differs from that of Hansen (1999) in which the transition is brutal since an individual can switch from one regime to another if the status of the threshold variable $q_{i,t}$ crosses up or down the threshold r. More precisely, if the threshold variable becomes smaller than the threshold ($q_{i,t} \leq r$), even so slightly, the process is described by the first defined regime with slope coefficients β_1. Conversely, when $q_{i,t} > r$ the process is described by the second regime with slope coefficients β_2. In other words, the regime indicator $R_{i,t}$ only depends on whether the threshold variable value is smaller or larger than r. In our proposed model, the regime indicator may depend on the past of the regime indicators infinitely far away. Indeed, in the $BTPD$ model, $R_{i,t}$ can be written as follows in the almost surely sense,

$$\begin{aligned} R_{i,t} &= \mathbf{1}_{(q_{i,t} \leq r_L)} + \mathbf{1}_{(r_L < q_{i,t} \leq r_U)} R_{i,t-1} \\ &= \mathbf{1}_{(q_{i,t} \leq r_L)} + \sum_{j=0}^{\infty} \prod_{m=0}^{j} \mathbf{1}_{(r_L < q_{i,t-m} \leq r_U)} \mathbf{1}_{(q_{i,t-j-1} \leq r_L)}. \end{aligned}$$

As already mentioned, an alternative approach to account for possible smooth and gradual transitions is the $PSTR$ model developed by Gonzàlez et al. (2017). Unlike our $BTPD$ model, the $PSTR$ model can be used to allow for a continuum of regimes, each one being characterized by a different value of the transition variable. The regression coefficients are bounded continuous functions of the transition variable. They thus fluctuate between extreme regimes.

We should note that this two-regime $BTPD$ model can be extended to a general K-regime model. Indeed, suppose there are K regimes and $K-1$ corresponding buffer zones $(r_{L,k}, r_{U,k}]$ with $k = 1, ..., K-1$, and

$$-\infty = r_{U,0} < r_{L,1} \leq r_{U,1} < r_{L,2} \leq r_{U,2} < \cdots < r_{L,K-1} \leq r_{U,K-1} < r_{L,K} = +\infty.$$

Then, the K-regime $BTPD$ model can be defined as follows

$$Y_{i,t} = \mu_i + \sum_{k=1}^{K} X_{i,t}\beta_k \mathbf{1}_{(R_{i,t}=k)} + \varepsilon_{i,t}, \tag{3.2}$$

where $\mathbf{1}_A$ is the indicator function of the set A and the regime indicator $R_{i,t}$ divides the observations into K regimes characterized by different regression slopes β_k, $k = 1, ..., K$. Note that the regime indicator $R_{i,t-1}$ is equal to k if $r_{U,k-1} < q_{i,t-1} \leq r_{L,k}$, and the data-generating process (DGP) will stay in the same regime at time t if the threshold variable $q_{i,t}$ falls into the buffer zones $(r_{L,k}, r_{U,k}]$ or $(r_{L,k-1}, r_{U,k-1}]$. When the threshold variable $q_{i,t}$ falls into the buffer zone $(r_{L,l}, r_{U,l}]$ with $l > k$ or $l < k-1$, the regime indicator $R_{i,t}$ is set respectively to l or $l+1$ (see Figure 1). More generally, we have

$$R_{i,t} = \begin{cases} l & \begin{array}{l} \text{if } (r_{U,l-1} < q_{i,t} \leq r_{L,l}) \\ \text{or } (r_{L,l} < q_{i,t} \leq r_{U,l},\ R_{i,t-1} = k \text{ and } l \geq k), \end{array} \\ l+1 & \text{if } r_{L,l} < q_{i,t} \leq r_{U,l},\ R_{i,t-1} = k \text{ and } l < k. \end{cases}$$

3.3 Least Squares Estimation

The first step of the estimation consists in eliminating the permanent differences that exist between individuals over the period and which could skew the estimate. Eliminating fixed individual effects involves removing specific individual means. This step is standard in

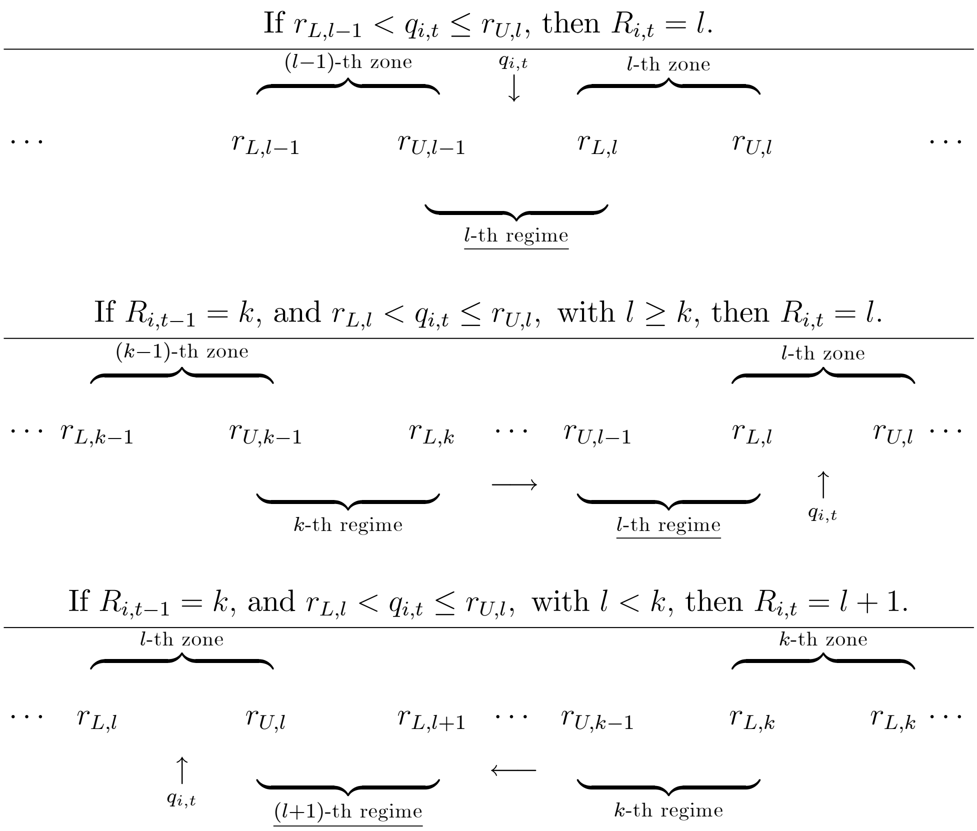

Figure 3.1: Diagram illustrating the buffered transition mechanism.

linear models (within transformation). However, it requires more careful processing in the context of threshold models and particularly in the $BTPD$ model. Indeed, the individual effects depend on our knowledge about the buffer zones and must therefore be recalculated for each candidate zones. Another problem can be posed when the values of the threshold variables for the first observations may fall within one of the buffer zones. This can make the identification of the regimes indicators difficult. To deal with this problem, we propose to set the regime indicator of each individual component i, for the first observation, as a realization of a uniform random variable on $\{k, k+1\}$ when its threshold variable value falls within the buffer zone $(r_{L,k}, r_{U,k}]$, i.e. when $r_{L,k} < q_{i,1} \leq r_{U,k}$. The regime indicators sequences $\{R_{i,t}, i = 1, ..., N \text{ and } t = 2, ..., T\}$ are subsequently generated.

For the individual effect elimination, we need the following compact representation of (3.2),

$$Y_{i,t} = \mu_i + \mathbf{X}_{i,t}\beta + \varepsilon_{i,t}, \tag{3.3}$$

where $\mathbf{X}_{i,t} = \left(X_{i,t}\mathbf{1}_{(R_{i,t}=1)}, X_{i,t}\mathbf{1}_{(R_{i,t}=2)}, ..., X_{i,t}\mathbf{1}_{(R_{i,t}=K)}\right)$ and $\beta = (\beta_1', \beta_2', ..., \beta_K')'$. As in Hansen (1999$a$), we take the averages of (3.3) over the time index t and obtain,

$$\overline{Y_i} = \mu_i + \overline{\mathbf{X}}_i\beta + \overline{\varepsilon}_i, \tag{3.4}$$

where

$$\overline{Y_i} = \frac{1}{T}\sum_{t=1}^{T} Y_{i,t},\ \overline{\varepsilon}_i = \frac{1}{T}\sum_{t=1}^{T} \varepsilon_{i,t},$$

and

$$\begin{aligned}\overline{\mathbf{X}}_i &= \frac{1}{T}\sum_{t=1}^{T}\mathbf{X}_{i,t} \\ &= \left(\frac{1}{T}\sum_{t=1}^{T} X_{i,t}\mathbf{1}_{(R_{i,t}=1)}, \frac{1}{T}\sum_{t=1}^{T} X_{i,t}\mathbf{1}_{(R_{i,t}=2)}, ..., \frac{1}{T}\sum_{t=1}^{T} X_{i,t}\mathbf{1}_{(R_{i,t}=K)}\right).\end{aligned}$$

Note that the means are only taken over the T observations for the individual component i, $i = 1, ..., N$. The difference between (3.3) and (3.4) gives

$$Y_{i,t}^* = \mathbf{X}_{i,t}^*\beta + \varepsilon_{i,t}^*,$$

where

$$Y_{i,t}^* = Y_{i,t} - \overline{Y_i},\ \mathbf{X}_{i,t}^* = \mathbf{X}_{i,t} - \overline{\mathbf{X}}_i \text{ and } \varepsilon_{i,t}^* = \varepsilon_{i,t} - \overline{\varepsilon}_i.$$

In Hansen (1999a), the stacked data and errors are taken for an individual, with one time period deleted.

Let

$$\mathbf{Y}_i^* = \begin{pmatrix} Y_{i,1}^* \\ Y_{i,2}^* \\ \vdots \\ Y_{i,T}^* \end{pmatrix}, \ \mathbf{X}_i^* = \begin{pmatrix} \mathbf{X}_{i,1}^* \\ \mathbf{X}_{i,2}^* \\ \vdots \\ \mathbf{X}_{i,T}^* \end{pmatrix}, \text{ and } \varepsilon_i^* = \begin{pmatrix} \varepsilon_{i,1}^* \\ \varepsilon_{i,2}^* \\ \vdots \\ \varepsilon_{i,T}^* \end{pmatrix}.$$

Collecting these terms gives the following transformed model

$$\mathbf{Y}^* = \mathbf{X}^*\beta + \varepsilon^*,$$

where $\mathbf{X}^* = (\mathbf{X}_1^{*\prime}, \mathbf{X}_2^{*\prime}, ...\mathbf{X}_N^{*\prime})'$ and $\varepsilon^* = (\varepsilon_1^{*\prime}, \varepsilon_2^{*\prime}, ..., \varepsilon_N^{*\prime})'$. This is a classical regression model. Hence, assuming that the number of regimes is known and, for a given configuration of the buffer zones $\gamma = (r_{L,1}, r_{U,1}, r_{L,2}, r_{U,2}, ..., r_{L,K-1}, \ \ r_{U,K-1})$, that the $NT \times Km$ matrix $\mathbf{X}^*(\gamma)$ has full column rank, the ordinary least squares (OLS) estimator of the slope coefficients is given by

$$\widehat{\beta}(\gamma) = \left((\mathbf{X}^*(\gamma))' \mathbf{X}^*(\gamma)\right)^{-1} (\mathbf{X}^*(\gamma))' \mathbf{Y}^*.$$

Once $\widehat{\beta}(\gamma)$ is obtained, the individual fixed effects are then estimated by

$$\widehat{\mu}_i = \overline{Y_i} - \overline{\mathbf{X}}_i(\gamma)\widehat{\beta}(\gamma), \text{ for } i = 1, ..., N.$$

We can also obtain the residual vector as follows

$$\widehat{\varepsilon}(\gamma) = \mathbf{Y}^* - \mathbf{X}^*(\gamma)\widehat{\beta}(\gamma).$$

This allows the calculation of the sum of the squared errors

$$\begin{aligned} S(\gamma) &= \widehat{\varepsilon}(\gamma)' \widehat{\varepsilon}(\gamma) \\ &= \mathbf{Y}^{*\prime}\mathbf{Y}^* - \left(\mathbf{Y}^{*\prime}\overline{\mathbf{X}}_i(\gamma)\right)\left(\overline{\mathbf{X}}_i(\gamma)'\overline{\mathbf{X}}_i(\gamma)\right)^{-1}\left(\overline{\mathbf{X}}_i(\gamma)'\mathbf{Y}^*\right). \end{aligned}$$

We thus can estimate the residual variance as follows

$$\widehat{\sigma}^2 (\gamma) = \frac{1}{NT - N} S (\gamma) .$$

As previously discussed, the estimation of the slope coefficients requires the knowledge of the buffer zones γ that we should estimate. To do this, we explore a set of buffer zones candidates constructed from quantiles of the observed threshold variable values. To ensure a minimum number of observations in each regime, we can take, for example when $K = 2$, quantiles from $a\%$ to $b\%$ of each threshold values sample with a carefully chosen interval $[a, b]$, and we generate all ordered 2-vectors of γ such as $\widehat{r}_{L,1} < \widehat{r}_{U,1}$. In general, we require all buffer zone parameters to lie in the bounded subset $[r_0, r_1]$ of the threshold variable sample space and we choose the $2(K-1)$-vector $\widehat{\gamma} = (\widehat{r}_{L,1}, \widehat{r}_{U,1}, \widehat{r}_{L,2}, \widehat{r}_{U,2}, ..., \widehat{r}_{L,K-1}, \widehat{r}_{U,K-1})$ which minimizes the sum of the squared errors, i.e.

$$\widehat{\gamma} = \arg \min_{\gamma \in \Gamma} (S (\gamma)) ,$$

where $\Gamma = \left\{ (r_{L,1}, r_{U,1}, ..., r_{L,K-1}, r_{U,K-1}) \in [r_0, r_1]^{2(K-1)} \middle| \, r_{L,1} \leq r_{U,1} < \cdots < r_{L,K-1} \leq r_{U,K-1} \right\}$. Note that our Γ is larger than the one used in Hansen (1999). Therefore, our exploration of the buffer zones candidates needs more computational efforts.

