International
Journal
Of Advanced Research in Engineering & Management (IJAREM)
ISSN: 2456-2033 || PP. 32-53

Modeling the Non-Linearity of the Moroccan Stock Market: Application of the Smooth Transition Autoregressive (STAR) Model

Eljoumani Najat,
PhD Researcher
Mohammed V University-Rabat-Morocco
Faculty of Legal, Economic and Social Sciences-Agdal

Abstract: Objective: The aim of this paper is to test the weak form of informational efficiency of the Moroccan stock market using a non-linear approach.

Methodology: For this purpose, we used the daily closing prices of the *MASI* index from the Casablanca Stock Exchange, covering the period from 03/01/2002 to 08/15/2023, totaling 5393 observations. We applied the Smooth Transition Autoregressive (*STAR*) model to the series of geometric returns of the *MASI* index to capture the non-linear property.

Results: Following a four-step specification procedure, our choice favored the Logistic Smooth Transition Autoregressive (*LSTAR*) model over the Exponential Smooth Transition Autoregressive (ESTAR) model. We estimated the *LSTAR* model and found that the parameters of the base and alternative parts of both linear and non-linear regimes, the transition parameter, and the transition threshold are all statistically significant at the 1% significance level. It is noteworthy that the estimated *LSTAR* model's transition parameter value suggests a relatively slow transition from one regime to another.

Conclusion: The conclusions derived from the estimation of the *LSTAR* model indicate the presence of non-linearity in the series of geometric returns of the *MASI* index. Thus, it can be concluded that the Moroccan stock market exhibits inefficiency according to the weak form.

Keywords: Informational efficiency, non-linearity, transition regimes, logistic and exponential STAR models, transition parameter, transition threshold.

JEL classification: C01, C13, C22, C24, C52, G14

1. Introduction

Several empirical studies have demonstrated that linear models fail to explain the stylized features of financial time series, such as leptokurtic distributions, volatility clustering, leverage effect, and asymmetry. Faced with this limitation of linear models, researchers have embarked on the development of non-linear time series models to account for the non-linear dynamics of the data.

In recent years, the application of regime-switching models in the fields of finance and economics has garnered increasing interest. Initially designed to represent expansion and recession phases in economic cycles, these regime-switching models have become particularly prevalent in the dynamic modeling of stock market returns.

An example of a model capable of capturing regime changes is the Smooth Transition Autoregressive (*STAR*) model, developed by Teräsvirta (1994) and Teräsvirta et al. (2002). The *STAR* model is an extension of the commonly used Autoregressive (*AR*) model, suggesting the existence of non-linear behavior within the time series. In this framework, the transition between regimes is governed by a transition function, typically described as a logistic or exponential function.

Unlike the Threshold Autoregressive (*TAR*) model, initially developed by Tong and Lim (1980) and extensively discussed in Tong (1983) and Tong (1990), which operates a sudden transition, Smooth Transition Regression (*STR*) models operate a transition smoothly, based on a transition variable.

The main added value of *STAR* models lies in the fact that variations observed in financial series are influenced by the evolving behavior of various participants. It is unlikely that all these participants react simultaneously to a specific economic signal.

In financial markets, characterized by the participation of numerous investors and speculators adjusting their positions at different times, the idea of a smooth transition or a continuum of states between extremes seems to better reflect reality. The diversity of investors' objectives stems from their varied investment horizons, diverse geographical locations, as well as distinct profiles and levels of risk aversion. Moreover, investors may exhibit different degrees of institutional inertia, depending, for example, on the efficiency of the stock markets in which they operate, and adjust their positions with distinct time lags. $STAR$ models provide precise flexibility to model such gradual changes while remaining adaptable enough to include discrete changes as a special case.

This study adopts the Smooth Transition Autoregressive ($STAR$) modeling framework to examine the non-linear dynamics of the Moroccan stock market. The structure of the document is as follows: Section 2 provides a literature review focusing on empirical works applying the $STAR$ model. Section 3 presents the data and the adopted methodology. Section 4 discusses and interprets the empirical results obtained. Finally, Section 5 concludes the study.

2. Literature Review

Over the past two decades, there has been a steady growth in interest in non-linear time series models. In the context of financial time series, models allowing for regime changes have been particularly favored, especially Smooth Transition Autoregressive ($STAR$) models. In this section, we review some empirical studies that have focused on the applications of $STAR$ models to financial time series.

Sarantis (1999) analyzed non-linearities in the real effective exchange rates of ten major G-10 industrial economies, applying Smooth Transition Autoregressive ($STAR$) models. Lagrange Multiplier (LM) linearity tests rejected linearity for eight exchange rates during the 1980s and 1990s. The results showed that exchange rates in three countries follow logistic $STAR$ models, while other exchange rates are governed by exponential $STAR$ models. Parameter estimates in $STAR$ models indicated a relatively slow transition speed from one exchange rate regime to another for all countries. Regarding out-of-sample forecasting performance, the authors found little difference between $STAR$ models and linear models. However, $STAR$ models outperformed the Markov regime-switching model.

Similarly, Tayyab et al. (2012) examined the adequacy of specifying Smooth Transition Autoregressive ($STAR$) models for modeling monthly real exchange rates in Pakistan from 1980 to 2010. The authors selected the $AR(1)$ autoregressive model as the best fit to the data. They found that the logistic $STAR$ model is preferred over the exponential $STAR$ model. The results indicated that the monthly real exchange rate market in Pakistan exhibits non-linearity during the study period.

Chien-Jen et al. (2013) applied Smooth Transition Autoregressive models with exogenous regressors ($STARX$) to assess the relationship between the Yuan/Dollar exchange rate and the stock returns of Shanghai A, Shanghai B, Shenzhen A, and Shenzhen B indices in China. The authors initially estimated the optimal linear regression model, where the dependent variable consists of the stock returns of Shanghai and Shenzhen, and the explanatory variables are lagged Yuan/Dollar exchange rates. The results suggested strong statistical significance for delayed exchange rates of one and four periods. Lagrange Multiplier (LM) tests revealed that the delay parameter, $d = 1$, was the best choice, and the null hypothesis of linearity in the LM test was rejected. The authors found that the logistic $STARX$ non-linear model provided the best forecasting performance for both Shanghai and Shenzhen stock markets. The results showed that Shanghai B and Shenzhen B indices have a higher transition speed than Shanghai A and Shenzhen A indices, possibly due to greater accessibility to foreign investors.

Si Mohammed et al. (2015) also investigated the real exchange rates of Algeria, modeling non-linearity using monthly data for the period M1:1994 to M4:2015, totaling 256 observations, applying Smooth Transition Autoregressive models. Test results rejected the null hypothesis of linearity in favor of the alternative hypothesis of non-linearity. Following Teräsvirta's (1994) strategy, the study suggested the use of a lagged real exchange rate of 1 as the transition variable and the logistic $STAR$ model as the suitable model for fitting the data.

Siti Rohani et al. (2015) applied the Smooth Transition Autoregressive ($STAR$) model to the daily stock returns of Malaysia Airlines (MAS) from August 29, 1996, to September 26, 2014, totaling 4450 observations. Following Teräsvirta's (1994) strategy, they first selected the AR(3) autoregressive model as the optimal model that best fits the data. Then, using Lagrange Multiplier (LM) tests, they found that the delay parameter, d=3, was the best choice, and the null hypothesis of linearity in the LM test was rejected. Applying the nested hypothesis sequence with $d = 3$, they found that the logistic $LSTAR$ model was preferred over the exponential $ESTAR$ model. After estimating the parameters of the chosen $LSTAR$ model, they conducted forecasts and compared them with other models. They found that the $LSTAR$ model outperformed the $ESTAR$ and $AR(3)$ models.

Similarly, Akintunde et al. (2016) applied the Smooth Transition Autoregressive ($STAR$) model to the monthly returns of the Botswana stock market from January 1987 to December 2012, covering three hundred months. They first selected the $AR(5)$ autoregressive model as the optimal model. Then, using Lagrange Multiplier (LM) tests, they found that the delay parameter, $d = 2$, was the best choice, and the null hypothesis of linearity in the LM test was rejected. They also found that the logistic $LSTAR$ autoregressive model was preferred over the exponential $ESTAR$ model and performed better in terms of forecasting.

Additionally, Daabaji (2018) conducted a study to test the efficiency hypothesis of the Moroccan stock market by applying the Smooth Transition Autoregressive ($STAR$) model to the daily returns of the $MASI$ index from the Casablanca Stock Exchange from January 1, 2004, to December 31, 2017. Based on the partial autocorrelation of the MASI returns series, the author chose the $AR(3)$ autoregressive model as the best-fitting model. The results of Lagrange Multiplier (LM) tests revealed that the delay parameter, $d = 3$, was the best choice, and the null hypothesis of linearity in the LM test was rejected. Additionally, the exponential logistic $ESTAR$ model was preferred over the exponential $LSTAR$ model, and the study concluded by estimating the parameters of the $ESTAR$ model.

Usman et al. (2018) compared the performances of Smooth Transition Autoregressive ($STAR$) models and linear Autoregressive (AR) models for monthly returns in Turkey and the FTSE Travel & Leisure index from April 1997 to August 2016. The $MSCI$ World index was used as a proxy for the global market. Linearity tests rejected the null hypothesis of linearity. However, the results showed that the logistic $STAR$ model did not improve out-of-sample forecasts compared to the linear AR model, indicating little gain in using the $LSTAR$ model in predicting the Travel & Leisure stock index.

3. Data and Methodology

3.1 Data

The data utilized in this study consist of the time series of daily closing prices, denoted as p_t, for the $MASI$ index covering the period from 03/01/2002 to 08/15/2023, totaling 5393 observations. Subsequently, these prices were transformed into geometric returns:

$$r_t = ln\left(\frac{p_t}{p_{t-1}}\right) \quad (1)$$

Where ln represents the natural logarithm.The data were collected from the investing.com website (www.investing.com).