3.4 Linearity test and determination of the number of regimes

We now deal with the issue of inference in $BTPD$ models. As a priority, we present the tests of linearity and determination of the number of regimes. These tests occupy a preponderant place and guide us in the choice of the specification in order to take into account the nonlinearity.

Consider again a $BTPD$ model with two regimes ($K = 2$) as given by (3.2). As in Hansen (1999a), one can use a likelihood ratio (LR) test to detect the effect of this nonlinearity. The hypothesis of linearity can be formulated as follows

$$H_0 : \beta_1 = \beta_2.$$

This test is well known in the literature and results in non-standard tests. Indeed, under H_0 the buffered threshold parameters $r_{L,1}$ and $r_{U,1}$ disappear, and the model is only identified under the alternative hypothesis ($BTPD$ model with two regimes)

$$H_1 : \beta_1 \neq \beta_2.$$

This testing problem corresponds to the famous Davies problem (1977, 1987) and has been investigated by Andrews and Ploberger (1994) and Hansen (1996). The bootstrap procedure suggested by Hansen (1996), and used in Hansen (1999a) to simulate the corresponding distribution of the LR test, can be adapted to our framework. Indeed, under the linearity hypothesis, the model is

$$Y_{i,t} = \mu_i + \beta_1^{'} X_{i,t} + \varepsilon_{i,t}.$$

After the within transformation, we obtain

$$Y_{i,t}^{*} = \beta_1^{'} X_{i,t}^{*} + \varepsilon_{i,t}^{*}.$$

Under H_0, the OLS method provides an estimation $\widehat{\beta}_1$ for β_1, the residuals $\tilde{\varepsilon}_{i,t}^{*}$ and the sum of squared errors $S_1 = \tilde{\varepsilon}^{*'} \tilde{\varepsilon}^{*}$. The LR test statistic is then defined by

$$F_{1,2} = \frac{S_1 - S_2\left(\widehat{\gamma}_2\right)}{\widehat{\sigma}_2^2\left(\widehat{\gamma}_2\right)},$$

where $\widehat{\sigma}_K^2\left(\widehat{\gamma}_K\right) = S_K\left(\widehat{\gamma}_K\right) / N\left(T-1\right)$, $S_K\left(\widehat{\gamma}_K\right) = \widehat{\varepsilon}\left(\widehat{\gamma}_K\right)^{'} \widehat{\varepsilon}\left(\widehat{\gamma}_K\right)$ and $\widehat{\gamma}_K = (\widehat{r}_{L,1}, \widehat{r}_{U,1}, \widehat{r}_{L,2}, \widehat{r}_{U,2}, ..., \widehat{r}_{L,K-1}, \widehat{r}_{U,K-1})$. $\widehat{\gamma}_K$ and $\widehat{\varepsilon}\left(\widehat{\gamma}_K\right)$ are respectively the estimates of the vector of the buffer zones γ and the residuals under the alternative hypothesis of a buffered model with K regimes.

For the special case $r_{L,1} = r_{U,1}$, it has been found (see Hansen, 1999a) that the asymptotic distribution of $F_{1,2}$ is non-standard, and strictly dominates the χ_m^2 distribution (where m represents the number of the explanatory variables). Furthermore, this asymptotic distribution depends, in general, upon moments of the sample, which are application-specific. Hence, the critical values and p-values cannot be tabulated. This problem can, however, be solved by referring to Hansen's (1996, 1999a) bootstrap methodology. The same logic can be applied to our buffered threshold modelling. The asymptotic p-value can be approximated from the following bootstrap procedure.

Algorithm 3 Bootstrap linearity test

1. Treat the regressors $X_{i,t}$ and the threshold jump variable $q_{i,t}$ as given, and their value remain fixed during repeated bootstrap simulations.
2. Recover the regression residuals obtained under the null hypothesis $\widehat{\varepsilon}^*_{i,t}$ and, group them by individual $\widehat{\varepsilon}^*_i = \left(\widehat{\varepsilon}^*_{i,1}, \widehat{\varepsilon}^*_{i,2}, \cdots, \widehat{\varepsilon}^*_{i,T}\right)$. Treat the sample $\{\widehat{\varepsilon}^*_1, \widehat{\varepsilon}^*_2, \cdots, \widehat{\varepsilon}^*_N\}$ as the empirical distribution to be used for the bootstrapping.
3. Draw (with replacement) a sample of size N from the empirical distribution and use these errors to create a bootstrap sample under H_0.
4. Using the bootstrap sample, estimate the model under the null hypothesis and the alternative one and calculate therefore the bootstrap value of $F_{1,2}$.
5. Repeat this procedure a large number of times and calculate the percentage of draws for which the simulated statistic exceeds the observed statistic $F_{1,2}$. This percentage represents the p-value of $F_{1,2}$ under H_0.

Once the linearity hypothesis is rejected, the question is whether all the nonlinearity of the observations has been taken into account or, in other words, should we use a buffered threshold modelling containing a greater number of regimes. The tests to determine the optimal number of regimes that are an extension of the linearity tests, allow us to answer this question. Indeed, to test whether the model has two regimes, i.e. $H_0 : \beta_3 = 0$, or at least three regimes, i.e. $H_1 : \beta_3 \neq 0$, the following LR test statistic must be applied

$$F_{2,3} = \frac{S_2\left(\widehat{\gamma}_2\right) - S_3\left(\widehat{\gamma}_3\right)}{\widehat{\sigma}^2_3\left(\widehat{\gamma}_3\right)}.$$

The null hypothesis of a single buffered zone is rejected in favor of at least two buffered zones, if the value of $F_{2,3}$ is greater than the critical value simulated by bootstrap. It is worth noting that the bootstrap approach to approximate the asymptotic critical values is very close to that presented in Algorithm 3. The only significant change is in the Step 2 in which the bootstrap errors used are no longer those obtained under the null hypothesis, but under the alternative hypothesis.

If the null hypothesis of a model with a single buffered zone is rejected, the process for determining the optimal number of regimes continues. It is then necessary to test the null hypothesis of two buffered zones against a model containing at least three buffered zones. In case of rejection of the null hypothesis, the specification must contain at least four regimes, and so forth until acceptation of the null hypothesis.

3.5 Simulation study

In this section, we conduct a Monte Carlo experiment to evaluate the performance of the OLS estimation under different settings. For the sake of simplicity, we only provide the $K = 2$ and $K = 3$ cases of the $BTPD$ model (3.2).

We generate samples from (3.2) using different values of γ and $\beta = (\beta_1', \beta_2', ..., \ \beta_K')'$, where $\beta_k = (\beta_{k,1}, \beta_{k,2}, ..., \beta_{k,m})'$. $X_{i,t}$, $q_{i,t}$ and $\varepsilon_{i,t}$ are generated from some independent distributions (normal distribution, log-normal distribution, uniform distribution). For the number of individuals N and the time period T, we use combinations of $T = 10, 20, 50, 100$, and $N = 10, 20, 50, 100$. For each set of generated sample observations, we calculate the OLS estimator. We do this 1000 times. With different values of model parameters and for different choices of N and T, finite sample properties of the estimators are summarized in Tables 3.1-3.4. For each case, we report the true values (True) of the parameters of each of the considered $BTPD$ model, empirical mean (Mean) and the empirical standard deviation (Std).

From Tables 3.1-3.4, we can observe that the estimates of the $BTPD$ structural coefficients based on the OLS method display reasonable biases, which decrease as either N or T becomes large. It can also be seen that the Std of some parameters are relatively large but they quickly decrease as either the number of individuals or the time period are increased. This thus empirically shows that the desirable consistency property of the OLS estimators is satisfied and that the proposed estimation procedure provides good results.

Let us now examine the empirical distribution of the estimator of the slope parameters. Figures 3.1-3.3 depict the sample histograms for the estimated first parameters ($\beta_{1,1}$ or $\beta_{2,1}$) in three configurations of the $BTPD$ model that differ by the distribution of the error $\varepsilon_{i,t}$. In the first case, the error term has a uniform distribution while in the second and third cases the distribution is normal. It is important to note that the $DGPs$ are the same and are given in Tables 3.2-3.4. From Jarque-Bera test (see Figures 3.1-3.3), it can be observed that when the cross-sectional and time dimensions of the panel are small, the empirical distribution is far from being normal. Indeed, the corresponding p-values are

		T	10		20		50		100	
N		True	Mean	Std	Mean	Std	Mean	Std	Mean	Std
	$\beta_{1,1}$	-0.5	-0.5158	0.1554	-0.5053	0.0932	-0.4994	0.0583	-0.4975	0.0405
	$\beta_{1,2}$	0.6	0.6244	0.1512	0.6027	0.0962	0.6004	0.0571	0.6017	0.0396
	$\beta_{2,1}$	0.3	0.3444	0.2263	0.3177	0.1370	0.2987	0.0798	0.2980	0.0541
10	$\beta_{2,2}$	-0.2	-0.2157	0.2276	-0.2130	0.1387	-0.1980	0.0785	-0.1971	0.0543
	$r_{L,1}$	0.25	0.1334	0.4536	0.2312	0.2417	0.2464	0.1105	0.2554	0.0583
	$r_{U,1}$	0.75	0.7656	0.3432	0.7551	0.1700	0.7483	0.0688	0.7494	0.0339
	σ^2	1	0.9183	0.1377	0.9680	0.1038	0.9937	0.0628	0.9968	0.0438
	$\beta_{1,1}$	-0.5	-0.5072	0.0981	-0.4971	0.0636	-0.4994	0.0406	-0.4987	0.0283
	$\beta_{1,2}$	0.6	0.6085	0.1002	0.5999	0.0659	0.5996	0.0408	0.5999	0.0280
	$\beta_{2,1}$	0.3	0.3073	0.1450	0.3002	0.0880	0.3015	0.0568	0.2964	0.0390
20	$\beta_{2,2}$	-0.2	-0.2214	0.1467	-0.2013	0.0943	-0.1972	0.0573	-0.1987	0.0380
	$r_{L,1}$	0.25	0.2460	0.2664	0.2493	0.1504	0.2513	0.0598	0.2506	0.0287
	$r_{U,1}$	0.75	0.7571	0.2040	0.7499	0.0832	0.7486	0.0357	0.7498	0.0213
	σ^2	1	0.9650	0.1036	0.9842	0.0732	0.9970	0.0453	1.0003	0.0320
	$\beta_{1,1}$	-0.5	-0.4996	0.0613	-0.4997	0.0417	-0.4981	0.0257	-0.4977	0.0179
	$\beta_{1,2}$	0.6	0.5988	0.0584	0.6004	0.0419	0.5987	0.0252	0.5981	0.0175
	$\beta_{2,1}$	0.3	0.2938	0.0862	0.2916	0.0591	0.2961	0.0351	0.2952	0.0257
50	$\beta_{2,2}$	-0.2	-0.2016	0.0880	-0.1960	0.0572	-0.1963	0.0345	-0.1969	0.0249
	$r_{L,1}$	0.25	0.2690	0.1415	0.2524	0.0751	0.2505	0.0249	0.2500	0.0158
	$r_{U,1}$	0.75	0.7534	0.0895	0.0645	0.0422	0.7497	0.0195	0.7499	0.0138
	σ^2	1	0.9900	0.0678	0.9988	0.0449	1.0021	0.0278	1.0028	0.0199
	$\beta_{1,1}$	-0.5	-0.4965	0.0411	-0.4970	0.0284	-0.4980	0.0185	-0.4976	0.0125
	$\beta_{1,2}$	0.6	0.5994	0.0421	0.5967	0.0291	0.5975	0.0174	0.5981	0.0127
	$\beta_{2,1}$	0.3	0.2887	0.0603	0.2959	0.0394	0.2950	0.0245	0.2961	0.0181
100	$\beta_{2,2}$	-0.2	-0.1931	0.0612	-0.1945	0.0376	-0.1953	0.0253	-0.1957	0.0183
	$r_{L,1}$	0.25	0.2620	0.0914	0.2525	0.0350	0.2495	0.0158	0.2499	0.0129
	$r_{U,1}$	0.75	0.7505	0.0419	0.7501	0.0240	0.7498	0.0150	0.7497	0.0133
	σ^2	1	1.0017	0.0471	1.0047	0.0330	1.0040	0.0204	1.0034	0.0142

Table 3.1: Results of a simulation study for a two-regime $BTPD$ model with $\varepsilon_{i,t} \sim \mathcal{N}(0,1)$, $X_{i,t} \sim \mathcal{N}(0, I_2)$ and $q_{i,t} \sim \mathcal{N}(0,2)$.

	T		10		20		50		100	
N		True	Mean	Std	Mean	Std	Mean	Std	Mean	Std
10	$\beta_{1,1}$	0.3	0.2987	0.1620	0.3039	0.1064	0.3011	0.0636	0.3000	0.0441
	$\beta_{1,2}$	-0.2	-0.2061	0.1632	-0.1991	0.1066	-0.1994	0.0619	-0.1996	0.0425
	$\beta_{1,3}$	1.5	1.4981	0.1602	1.5020	0.1090	1.4943	0.0608	1.5017	0.0425
	$\beta_{1,4}$	1.3	1.3092	0.1659	1.3000	0.1108	1.3005	0.0630	1.2985	0.0442
	$\beta_{1,5}$	-0.2	-0.1907	0.1582	-0.1999	0.1126	-0.2063	0.0608	-0.1988	0.0465
	$\beta_{2,1}$	0.5	0.4927	0.3071	0.5033	0.1650	0.5040	0.1023	0.4964	0.0715
	$\beta_{2,2}$	0.2	0.2244	0.3066	0.1985	0.1635	0.2004	0.1020	0.1981	0.0711
	$\beta_{2,3}$	0.9	0.9115	0.3241	0.9073	0.1570	0.9023	0.0997	0.9036	0.0719
	$\beta_{2,4}$	-0.8	-0.7849	0.3284	-0.7999	0.1590	-0.7912	0.1053	-0.7887	0.0711
	$\beta_{2,5}$	-0.5	-0.5081	0.3167	-0.5070	0.1518	-0.5029	0.1058	-0.4980	0.0715
	$r_{L,1}$	0.25	0.2120	0.2550	0.2432	0.1525	0.2487	0.0575	0.2514	0.0285
	$r_{U,1}$	0.75	0.7500	0.1396	0.7414	0.0971	0.7506	0.0263	0.7495	0.0168
	σ^2	1.33	1.1623	0.1356	1.3147	0.0708	1.3180	0.0607	1.3301	0.0407
20	$\beta_{1,1}$	0.3	0.3065	0.1078	0.3061	0.0671	0.3006	0.0435	0.3008	0.0310
	$\beta_{1,2}$	-0.2	-0.1980	0.1085	-0.1970	0.0726	-0.1989	0.0427	-0.1976	0.0309
	$\beta_{1,3}$	1.5	1.4997	0.1056	1.4970	0.0708	1.4982	0.0449	1.4983	0.0298
	$\beta_{1,4}$	1.3	1.2981	0.1083	1.2942	0.0744	1.2995	0.0441	1.2950	0.0313
	$\beta_{1,5}$	-0.2	-0.2005	0.1064	-0.2032	0.0712	-0.2002	0.0428	-0.2024	0.0311
	$\beta_{2,1}$	0.5	0.5026	0.1887	0.4998	0.1231	0.4991	0.0711	0.5003	0.0508
	$\beta_{2,2}$	0.2	0.1903	0.1841	0.1930	0.1275	0.1981	0.0752	0.2013	0.0500
	$\beta_{2,3}$	0.9	0.8912	0.1860	0.9080	0.1203	0.9073	0.0738	0.9059	0.0485
	$\beta_{2,4}$	-0.8	-0.7838	0.1888	-0.7819	0.1216	-0.7788	0.0723	-0.7866	0.0522
	$\beta_{2,5}$	-0.5	-0.4958	0.1926	-0.4988	0.1187	-0.4949	0.0718	-0.4969	0.0510
	$r_{L,1}$	0.25	0.2618	0.1600	0.2585	0.0936	0.2501	0.0286	0.2510	0.0155
	$r_{U,1}$	0.75	0.7513	0.0749	0.7506	0.0364	0.7503	0.0167	0.7501	0.0131
	σ^2	1.33	1.2754	0.1050	1.3174	0.0742	1.3390	0.0438	1.3413	0.0298
50	$\beta_{1,1}$	0.3	0.3035	0.0655	0.2991	0.0440	0.2998	0.0280	0.2998	0.0188
	$\beta_{1,2}$	-0.2	-0.2000	0.0664	-0.1966	0.0439	-0.1985	0.0279	-0.1994	0.0194
	$\beta_{1,3}$	1.5	1.4927	0.0648	1.4962	0.0450	1.4983	0.0279	1.4978	0.0194
	$\beta_{1,4}$	1.3	1.2870	0.0661	1.2917	0.0444	1.2938	0.0278	1.2954	0.0208
	$\beta_{1,5}$	-0.2	-0.2000	0.0666	-0.1999	0.0448	-0.1993	0.0272	-0.2010	0.0188
	$\beta_{2,1}$	0.5	0.4998	0.1135	0.4955	0.0726	0.4982	0.0455	0.4985	0.0310
	$\beta_{2,2}$	0.2	0.1940	0.1109	0.1876	0.0748	0.1967	0.0465	0.1957	0.0326
	$\beta_{2,3}$	0.9	0.9119	0.1129	0.9065	0.0763	0.9047	0.0431	0.9057	0.0319
	$\beta_{2,4}$	-0.8	-0.7661	0.1182	-0.7733	0.0803	-0.7812	0.0472	-0.7845	0.0346
	$\beta_{2,5}$	-0.5	-0.4979	0.1133	-0.4946	0.0769	-0.4978	0.0453	-0.4987	0.0323
	$r_{L,1}$	0.25	0.2663	0.0902	0.2547	0.0369	0.2508	0.0135	0.2500	0.0104
	$r_{U,1}$	0.75	0.7519	0.0355	0.7509	0.0178	0.7510	0.0130	0.7498	0.0111
	σ^2	1.33	1.3465	0.0718	1.3502	0.0448	1.3490	0.0275	1.3489	0.0198

Table 3.2: Results of a simulation study for a two-regime $BTPD$ model with $\varepsilon_{i,t} \sim \mathcal{U}[-2,2]$, $X_{i,t} \sim \mathcal{N}(0, I_5)$ and $q_{i,t} \sim \mathcal{N}(0,2)$

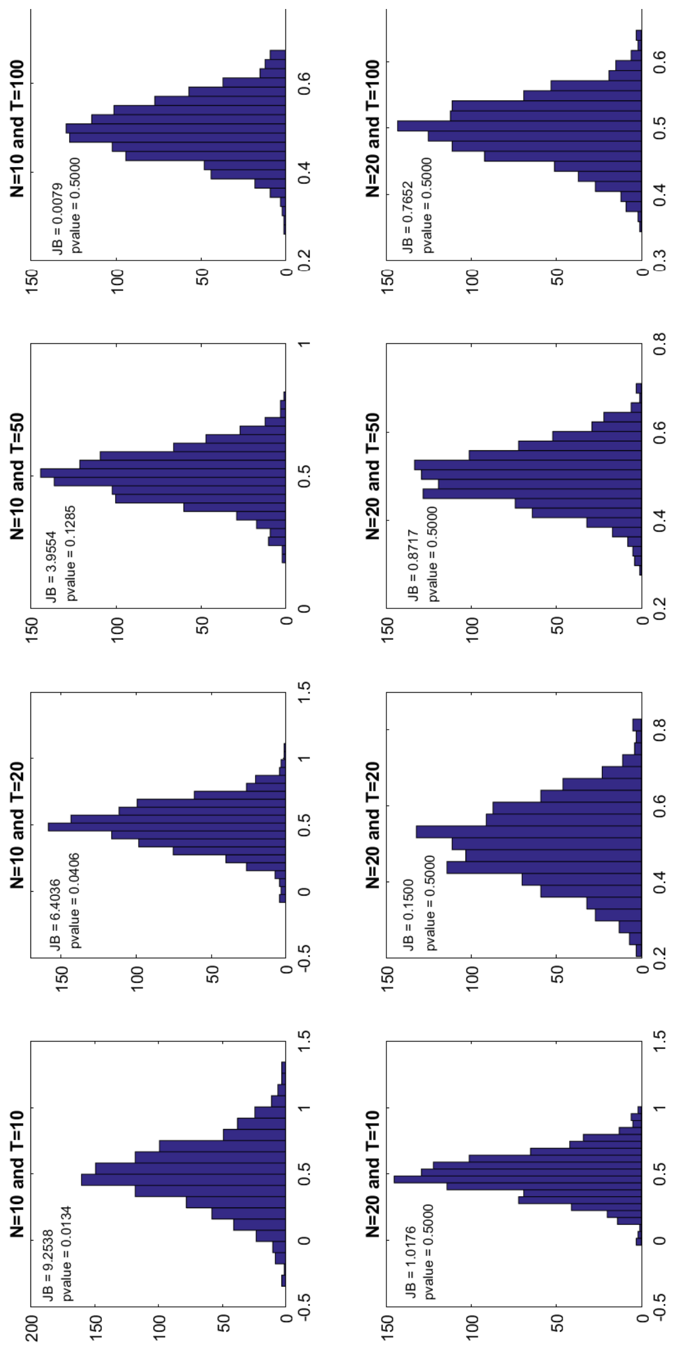

Figure 3.2: Empirical distributions of the estimator of $\beta_{2,1} = 0.5$. The samples are generated from a three-regime $BTPD$ as defined in Table 2.

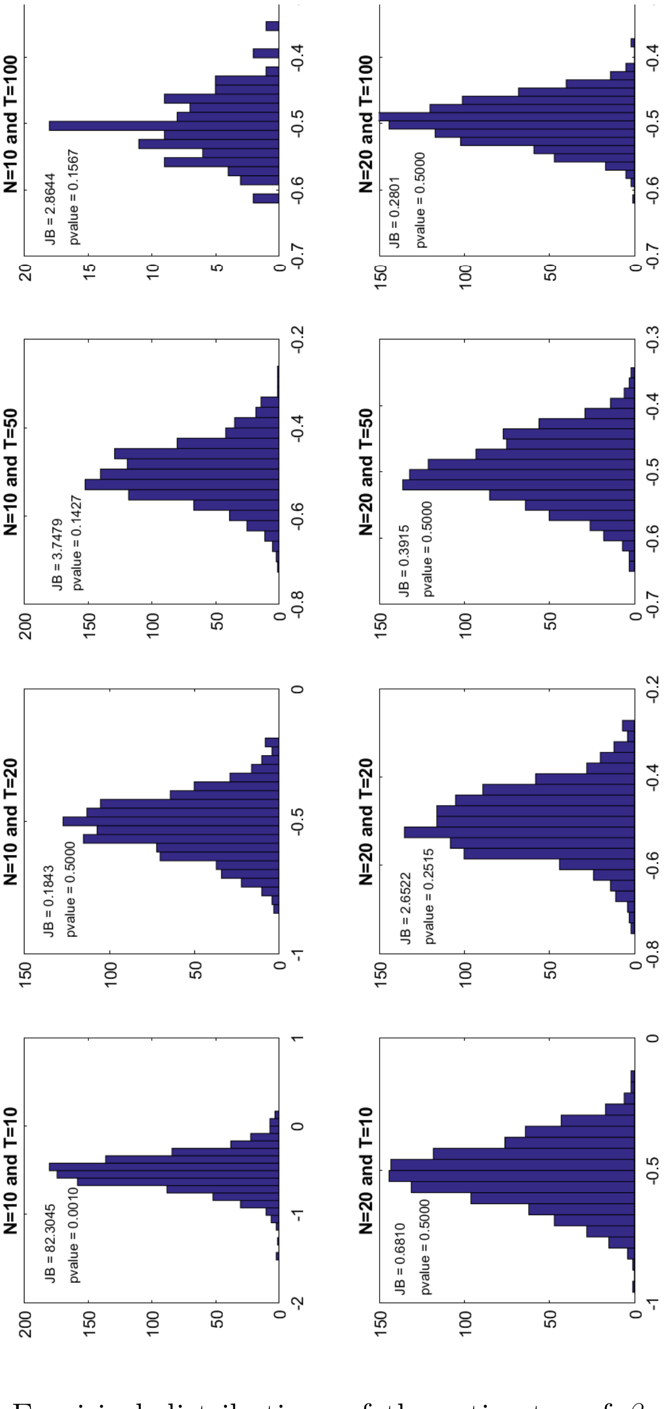

Figure 3.3: Empirical distributions of the estimator of $\beta_{1,1} = -0.5$. The samples are generated from a three-regime $BTPD$ as defined in Table 3.

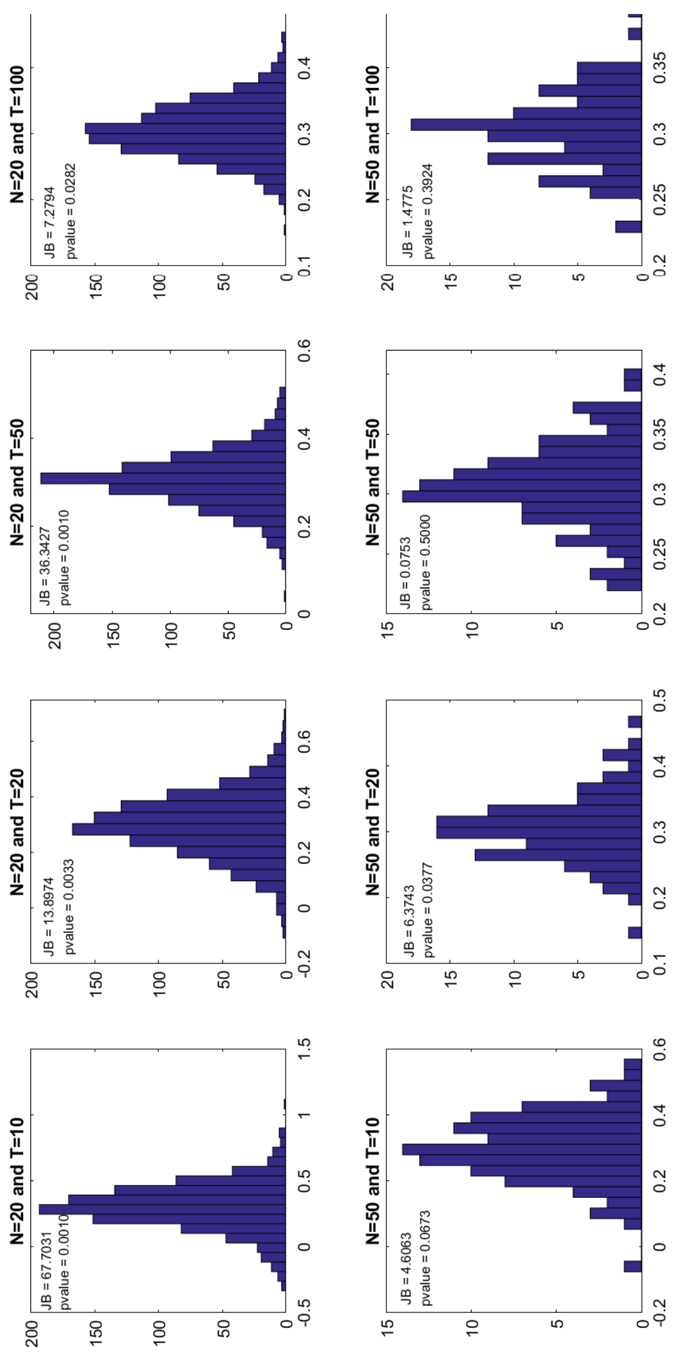

Figure 3.4: Empirical distributions of the estimator of $\beta_{1,1} = 0.3$. The samples are generated from a three-regime $BTPD$ as defined in Table 4.