3.2 Methodology

In this section, we will present the Smooth Transition Autoregressive model, denoted as $STAR$, developed by Teräsvirta (1994) and Van Dijk et al. (2002). However, we will start by introducing the Threshold Autoregressive model (abrupt transition), denoted as TAR, initially proposed by Tong and Lim (1980) and extensively discussed in the works of Tong, H. (1983) and Tong (1990). This model posits that the dynamic behavior of economic and financial time series depends on different regimes over time and can be adequately described by a linear AR model in each of these regimes. Our focus will be on models that are limited to two regimes.

3.2.1 Threshold Autoregressive Model (TAR)

A stochastic process (y_t) is said to satisfy a Threshold Autoregressive model of order p, denoted as $TAR(p)$, if it can be expressed in the form: for all $t > p$

$$y_t = (\phi_{10} + \phi_{11}.y_{t-1} + \cdots + \phi_{1p}.y_{t-p}).\mathbb{I}(s_t \leq c) + (\phi_{20} + \phi_{21}.y_{t-1} + \cdots + \phi_{2p}.y_{t-p}).\mathbb{I}(s_t > c) + \varepsilon_t \quad (2)$$

or alternatively:

$$y_t = \Phi_1^{'}Z_t.\mathbb{I}(s_t \leq c) + \Phi_2^{'}Z_t.\mathbb{I}(s_t > c) + \varepsilon_t \quad (3)$$

with :

- ϕ_{ij} are the real coefficients to be estimated in the model for $1 \leq i \leq 2$ and $1 \leq j \leq p$, grouped in the single-column matrix:

$$\Phi_i = (\phi_{i0}\phi_{i1} \cdots \phi_{ip})^{'} \quad (4)$$

- c is the threshold parameter.
- s_t is the transition variable.

- $\mathbb{I}(.)$ is the indicator function, where $\mathbb{I}(A) = 1$ if event A occurs and $\mathbb{I}(A) = 0$ if event A does not occur.
- (ε_t) is a conditionally i.i.d white noise given the set $\Im_{t-1} = \{y_{t-1}, y_{t-2}, \cdots, y_{t-p}\}$ of information up to date $t-1$, meaning $E(\varepsilon_t/\Im_{t-1}) = 0$ and $E(\varepsilon_t^2/\Im_{t-1}) = \sigma^2$.
- The matrix Z_t is defined by:

$$Z_t = \left(1 \quad y_{t-1} y_{t-2} \cdots y_{t-p}\right)' \tag{5}$$

The model (3) can also be written as: for all $t > p$

$$y_t = \Phi_1' Z_t . \left(1 - \mathbb{I}(s_t > c)\right) + \Phi_2' Z_t . \mathbb{I}(s_t > c) + \varepsilon_t \tag{6}$$

When the transition variable s_t is taken as a lagged value of the time series itself, i.e., $s_t = y_{t-d}$ for a certain integer $d > 0$, the $TAR(p)$ model is called a Self-Exciting Threshold Autoregressive model, denoted as $SETAR(p)$for Self Excited Threshhold Autorgressive. In this case, we can write:

$$E(y_t/\Im_{t-1}) = \begin{cases} \Phi_1' Z_t & si \ y_{t-d} \le c \\ \Phi_2' Z_t & si \ y_{t-d} > c \end{cases} \tag{7}$$

The main limitation of the *TAR* model lies in the fact that the transition between the two regimes occurs instantly at the threshold value c. Therefore, if the transition from one regime to another happens gradually over time rather than instantly when the value c is reached, the *TAR* model fails to adequately capture this gradual transition between the two regimes.

3.2.2 Smooth Transition Autoregressive Model (STAR)

To overcome the inherent limitation of the *TAR* model, one solution is to adopt the Smooth Transition Autoregressive model, denoted as *STAR*, developed by Teräsvirta (1994) and Van Dijk et al. (2002).

A more gradual transition between different regimes can be achieved by substituting the indicator function $\mathbb{I}(s_t > c)$in equation (6) with a continuous function $G(s_t, \gamma, c)$, which smoothly varies from 0 to 1 ass_t increases. The resulting model is called a Smooth Transition Autoregressive model (*STAR*) and is given by:

$$\begin{aligned} y_t = & \left(\phi_{10} + \phi_{11} . y_{t-1} + \cdots + \phi_{1p} . y_{t-p}\right) . \left(1 - G(s_t, \gamma, c)\right) \\ & + \left(\phi_{20} + \phi_{21} . y_{t-1} + \cdots + \phi_{2p} . y_{t-p}\right) . G(s_t, \gamma, c) + \varepsilon_t \end{aligned} \tag{8}$$

or alternatively using the matrices Φ_1 and Φ_2:

$$y_t = \Phi_1' Z_t . \left(1 - G(s_t, \gamma, c)\right) + \Phi_2' Z_t . G(s_t, \gamma, c) + \varepsilon_t \tag{9}$$

where $G(s_t, \gamma, c)$ is the transition function that takes values between 0 and 1.

An alternative formulation of the model is presented by:

$$y_t = \Phi_1' Z_t + (\Phi_2 - \Phi_1)' Z_t . G(s_t, \gamma, c) + \varepsilon_t \tag{10}$$

In practice, the transition function G is most often of logistic or exponential form, and the transition variable s_t is a lagged endogenous variable (with lag d):

$$s_t = y_{t-d} \tag{11}$$

The logistic transition function is defined by:

$$G(x, \gamma, c) = \frac{1}{1 + exp\left(-\gamma(x - c)\right)} \tag{12}$$

The exponential transition function is defined by:

$$G(x, \gamma, c) = 1 - exp(-\gamma(x - c)^2) \tag{13}$$

Where c is the transition parameter, and $\gamma > 0$.

When the transition function G is logistic, the parameter γ determines the smoothness of the transition between the two regimes. When γ is high, the function G approaches an indicator function. In this case, the logistic $STAR(p)$ model tends to resemble the $TAR(p)$ model. Conversely, the smaller γ is, the slower the transition between the two regimes. When $\gamma \to 0$, the function G approaches 1/2, and the model approaches a linear autoregressive model $AR(p)$.

When the transition function G is exponential, as $\gamma \to 0$, G approaches 0, and as $\gamma \to \infty$, G approaches 1. In these cases, the model behaves like a linear autoregressive model $AR(p)$. The exponential transition function becomes zero at the value c, i.e., $G(c, \gamma, c) = 0$. As $x \to \pm\infty$, the transition function G approaches 1.

3.2.3 Specification Procedure

To specify a *STAR* model fitted to the data, we follow the procedure proposed by Teräsvirta (1994):

1) We specify the most appropriate linear autoregressive model that fits the series of data under study.
2) We test linearity for different values of the lag parameter d, and determine the value of d if the linearity test is rejected.
3) We choose between the *LSTAR* and *ESTAR* models.
4) We estimate the chosen *STAR* model.

1) Specification of the Linear Autoregressive Model

In this step, we specify the most appropriate $AR(p)$ autoregressive model of order p for the series of data under investigation. Teräsvirta (1994) suggests determining the lag p that minimizes the Akaike Information Criterion, denoted as *AIC*, based on autoregressive $AR(p)$ models fitted to the data.

In this paper, we have selected the $AR(p)$ model whose order p minimizes the Akaike Information Criterion and exhibits non-autocorrelated residuals.

2) Testing the Null Hypothesis of Linearity against the Alternative of STAR Non-Linearity

Before applying a non-linear *STAR* model, we need to validate the non-linearity of the series of data under study using appropriate tests.

The null hypothesis of linearity H_0 can be formulated as the equality of the autoregressive parameters Φ_1 and Φ_2 in both regimes of the *STAR* model (10), i.e., $H_0 : \Phi_1 = \Phi_2$. The alternative hypothesis is written as $H_1 : \Phi_{1j} \neq \Phi_{2j}$ for at least one $j \in \{0,1,\cdots,p\}$.

The problem of the linearity test is complicated by the presence of unidentified nuisance parameters under the null hypothesis. Informally, the *STAR* model contains parameters that are not restricted by the null hypothesis, but we cannot learn anything about them from the data when the null hypothesis is true. For example, the null hypothesis does not restrict the parameters γ and c in the transition function.

The linearity test can also be formulated differently:

$$\begin{cases} H_0^{'} : \gamma = 0 \\ H_1^{'} : \gamma \neq 0 \end{cases} \tag{14}$$

Under the null hypothesis $H_0^{'}$, the *STAR* model (10) is reduced to a linear autoregressive model, and the parameter c and the parameters Φ_1 and $\Phi_2$are unidentified.

Escribano and Jordá (2001) stated that the parameters Φ_1 and $\Phi_2$can take any values as long as their average remains the same.

The problem of unidentified nuisance parameters under the null hypothesis was initially considered by Davies (1987) and occurs in many testing problems. The main consequence of the presence of such nuisance parameters is that conventional statistical theory is not available to obtain the asymptotic distribution of test statistics. Instead, test statistics tend to have non-standard distributions for which analytical expressions are often not available. This implies that critical values must be determined through simulations.

Due to the identification problem, normal testing procedures such as the likelihood ratio test, Lagrange multiplier test, and *Wald* test will produce undesirable parameter estimates. Instead, Luukonen et al. (1988b) suggest approximating the alternative model by adopting a Taylor series expansion of the transition function to circumvent the identification problem.