	T	10		20		50		100		
N		True	Mean	Std	Mean	Std	Mean	Std	Mean	Std
	$\beta_{1,1}$	0.3	0.3021	0.0453	0.3008	0.0320	0.2991	0.0203	0.3001	0.0140
	$\beta_{1,2}$	-0.2	-0.2003	0.0445	-0.1980	0.0316	-0.1989	0.0192	-0.1999	0.0134
	$\beta_{1,3}$	1.5	1.4976	0.0472	1.4977	0.0320	1.4977	0.0195	1.4989	0.0137
	$\beta_{1,4}$	1.3	1.2883	0.0501	1.2923	0.0317	1.2946	0.0204	1.2954	0.0137
	$\beta_{1,5}$	-0.2	-0.2004	0.0475	-0.2008	0.0313	-0.1999	0.0191	-0.2007	0.0139
	$\beta_{2,1}$	0.5	0.4960	0.0798	0.4962	0.0540	0.4966	0.0317	0.4981	0.0217
100	$\beta_{2,2}$	0.2	0.1887	0.0774	0.1941	0.0521	0.1967	0.0306	0.1963	0.0226
	$\beta_{2,3}$	0.9	0.9183	0.0785	0.9089	0.0521	0.9054	0.0314	0.9038	0.0222
	$\beta_{2,4}$	-0.8	-0.7502	0.0829	-0.7666	0.0569	-0.7783	0.0363	-0.7825	0.0252
	$\beta_{2,5}$	-0.5	-0.4907	0.0773	-0.4964	0.0539	-0.4964	0.0318	-0.4976	0.0233
	$r_{L,1}$	0.25	0.2624	0.0512	0.2529	0.0224	0.2503	0.0107	0.2498	0.0087
	$r_{U,1}$	0.75	0.7505	0.0223	0.7506	0.0137	0.7498	0.0114	0.7493	0.0103
	σ^2	1.33	1.3711	0.0537	1.3620	0.0349	1.3545	0.0198	1.3510	0.0151

Table 3.3: Continued

lower than 0.1. We thus reject the normality hypothesis at level 1%. However, it turns out that when the number of individuals and/or the time period is at least equal to 50 we approach the asymptotic normality. Indeed and according to the Jarque-Bera test, we do not reject the normality of the empirical distribution at any reasonable level. It is worth noting that the same kinds of results hold for the rest of the slope parameters.

In Section 3.4, we have proposed a useful procedure for testing the linearity in our $BTPD$ framework. Hereafter, we investigate its finite sample performance. For that purpose, we do consider two $DGPs$. The first one is used to assess the size of the test while the second is used to study its power. The number of replications is fixed to 1000. In each replication, the model is estimated and analyzed using the proposed bootstrap-based test of linearity.

The rejection frequencies at the nominal levels 1, 5 and 10 per cent are presented in Table 3.7. It clearly appears that the rejection frequencies are quite close to the nominal sizes. Furthermore, it turns out that the power of the test increases with the cross-sectional and time dimensions of the panel. In summary, our extension of the bootstrap-based test, proposed by Hansen (1999a), to investigate linearity seems to work very well in our framework.

		T	10		20		50		100	
N		True	Mean	Std	Mean	Std	Mean	Std	Mean	Std
	$\beta_{1,1}$	0.3	0.3351	0.2801	0.3101	0.1814	0.3067	0.0916	0.3031	0.0591
	$\beta_{1,2}$	-0.2	-0.1205	0.3100	-0.1743	0.1895	-0.1937	0.0971	-0.1932	0.0616
	$\beta_{2,1}$	0.5	0.4980	0.6182	0.5102	0.1978	0.5005	0.0617	0.4965	0.0380
	$\beta_{2,2}$	0.2	0.0744	0.6573	0.1836	0.2236	0.2023	0.0697	0.1970	0.0371
	$\beta_{3,1}$	0.3	0.2909	0.2510	0.2991	0.0936	0.3009	0.0423	0.3012	0.0279
10	$\beta_{3,2}$	-0.4	-0.4013	0.2327	-0.4049	0.0876	-0.4019	0.0428	-0.3976	0.0288
	$r_{L,1}$	0.3	0.4915	0.2479	0.3717	0.1539	0.3113	0.0473	0.3021	0.0200
	$r_{U,1}$	0.5	0.6898	0.3283	0.5591	0.1950	0.5112	0.1015	0.5037	0.0708
	$r_{L,2}$	0.8	0.9627	0.4588	0.8361	0.2258	0.8029	0.0740	0.8013	0.0264
	$r_{U,2}$	1	1.3702	0.7407	1.0975	0.3838	1.0079	0.1130	0.9989	0.0338
	σ^2	2	1.7494	0.2777	1.9023	0.1952	1.9825	0.1291	2.0028	0.0941
	$\beta_{1,1}$	0.3	0.3062	0.1772	0.2978	0.1130	0.3058	0.0622	0.3056	0.0415
	$\beta_{1,2}$	-0.2	-0.1582	0.2048	-0.1851	0.1171	-0.1949	0.0616	-0.1937	0.0414
	$\beta_{2,1}$	0.5	0.5055	0.2046	0.5053	0.0725	0.4996	0.0391	0.4973	0.0254
	$\beta_{2,2}$	0.2	0.1654	0.2611	0.2014	0.0812	0.1958	0.0404	0.1933	0.0263
	$\beta_{3,1}$	0.3	0.2979	0.0868	0.3007	0.0485	0.3002	0.0275	0.3012	0.0191
20	$\beta_{3,2}$	-0.4	-0.4000	0.1132	-0.4008	0.0493	-0.3991	0.0280	-0.3970	0.0204
	$r_{L,1}$	0.3	0.3907	0.1623	0.3188	0.0593	0.3036	0.0249	0.3021	0.0156
	$r_{U,1}$	0.5	0.5842	0.2107	0.5097	0.1122	0.5001	0.0739	0.4962	0.0452
	$r_{L,2}$	0.8	0.8492	0.2269	0.8014	0.0616	0.8042	0.0387	0.8001	0.0205
	$r_{U,2}$	1	1.1171	0.4104	1.0073	0.1334	1.0007	0.0565	1.0002	0.0249
	σ^2	2	1.8779	0.1899	1.9556	0.1500	2.0205	0.0631	2.0056	0.0611
	$\beta_{1,1}$	0.3	0.3001	0.1040	0.3038	0.0529	0.3075	0.0363	0.3032	0.0293
	$\beta_{1,2}$	-0.2	-0.2095	0.0946	-0.2022	0.0587	-0.1980	0.0359	-0.1874	0.0239
	$\beta_{2,1}$	0.5	0.4972	0.0599	0.4975	0.0381	0.4973	0.0252	0.4953	0.0174
	$\beta_{2,2}$	0.2	0.2042	0.0575	0.1977	0.0391	0.1914	0.0267	0.1898	0.0181
	$\beta_{3,1}$	0.3	0.3004	0.0418	0.3081	0.0273	0.3011	0.0171	0.3005	0.0116
50	$\beta_{3,2}$	-0.4	-0.3953	0.0492	-0.3974	0.0298	-0.4000	0.0177	-0.3980	0.0116
	$r_{L,1}$	0.3	0.3114	0.0445	0.2985	0.0180	0.2998	0.0153	0.3027	0.0161
	$r_{U,1}$	0.5	0.4856	0.0812	0.4962	0.0692	0.4991	0.0387	0.4961	0.0267
	$r_{L,2}$	0.8	0.7951	0.0394	0.8004	0.0292	0.7998	0.0199	0.8000	0.0189
	$r_{U,2}$	1	0.9911	0.0577	0.9976	0.0347	0.9993	0.0212	1.0005	0.0196
	σ^2	2	1.9701	0.1322	1.9811	0.1026	2.0239	0.0608	2.0303	0.0394
	$\beta_{1,1}$	0.3	0.2999	0.0539	0.3050	0.0493	0.3098	0.0272	0.3097	0.0217
	$\beta_{1,2}$	-0.2	-0.1907	0.0639	-0.1931	0.0487	-0.1881	0.0292	-0.1831	0.0203
	$\beta_{2,1}$	0.5	0.4992	0.0357	0.4929	0.0251	0.4927	0.0155	0.4954	0.0114
	$\beta_{2,2}$	0.2	0.1993	0.0385	0.1979	0.0259	0.1924	0.0144	0.1912	0.0127
	$\beta_{3,1}$	0.3	0.2984	0.0292	0.3000	0.0202	0.3023	0.0117	0.3008	0.0073
100	$\beta_{3,2}$	-0.4	-0.3969	0.0268	-0.3973	0.0228	-0.3986	0.0140	-0.3967	0.0096
	$r_{L,1}$	0.3	0.3014	0.0211	0.3016	0.0195	0.3031	0.0159	0.3060	0.0158
	$r_{U,1}$	0.5	0.5052	0.0737	0.5012	0.0353	0.4984	0.0222	0.5022	0.0164
	$r_{L,2}$	0.8	0.8011	0.0260	0.8030	0.0200	0.7987	0.0212	0.8011	0.0200
	$r_{U,2}$	1	1.0082	0.0394	1.0018	0.0252	0.9986	0.0205	1.0002	0.0166
	σ^2	2	2.0036	0.0904	2.0251	0.0692	2.0331	0.0399	2.0378	0.0298

Table 3.4: Results of a simulation study for a three-regime $BTPD$ model with $\varepsilon_{i,t} \sim \mathcal{N}(0,2)$, $\log(X_{i,t}) \sim \mathcal{N}(0, I_2)$ and $\log(q_{i,t}) \sim \mathcal{N}(0,2)$

	T		10		20		50		100	
N		True	Mean	Std	Mean	Std	Mean	Std	Mean	Std
	$\beta_{1,1}$	-0.5	-0.5145	0.1988	-0.5122	0.1108	-0.4998	0.0646	-0.5044	0.0478
	$\beta_{1,2}$	0.6	0.6179	0.1899	0.6062	0.1145	0.6012	0.0673	0.5963	0.0450
	$\beta_{1,3}$	1.2	1.2051	0.1992	1.1917	0.1074	1.2026	0.0663	1.2036	0.0434
	$\beta_{1,4}$	1.5	1.5021	0.1869	1.5027	0.1119	1.5011	0.0701	1.5032	0.0480
	$\beta_{1,5}$	0.2	0.2114	0.1945	0.2004	0.1150	0.1988	0.0648	0.1922	0.0459
	$\beta_{2,1}$	0.3	0.2764	0.5906	0.3053	0.1967	0.2904	0.1077	0.2932	0.0723
	$\beta_{2,2}$	-0.2	-0.1877	0.5329	-0.2029	0.1941	-0.1946	0.1078	-0.1890	0.0636
	$\beta_{2,3}$	1.5	1.5075	0.8154	1.4991	0.1918	1.4905	0.1075	1.4853	0.0645
	$\beta_{2,4}$	0.9	0.9473	0.5113	0.8804	0.1985	0.9000	0.1130	0.8963	0.0663
10	$\beta_{2,5}$	-0.7	-0.6928	0.6169	-0.7005	0.2058	-0.6988	0.1081	-0.6954	0.0752
	$\beta_{3,1}$	-0.4	-0.4005	0.2695	-0.4002	0.1427	-0.3967	0.0852	-0.3798	0.0597
	$\beta_{3,2}$	-0.3	-0.3103	0.2699	-0.2951	0.1493	-0.2982	0.0868	-0.3004	0.0703
	$\beta_{3,3}$	0.9	0.9090	0.2625	0.9052	0.1529	0.8989	0.0850	0.9111	0.0623
	$\beta_{3,4}$	-0.5	-0.5024	0.2800	-0.5035	0.1482	-0.4922	0.0871	-0.4864	0.0588
	$\beta_{3,5}$	-0.8	-0.8176	0.2673	-0.7993	0.1499	-0.7996	0.0875	-0.8070	0.0591
	$r_{L,1}$	-0.2	-0.3261	0.3386	-0.2104	0.1183	-0.2023	0.0575	-0.1920	0.0365
	$r_{U,1}$	0.1	0.0700	0.2503	0.1021	0.1049	0.0965	0.0571	0.0953	0.0370
	$r_{L,2}$	0.6	0.4744	0.2554	0.5554	0.1360	0.5914	0.0711	0.5952	0.0509
	$r_{U,2}$	0.8	0.8441	0.2732	0.8194	0.1099	0.7993	0.0518	0.8022	0.0440
	σ^2	1	0.7856	0.1295	0.9113	0.0982	0.9857	0.0639	1.0107	0.0457
	$\beta_{1,1}$	-0.5	-0.5085	0.1174	-0.4987	0.0746	-0.4966	0.0488	-0.4958	0.0325
	$\beta_{1,2}$	0.6	0.5997	0.1196	0.5974	0.0735	0.5969	0.0462	0.5946	0.0326
	$\beta_{1,3}$	1.2	1.2010	0.1130	1.1998	0.0795	1.2010	0.0464	1.2011	0.0330
	$\beta_{1,4}$	1.5	1.4981	0.1145	1.4983	0.0747	1.4969	0.0455	1.4971	0.0325
	$\beta_{1,5}$	0.2	0.1999	0.1116	0.1980	0.0772	0.1941	0.0483	0.1955	0.0334
	$\beta_{2,1}$	0.3	0.3175	0.2106	0.2955	0.1271	0.2822	0.0737	0.2784	0.0548
	$\beta_{2,2}$	-0.2	-0.2036	0.2019	-0.1986	0.1207	-0.1912	0.0735	-0.1881	0.0517
	$\beta_{2,3}$	1.5	1.4954	0.2039	1.4944	0.1194	1.4888	0.0749	1.4885	0.0512
	$\beta_{2,4}$	0.9	0.9015	0.2133	0.8987	0.1282	0.8912	0.0759	0.8920	0.0559
20	$\beta_{2,5}$	-0.7	-0.7032	0.2040	-0.6916	0.1199	-0.6912	0.0761	-0.6893	0.0519
	$\beta_{3,1}$	-0.4	-0.4027	0.1621	-0.3989	0.0990	-0.3948	0.0589	-0.3952	0.0418
	$\beta_{3,2}$	-0.3	-0.3020	0.1498	-0.3001	0.1000	-0.2983	0.0609	-0.2987	0.0426
	$\beta_{3,3}$	0.9	0.8952	0.1543	0.8997	0.1017	0.9013	0.0613	0.9070	0.0424
	$\beta_{3,4}$	-0.5	-0.4977	0.1511	-0.4932	0.0980	-0.4893	0.0634	-0.4888	0.0455
	$\beta_{3,5}$	-0.8	-0.7923	0.1426	-0.7987	0.0976	-0.8000	0.0577	-0.7985	0.0427
	$r_{L,1}$	-0.2	-0.2199	0.1308	-0.2028	0.0702	-0.1994	0.0413	-0.1999	0.0333
	$r_{U,1}$	0.1	0.0975	0.1201	0.0991	0.0648	0.0995	0.0378	0.0965	0.0308
	$r_{L,2}$	0.6	0.5516	0.1420	0.5883	0.0824	0.6000	0.0518	0.6014	0.0396
	$r_{U,2}$	0.8	0.8152	0.1166	0.8062	0.0602	0.8026	0.0409	0.7989	0.0367
	σ^2	1	0.9103	0.1024	0.9736	0.0720	1.0117	0.0488	1.0220	0.0350