Testing the linearity hypothesis against the *LSTAR* non-linearity alternativeLuukkonen et al. (1988a) suggest approximating the logistic function $G(y_{t-d}, \gamma, c)$ by the Taylor formula around $\gamma = 0$ to the order 3:

$$G(y_{t-d}, \gamma, c) = G(y_{t-d}, 0, c) + \frac{\gamma}{1!} \frac{\partial G(y_{t-d}, \gamma, c)}{\partial \gamma}\bigg|_{\gamma=0} + \frac{\gamma^2}{2!} \frac{\partial^2 G(y_{t-d}, \gamma, c)}{\partial \gamma^2}\bigg|_{\gamma=0}$$
$$+ \frac{\gamma^3}{3!} \frac{\partial^3 G(y_{t-d}, \gamma, c)}{\partial \gamma^3}\bigg|_{\gamma=0} + R_3(y_{t-d}, \gamma, c)$$

where $R_3(y_{t-d}, \gamma, c)$ is the remainder of the Taylor formula to order 3. We have:

$$G(y_{t-d}, 0, c) = \frac{1}{2} \tag{15}$$

Calculating the first partial derivative of $G(y_{t-d}, \gamma, c)$ yields:

$$\frac{\partial G(y_{t-d}, \gamma, c)}{\partial \gamma} = (y_{t-d} - c)\left(exp\ \ (-\gamma(y_{t-d} - c))\right)\left(1 + exp\ \ (-\gamma(y_{t-d} - c))\right)^{-2}$$

We deduce:

$$\left.\frac{\partial G\,(y_{t-d},\gamma,c)}{\partial\gamma}\right|_{\gamma=0}=\frac{1}{4}(y_{t-d}-c) \tag{16}$$

The calculation of the second partial derivative of $G(y_{t-d},\gamma,c)$ yields:

$$\begin{aligned}\frac{\partial^2 G(y_{t-d},\gamma,c)}{\partial\gamma^2} &= (y_{t-d}-c)\Big[-(y_{t-d}-c)exp\ \ (-\gamma(y_{t-d}-c))\big(1+exp\ \ (-\gamma(y_{t-d}-c))\big)^{-2}\\ &+2exp\ \ (-\gamma(y_{t-d}-c))exp\ \ (-\gamma(y_{t-d}-c))(y_{t-d}\\ &-c)\big(1+exp\ \ (-\gamma(y_{t-d}-c))\big)^{-3}\Big]\Leftrightarrow\end{aligned}$$

$$\begin{aligned}\frac{\partial^2 G(y_{t-d},\gamma,c)}{\partial\gamma^2} &= (y_{t-d}-c)^2 exp\ \ (-\gamma(y_{t-d}-c))\big(1+exp\ \ (-\gamma(y_{t-d}-c))\big)^{-2}.\Big[-1\\ &+2exp\ \ (-\gamma(y_{t-d}-c))\big(1+exp\ \ (-\gamma(y_{t-d}-c))\big)^{-1}\Big]\\ &=(y_{t-d}-c).\frac{\partial G\,(y_{t-d},\gamma,c)}{\partial\gamma}\big(-1+2exp\ \ (-\gamma(y_{t-d}-c)).G(y_{t-d},\gamma,c)\big)\end{aligned}$$

We deduce :

$$\left.\frac{\partial^2 G(y_{t-d},\gamma,c)}{\partial\gamma^2}\right|_{\gamma=0}=0 \tag{17}$$

The calculation of the third partial derivative of $G(y_{t-d},\gamma,c)$ yields:

$$\begin{aligned}\frac{\partial^3 G(y_{t-d},\gamma,c)}{\partial\gamma^3} &= (y_{t-d}-c).\Bigg[\frac{\partial^2 G(y_{t-d},\gamma,c)}{\partial\gamma^2}.\big(-1+2ex\ p(-\gamma(y_{t-d}-c)).G(y_{t-d},\gamma,c)\big)\\ &+\frac{\partial G\,(y_{t-d},\gamma,c)}{\partial\gamma}.2.\Big(-(y_{t-d}-c)exp\ \ (-\gamma(y_{t-d}-c))G(y_{t-d},\gamma,c)\\ &+exp\ \ (-\gamma(y_{t-d}-c)).\frac{\partial G\,(y_{t-d},\gamma,c)}{\partial\gamma}\Big)\Bigg]\end{aligned}$$

We deduce :

$$\left.\frac{\partial^3 G(y_{t-d},\gamma,c)}{\partial\gamma^3}\right|_{\gamma=0}=-\frac{1}{8}(y_{t-d}-c)^3 \tag{18}$$

We deduce the third-order Taylor formula of the logistic transition function:

$$G(y_{t-d},\gamma,c)=\frac{1}{2}+\frac{\gamma}{4}(y_{t-d}-c)-\frac{\gamma^3}{48}(y_{t-d}-c)^3+R_3(y_{t-d},\gamma,c) \tag{19}$$

By replacing the logistic transition function $G(y_{t-d},\gamma,c)$ with its Taylor formula in the *STAR* model (10), we obtain:

$$\begin{aligned}y_t=\Bigg(&\bigg(\phi_{10}+\frac{1}{2}(\phi_{20}-\phi_{10})-\frac{\gamma.c}{4}(\phi_{20}-\phi_{10})+\cdots+\frac{c^3\gamma^3}{48}(\phi_{20}-\phi_{10})\bigg)\\ &+\bigg(\phi_{11}+\frac{1}{2}(\phi_{21}-\phi_{11})-\frac{\gamma.c}{4}(\phi_{21}-\phi_{11})+\cdots+\frac{c^3\gamma^3}{48}(\phi_{21}-\phi_{11})\bigg).y_{t-1}+\cdots\\ &+\bigg(\phi_{1p}+\frac{1}{2}(\phi_{2p}-\phi_{1p})-\frac{\gamma.c}{4}(\phi_{2p}-\phi_{1p})+\cdots+\frac{c^3\gamma^3}{48}(\phi_{2p}-\phi_{1p})\bigg).y_{t-p}\Bigg)\end{aligned}$$

$$+\left((\phi_{20}-\phi_{10})+(\phi_{21}-\phi_{11}).y_{t-1}+\cdots+(\phi_{2p}-\phi_{1p}).y_{t-p}\right).\left(\frac{\gamma}{4}-\frac{c^2\gamma^3}{16}\right).y_{t-d}$$
$$+\left((\phi_{20}-\phi_{10})+(\phi_{21}-\phi_{11}).y_{t-1}+\cdots+(\phi_{2p}-\phi_{1p}).y_{t-p}\right).\frac{c\gamma^3}{16}.y_{t-d}^2$$
$$+\left((\phi_{20}-\phi_{10})+(\phi_{21}-\phi_{11}).y_{t-1}+\cdots+(\phi_{2p}-\phi_{1p}).y_{t-p}\right).\left(-\frac{\gamma^3}{48}\right)y_{t-d}^3$$
$$+(\Phi_2-\Phi_1)'Z_t.R_3(y_{t-d},\gamma,c)+\varepsilon_t$$

We define for $0\leq j\leq p$:

$$\begin{aligned}
\psi_{0j}&=\phi_{1j}+\frac{1}{2}(\phi_{2j}-\phi_{1j})-\frac{\gamma.c}{4}(\phi_{2j}-\phi_{1j})+\cdots+\frac{c^3\gamma^3}{48}(\phi_{2j}-\phi_{1j})\\
\psi_{1j}&=\left(\frac{\gamma}{4}-\frac{c^2\gamma^3}{16}\right)(\phi_{2j}-\phi_{1j})\\
\psi_{2j}&=\frac{c\gamma^3}{16}(\phi_{2j}-\phi_{1j})\\
\psi_{3j}&=-\frac{\gamma^3}{48}(\phi_{2j}-\phi_{1j})
\end{aligned}\tag{20}$$

We define for $0\leq i\leq 3$:

$$\psi_i'=(\psi_{i0}\psi_{i1}\cdots\psi_{ip})'$$

The auxiliary regression of y_t on $Z_t, Z_t.y_{t-d}, Z_t.y_{t-d}^2$, and $Z_t.y_{t-d}^3$ is then written as:

$$y_t=\psi_0'Z_t+\psi_1'Z_t.y_{t-d}+\psi_2'Z_t.y_{t-d}^2+\psi_3'Z_t.y_{t-d}^3+\eta_t\tag{21}$$

with :

$$\eta_t=(\Phi_2-\Phi_1)'Z_t.R_3(y_{t-d},\gamma,c)+\varepsilon_t\tag{22}$$
$$Z_t=(1\ \ y_{t-1}y_{t-2}\cdots y_{t-p})'\tag{23}$$

Note that the parameters ψ_0', ψ_1', ψ_2', and ψ_3' in the auxiliary regression (21) are functions of the parameters Φ_1, Φ_2, γ and c.

The terms $\psi_{i0}.y_{t-d}^i$ should be excluded from the auxiliary regression (21) to avoid perfect multicollinearity.

The new auxiliary regression of y_t on $Z_t, \tilde{Z}_t.y_{t-d}, \tilde{Z}_t.y_{t-d}^2$, and $\tilde{Z}_t.y_{t-d}^3$ is then written as:

$$y_t=\psi_0'Z_t+\tilde{\psi}_1'\tilde{Z}_t.y_{t-d}+\tilde{\psi}_2'\tilde{Z}_t.y_{t-d}^2+\tilde{\psi}_3'\tilde{Z}_t.y_{t-d}^3+\eta_t\tag{24}$$

with :

$$\tilde{\psi}_i'=(\psi_{i1}\psi_{i2}\cdots\psi_{ip})'\tag{25}$$
$$\tilde{Z}_t=(y_{t-1}y_{t-2}\cdots y_{t-p})'\tag{26}$$

Under the null hypothesis $H_0':\gamma=0$, we have $R_3(y_{t-d},\gamma,c)=0$ and $\eta_t=\varepsilon_t$. Therefore, the remainder R_3does not affect the properties of the error under the null hypothesis and the asymptotic theory of the distribution.

The null hypothesis $H_0':\gamma=0$ translates to $\tilde{\psi}_i=0$ for $i=1,2,3$. Consequently, the new null and alternative hypotheses are written as:

$$\begin{cases}H_0'':\tilde{\psi}_i=0\ pour i\quad =1,2,3\\ H_1'':\tilde{\psi}_i\neq 0\ pour au moins un i\end{cases}\tag{27}$$

The linearity can be tested using a Lagrange Multiplier (LM) statistic that follows a standard asymptotic Chi-square distribution under the null hypothesis. Under the null hypothesis of linearity, the LM test statistic is given by:

$$LM=\frac{T(SSR_0-SSR_1)}{SSR_0}\tag{28}$$

which follows a χ^2 distribution with $3p$ degrees of freedom. Here:
T: is the sample size.
SSR_0: is the sum of squares of residuals from the linear regression model under the null hypothesis, i.e., regressing y_t on Z_t.

SSR_1: is the sum of squares of residuals from the *LSTAR* non-linear regression model, i.e., regressing y_t on Z_t, $\tilde{Z}_t.y_{t-d}$, $\tilde{Z}_t.y_{t-d}^2$, and $\tilde{Z}_t.y_{t-d}^3$.

Another version of the *LM* test based on the F-distribution involves using the statistic:

$$LM = \frac{(SSR_0 - SSR_1)/(3p)}{SSR_0/(T-4p)} \quad (29)$$

which follows a Fisher distribution with $3p$ and $T-4p$ degrees of freedom.

During the execution of the *LM*-type linearity test, the lag parameter d is fixed. To determine this lag parameter d, the *LM*-type test is performed for different values of d in the range $1 \leq d \leq D$. If the null hypothesis is rejected for at least one value of d, then, to determine the appropriate value of d, we choose the one associated with the smallest p-value.

3) Choice between LSTAR and ESTAR models

After rejecting the null hypothesis of linearity in favor of *STAR*-type non-linearity and choosing the lag parameter d, this step involves choosing between *LSTAR* and *ESTAR* models through a sequence of nested null hypotheses in the auxiliary regression (24):

$$\begin{cases} H_{01}: \tilde{\psi}_3 = 0 \\ H_{02}: \tilde{\psi}_2 = 0 | \tilde{\psi}_3 = 0 \\ H_{03}: \tilde{\psi}_1 = 0 | \tilde{\psi}_2 = \tilde{\psi}_3 = 0 \end{cases} \quad (30)$$

Teräsvirta (1994) suggested decision rules for choosing between *LSTAR* and *ESTAR* models:

1. Rejection of hypothesis H_{01} implies the acceptance of the *LSTAR* model.
2. Rejection of hypothesis H_{02} implies the acceptance of the *LSTAR* model.
3. Acceptance of hypothesis H_{03} after the rejection of hypothesis H_{02} implies the choice of the *ESTAR* model.

Teräsvirta (1994) also proposed a more practical method for choosing between the two models by comparing the significance levels of the three Fisher tests; if the p-value of the H_{02} test is the smallest among the three, we choose the *ESTAR* model; otherwise, we choose the *LSTAR* model.

4) Estimation of the STAR Model

Once the transition variable y_{t-d} and the transition function $G(y_{t-d}, \gamma, c)$ have been selected, the modeling process's next step is the estimation of parameters in the *STAR* model. The parameter estimation $\theta = (\Phi_1' \Phi_2' \gamma c)$ in the *STAR* model is performed using the non-linear least squares method:

$$\hat{\theta} = \underset{\theta}{\text{Argmin}} \sum_{t=1}^{T} (y_t - H(Z_t, \theta)) \quad (31)$$

$$H(Z_t, \theta) = \Phi_1' Z_t.(1 - G(s_t, \gamma, c)) + \Phi_2' Z_t.G(s_t, \gamma, c) \quad (32)$$

4. Empirical results

4.1 Preliminary data analysis

4.1.1 Graphical representation of data

Figures 1 and 2 display the time series of daily closing prices of *MASI* and the time series of daily geometric returns of MASI.

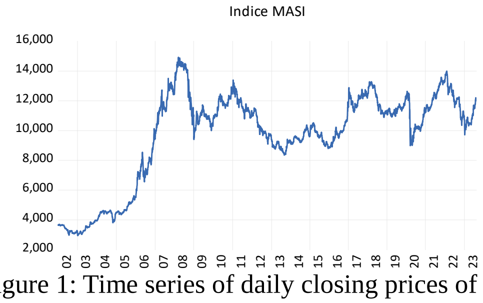


Figure 1: Time series of daily closing prices of the *MASI* index

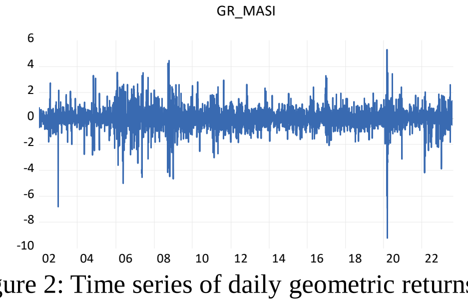
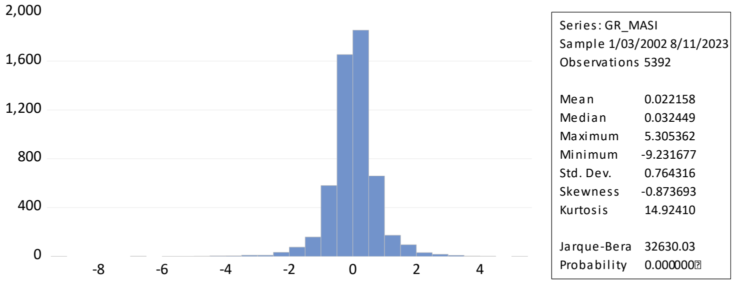


Figure 2: Time series of daily geometric returns of the *MASI* index

4.1.2. Descriptive Statistics

Table 1 presents the descriptive statistics of the geometric return of the *MASI* index. This table includes the histogram, mean, median, minimum, maximum, standard deviation, skewness, kurtosis, and Jarque-Bera statistics.

Table 1: Histogram and descriptive statistics of the geometric returns series of *MASI*



The preceding table shows a negative skewness of -0.87, indicating that the distribution of the geometric returns of *MASI* has a long left tail. The excessively high kurtosis value (14.92 > 3) suggests thick-tailed characteristics of the distribution. The elevated value of the Jarque-Bera (*JB*) statistics implies rejecting the null hypothesis of normality.

Confirmation of the non-normality of the distribution of geometric returns of *MASI* will be done through the quantile-quantile (QQ-Plot) graph that follows.

4.1.3. Normality Test: Quantile-Quantile graph (QQ Plot)

We use the quantile-quantile graph (QQ-Plot) test to assess the conformity of the geometric return series of *MASI* to a normal distribution. If the empirical distribution and the theoretical (normal) distribution are equivalent, the QQ-Plot should align points on a straight line at 45 degrees. Figure 3 below presents the QQ-Plot illustrating the empirical distribution of geometric returns of *MASI*.

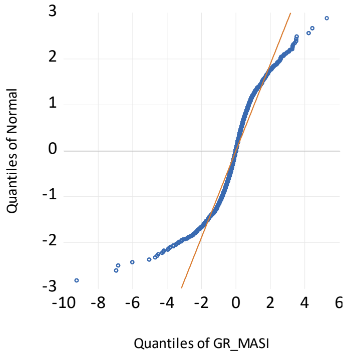


Figure 3: QQ-Plot of the geometric returns series of the *MASI*index

The analysis of Figure 3 reveals that the empirical distribution of the geometric returns series of *MASI* deviates from normality, showing thicker tails than those of a normal distribution. The QQ-Plot does not conform to a straight line and adopts an S-shaped form. This observation confirms the non-normality previously identified by the Jarque-Bera (*JB*) statistic.

4.1.4 Non-stationarity Test (Unit Root Test): Augmented Dickey-Fuller Test (ADF)

Analyzing the previous Figures 1 and 2, we can conclude that the geometric returns series of the *MASI* index appears to be a manifestation of a stationary process.

The Augmented Dickey-Fuller (*ADF*) test was applied to the geometric returns series of the *MASI* index, and the results of this test are presented in Table 2.

Table 2: Results of the Augmented Dickey-Fuller test applied to the geometric returns series of the *MASI*index

Null Hypothesis: GR_MASI has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Automatic - based on SIC, maxlag=32)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-56.97702	0.0000
Test critical values:	1% level	-3.959739	
	5% level	-3.410638	
	10% level	-3.127099	

*MacKinnon (1996) one-sided p-values.

In Table 2, the *ADF* statistic is evaluated at -56.97702, with an associated p-value of 0.0000. It is noteworthy that the statistic is below the critical values at the 1%, 5%, and 10% levels. Therefore, we reject the null hypothesis of a unit root, indicating that the geometric returns series of *MASI* is generated by a stationary process.

4.2 Application of the nonlinear *STAR*model to the *MASI*index

4.2.1 Specification of the linear autoregressive model

In this step, we specify a lagged autoregressive AR (p) model of the most appropriate order p for the series under investigation. As suggested by Teräsvirta (1994), we choose the lag p that minimizes the Akaike Information Criterion (*AIC*) of the autoregressive linear regression of order p. We also used the Ljung-Box test for the correlation of autocorrelation and partial autocorrelation functions of the AR (p) models.

We begin by presenting the results of the Ljung-Box test for serial correlation of autocorrelation and partial autocorrelation functions of the geometric returns series of *MASI*. Table 3 displays the results of this test.