Table 3.5: Results of a simulation study for a three-regime $BTPD$ model with $\varepsilon_{i,t} \sim \mathcal{N}(0,1)$, $X_{i,t} \sim \mathcal{N}(0, I_5)$ and $q_{i,t} \sim \mathcal{N}(0,2)$

	T		10		20		50		100	
N		True	Mean	Std	Mean	Std	Mean	Std	Mean	Std
	$\beta_{1,1}$	-0.5	-0.5187	0.0661	-0.4942	0.0477	-0.4940	0.0293	-0.4982	0.0189
	$\beta_{1,2}$	0.6	0.6032	0.0705	0.5993	0.0446	0.5949	0.0315	0.5954	0.0195
	$\beta_{1,3}$	1.2	1.1992	0.0686	1.1982	0.0483	1.2048	0.0290	1.2040	0.0232
	$\beta_{1,4}$	1.5	1.4961	0.0658	1.4964	0.0523	1.4937	0.0324	1.5012	0.0182
	$\beta_{1,5}$	0.2	0.2149	0.0740	0.1987	0.0416	0.1956	0.0323	0.1971	0.0212
	$\beta_{2,1}$	0.3	0.2820	0.1035	0.2908	0.0768	0.2779	0.0485	0.2723	0.0367
	$\beta_{2,2}$	-0.2	-0.1894	0.1131	-0.1877	0.0785	-0.1954	0.0507	-0.1926	0.0351
	$\beta_{2,3}$	1.5	1.4794	0.1140	1.4852	0.0707	1.4835	0.0466	1.4829	0.0316
	$\beta_{2,4}$	0.9	0.9172	0.1141	0.8904	0.0779	0.8851	0.0515	0.8791	0.0427
50	$\beta_{2,5}$	-0.7	-0.6903	0.1113	-0.6899	0.0843	-0.6814	0.0489	-0.6852	0.0311
	$\beta_{3,1}$	-0.4	-0.4019	0.0869	-0.4007	0.0612	-0.3995	0.0394	-0.3950	0.0249
	$\beta_{3,2}$	-0.3	-0.2962	0.0893	-0.2922	0.0561	-0.2961	0.0397	-0.3010	0.0297
	$\beta_{3,3}$	0.9	0.8996	0.0911	0.8911	0.0631	0.9055	0.0366	0.9033	0.0243
	$\beta_{3,4}$	-0.5	-0.4858	0.0862	-0.4896	0.0676	-0.4864	0.0438	-0.4903	0.0309
	$\beta_{3,5}$	-0.8	-0.8103	0.0844	-0.7983	0.0578	-0.8072	0.0402	-0.7995	0.0260
	$r_{L,1}$	-0.2	-0.2127	0.0746	-0.1955	0.0383	-0.1959	0.0290	-0.2016	0.0288
	$r_{U,1}$	0.1	0.0833	0.0595	0.0947	0.0369	0.0998	0.0301	0.0939	0.0270
	$r_{L,2}$	0.6	0.5825	0.0672	0.6019	0.0554	0.5975	0.0383	0.6067	0.0357
	$r_{U,2}$	0.8	0.7958	0.0590	0.8021	0.0413	0.8049	0.0388	0.8053	0.0362
	σ^2	1	0.9924	0.0679	1.0151	0.0532	1.0249	0.0283	1.0350	0.0241
	$\beta_{1,1}$	-0.5	-0.4995	0.0476	-0.4956	0.0365	-0.4974	0.0196	-0.4966	0.0152
	$\beta_{1,2}$	0.6	0.6047	0.0536	0.5988	0.0366	0.5959	0.0213	0.5952	0.0164
	$\beta_{1,3}$	1.2	1.2025	0.0461	1.2089	0.0348	1.2000	0.0196	1.2004	0.0148
	$\beta_{1,4}$	1.5	1.5037	0.0440	1.4923	0.0312	1.5009	0.0202	1.4969	0.0160
	$\beta_{1,5}$	0.2	0.2027	0.0478	0.1903	0.0308	0.1960	0.0197	0.1935	0.0141
	$\beta_{2,1}$	0.3	0.2765	0.0786	0.2823	0.0478	0.2786	0.0351	0.2685	0.0274
	$\beta_{2,2}$	-0.2	-0.2053	0.0771	-0.1985	0.0521	-0.1957	0.0341	-0.1926	0.0252
	$\beta_{2,3}$	1.5	1.4926	0.0757	1.4927	0.0516	1.4847	0.0385	1.4808	0.0221
	$\beta_{2,4}$	0.9	0.8907	0.0840	0.8781	0.0576	0.8897	0.0372	0.8757	0.0358
100	$\beta_{2,5}$	-0.7	-0.6883	0.0839	-0.7034	0.0521	-0.6877	0.0356	-0.6934	0.0235
	$\beta_{3,1}$	-0.4	-0.3979	0.0669	-0.3938	0.0387	-0.3928	0.0270	-0.3882	0.0201
	$\beta_{3,2}$	-0.3	-0.3031	0.0699	-0.3013	0.0421	-0.2987	0.0243	-0.3010	0.0183
	$\beta_{3,3}$	0.9	0.9037	0.0656	0.9089	0.0419	0.9066	0.0284	0.9066	0.0215
	$\beta_{3,4}$	-0.5	-0.4908	0.0610	-0.4918	0.0441	-0.4826	0.0341	-0.4845	0.0263
	$\beta_{3,5}$	-0.8	-0.7883	0.0626	-0.7939	0.0436	-0.7995	0.0233	-0.8005	0.0195
	$r_{L,1}$	-0.2	-0.1938	0.0404	-0.1978	0.0321	-0.1997	0.0289	-0.1899	0.0250
	$r_{U,1}$	0.1	0.0949	0.0392	0.0990	0.0353	0.0989	0.0261	0.0910	0.0192
	$r_{L,2}$	0.6	0.6041	0.0486	0.6012	0.0392	0.6031	0.0371	0.6112	0.0346
	$r_{U,2}$	0.8	0.8049	0.0422	0.8067	0.0359	0.8034	0.0348	0.8056	0.0332
	σ^2	1	1.0047	0.0479	1.0211	0.0307	1.0331	0.0249	1.0365	0.0179

Table 3.6: Continued

Empirical size

True model and H_0	H_1	N	α	$T = 10$	$T = 20$	$T = 50$	$T = 100$
			0.01	6	9	8	9
		10	0.05	53	44	56	50
			0.10	103	94	98	98
			0.01	9	17	14	10
Linear	Two-regime $BTPD$	20	0.05	54	54	53	47
			0.10	101	89	96	95
			0.01	13	12	12	10
		50	0.05	49	51	44	49
			0.10	104	98	87	90

Empirical power

True model and H_1	H_0	N	α	$T = 10$	$T = 20$	$T = 50$	$T = 100$
			0.01	829	1000	1000	1000
		10	0.05	935	1000	1000	1000
			0.10	968	1000	1000	1000
			0.01	1000	1000	1000	1000
Two-regime $BTPD$	Linear	20	0.05	1000	1000	1000	1000
			0.10	1000	1000	1000	1000
			0.01	1000	1000	1000	1000
		50	0.05	1000	1000	1000	1000
			0.10	1000	1000	1000	1000

Table 3.7: Rejection frequencies of the bootstrap based test for linearity.

Algeria	Cameroon	Ecuador	Kuwait	Qatar
Australia	Canada	Egypt, Arab Rep	Nigeria	Saudi Arabia
Bolivia	Colombia	Gabon	Norway	Vietnam
Brunei Darussalam	Côte d'Ivoire	Indonesia	Oman	

Table 3.8: Countries of the sample.

3.6 Application

The literature generally assumes the linearity in the dynamics to deal with the relationship between natural resources dependence, economic growth and quality of institutions. However, Leite and Weidmann (1999) and Sala-i-Martin and Subramanian (2013) show that the econometric analysis measuring the impact of natural resources and the quality of institutions on growth are not linear. This nonlinearity has also been analyzed by Belarbi et al. (2016) who validate a two-regime $PSTR$ model. In this section, we propose an application of our model to study the combined interaction effects of natural resources dependence and quality of institutions on economic growth for a panel of 19 countries for the period 1996-2017. Through the buffered regime switching mechanism, we analyze the heterogeneity in the studied panel and how the interaction between natural resources and quality of institutions impacts the economic growth of rentier States. The countries of our sample study are given in Table 3.8.

Variable	Description
GDPG	Growth rate of GDP.
QINST	Rule of law is a governance indicator developed by the World Bank. It includes several indicators that measure the confidence and respect of the laws and rules of society. Its value varies between -2.5 and 2.5. A high value indicates a favorable institutional environment and vice versa.
DEP	Dependence on natural resources is represented by the variable oil rents as a % of GDP. Oil rents are the difference between the value of crude oil production at world prices and total costs of production.
INFL	Macroeconomic stability as measured by the inflation rate.
INVEST	Gross fixed capital formation (GFCF) as a % of GDP.
OPEN	Trade openness: sum of exports and imports of goods and services relative to GDP.
POPG	Population growth as the annual rate of population growth.

Table 3.9: Description of used variables.

To control for dependence on natural resources and quality of institution effects, we introduce, respectively, the variables 'oil rents' and 'rule of law'. The interaction effect can be analyzed by using these variables as explanatory and transition variables at the same time. We add to our econometric specification some other variables used in the traditional empirical literature on the macroeconomic determinants of growth. We have chosen the most used ones in the explanation of growth in the literature. According to several studies (e.g. Barro, 1991, Barro and Sala-i-Martin, 2003 and Jones, 2001), trade openness, fixed investment, moderate inflation and output volatility, and a better educated workforce have helped countries achieve a higher rate of growth. All the variables used in our study are taken from the World Development Indicators database, the World Government based Indicators and the International Financial Statistics. A short description of these variables is provided in Table 3.9.

Table 3.10 provides some descriptive statistics of the used data. It appears that the mean of growth rate of GDP in our sample has decreased between 1996 and 2017. The indicator of the quality of institutions (rule of law) has very slightly increased. Moreover, we observe a significant decrease in the mean level of oil rents as a percentage of GDP, from 14.37% to 8.52%. Inflation has also significantly decreased, approximately from 7.95% to 4.16%. However, the means of investment and growth of population have increased during this period in our sample while openness has very slightly decreased. These tendencies are globally clearly observed in comparing the maximum and minimum

	Mean		Max		Min		St.dev	
	1996	2017	1996	2017	1996	2017	1996	2017
GDPG	4.13	2.337	9.34	7.70	0.61	-2.87	2.23	2.55
QINST	-0.026	0.019	1.92	2.02	-1.44	-1.20	1.00	0.97
DEP	14.37	8.52	40.96	36.61	0.60	0.21	14.25	9.86
INFL	7.95	4.16	29.27	29.50	0.496	-0.83	8.84	7.17
INVEST	23.45	27.26	41.31	48.40	12.11	15.27	8.14	8.77
OPEN	0.67	0.66	1.206	2.00	0.36	0.23	0.231	0.39
POPG	1.84	1.89	3.08	4.67	0.51	0.84	0.628	0.90

Source : Constructed using World Bank and IMF datasets

Table 3.10: Descriptive statistics (1996-2017).

values between these two years.

In our application, we investigate two models to explain the interaction between natural resources and quality of institutions and their impact on GDP growth in our sample. In the first model, we consider the quality of institutions as a threshold variable and the dependence on natural resources as an explanatory variable. In the second model, the dependence on natural resources is considered as the threshold variable and the quality of institutions as an explanatory one.

We first test for the linearity against a two-regime $BTPD$, and thus a two-regime model against a three-regime model.[2] The results of these tests are reported in Table 3.11. It clearly appears that we reject the linearity, in our sample, in both models at 1%. We also reject the two-regime model in favor of the three-regime model at 1% level. We hereafter discuss the implications of this model in explaining the evolution of the growth for these countries that depend on oil rents. The results of our regressions are given in Table 3.12.

Let us first discuss the Model (I) in which the threshold variable is the quality of institutions. Three regimes are clearly identified.[3] In the lower one (regime 1 with low values of the rule of law), it appears that oil rents and openness have no significant effect on the growth of the economy while inflation has a negative impact and the investment

[2] It is important to notice that in practice it is sufficient to consider only the cases $K = 2$ and $K = 3$ to capture the nonlinearities due to regime switching. This is why our analysis is limited to a model with at most three regimes. In our case, this is mainly due to the computational costs of the estimation. Indeed, for our case we have $N \times T = 19 \times 22 = 418$ observations and if the values of a and b are the $10th$ and $90th$ percentiles of the data respectively, we need to estimate 324540216 candidate models when $K = 4$. Note that the creation of arrays with this dimension (324540216×4) causes Matlab to become unresponsive.

[3] Figures 3.5 and 3.6 give the different regimes of the countries in both models. When the number is in black, it means that the country is clearly in this regime. When the color is different, the country is in a buffer zone (i.e. in a possible transition to another regime).