Table 3: Autocorrelation and partial autocorrelation functions of the geometric returns series of *MASI*

Sample: 1/03/2002 8/11/2023
Included observations: 5392

Lag	AC	PAC	Q-Stat	Prob
1	0.249	0.249	333.50	0.000
2	0.070	0.009	359.96	0.000
3	-0.010	-0.031	360.51	0.000
4	-0.012	-0.003	361.26	0.000
5	0.013	0.020	362.17	0.000
6	-0.016	-0.025	363.56	0.000
7	0.009	0.018	364.03	0.000
8	-0.011	-0.016	364.72	0.000

The autocorrelation function shows a peak at lag 2, and the partial autocorrelation function shows a peak at lag 1. For the autoregressive model selection, we explored up to an order p equal to 3.

To this end, we estimated the AR (1), AR (2)and AR (3) models. The *AIC* criteria associated with the three models are provided in the following Table 4:

Table 4: Akaike Information Criteria for the 3 estimated models AR (1), AR (2)and AR (3)

	AR (1)	AR (2)	AR (3)
AIC	2.237442	2.237737	2.237116

We observe that the AR (3) model has the lowest *AIC* criterion. Therefore, we choose the AR (3) model, and its estimation is provided in the following Table 5.

Table 5: Estimation of the Autoregressive Model of Order 3

Dependent Variable: GR_MASI
Method: ARMA Maximum Likelihood (OPG - BHHH)
Sample: 1/03/2002 8/11/2023
Included observations: 5392

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.022212	0.013745	1.616040	0.1061
AR(1)	0.246750	0.006931	35.60190	0.0000
AR(2)	0.016487	0.005782	2.851171	0.0044
AR(3)	-0.031491	0.006828	-4.611952	0.0000
SIGMASQ	0.547371	0.004075	134.3356	0.0000

The Table 6 below displays the results of the Ljung-Box test for autocorrelation and partial autocorrelation functionsapplied to the residual series of the estimated autoregressive model of order 3.

Table 6: Results of the Ljung-Box test for the residual series of $AR\ (3)$

Sample: 1/03/2002 8/11/2023
Q-statistic probabilities adjusted for 3 ARMA terms

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.000	-0.000	6.E-05	
		2	0.001	0.001	0.0025	
		3	-0.000	-0.000	0.0028	
		4	-0.007	-0.007	0.2593	0.611
		5	0.024	0.024	3.3550	0.187
		6	-0.023	-0.023	6.2148	0.102
		7	0.017	0.017	7.8390	0.098
		8	-0.020	-0.020	10.000	0.075

From Table 6, we deduce that the residuals of the autoregressive model of order 3 are non-autocorrelated.

4.2.2 Test of the null hypothesis of linearity and choice between LSTAR and ESTAR models

We will test linearity for different values of the lag parameter d against the alternative of *LSTAR*-type non-linearity by applying the strategy of Luukkonen et al. (1988a), which suggests approximating the logistic function $G(y_{t-d}, \gamma, c)$ by the Taylor series expansion around $\gamma = 0$ up to the 3rd order.

We estimated several $LSTAR\ (p_1, p_2)$ models with $1 \leq p_1 \leq 3$ as the order of the linear regime and $1 \leq p_2 \leq 3$ as the order of the non-linear regime.

We kept only the $LSTAR\ (p_1, p_2)$ models whose lag polynomials of both linear and non-linear regimes satisfy the following conditions:

- Both lag polynomials contain the constant term.
- Both lag polynomials contain at least two lag terms of orders between 1 and 3.

There are a total of 16 possible configurations. For example, in the model:

$$y_t = (\phi_{10} + \phi_{11}.y_{t-1} + \phi_{12}.y_{t-p} + \phi_{13}.y_{t-p}).(1 - G(s_t, \gamma, c)) \\ + (\phi_{20} + \phi_{21}.y_{t-1} + \phi_{23}.y_{t-3}).G(s_t, \gamma, c) + \varepsilon_t$$

- The first lag polynomial contains 3 lag terms $AR\ (1)$, $AR\ (2)$ and $AR\ (3)$.
- The second lag polynomial contains 2 lag terms $AR(1)$ and $AR\ (3)$.

We will denote this model as *LSTAR* C123-C13.The various configurations of the estimated models are provided in the following Table 7:

Table 7: Estimated models with different configurations of lag polynomials for both linear and non-linear regimes

*LSTAR*C123-C123	*LSTAR*C123-C12	*LSTAR*C123-C13	*LSTAR*C123-C23
*LSTAR*C12-C123	*LSTAR*C12-C12	*LSTAR*C12-C13	*LSTAR*C12-C23
*LSTAR*C13-C123	*LSTAR*C13-C12	*LSTAR*C13-C13	*LSTAR*C13-C23
*LSTAR*C23-C123	*LSTAR*C23-C12	*LSTAR*C23-C13	*LSTAR*C23-C23

For each model, we selected the optimal lag value d for the transition variable y_{t-d}. We kept only the *LSTAR* models whose coefficients of the two lag polynomials and the parameters of the logistic transition function are statistically significant.

We found that only two models satisfy these conditions: the *LSTAR* C123-C23 model and the *LSTAR* C23-C123 model. The two tables, 8 and 9, below display the estimates of the coefficients of the lag polynomials for both linear and non-linear regimes, the parameters of the logistic transition function, and the values of the three information criteria *AIC*, *SC* and *HQC* associated with these two models.

Table 8: Estimation of the model *LSTAR* C123-C23

Dependent Variable: GR_MASI
Method: Smooth Threshold Regression
Transition function: Logistic
Sample (adjusted): 1/08/2002 8/11/2023
Included observations: 5389 after adjustments
Threshold variable: GR_MASI(-2)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Threshold Variables (linear part)				
C	3.279235	0.502168	6.530153	0.0000
GR_MASI(-2)	0.900801	0.106784	8.435759	0.0000
GR_MASI(-3)	-0.347778	0.068591	-5.070301	0.0000
GR_MASI(-1)	0.250626	0.013484	18.58679	0.0000
Threshold Variables (nonlinear part)				
C	-3.272773	0.503891	-6.495005	0.0000
GR_MASI(-2)	-0.875131	0.106343	-8.229356	0.0000
GR_MASI(-3)	0.342567	0.071109	4.817509	0.0000
Slopes				
SLOPE	3.005118	0.965063	3.113910	0.0019
Thresholds				
THRESHOLD	-2.588397	0.203235	-12.73597	0.0000

Tableau 9 : Estimation du modèle *LSTAR*C23-C123

Dependent Variable: GR_MASI
Method: Smooth Threshold Regression
Transition function: Logistic
Sample (adjusted): 1/08/2002 8/11/2023
Included observations: 5389 after adjustments
Threshold variable: GR_MASI(-2)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Threshold Variables (linear part)				
C	3.187877	0.490355	6.501157	0.0000
GR_MASI(-2)	0.920972	0.104398	8.821773	0.0000
GR_MASI(-3)	-0.343162	0.070902	-4.839963	0.0000
Threshold Variables (nonlinear part)				
C	-3.180870	0.491503	-6.471723	0.0000
GR_MASI(-2)	-0.900781	0.104321	-8.634748	0.0000
GR_MASI(-3)	0.337450	0.073378	4.598793	0.0000
GR_MASI(-1)	0.263700	0.014222	18.54235	0.0000
Slopes				

SLOPE	2.956182	0.865138	3.417007	0.0006
		Thresholds		
THRESHOLD	-2.636624	0.183024	-14.40593	0.0000

The following Table 10 consolidates the three information criteria: Akaike Information Criterion (*AIC*), Schwarz Criterion (*SC*), and Hannan-Quinn Criterion (*HQC*) for the two models, LSTAR C123-C23 and LSTAR C23-C123.

Table 10: Information Criteria (*AIC*, *SC*, *HQC*) for the models*LSTAR* C123-C23 and C23-C123

	*LSTAR*C123-C23	*LSTAR*C23-C123
Akaike info criterion	2.215560	2.215601
Schwarz criterion	2.226570	2.226610
Hannan-Quinn criterion	2.219404	2.219445

We observe that the model *LSTAR* C123-C23 is the model that minimizes all three information criteria. We will choose it as the *STAR* model to be tested against the linear autoregressive model in the linearity test. The following Table 11 presents the results of the linearity test and the test for choosing between *LSTAR* and *ESTAR*.

Table 11: Results of the Linearity Test and the Test for Choosing Between *LSTAR* and *ESTAR*

Smooth Threshold Linearity Tests
Sample: 1/03/2002 8/11/2023
Included observations: 5389
Test for nonlinearity using GR_MASI(-2) as the threshold variable
Taylor series alternatives: b0 + b1*s [+ b2*s^2 + b3*s^3 + b4*s^4]

Linearity Tests			
Null Hypothesis	F-statistic	d.f.	p-value
H04: b1=b2=b3=b4=0	32.07444	(8, 5377)	0.0000
H03: b1=b2=b3=0	24.03130	(6, 5379)	0.0000
H02: b1=b2=0	24.15503	(4, 5381)	0.0000
H01: b1=0	47.60666	(2, 5383)	0.0000

The H0i test uses the i-th order Taylor expansion (bj=0 for all j>i).

Teräsvirta Sequential Tests			
Null Hypothesis	F-statistic	d.f.	p-value
H3: b3=0	23.38196	(2, 5379)	0.0000
H2: b2=0 | b3=0	0.708555	(2, 5381)	0.4924
H1: b1=0 | b2=b3=0	47.60666	(2, 5383)	0.0000

All tests are based on the third-order Taylor expansion (b4=0).
Linear model is rejected at the 5% level using H03.
Recommended model: first-order logistic.
. Pr(H3) <= Pr(H2) or Pr(H1) <= Pr(H2)

Escribano-Jorda Tests			
Null Hypothesis	F-statistic	d.f.	p-value
H0L: b2=b4=0	36.74289	(3, 5377)	0.0000
H0E: b1=b3=0	9.229704	(2, 5377)	0.0001

All tests are based on the fourth-order Taylor expansion.
Linear model is rejected at the 5% level using H04.
Recommended model: exponential with nonzero threshold.
. Pr(H0L) < Pr(H0E) with Pr(H0E) < .05

On Table 11, we observe that the sequential tests by Teravista reject the hypothesis of the linear model at a 5% significance level and recommend the *LSTAR* model over the *ESTAR* model. Similarly, the Escribano-Jorda tests reject the hypothesis of the linear model at a 5% significance level but recommend the *ESTAR* model over the *LSTAR* model.