	The linearity against a 2-regime $BTPD$.	
	Model (I)	Model (II)
$\widehat{\gamma}_2$	$[0.1490, 0.9584]$	$[0.9941, 6.3392]$
$S(\widehat{\gamma}_2)$	3503.1316	3760.4068
$F_{1,2}$	55.5199	52.6517
p-value	0.0020	0.0000
	2-regime $BTPD$ against a 3-regime $BTPD$ model.	
	Model (I)	Model (II)
$\widehat{\gamma}_3$	$[-1.0506, 0.0720, 0.1490, 0.9584]$	$[0.9941, 6.3392, 35.9635, 36.6427]$
$S(\widehat{\gamma}_3)$	3263.9008	3324.4110
$F_{2,3}$	29.2451	52.3288
p-value	0.0020	0.0000

Table 3.11: Results of the tests in the $BTPD$ modelling.

Model	Model (I)			Model (II)		
Endogenous variable	GDPG					
Threshold variable	QINST			DEP		
	Lower	Middle	Upper	Lower	Middle	Upper
QINST	–	–	–	3.0868^{**} (2.0027)	-3.3466^{***} (3.7693)	4.3239^{***} (2.5206)
DEP	0.0550 (1.0948)	0.1389^{***} (3.5044)	0.3144^{***} (5.4706)	–	–	–
OPEN	-1.5412 (0.7802)	-0.5650 (0.3474)	5.6116^{**} (2.0016)	1.7676 (0.5819)	2.5070^{*} (1.7458)	5.2574^{*} (1.7982)
INF	-0.2442^{***} (5.3949)	-0.0093 (0.4126)	-0.5179^{**} (2.1196)	-0.3091^{***} (5.6253)	-0.0142 (0.6850)	-0.0321 (0.1759)
INVEST	0.0798^{*} (1.9024)	0.1457^{***} (3.1595)	-0.0069 (0.1090)	0.2213^{***} (3.5098)	-0.0025 (0.0847)	0.1247 (1.2769)
POPG	3.0376^{***} (3.7167)	0.4452^{***} (3.4131)	-1.2796^{***} (3.9515)	1.1948 (1.3368)	0.9298^{***} (7.4040)	-0.6637^{***} (2.7045)
$\widehat{\gamma}_3$	$[-1.0506, 0.0720, 0.1490, 0.9584]$			$[0.9941, 6.3392, 35.9635, 36.6427]$		
AIC	2.1783			2.1966		
BIC	2.3327			2.3511		

The t-statistics are given in parentheses.

***Significant at the 0.01 level, **at the 0.05 level, *at the 0.10 level.

Table 3.12: GDP growth, quality of institutions and natural resources dependence: estimated three-regime $BTPD$ model.

Pays /Année	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Algeria	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Australia	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
Bolivia	2	2	2	2	2	2	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1
Brunei Darussalam	3	3	3	3	3	3	3	3	3	3	2	2	2	2	2	2	2	2	2	2	2	2
Cameroon	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Canada	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
Colombia	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Cote d'Ivoire	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Ecuador	2	2	2	2	2	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1
Egypt	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
Gabon	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Indonesia	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Kuwait	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	2	2	2	2
Nigeria	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Qatar	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3
Norway	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
Oman	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
Saudi Arabia	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
Vietnam	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2

Figure 3.5: GDP growth and quality of institutions: the regime indicator $R_{i,t}$ values obtained from the estimated three-regime $BTPD$ model.

	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Algeria	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
Australia	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Bolivia	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	1	1
Brunei Darussalam	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
Cameroon	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
Canada	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Colombia	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2
Cote d'Ivoire	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Ecuador	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
Egypt	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
Gabon	2	2	2	2	3	2	2	2	2	3	3	2	3	2	2	2	2	2	2	2	2	2
Indonesia	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1	1	1
Kuwait	3	3	2	2	3	3	2	3	3	3	3	3	3	3	3	3	3	3	3	3	2	2
Nigeria	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
Qatar	2	2	2	2	3	2	2	2	2	3	2	2	2	2	2	2	2	2	2	2	2	2
Norway	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
Oman	3	2	2	2	3	3	2	2	3	3	3	3	3	2	2	3	3	3	2	2	2	2
Saudi Arabia	2	2	2	2	3	2	2	3	3	3	3	3	3	2	3	3	3	3	3	2	2	2
Vietnam	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2

Figure 3.6: GDP growth and natural resources dependence: the regime indicator $R_{i,t}$ values obtained from the estimated three-regime $BTPD$ model.

and the population growth have a positive one. In the middle regime (regime 2 with intermediary values of the rule of law), the oils rents have a positive and significant impact on growth as the investment and the population growth. Openness has still no significant effect. In the upper regime (regime 3 with high values of the rule of law), the oils rents have an even higher positive and significant effect on growth. Moreover, openness has a strong and significant positive effect while the population growth and inflation have a negative impact. It thus seems that the higher the rule of law, the higher the positive impact of oils rents on economic growth.

In practice, our modelling shows that most of the countries (14 of 19) did not experience a switch in the growth regime during the period of study. We can classify the countries of our sample into the three regimes as follows. Regime 1: Algeria, Cameroon, Colombia, Côte d'Ivoire, Gabon, Indonesia and Nigeria. Regime 2: Egypt, Oman, Saudi Arabia and Vietnam. Regime 3: Australia, Canada and Norway. However, five countries of our sample have experienced a change in their growth regime. The situation of four of them has worsened: Bolivia and Ecuador (from regime 2 to regime 1), Brunei Darussalam and Kuwait (from regime 3 to regime 2). Only Qatar has improved its situation (from regime 2 to regime 3). It is worth noting that for the most part of the period, these countries were in the buffer zones (see Figure 3.5). They can be considered as being still in transition and the change in the regime is not definitive. More generally, it is important to mention that except the countries of Regime 3 that are very stable during the whole period of study, all the other countries have experienced, at different degrees, a transition in the buffer zones.

For the Model (II), the three regimes show very different impacts of the quality of institutions on the economic growth. In the lower and upper regimes (regimes 1 and 3), the quality of institutions has a positive and significant effect on growth (even though the impact is higher in the upper regime in which the oil rents are very high). In the middle regime (regime 2 in which the oils rents have intermediate values), the effect of the quality of institutions is negative and significant. It means that enhancing the rule of law damages the economic growth for the countries in this regime. One can consider that these countries experience somehow a trap to oil dependence. In the regime 1, we have Australia, Canada and Côte d'Ivoire (see Figure 3.6). The countries belonging to the regime 2 are Algeria, Brunei Darussalam, Cameroon, Ecuador, Egypt, Nigeria and Norway. All the other countries have experienced switching from a regime to another one during the period of study. For example, Colombia (in 2011) and Vietnam (in 2000)

The linearity against a one-threshold model.		
	Model (I)	Model (II)
$\widehat{\gamma}_2 = \widehat{r}_1$	0.4079	36.7892
$S(\widehat{\gamma}_2)$	3798.6519	3855.9799
$F_{1,2}$	20.1600	41.4572
p-value	0.0360	0.0000
One-threshold model against a 2-threshold model.		
	Model (I)	Model (II)
$\widehat{\gamma}_3 = (\widehat{r}_1, \widehat{r}_2)$	$[0.0757, 0.1037]$	$[27.3327, 28.6872]$
$S(\widehat{\gamma}_3)$	3376.3984	3441.7750
$F_{2,3}$	52.4932	48.0182
p-value	0.0000	0.0000

Table 3.13: Results of the tests in the threshold modelling.

have moved from regime 1 to regime 2. In these countries, the quality of institutions exert a negative impact on economic growth (from the transition year) while their oils rents sufficiently increased. Other countries have experienced back and forth changes: between regimes 1 and 2 (Bolivia and Indonesia), and between regimes 2 and 3 (Gabon, Kuwait, Qatar, Oman and Saudi Arabia). It is worth noting that no switch between the extreme regimes 1 and 3 happened during the whole period of study.

For the sake of comparison, we did the same kind of analysis with simple Threshold Regression models of Hansen (1999a). Table 13.131 provides the results of the tests of linearity vs. one-threshold model and one-threshold vs. two-threshold model. Our results clearly reject the linearity against the existence of a threshold at the 5% level. Moreover, the one-threshold model is rejected in favor of the two-threshold model at the 1% level. Table 3.14 provides the results of the estimation of a two-threshold model of Hansen (1999a) with a 0.1% step.

Concerning the Model (I), in which the transition variable is the quality of institutions, the three regimes provide results that differ substantially from ours in particular for the impact of the natural resources dependence on economic growth. Indeed, in regimes 1 and 3, the impact is positive and significant while in regime 2 the impact is negative and significant. However, it is worth noting that regime 2 is very rarely reached in our sample (only five times during the whole period of study, see Figure 3.7). Moreover, it appears that there were some switching from the extreme regimes 1 and 3 contrary to our model. This is explained by the presence of the buffer zone which makes the transition between the regimes smoother. More generally, it appears that our modelling allows us to clearly classify the countries into three categories while this is not the case with simple Threshold

Model	Model (I)			Model (II)		
Endogenous variable	GDPG					
Threshold variable	QINST			DEP		
	Lower	Middle	Upper	Lower	Middle	Upper
QINST	–	–	–	0.4499 (0.5717)	-99.9189^{***} (2.6600)	2.2351 (1.5584)
DEP	0.1035^{***} (2.4978)	-4.3509^{***} (3.4777)	0.1884^{***} (4.0568)	–	–	–
OPEN	-1.7152 (1.1812)	152.2385^{***} (2.5700)	-3.0108 (0.9635)	-0.3238 (0.2492)	781.1823^{***} (3.0363)	-0.1336 (0.0579)
INF	-0.0576^{***} (2.9157)	4.3268^{***} (5.6810)	-0.0037 (0.0319)	-0.0636^{***} (3.2244)	21.0582^{***} (4.0837)	0.2342 (1.6050)
INVEST	0.1016^{***} (2.7473)	-3.2426^{***} (2.5130)	0.0731 (1.6057)	0.0347 (1.1076)	-21.3537^{***} (3.0790)	0.0647 (0.9093)
POPG	0.6112 (0.9315)	50.9949^{***} (3.4523)	0.6168^{***} (4.4795)	0.4377^{*} (1.7740)	13.2968^{***} (3.6394)	0.4620^{***} (2.7784)
$\widehat{\gamma}_3 = (\widehat{r}_1, \widehat{r}_2)$	$[0.0757, 0.1037]$			$[27.3327, 28.6872]$		
AIC	2.2122			2.2313		
BIC	2.3666			2.3858		

Table 3.14: GDP growth, quality of institutions and natural resources dependence (step = 0.1%): estimated three-regime threshold model.

Regression models. The same kind of results are found in the Model (II). The regime 2 is very different from the others but is very rarely reached (only five times, see Figure 3.8). It even turns out that in threshold regressions with a 1% step, the regime 2 is not significant. Our model thus better captures the evolution of the dynamics in the studied sample.

Fundamentally, it is worth noting that from formulation (3.1) we can obtain the classical panel threshold model as a special case by setting $r_{L,k} = r_{U,k} = r_k$ for all $k = 1, ..., K-1$ (i.e. the buffering regions are absent). Therefore, when we choose the vector $\widehat{\gamma}$ which minimizes the sum of the squared errors, we do implicitly consider the classical threshold models as candidates. Indeed, these two classes of models are nested. So minimizing the sum of the squares allows us to implicitly indicate whether it is better to choose our buffered model or the classical threshold model.

We can go further in the interpretation of our results. The minimization of $S(\gamma)$ indicates that the optimal transition mechanism is the one given by buffered modelling (see Tables 3.12 and 3.14). Moreover, we can strengthen our choice strategy between the two classes of models by calculating certain selection criteria known in the literature such as the AIC and BIC criteria that select models via the optimization of a penalized objective function. It turns out that based on the minimum AIC and BIC selection criteria and for a fixed K, the buffered models are preferred to the classical threshold

Pays /Année	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Algeria	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Australia	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
Bolivia	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Brunei Darussalam	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
Cameroon	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Canada	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
Colombia	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Cote d'Ivoire	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Ecuador	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Egypt	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Gabon	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Indonesia	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Kuwait	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	1	1	1	2
Nigeria	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Qatar	1	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
Norway	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
Oman	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
Saudi Arabia	3	2	1	1	1	1	1	3	1	1	1	2	1	1	3	1	3	3	3	3	3	2
Vietnam	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Figure 3.7: GDP growth and quality of institutions: the regime indicator $R_{i,t}$ values obtained from the estimated threshold model.

Pays /Année	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Algeria	1	1	1	1	1	1	1	1	1	3	3	2	3	1	1	1	1	1	1	1	1	1
Australia	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Bolivia	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Brunei Darussalam	1	1	1	1	1	1	1	1	1	1	3	3	1	1	1	1	1	1	1	1	1	1
Cameroon	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Canada	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Colombia	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Cote d'Ivoire	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Ecuador	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Egypt	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Gabon	3	1	1	1	3	1	1	1	1	3	3	3	3	1	3	3	3	3	1	1	1	1
Indonesia	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Kuwait	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
Nigeria	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Qatar	3	2	1	3	3	3	2	3	3	3	3	3	3	1	2	3	3	1	1	1	1	1
Norway	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Oman	3	3	1	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	1	1	1
Saudi Arabia	3	3	1	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	1	1	1
Vietnam	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Figure 3.8: GDP growth and natural resources dependence: the regime indicator $R_{i,t}$ values obtained from the estimated threshold model.

ones.

As we have already mentioned it, an alternative approach to tackle sudden jumps is $PSTR$ modelling (Gonzales et al., 2017). We thus propose a modelling of our data using this approach. More precisely, we consider the following model: $y_{i,t} = \mu_i + X_{i,t}\beta_0 + \sum_{j=1}^{r} X_{i,t}\beta_j \left(1 + \exp\left(-\gamma_j \left(q_{i,t} - c_j\right)\right)\right)^{-1} + u_{i,t}$. Table 3.15 provides the results of our tests of the different number of regimes. While using the quality of institutions as a threshold variable, we do not reject the linearity, Model (I). So a $PSTR$ modelling is not suitable for our data. Indeed, as our $BTPD$ (and even a PTR) modelling rejects linearity, we can say that our approach better fits the data than a $PSTR$ one. For the Model (II) in which the threshold variable is the dependence on natural resources, we accept a three-regime $PSTR$ model. Let us now compare the two models: three-regime $BTPD$ and three-

The linearity against a 2-regime $PSTR$.		
	Model (I)	Model (II)
$\widehat{\gamma}$	70.2885	2.0617
$\widehat{c}$	0.3944	37.6029
SSR_2	3938.0969	3987.9135
$\mathrm{LM}_{1,2}$	5.5343	26.3876
p-value	0.3542	0.0001
2-regime $PSTR$ against a 3-regime $PSTR$ model.		
	Model (I)	Model (II)
$\widehat{\gamma} = (\widehat{\gamma}_1, \widehat{\gamma}_2)$	–	$(25.8914, 2.1034)$
$\widehat{c} = (\widehat{c}_1, \widehat{c}_2)$	–	$(3.9882, 37.5683)$
SSR_3	–	3463.8230
$\mathrm{LM}_{2,3}$	–	42.2983
p-value	–	0.0000

Table 3.15: Results of the tests in the $PSTR$ modelling. The $LM_{k,k+1}$ statistics is defined by $LM_{k,k+1} = TN\,(SSR_k - SSR_{k+1})\,/SSR_k$, where SSR_k is the panel sum of squared residuals under H_0 ($PSTR$ model with k regimes) and SSR_{k+1} is the panel sum of squared residuals under H_1 ($PSTR$ model $(k+1)$ regimes).

Model	Model (I)			Model (II)		
Endogenous variable	GDPG					
Threshold variable	QINST			DEP		
	β_0	β_1	β_2	β_0	β_1	β_2
QINST		–	–	-0.7567 (-0.7515)	-0.7750 (-1.5072)	8.6892^{***} (3.6096)
DEP	0.1639^{***} (5.2576)	–	–	–	–	–
OPEN	-2.7269^{*} (-1.9166)	–	–	-0.8193 (-0.7292)	1.2486 (1.1822)	3.5827 (1.0384)
INF	-0.0509^{**} (-2.4788)	–	–	-0.2536^{***} (-3.7430)	0.2402^{***} (3.3965)	-0.0595 (-0.2948)
INVEST	0.0908^{***} (2.9265)	–	–	0.0882^{**} (2.1498)	-0.0747^{*} (-1.6931)	0.1304 (0.9377)
POPG	0.4829^{***} (3.9210)	–	–	1.1144^{**} (2.5139)	-0.2132 (-0.4643)	-1.6709^{***} (-4.0201)
$\widehat{\gamma} = (\widehat{\gamma}_1, \widehat{\gamma}_2)$	–			$(25.8914, 2.1034)$		
$\widehat{c} = (\widehat{c}_1, \widehat{c}_2)$	–			$(3.9882, 37.5683)$		
AIC	2.2922			2.2797		
BIC	2.3404			2.4631		

Table 3.16: GDP growth, quality of institutions and natural resources dependence: estimated $PSTR$ Model.

regime $PSTR$. The estimation results of the latter are provided in Table 3.16. In this modelling, we have two location parameters $c_1 = 3.98$ and $c_2 = 37.56$. They represent the threshold values that break up the relationship within the different regimes. However, the interpretation of the parameters is less easy than in the $BTPD$ case. Indeed, in the latter the beta coefficients correspond to the effect of the exogenous variables in each regime (lower, middle and upper) identified by the threshold values. In the $PSTR$ case, the beta coefficients should be combined with the threshold functions in order to determine the effect of the exogenous variables on the dependent variable. Roughly speaking, we can for sure determine two extreme regimes: for very low natural resources dependence the effect of the exogenous variables is β_0; for very high dependence rates this effect is given by $\beta_0 + \beta_1 + \beta_2$. In between, the impact of the exogenous variables varies continuously. In practice, the interpretation of the marginal effect is more interesting (e.g. Belarbi et al., 2016). The $BTPD$ modelling thus offers a more precise classification while giving, in the same time, a smooth transition between the different regimes. The effect of the quality of institutions is highly statistically significant in the three regimes while only β_2 is significant in the $PSTR$ model. The $BTPD$ approach thus better captures the heterogeneity of the growth between the countries of our sample.

Conclusion

This thesis contributed to the nonlinarity study in econometric modelling using *TP*. We have considered the *TP* for two different issues.

Through the first chapter, we contribute to the model selection field in nonlinear $SETAR$ framework. We have proposed to extend the Predictive Density Criterion first introduced by Djuric and Kay (1992) for the linear autoregressive model to the $SETAR$ model. More specifically, we have used the arranged autoregression representation (Tsay, 1989) to overcome the difficulty of evaluating the predictive likelihood function in the $SETAR$ case. We have focused on the case in which the number of regimes and the threshold parameters are known. In the future, we plan to extend our PDC criterion to the case in which K, $(r_1, \cdots, r_{K-1})$, d and $\mathbf{p}$ are all unknown. From simulation experiments in Section 1.4, we can conclude that PDC outperforms the criteria AIC, $AICc$ and BIC. Furthermore, the AIC and $AICc$ have proven not reliable selection criteria while BIC and PDC are consistent and when the sample size increases, their frequencies of selection of the true $\mathbf{p}$ and/or d become approximately equal. Also, we showed that PDC performs much better than $AICc$ and BIC when the sample size is small. This performance is supported also by the application of our criterion on two real data examples, namely the annual Wolf's sunspot and the GNP growth of US series. This motivates us to analyze other time series using our criterion and hence, we recommend the use of PDC in practice.

The second aspect from witch we have analyzed the *TP*, is the sudden jump in the theorical structure that its classical definition involves. By considering the definition of a new flexible regime switching defined in Li et al. (2015), we extended this idea to deal with the same problem in the Panel data and the financial time series analysis. This new regime switching is based on the introduction of buffer zone instead of taking a point threshold parameters.

In the second chapter, we have proposed a new formulation to capture and explain the asymmetry and the leverage effect encountered during the analysis of several financial time series. By using the same regime switching mechanism introduced by Li et al. (2015), we defined the BSV model to deal with these stylized facts. After the definition of our model, we proposed an EM algorithm with particle filters and smoothers for the estimation of the model parameters. The performance of the proposed method has been analyzed through a simulation study and the empirical consistency has been shown. This performence has been also demonstrated in the studied real data example where our BSV model proved to be globally better than the $BGARCH$, SV and TSV to fit the HON index series. For a future consideration, an interesting issue is to extend the BSV to take into account the nonlinearity in both the conditional mean and the conditional variance both. Probabilistic properties of BSV are under investigation and we expect that the stationarity conditions are the same obtained within the TSV frawework.

Finally, we proposed an alternative model to deal with nonlinearity in the panel data framework. More precisely, our $BTPD$ is suitable to account for the problem of sudden jumps in PTR models. It turns out to be a very useful tool to study smooth transition between different regimes. It can be considered as a promising alternative to $PSTR$ modeling and provides results that are more easily interpretable. We have provided a statistical toolbox starting with an OLS estimation method, arriving at the adequacy tests where we have proposed bootstrap procedures to detect the regime change nonlinearity in the dynamics of the underlying variable related to the existence of a buffer zone as well as the number of regimes to which the phenomenon under study is obeying. Intensive simulations experiments have been realised to asses the perfomance of the proposed precedures, where the empirical consistency of the presented methods has been shown. Moreover, our application shows that it can provide a richer and more precise description of the evolution of growth of rentier States while taking into account the interaction between natural resources dependence and quality of institutions than PTR and $PSTR$ modellings. Our $BTPD$ formulation better highlights the impact of natural resources dependence and the quality of institutions on economic growth. Indeed, our modeling allows us to clearly classify the countries into three categories while this is not the case with classical PTR and $PSTR$ models. We do think that our model offers a wide perspective in terms of applications to different frameworks and could enhance our understanding of the dynamics of evolution of various phenomena, in particular economic ones. Our work can also be developed in many directions. Firstly, the study of the theoretical properties of our estimator (consistency and asymptotic normality) deserves to be dealt carefully and analyzed in detail. Secondly, an interesting perspective of research would be the development of a dynamic version of our $BTPD$ model. Finally, the estimation of the buffer zones should be thoroughly investigated in order to reduce the algorithmic complexity generated by the estimation step.

Bibliography

[1] Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In: Proceedings of the Second International Symposium on Information Theory. Akademiai Kiadó, Budapest, 267-281.

[2] Andrews, D. W. (1993). Tests for parameter instability and structural change with unknown change point. *Econometrica: Journal of the Econometric Society*, 821-856.

[3] Andrews, D. W. (2003). Tests for parameter instability and structural change with unknown change point: A corrigendum. *Econometrica*, 395-397.

[4] Andrews, D. W., & Ploberger, W. (1994). Optimal tests when a nuisance parameter is present only under the alternative. *Econometrica: Journal of the Econometric Society*, 1383-1414.

[5] Bai, J. (1997). Estimating multiple breaks one at a time. *Econometric theory*, 315-352.

[6] Bai, J., Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica*, 47-78.

[7] Balestra, P., & Nerlove, M. (1966). Pooling cross section and time series data in the estimation of a dynamic model: The demand for natural gas. *Econometrica: Journal of the econometric society*, 585-612.

[8] Baltagi, B. (2008). Econometric analysis of panel data. John Wiley & Sons.

[9] Barro, R. J. (1991). Economic Growth in a Cross Section of Countries. *The Quarterly Journal of Economics*, 106 (2), 407–443.

[10] Barro, R., Sala-i-Martin, X. (2003). *Economic growth*. Cambridge, MA: The MIT Press.

[11] Baum, L. E., & Petrie, T. (1966). Statistical inference for probabilistic functions of finite state Markov chains. *The annals of mathematical statistics*, 37(6), 1554-1563.

[12] Beaudry, P., & Koop, G. (1993). Do recessions permanently change output?. *Journal of Monetary economics*, *31*(2), 149-163.

[13] Bekaert, G., & Wu, G. (2000). Asymmetric volatility and risk in equity markets. *The review of financial studies*, **13**(1), 1-42.

[14] Belarbi, Y. , Sami L., Souam S.(2016) The effects of institutions and natural resources in heterogeneous growth regimes,*Middle East Development Journal*, 8 (2), 248-265.

[15] Belke, A., Kronen, D. (2019). Exchange Rate Bands of Inaction and Hysteresis in EU Exports to the Global Economy: The Role of Uncertainty. *Journal of Economic Studies*, 46 (2), 335-355.

[16] Bentarzi, M., Guerbyenne, H., & Hemis, R. (2008). Predictive Density Order Selection of Periodic *AR* Models. *Communications in Statistics - Simulation and Computation*, **37**, 1167-1182.

[17] Black, F. (1976). Studies of stock market volatility changes. *Proceedings of the Business and Economics Section of the American Statistical Association*, 177–181.

[18] Borgersen, T.R., Göcke, M. (2007). Exchange rate overshooting and path-dependence in international trade. *Macroeconomic Dynamics*, 11 (3), 295-317.

[19] Boussaha, N., & Hamdi, F. (2018). On periodic autoregressive stochastic volatility models: structure and estimation. *Journal of Statistical Computation and Simulation*, 88(9), 1637-1668.

[20] Boussaha, N., Hamdi, F., & Souam, S. (2018). Multivariate Periodic Stochastic Volatility Models: Applications to Algerian dinar exchange rates and oil prices modeling (No. 2018-14). University of Paris Nanterre, EconomiX.

[21] Breidt, F. J. (1996). A threshold autoregressive stochastic volatility model. In VI Latin American Congress of Probability and Mathematical Statistics (CLAPEM), Valparaiso, Chile.

[22] Brooks, C. (2001). A Double-threshold GARCH Model for the French Franc/Deutschmark exchange rate. *Journal of Forecasting*, **20**(2), 135-143.

[23] Caner, M., Hansen, B. (1998). Threshold autoregressions with a near unit root. Department of Economics, Bilkent University.

[24] Carnero, M. A., Peña, D., Ruiz, E. (2004). Persistence and kurtosis in GARCH and stochastic volatility models. *Journal of Financial Econometrics*, **2**(2), 319-342.

[25] Chan, K. S., Tong, H. (1986). On estimating thresholds in autoregressive models. *Journal of time series analysis*, **7**(3), 179-190.

[26] Chan, K. S. (1990). Testing for threshold autoregression. *The Annals of Statistics*, 1886-1894.

[27] Chan, K. S. (1991). Percentage points of likelihood ratio tests for threshold autoregression. *Journal of the Royal Statistical Society: Series B (Methodological)*, **53**(3), 691-696.

[28] Chan, K.S. (1993). Consistency and limiting distribution of the least squares estimator of a threshold autoregressive model. *The Annals of Statistics*, **21**, 520-533.

[29] Chan, K.S., Petruccelli, J.D., Tong, H., & Woolford, S.W. (1985). A multiple-threshold AR (1) model. *Journal of applied probability*, **22**, 267-279.

[30] Chan, K.S., & Tong, H. (1985). On the use of the deterministic Lyapunov function for the ergodicity of stochastic difference equations. *Advances in Applied Probability*, **17**, 666-678.

[31] Chen, C. W., Chiang, T. C., & So, M. K. (2003). Asymmetrical reaction to US stock-return news: evidence from major stock markets based on a double-threshold model. *Journal of economics and business*, **55**(5-6), 487-502.

[32] Chen, C. W., So, M. K., & Gerlach, R. H. (2005). Assessing and testing for threshold nonlinearity in stock returns. *Australian & New Zealand Journal of Statistics*, **47**(4), 473-488.

[33] Chen, C. W., & So, M. K. (2006). On a threshold heteroscedastic model. *International Journal of Forecasting*, **22**(1), 73-89.

[34] Chen, C. W., Liu, F. C., & So, M. K. (2008). Heavy-tailed distributed threshold stochastic volatility models in financial time series. *Australian & New Zealand Journal of Statistics*, **50**(1), 29-51.