As with the estimated *LSTAR* models, we will estimate several $ESTAR\ (p_1, p_2)$ models with $1 \leq p_1 \leq 3$ as the order of the linear regime and $1 \leq p_2 \leq 3$ as the order of the non-linear regime. We only kept the models $ESTAR\ (p_1, p_2)$ whose lag polynomials of both linear and non-linear regimes satisfy the following conditions:

- Both lag polynomials contain the constant term.
- Both lag polynomials contain at least two lag terms of orders between 1 and 3.

Similar to the *LSTAR* models, there are a total of 16 possible configurations for the *ESTAR* models.

For each model, we selected the optimal lag value d for the transition variable y_{t-d}. We kept only the *ESTAR* models whose coefficients of the two lag polynomials and the parameters of the logistic transition function are statistically significant.

We found that only three *ESTAR* models satisfy these conditions: *ESTAR* C13-C12, *ESTAR* C23-C123, and *ESTAR* C23-C12. The three tables 12, 13, and 14 below display the estimates of the coefficients of the lag polynomials for both linear and non-linear regimes, the parameters of the logistic transition function, and the values of the three information criteria *AIC*, *SC*, and *HQC* associated with these three *ESTAR* models.

Table 12: Estimation of the model *ESTAR* C13-C12

Dependent Variable: GR_MASI
Method: Smooth Threshold Regression
Transition function: Exponential
Sample (adjusted): 1/08/2002 8/11/2023
Included observations: 5389 after adjustments
Threshold variable chosen: GR_MASI(-2)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Threshold Variables (linear part)				
C	0.127881	0.028888	4.426818	0.0000
GR_MASI(-1)	0.304583	0.030331	10.04188	0.0000
GR_MASI(-3)	-0.029188	0.013585	-2.148622	0.0317
Threshold Variables (nonlinear part)				
C	-0.214404	0.042081	-5.095084	0.0000
GR_MASI(-1)	-0.089765	0.039970	-2.245808	0.0248
GR_MASI(-2)	-0.052232	0.021966	-2.377832	0.0174
Slopes				
SLOPE	1.151009	0.497802	2.312182	0.0208
Thresholds				
THRESHOLD	0.909902	0.098331	9.253448	0.0000

For this model *ESTAR* C13-C12, the chosen transition variable is GR_MASI(-2).

Table 13: Estimation of the model *ESTAR* C23-C123

Dependent Variable: GR_MASI
Method: Smooth Threshold Regression
Transition function: Exponential
Sample (adjusted): 1/08/2002 8/11/2023
Included observations: 5389 after adjustments
Threshold variable chosen: GR_MASI(-2)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Threshold Variables (linear part)				
C	9.627652	3.428617	2.808028	0.0050
GR_MASI(-2)	6.097564	2.160881	2.821796	0.0048
GR_MASI(-3)	-0.371555	0.130570	-2.845651	0.0044
Threshold Variables (nonlinear part)				
C	-9.614251	3.428392	-2.804303	0.0051
GR_MASI(-2)	-6.078776	2.161219	-2.812660	0.0049
GR_MASI(-3)	0.347945	0.131877	2.638408	0.0084
GR_MASI(-1)	0.259482	0.013898	18.66993	0.0000
Slopes				
SLOPE	35.51427	12.04093	2.949463	0.0032
Thresholds				
THRESHOLD	-1.587642	0.025083	-63.29657	0.0000

For this model *ESTAR* C23-C123, the chosen transition variable is GR_MASI(-2).

Table 14: Estimation of the model *ESTAR* C23-C12

Dependent Variable: GR_MASI
Method: Smooth Threshold Regression
Transition function: Exponential
Sample (adjusted): 1/08/2002 8/11/2023
Included observations: 5389 after adjustments
Threshold variable: GR_MASI(-1)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Threshold Variables (linear part)				
C	-8.777471	1.437011	-6.108142	0.0000
GR_MASI(-2)	-3.960493	1.161524	-3.409740	0.0007
GR_MASI(-3)	-0.029982	0.013508	-2.219618	0.0265
Threshold Variables (nonlinear part)				
C	8.796976	1.436950	6.121979	0.0000
GR_MASI(-2)	3.976414	1.161792	3.422656	0.0006
GR_MASI(-1)	0.241141	0.013529	17.82450	0.0000
Slopes				
SLOPE	26449.72	8982.972	2.944429	0.0032

Thresholds

THRESHOLD	-1.400001	0.000890	-1573.216	0.0000

For this model *ESTAR* C23-C12, the chosen transition variable is GR_MASI(-1).

The following Table 15 consolidates the three information criteria: Akaike Information Criterion (*AIC*), Schwarz Criterion (*SC*), and Hannan-Quinn Criterion (*HQC*) for the three models, *ESTAR* C13-C12, *ESTAR* C23-C123 and *ESTAR* C23-C12.

Table 15: Information Criteria (*AIC*, *SC*, *HQC*) for the three models *ESTAR* C13-C12, *ESTAR* C23-C123 and *ESTAR* C23-C12

	ESTAR C13-C12	*ESTAR* C23-C123	*ESTAR* C23-C12
Akaike info criterion	2.230579	2.229013	2.220998
Schwarz criterion	2.240365	2.240022	2.230784
Hannan-Quinn criterion	2.233996	2.232857	2.224414

We observe that the model *ESTAR* C23-C12 is the model that minimizes all three information criteria. However, we notice that the transition parameter $\gamma = 26449.72$ has a very high value for this model. It has been observed that when $\gamma \to \infty$, the exponential transition function approaches 1, and in such cases, the model *ESTAR* C23-C12 behaves like a linear autoregressive model of order 3.

The transition parameter for the model *ESTAR* C13-C12 is $\gamma = 1.151009$, and for the model *ESTAR* C23-C123, it is $\gamma = 35.51427$. Both of these models are nonlinear. The model *ESTAR* C23-C123 is better than the model *ESTAR* C13-C12 in terms of minimizing all three information criteria.

Now, we will compare the values of the three information criteria for the nonlinear model *LSTAR* C123-C23, the linear model *ESTAR* C23-C12, and the nonlinear model *ESTAR* C23-C123 (see Table 16 below).

Table 16: Values of the three Information Criteria (*AIC*, *SC*, *HQC*) for the models *LSTAR* C123-C23, *ESTAR* C23-C123 and *ESTAR* C23-C12

	*LSTAR*C123-C23	*ESTAR*C23-C123	*ESTAR*C23-C12
Akaike info criterion	2.215560	2.229013	2.220998
Schwarz criterion	2.226570	2.240022	2.230784
Hannan-Quinn criterion.	2.219404	2.232857	2.224414

As we can see in Table 16, it is the model *LSTAR* C123-C23 that minimizes all three information criteria. For further comparison, we will analyze the three models *LSTAR* C123-C23, *ESTAR* C23-C12 and *ESTAR* C23-C123 in sections 4.2.3, 4.2.4, and 4.2.5 respectively.

4.2.3 Estimation of the model *LSTAR* C123-C23

The selected model *LSTAR* C123-C23 is expressed as follows:

$$y_t = (\phi_{10} + \phi_{11}.y_{t-1} + \phi_{12}.y_{t-2} + \phi_{13}.y_{t-3}).\left(1 - \frac{1}{1 + exp\left(-\gamma(y_{t-2} - c)\right)}\right) + (\phi_{20} + \phi_{22}.y_{t-2} + \phi_{23}.y_{t-3}).\frac{1}{1 + exp\left(-\gamma(y_{t-2} - c)\right)} \quad (33)$$

or alternatively:

$$y_t = (\phi_{10} + \phi_{11}.y_{t-1} + \phi_{12}.y_{t-2} + \phi_{13}.y_{t-3}) + \Big((\phi_{20} - \phi_{10}) - \phi_{11}.y_{t-1} + (\phi_{22} - \phi_{12}).y_{t-2} + (\phi_{23} - \phi_{13}).y_{t-3}\Big).\frac{1}{1 + exp\left(-\gamma(y_{t-2} - c)\right)} \quad (34)$$

The estimates of the parameters ϕ_{1j}, ϕ_{2j} for both regimes, the transition parameter γ, and the transition threshold c of the model *LSTAR* C123-C23 are provided in the above Table 8.

From Table 8, we can observe that the parameters ϕ_{1j}, ϕ_{2j} for both regimes, the transition parameter γ, and the transition threshold c of the model *LSTAR* C123-C23 are statistically significant at a 1% significance level.