[35] Chen, C. W., Gerlach, R. H., & Lin, A. M. (2010). Falling and explosive, dormant, and rising markets via multiple-regime financial time series models. *Applied Stochastic Models in Business and Industry*, **26**(1), 28-49.

[36] Chen, C. W., Liu, F. C., & So, M. K. (2011). A review of threshold time series models in finance. *Statistics and its Interface*, **4**(2), 167-181.

[37] Chen, C. W., Truong, B. C. (2016). On double hysteretic heteroskedastic model. *Journal of Statistical Computation and Simulation*, **86** (13), 2684-2705.

[38] Clements, M.P., & Smith, J. (1997). The performance of alternative forecasting methods for $SETAR$ models. *International Journal of Forecasting*, **13**, 463-475.

[39] Clements, M.P., & Smith, J. (2000). Evaluating the forecast densities of linear and non-linear models: applications to output growth and unemployment. *Journal of Forecasting*, **19**, 255-276.

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

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

[42] De Gooijer, J.G. 2001. Cross-validation criteria for $SETAR$ model selection. *Journal of Time Series Analysis*, **22**, 267-281.

[43] Dijk, D. V., Teräsvirta, T., & Franses, P. H. (2002). Smooth transition autoregressive models—a survey of recent developments. *Econometric reviews*, **21**(1), 1-47.

[44] Djuric, P.M., & Kay, S.M. (1992). Order Selection of Autoregressive Models, *IEEE Tranaction on Signal Processing*, **40**, 2829-2833.

[45] Enns, R. H. (2010). It's a nonlinear world. Springer Science & Business Media.

[46] Fenga, L., & Politis, D.N. (2015). Bootstrap order selection for $SETAR$ models. *Journal of Statistical Computation and Simulation*, **85**, 235-250.

[47] Fok, D., Van Dijk, D., & Franses, P. H. (2005). A multi-level panel STAR model for US manufacturing sectors. *Journal of Applied Econometrics*, 20(6), 811-827.

[48] Franses, P. H., & Van Dijk, D. (2000). Non-linear time series models in empirical finance. Cambridge University Press.

[49] Galeano, P., & Peña, D. 2007. Improved model selection criteria for $SETAR$ time series models. *Journal of statistical planning and inference*, **137**, 2802-2814.

[50] Geweke, J., & Terui, N. (1993). Bayesian threshold autoregressive models for nonlinear time series. *Journal of Time Series Analysis*, **14**, 441-454.

[51] Göcke, M. (2002). Various concepts of hysteresis applied in economics. *Journal of Economic Surveys*, **16** (2), 167-188.

[52] Goldfeld, S., & Quandt, R. (1973). The estimation of structural shifts by switching regressions. In *Annals of Economic and Social Measurement*, Volume 2, number 4 (pp. 475-485). NBER.

[53] Gonzalez, A., Teräsvirta, T., Van Dijk, D., & Yang, Y. (2017). Panel smooth transition regression models.

[54] Gourieroux, C., & Monfort, A. (1992). Qualitative threshold ARCH models. *Journal of Econometrics*, **52**(1-2), 159-199.

[55] Hall, P. (1985). Resampling a coverage pattern. *Stochastic processes and their applications*, **20**(2), 231-246.

[56] Hamdi, F., & Khalfi, A. (2019). Predictive density criterion for SETAR models. *Communications in Statistics-Simulation and Computation*, 1-17

[57] 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.

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

[59] Hansen, B. E. (1997). Inference in TAR model. *Studies in Nonlinear Dynamics and Econometrics* . **1**(2), 1-14.

[60] Hansen, B. E. (1999a). Threshold effects in non-dynamic panels: Estimation, testing, and inference. *Journal of econometrics*, 93(2), 345-368.

[61] Hansen, B. (1999b). Testing for linearity. *Journal of economic surveys*, **13**(5), 551-576.

Hansen, B. E. (2000). Sample splitting and threshold estimation. *Econometrica*, **68**(3), 575-603.

[62] Hansen, B. E. (2002). Tests for parameter instability in regressions with I (1) processes. *Journal of Business Economic Statistics*, **20**(1), 45-59.

[63] Hansen, B. E. (2011). Threshold Autoregressions in Economics, *Statistics and its Interface*, **4**, 123-127.

[64] Harvey, A. C., & Shephard, N. (1996). Estimation of an asymmetric stochastic volatility model for asset returns. *Journal of Business & Economic Statistics*, **14**(4), 429-434.

[65] Hili, O. (1992). Sur les modèles autorégressifs à seuils. *Comptes rendus de l'Académie des sciences. Série 1, Mathématique*, **314**, 573-576.

[66] Hsiao, C. (2003). Analysis of Panel Data (Econometric Society Monographs, 34). Cambridge Univ.

[67] Hsiao, C. (2007). Panel data analysis—advantages and challenges. *Test*, 16(1), 1-22.

[68] Hsiao, C. (2014). Analysis of panel data (No. 54). Cambridge university press.

[69] Hurvich, C. M., Shumway, R., & Tsai, C. L. (1990). Improved estimators of Kullback–Leibler information for autoregressive model selection in small samples. *Biometrika*, **77**(4), 709-719.

[70] Jacquier, E., Polson, N. G., & Rossi, P. E. (2004). Bayesian analysis of stochastic volatility models with fat-tails and correlated errors. *Journal of Econometrics*, **122**(1), 185-212.

[71] Jones, C. I. (2001). Was an Industrial Revolution Inevitable? Economic Growth Over the Very Long Run. *Advances in Macroeconomics*, **1** (2).

[72] Kim, J., & Stoffer, D. S. (2008). Fitting stochastic volatility models in the presence of irregular sampling via particle methods and the EM algorithm. *Journal of time series analysis*, 29(5), 811-833.

[73] Koutmos, G. (1999). Asymmetric price and volatility adjustments in emerging Asian stock markets. Journal of Business Finance & Accounting, 26(1-2), 83-101.

[74] Kuo, B. S. (1998). Test for partial parameter instability in regressions with I (1) processes. *Journal of Econometrics*, **86**(2), 337-368.

[75] Lee, J. (2000). The robustness of Okun's law: Evidence from OECD countries. *Journal of macroeconomics*, 22(2), 331-356.

[76] Leeuw, J. (1994). The covariance matrix of ARMA errors in closed form. *Journal of econometrics*, **63**(2), 397-405.

[77] Leite, C., Weidmann, J. (1999). Does Mother Nature corrupt? Natural resources, corruption and economic growth (Working Paper of the International Monetary Fund, IMF Working Paper WP/99/85), Washington, DC.

[78] Li, W.K. (1988). The Akaike information criterion in threshold modelling: some empirical evidences. In *Nonlinear time series and signal processing* (pp. 88-96). Springer, Berlin, Heidelberg.

[79] Li, W. K. (2009). The Threshold Approach in Volatility Modelling. In Exploration of a Nonlinear World: An Appreciation of Howell Tong's Contributions to Statistics (pp. 95-100).

[80] Li, W. K., & Lam, K. (1995). Modelling asymmetry in stock returns by a threshold autoregressive conditional heteroscedastic model. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 44(3), 333-341.

[81] Li, C. W., & Li, W. K. (1996). On a double-threshold autoregressive heteroscedastic time series model. Journal of applied econometrics, 11(3), 253-274.

[82] Li, G., Guan, B., Li, W. K., & Yu, P. L. (2015). Hysteretic autoregressive time series models. *Biometrika*, **102**(3), 717-723.

[83] Ling, S. (1999). On the probabilistic properties of a double threshold $ARMA$ conditional heteroskedastic model. *Journal of Applied Probability*, **36**, 688-705.

[84] Liu, J., Li, W. K., & Li, C. W. (1997). On a threshold autoregression with conditional heteroscedastic variances. *Journal of Statistical Planning and Inference*, **62**(2), 279-300.

[85] Lo, P. H., Li, W. K., Yu, P. L., Li, G. (2016). On buffered threshold GARCH models. *Statistica Sinica*, 1555-1567.

[86] Malmsten, H., Teräsvirta, T. (2004). Stylized facts of financial time series and three popular models of volatility. *SSE/EFI Working Paper Series in Economics and Finance*, 563, 1-44.

[87] Mao, X., Ruiz, E., & Veiga, H. (2017). Threshold stochastic volatility: Properties and forecasting. *International Journal of Forecasting*, **33**(4), 1105-1123.

[88] McCulloch, R. E., & Tsay, R. S. (1994). Statistical analysis of economic time series via Markov switching models. *Journal of time series analysis*, 15(5), 523-539.

[89] McQuarrie, A., Shumway, R., & Tsai, C.L. (1997). The model selection criterion *AICu*. *Statistics & probability letters*, **34**, 285-292.

[90] Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*, 347-370.

[91] Pakes, A., & Griliches, Z. (1984). Estimating distributed lags in short panels with an application to the specification of depreciation patterns and capital stock constructs. The Review of Economic Studies, 51(2), 243-262.

[92] Pesaran, M. H., & Potter, S. M. (1997). A floor and ceiling model of US output. *Journal of Economic Dynamics and Control*, **21**(4-5), 661-695.

[93] Petruccelli, J.D., & Woolford, S.W. (1984). A threshold $AR(1)$ model. *Journal of Applied Probability*, **21**, 270-286.

[94] Potter, S. M. (1995). A nonlinear approach to US GNP. *Journal of applied econometrics*, **10**(2), 109-125.

[95] Quandt, R. E. (1958). The estimation of the parameters of a linear regression system obeying two separate regimes. *Journal of the american statistical association*, 53(284), 873-880.

[96] Quandt, R. E. (1960). Tests of the hypothesis that a linear regression system obeys two separate regimes. *Journal of the American statistical Association*, **55**(290), 324-330.

[97] Rabemananjara, R., & Zakoian, J. M. (1993). Threshold ARCH models and asymmetries in volatility. *Journal of Applied Econometrics*, **8**(1), 31-49.

[98] Rodríguez, M. J., & Ruiz, E. (2012). Revisiting several popular GARCH models with leverage effect: Differences and similarities. *Journal of Financial Econometrics*, **10**(4), 637-668.

[99] Sala-i-Martin, X., Subramanian, A. (2013). Addressing the natural resource curse: An illustration from Nigeria. *Journal of African Economies*, **22** (4), 570-615.

[100] Schwarz, G. (1978). Estimating the dimension of a model. *The annals of statistics*, **6**, 461-464.

[101] Seo, B. (1998). Tests for structural change in cointegrated systems. *Econometric Theory*, 222-259.

[102] Shephard, N. (1996). Statistical aspects of ARCH and stochastic volatility. *Monographs on Statistics and Applied Probability*, 65, 1-68.

[103] So, M. K., Li, W. K., & Lam, K. (2002). A threshold stochastic volatility model. Journal of Forecasting, 21(7), 473-500.

[104] So, M. K., Choi, C. Y. (2009). A threshold factor multivariate stochastic volatility model. *Journal of Forecasting*, **28**(8), 712-735.

[105] Stinchcombe, M. B., White, H. (1998). Consistent specification testing with nuisance parameters present only under the alternative. *Econometric theory*, 295-325.

[106] Taylor, S. J. (1986), Modelling Financial Time Series. Chichester, New York, Wiley.

[107] Tiao, G. C., & Tsay, R. S. (1994). Some advances in non-linear and adaptive modelling in time-series. *Journal of forecasting*, **13**(2), 109-131.

[108] Tong, H. (1977). Discussion of a paper by AJ Lawrance and NT Kottegoda. *Journal of Royal Statistical Society: Series A*, **140**, 34-35.

[109] Tong, H., & Lim, K.S. (1980). Threshold autoregression, limit cycles and cyclical data (with discussion). *Journal of the Royal Statistical Society. Series B (Methodological)*, 245-292.

[110] Tong, H., (1983). *Threshold models in non-linear time series analysis.* Springer Lecture Notes in Statistics, Vol. 21. New York: Springer.

[111] Tong, H. (1978). On a threshold model. In Pattern Recognition and Signal Processing (ed. C. H. Chen). Amsterdam: Sijthoff and Noordhoff, 575-586.

[112] Tong, H. (1990). Non-Linear Time Series: A Dynamical System Approach. Oxford Statistical Science Series, 6.

[113] Tong, H. (2009). A personal overview of non-linear time series analysis from a chaos perspective. *In Exploration Of A Nonlinear World*: An Appreciation of Howell Tong's Contributions to Statistics (pp. 183-229).

[114] Tsay, R.S. (1989). Testing and Modeling Threshold Autoregressive Processes. *Journal of the American Statistical Association*, **84**, 231-240.

[115] White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica: journal of the Econometric Society*, 817-838.

[116] Wong, C.S., & Li, W.K. (1998). A note on the corrected Akaike information criterion for the threshold autoregressive models. *Journal of Time Series Analysis*, **19**, 113-124.

[117] Wong, C. S., & Li, W. K. (2000). On a mixture autoregressive model. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **62**(1), 95-115.

[118] Wooldridge, J. M. (2010). Econometric analysis of cross section and panel data. MIT press.

[119] Wu, S., & Chen, R. (2007). Threshold variable determination and threshold variable driven switching autoregressive models. *Statistica Sinica*, **17**(1), 241-S38.

[120] Xu, D. (2012). Examining realized volatility regimes under a threshold stochastic volatility model. *International Journal of Finance Economics*, **17**(4), 373-389.

[121] Yule, G. U. (1927). On a method of investigating periodicities in disturbed series with specia1 reference to Wolfer's sunspot numbers, *Philos. Trans. Roy. Soc., London, Series A*, **226**, 267-298.

[122] Zakoïan, J. M. (1994). Threshold heteroskedastic models. *Journal of Economic Dynamics and control*, **18**(5), 931-955.

[123] Zakoïan, J. M. (1991). Threshold heteroskedastic models. D. P. INSEE.

[124] Zhu, K., Yu, P. L., & Li, W. K. (2014). Testing for the buffered autoregressive processes. *Statistica Sinica*, 971-984.

[125] Zhu, K., Li, W. K., Yu, P. L. (2017). Buffered autoregressive models with conditional heteroscedasticity: An application to exchange rates. *Journal of Business & Economic Statistics*, **35** (4), 528-542.