If we substitute the coefficients with their estimated values in the model *LSTAR* C123-C23, we obtain:

$$y_t = (3.279235 + 0.250626.y_{t-1} + 0.900801.y_{t-2} - 0.347778.y_{t-3}).\left(1 - \frac{1}{1 + exp\ \left(-3.005118.(y_{t-2} + 2.588397)\right)}\right) + (-3.272773 - 0.875131.y_{t-2} + 0.342567.y_{t-3}).\frac{1}{1 + exp\ \left(-3.005118.(y_{t-2} + 2.588397)\right)} \quad (35)$$

or alternatively:

$$y_t = (3.279235 + 0.250626.y_{t-1} + 0.900801.y_{t-2} - 0.347778.y_{t-3}) + \left((-3.272773 - 3.279235) - 0.250626.y_{t-1} + (-0.875131 - 0.900801).y_{t-2} + (0.342567 + 0.347778).y_{t-3}\right).\frac{1}{1 + exp\ \left(-3.005118.(y_{t-2} + 2.588397)\right)}$$

$$y_t = (3.279235 + 0.250626.y_{t-1} + 0.900801.y_{t-2} - 0.347778.y_{t-3}) + (-6{,}552008 - 0.250626.y_{t-1} - 1{,}775932.y_{t-2} + 0{,}690345.y_{t-3}).\frac{1}{1 + exp\ \left(-3.005118.(y_{t-2} + 2.588397)\right)} \quad (36)$$

The Figure 4 below displays the graphical representation of the estimated logistic transition function of the model *LSTAR* C123-C23.

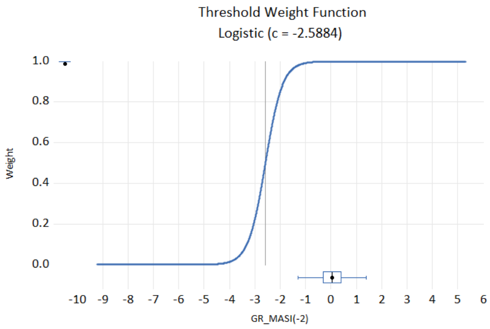


Figure 4: Logistic transition function curve of the model *LSTAR* C123-C23

The estimated value of $\gamma = 3.005118$ suggests that the transition from one regime to another is quite slow, as illustrated in Figure 4.

Table 17 below shows the p-values of the Ljung-Box test for the residuals series of the estimated model *LSTAR* C123-C23.

Table 17: Results of the Ljung-Box test applied to the residuals of *LSTAR* C123-C23

Date: 01/31/24 Time: 08:50
Sample (adjusted): 1/08/2002 8/11/2023
Q-statistic probabilities adjusted for 3 dynamic regressors

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.008	-0.008	0.3681	0.544
		2	0.005	0.005	0.4886	0.783
		3	0.001	0.001	0.4977	0.919
		4	-0.003	-0.003	0.5365	0.970
		5	0.024	0.024	3.7490	0.586
		6	-0.008	-0.007	4.0693	0.667
		7	0.004	0.003	4.1455	0.763
		8	-0.022	-0.022	6.6734	0.572

The examination of the p-values of the Q-statistics from the Ljung-Box test for serial correlation shows that the residuals of the estimated model *LSTAR* C123-C23 are not autocorrelated.

We also applied the Ljung-Box test to the series of squared residuals of the model *LSTAR* C123-C23 and found that the squared residuals are autocorrelated.

4.2.4 Estimation of the model *ESTAR* C23-C123

The model *ESTAR* C23-C123 is expressed as follows:

$$y_t = (\phi_{10} + \phi_{12}.y_{t-2} + \phi_{13}.y_{t-3}).\exp\ (-\gamma(x-c)^2) + (\phi_{20} + \phi_{21}.y_{t-1} + \phi_{22}.y_{t-2} + \phi_{23}.y_{t-3}).\left(1 - \exp\ (-\gamma(x-c)^2)\right) \quad (37)$$

or alternatively:

$$y_t = (\phi_{10} + \phi_{12}.y_{t-2} + \phi_{13}.y_{t-3}) + \left((\phi_{20} - \phi_{10}) + \phi_{21}.y_{t-1} + (\phi_{22} - \phi_{12}).y_{t-2} + (\phi_{23} - \phi_{13}).y_{t-3}\right).\exp\ (-\gamma(x-c)^2) \quad (38)$$

The estimates of the parameters ϕ_{1j}, ϕ_{2j} for both regimes, the transition parameter γ, and the transition threshold c of the model *ESTAR* C23-C123 are provided in the above able 13.

From Table 13, we can observe that the parameters ϕ_{1j}, ϕ_{2j} for both regimes, the transition parameter γ, and the transition threshold c of the model *ESTAR* C23-C123 are statistically significant at a 1% significance level.

If we substitute the coefficients with their estimated values in the model *ESTAR* C23-C123, we obtain:

$$y_t = (9.627652 + 6.097564.y_{t-2} - 0.371555.y_{t-3}).\exp\ (-35.51427(x + 1.587642)^2) + (-9.614251 + 0.259482.y_{t-1} - 6.078776.y_{t-2} + 0.347945.y_{t-3}).\left(1 - \exp\ (-35.51427(x + 1.587642)^2)\right) \quad (39)$$

or alternatively:

$$y_t = (9.627652 + 6.097564.y_{t-2} - 0.371555.y_{t-3}) + \left((-9.614251 - 9.627652) + 0.259482.y_{t-1} + (-6.078776 - 6.097564).y_{t-2} + (0.347945 + 0.371555).y_{t-3}\right).\exp\ (-35.51427(x + 1.587642)^2)$$

$$y_t = (9.627652 + 6.097564.y_{t-2} - 0.371555.y_{t-3}) + (-19{,}241903 + 0.259482.y_{t-1} - 12{,}17634.y_{t-2} + 0{,}7195.y_{t-3}).\exp\ (-35.51427(x + 1.587642)^2) \quad (40)$$

The Figure 5 below displays the graphical representation of the estimated exponential transition function of the model *ESTAR* C23-C123.

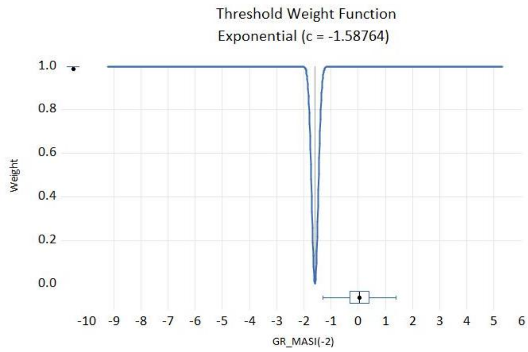


Figure 5: Curve of the estimated exponential transition function of the model *ESTAR* C23-C123

Table 18 below displays the p-values of the Ljung-Box test for the residuals series of the estimated model *ESTAR* C23-C123.

Table 18: Results of the Ljung-Box test applied to the residuals of the model *ESTAR* C23-C123

Date: 01/31/24 Time: 16:53
Sample (adjusted): 1/08/2002 8/11/2023
Q-statistic probabilities adjusted for 3 dynamic regressors

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.006	-0.006	0.2215	0.638
		2	0.003	0.003	0.2642	0.876
		3	-0.000	-0.000	0.2654	0.966
		4	-0.009	-0.009	0.7170	0.949
		5	0.025	0.025	4.0755	0.539
		6	-0.020	-0.020	6.3307	0.387
		7	0.015	0.015	7.5757	0.371
		8	-0.018	-0.018	9.3858	0.311

The examination of the p-values of the Q-statistics from the Ljung-Box test for serial correlation shows that the residuals of the estimated model*ESTAR* C23-C123 are not autocorrelated.

We also applied the Ljung-Box test to the series of squared residuals and found that the squared residuals of the model *ESTAR* C23-C123 are autocorrelated.

4.2.5 Estimation of the model *ESTAR* C23-C12

The model *ESTAR* C23-C12 is expressed as follows:

$$y_t = (\phi_{10} + \phi_{12}.y_{t-2} + \phi_{13}.y_{t-3}).exp\left(-\gamma(x - c)^2\right) + (\phi_{20} + \phi_{21}.y_{t-1} + \phi_{22}.y_{t-2}).\left(1 - exp\left(-\gamma(y_{t-1} - c)^2\right)\right) \quad (41)$$

or alternatively:

$$y_t = (\phi_{10} + \phi_{12}.y_{t-2} + \phi_{13}.y_{t-3}) + \left((\phi_{20} - \phi_{10}) + \phi_{21}.y_{t-1} + (\phi_{22} - \phi_{12}).y_{t-2} - \phi_{13}.y_{t-3}\right).exp\left(-\gamma(y_{t-1} - c)^2\right) \quad (42)$$

The estimates of the parameters ϕ_{1j}, ϕ_{2j} for both regimes, the transition parameter γ, and the transition threshold c of *ESTAR* C23-C12 are provided in the above Table 14.

From Table 14, we can observe that the parameters ϕ_{1j}, ϕ_{2j} for both regimes, the transition parameter γ, and the transition threshold c of the model *ESTAR* C23-C12 are statistically significant at a 1% significance level.

If we substitute the coefficients with their estimated values in the model *ESTAR* C23-C12, we obtain:

$$y_t = (-8.777471 - 3.960493.y_{t-2} - 0.029982.y_{t-3}).exp\left(-26449.72(y_{t-1} + 1.400001)^2\right) + (8.796976 + 0.241141.y_{t-1} + 3.976414.y_{t-2}).\left(1 - exp\left(-26449.72(y_{t-1} + 1.400001)^2\right)\right) \quad (43)$$

or alternatively:

$$y_t = (-8.777471 - 3.960493.y_{t-2} - 0.029982.y_{t-3}) + \left((8.796976 + 8.777471) + 0.241141.y_{t-1} + (3.976414 + 3.960493).y_{t-2} - -0.029982.y_{t-3}\right).exp\left(-26449.72(y_{t-1} + 1.400001)^2\right)$$

$$y_t = (-8.777471 - 3.960493.y_{t-2} - 0.029982.y_{t-3}) + (17{,}574447 + 0.241141.y_{t-1} + 7{,}936907.y_{t-2} + 0.029982.y_{t-3}).exp\left(-26449.72(y_{t-1} + 1.400001)^2\right) \quad (44)$$

The Figure 6 below displays the graphical representation of the estimated exponential transition function of the model *ESTAR* C23-C12.

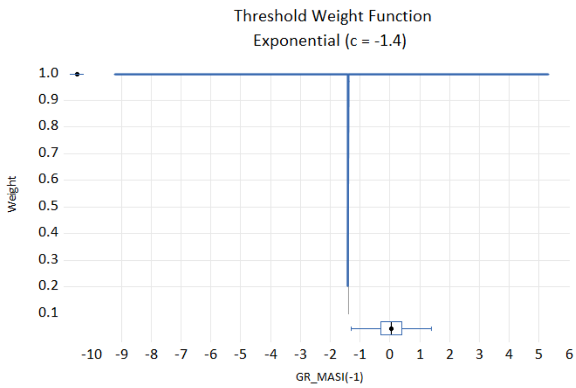


Figure 6: Curve of the estimated exponential transition function of the model *ESTAR* C23-C12

We can see clearly that this curve is reduced to 1, and therefore, the model *ESTAR* C23-C12 behaves like a linear autoregressive model of order 3.

Table 19 below shows the p-values of the Ljung-Box test for the residuals series of the estimated model *ESTAR* C23-C12.

Table 19: Results of the Ljung-Box test applied to the residuals of the model *ESTAR* C23-C12

Date: 01/31/24 Time: 15:01
Sample (adjusted): 1/08/2002 8/11/2023
Q-statistic probabilities adjusted for 3 dynamic regressors

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.005	0.005	0.1142	0.735
		2	0.001	0.001	0.1166	0.943
		3	0.000	0.000	0.1175	0.990
		4	-0.008	-0.008	0.5040	0.973
		5	0.024	0.024	3.7058	0.593
		6	-0.019	-0.019	5.6547	0.463
		7	0.016	0.017	7.1184	0.417
		8	-0.019	-0.019	9.0707	0.336

The examination of the p-values of the Q-statistics from the Ljung-Box test for serial correlation shows that the residuals of the estimated model *ESTAR* C23-C12 are not autocorrelated.

We also applied the Ljung-Box test to the series of squared residuals and found that the squared residuals of the model *ESTAR* C23-C12 are autocorrelated.

5. Conclusion

In this study, we tested the hypothesis of weak-form efficiency in the Moroccan stock market. To do so, we analyzed the daily geometric returns series of the *MASI* index on the Casablanca Stock Exchange from 03/01/2002 to 15/08/2023, totaling 5393 observations. In this regard, we have applied the Smooth Transition Autoregressive (*STAR*)model to the daily geometric returns series of the *MASI* index.

The application of the Augmented Dickey-Fuller test rejected the null hypothesis of a unit root, indicating that the series of geometric returns of the *MASI* follows a stationary process.

We then initiated the model specification procedure based on Teräsvirta's strategy (1994). In the first step, we selected the most appropriate autoregressive model $AR\ (p)$ of order p for the data series under investigation. Following Teräsvirta's suggestion, we chose the lag p that minimizes the Akaike Information Criterion (*AIC*) of the autoregressive linear regression of order p. We found that the $AR\ (3)$ model is the one minimizing the Akaike Information Criterion, and the analysis of the Ljung-Box test showed that the residuals of the $AR\ (3)$ model are non-autocorrelated.

Next, we estimated several $LSTAR\ (p_1, p_2)$ models with $1 \leq p_1 \leq 3$ as the order of the linear regime and $1 \leq p_2 \leq 3$as the order of the non-linear regime. We retained only those $LSTAR\ (p_1, p_2)$ models where the lag polynomials of both linear and non-linear regimes meet the following conditions:

- Both lag polynomials contain the constant term.
- Both lag polynomials contain at least two lag monomials of orders between 1 and 3.

With these conditions, we had a total of 16 possible configurations. For each model, we selected the optimal lag value d for the transition variable y_{t-d}. We only kept the *LSTAR* models whose coefficients of both lag polynomials and logistic transition function parameters are statistically significant.

We found only two models that meet the conditions: the model *LSTAR* C123-C23 (Base part: $Constant + AR(1) + AR(2) + AR(3)$; Alternative part: $Constant + AR(2) + AR(3)$) and the model *LSTAR* C23-C123 (Base part: $Constant + AR(2) + AR(3)$; Alternative part: $Constant + AR(1) + AR(2) + AR(3)$).

We compared the three Akaike Information Criterion, Schwarz Criterion, and Hannan-Quinn Criterion for the two models *LSTAR* C123-C23 and *LSTAR*C23-C123, and we found that the *LSTAR* C123-C23 model is the one minimizing all three criteria.

We then applied the linearity test to choose between the *STAR* model and the linear autoregressive model. The Teravista sequential tests rejected the hypothesis of the linear model at a significance level of 5% and recommended the *LSTAR* model instead of the *ESTAR* model. Similarly, the Escribano-Jorda tests rejected the hypothesis of the linear model at a significance level of 5% but recommended the *ESTAR* model instead of the *LSTAR* model.

As with the estimated *LSTAR* models, we estimated 16 $ESTAR\ (p_1, p_2)$ models with $1 \leq p_1 \leq 3$ as the order of the linear regime and $1 \leq p_2 \leq 3$ as the order of the non-linear regime.

We found only three *ESTAR* models that meet the required conditions, namely the model *ESTAR* C13-C12, the model *ESTAR* C23-C123, and the model*ESTAR* C23-C12.

The analysis of the three Akaike Information Criterion, Schwarz Criterion, and Hannan-Quinn Criterion for the three *ESTAR* models revealed that the model *ESTAR* C23-C12 is the one minimizing all three criteria. However, we noticed that the transition parameter for the model *ESTAR* C23-C12 is $\gamma = 26449.72$, a very high value for which the exponential transition function approaches 1. In such cases, the model *ESTAR* C23-C12 behaves like a linear autoregressive model of order 3.

Comparing the three information criteria of the nonlinear model *LSTAR* C123-C23, the linear model *ESTAR* C23-C12, and the nonlinear model *ESTAR* C23-C123 revealed that the *LSTAR* C123-C23 model is the one minimizing all three criteria. We selected this model as our choice; however, we still estimated all three models for comparison.

The objective of detecting whether the daily geometric returns series of MASI exhibits nonlinearity has been achieved. Indeed, the application of the Smooth Transition Autoregressive (*STAR*) model has captured nonlinearity in the daily geometric returns series of the *MASI* index over the studied period. This finding leads us to conclude that the Moroccan stock market demonstrates inefficiency in its weak-form.

References

[1]. Akintunde M.O., Kgosi P.M., Olalude, G.A. (2016). Examination of Botswana stock marketsusing regime switching models. *Journal of Statistical and Econometric Methods, vol.5, no.1, 2016, 49-58.*

[2]. Chien-Jen W., Po-Chin W., Huei-Hsieh L. (2013). The relationship between stock returns and foreign exchange rates in China using smooth regime switching approach. *Investment Management and Financial Innovations, 10 (3).*

[3]. Daabaji A. (2018). Non-linear modeling of the Moroccan stock market. Revue du Contrôle de la Comptabilité et de l'Audit. Numéro 6 : Septembre 2018. Page 878.

[4]. Davies R. B. (1987). Hypothesis Testing when a Nuisance Parameter is Present Only Under the Alternatives. Biometrika. Vol. 74, No. 1 (Mar.,1987), pp. 33-43.

[5]. Escribano A., Jordá O. (2001). Testing nonlinearity: Decision rules for selecting between logistic and exponential STAR models, *Spanish Economic Review, Springer;Spanish Economic* Association, vol. 3(3), pages 193-209.

[6]. Luukkonen R., Saikkonen P., Teräsvirta T. (1988a), Testing linearity against smooth transition autoregressive models. *Biometrika 75, 491-499*.

[7]. Luukkonen R., Saikkonen P., Teräsvirta T. (1988b). Testing linearity in univariate time series models. *Scandinavian Journal of Statistics 16, 161-175*.

[8]. Sarantis N. (1999). Modeling non-linearities in real effective exchange rates. Journal of *International Money and Finance. 18 (1999). 27-45.*

[9]. Si Mohammed K., Mouslim-Dib M., Zeddoun D., Benameur A. (2015). Application of Smooth Transition Autoregressive (STAR) Models for the Real Exchange Rate in Algeria. International *Journal of Business and Social Science Vol. 6, No. 11; November 2015.*

[10]. Siti Rohani M. N., Fadhilah Y., Ibrahim L. K. (2015). Nonlinear smooth transition autoregressive (star)–type modelling and forecasting on malaysia airlines (mas) stock returns. *Jurnal Teknologi (Sciences & Engineering) 74:11 (2015) 137–145*.

[11]. Teräsvirta T. (1994). Specification, Estimation and Evaluation of Smooth Transition Autoregressive Models. *Journal of the American Statistical Association,Vol.89.*

[12]. Tayyab M., Ayesha T., Riaz M. (2012). Application of Smooth Transition autoregressive (STAR) models for Exchange Rate. *Mathematical Theory and Modeling. Vol.2, No.9, 2012.*

[13]. Tong H., Lim K.S. (1980) Threshold Autoregression, Limit Cycles and Data. *Journal of the Royal Statistical Society: Serie B, 42, 245-292.*

[14]. Tong H. (1983). Threshold Models in Nonlinear Time-Series Analysis. *Springer-Verlag, New York.*

[15]. Tong H. (1990) Nonlinear Time Series Analysis. *Oxford University Press, Oxford, 564 p.*

[16]. Usman M. U., Tuba S., Guven S. (2018). Forecasting performance of smooth transition autoregressive (STAR) model on travel and leisure stock index. *The Journal of Finance and Data Science 4 (2018) 90-100.*

[17]. Van Dijk D., Teräsvirta T., Franses H. (2002). Smooth Transition Autoregressive Models: A Survey of Recent Developments. *Econometric Reviews 21: 1-47*.