# NON-LINEAR PREDICTABILITY OF STOCK MARKET RETURNS WITH MACROECONOMIC VARIABLES: EVIDENCE FROM THAILAND

BY

MR. SARUN PANMANOTHAM

AN INDEPENDENT STUDY SUBMITTED IN PARTIAL
FULFILLMENT OF THE REQUIREMENTS FOR
THE DEGREE OF MASTER OF SCIENCE
PROGRAM IN FINANCE (INTERNATIONAL PROGRAM)
FACULTY OF COMMERCE AND ACCOUNTANCY
THAMMASAT UNIVERSITY
ACADEMIC YEAR 2016

# NON-LINEAR PREDICTABILITY OF STOCK MARKET RETURNS WITH MACROECONOMIC VARIABLES: EVIDENCE FROM THAILAND

**BY**

**MR. SARUN PANMANOTHAM**

**AN INDEPENDENT STUDY SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE PROGRAM IN FINANCE (INTERNATIONAL PROGRAM)**
**FACULTY OF COMMERCE AND ACCOUNTANCY**
**THAMMASAT UNIVERSITY**
**ACADEMIC YEAR 2016**

THAMMASAT UNIVERSITY
FACULTY OF COMMERCE AND ACCOUNTANCY

INDEPENDENT STUDY

BY

MR. SARUN PANMANOTHAM

ENTITLED

NON-LINEAR PREDICTABILITY OF STOCK MARKET RETURNS WITH MACROECONOMIC VARIABLES: EVIDENCE FROM THAILAND

was approved as partial fulfillment of the requirements for
the degree of master of science (Finance)

on 01 MAY 2017

Chairman P. Chaiyuth

(Assistant Professor Chaiyuth Padungsaksawasdi, Ph.D.)

Member and advisor Chaleampong Kongcharoen

(Assistant Professor Chaleampong Kongcharoen, Ph.D.)

Dean P. Udom

(Associate Professor Pipop Udom, Ph.D.)

| | |
|---|---|
| Independent study title | NON-LINEAR PREDICTABILITY OF STOCK MARKET RETURNS WITH MACROECONOMIC VARIABLES: EVIDENCE FROM THAILAND |
| Author | Mr. Sarun Panmanotham |
| Degree | Master of Science (Finance) |
| Major field/Faculty/University | Master of Science Program in Finance, (International Program) Faculty of Commerce and Accountancy, Thammasat University |
| Independent study advisor | Assistant Professor Chaleampong Kongcharoen, Ph.D. |
| Academic year | 2016 |

# ABSTRACT

This paper examines in the non-linear relationship between macroeconomic factors and stock returns but mainly focus on the ability to forecast from each model including Autoregressive (AR) model, Threshold Autoregressive (TAR) model and Smooth transition autoregressive (STAR) model. The monthly data ranging from April 2004 to October 2016 are used. The results show that the distribution of the stock market return is non-linear. Moreover, the ability to forecast of Autoregressive (AR) model is better than Smooth transition autoregressive (STAR) model and Threshold (TAR) model.

**Keywords**: Non-linear relationship, STAR model, TAR model, AR model, Stock returns, Forecasting performanc

# ACKNOWLEDGEMENTS

I would like to express my sincere thanks to my independent study advisor, Assistant Professor Chaleampong Kongcharoen, Ph.D. for his help and encouragement throughout the course of this research. I would not have achieved this far and this independent study would not have been completed without his support.

Finally, I most thankfully acknowledge to my friends and my parents for all their support throughout the period of this research.

Mr. Sarun Panmanotham

# TABLE OF CONTENTS

# LIST OF TABLES

# CHAPTER 1
# INTRODUCTION

## 1.1 Introduction

The stock exchange basically serve as secondary markets which stock brokers and traders can buy or sell securities including stocks, bonds, and other securities. However, some actions in the stock exchange must be done in the primary market such as the initial public offering of stocks and bonds. Generally, the major type of the investors in the country is the retail investor. Retail investor love to speculate and they always use their emotion and feeling to judge their actions in the market which make they have a higher chance to get a capital loss. So, the fundamental analysis and the technical analysis are come. Those concepts can help investors to have a higher chance to take profit in the market.

From past to present, we have many variables those impact to the stock market index such as exchange rate, unemployment rate, inflation, interest rate, national income and so on. which will affect the stock market and stock price in the long-term based on linear relationships. Many of researchers use the regression analysis to find the relationship between macroeconomic variables and stock return which include stationary test to avoid spurious problem, Vector Autoregressive model and Granger Causality test to find the lag term relationship between variables, and GARCH or another method that adapt from ARCH-GARCH model to predict the return volatility.

However, many researchers have found evidence of non-linear relationships in the stock market. *Kyle, 1985* and *Black, 1986* found the relationship between noise and informed trader in the stock market. They found that noise and informed trader will push stock prices away from equilibrium values. *Cootner, 1962* claimed that the noise traders will cause prices to swing around equilibrium, and trigger arbitrage activities by informed traders which push prices back to equilibrium. Moreover, the barrier is likely depending on the transactions costs. *Ali Siad, 2001* found that the distribution of the daily return on the SET index is non- normal and leptokurtic. The results of the study also suggest that non-linear processes play a significant role in stock market behavior. *McMillan, 2000* suggest that a non-linear relationship does indeed exist between returns and interest rates in UK stock market returns. *Çil Yavuz, 2010* examined the dynamic relationship between stock returns and inflation in Turkey. He found that the presence of nonlinear cointegration between stock returns and inflation. All of these evidences show that stock market should have the non-linear behavior. If we still use the tools and the assumptions those based on linear relationships, the result will not be realistic and maybe it cannot tell real behavior of the stock market.

To find non-linear relationships, we need the different tools to examine relationships instead of the tools from linear relationships model. Regime-switching models can be useful in this situation. The models can be divided into two groups, Threshold model and Markov-switching model. The classic Threshold model was presented by Schelling, Axelrod, and Granovetter to model collective behavior. *Tong, 1978* introduced the Threshold autoregressive model (TAR) by assuming that regime shifts are triggered by observed variables. By the way, a criticism of the TAR model is that its conditional mean equation is not continuous. *Hamilton, 1989* developed the Markov-switching autoregressive model (MSA). This model can characterize the time-series behaviors in different regimes. In 1996, Gray studied about the various ARCH and GARCH models with Markov-switching. *McCulloch and TSAY, 1994* also consider a Markov Chain Monte Carlo (MCMC) to estimate a general MSA model. We have several researchers who developed this model such as *Hamilton and Lin, 1996, Dueker, 1997, Lam and Li, 1998, Susmel, 1998, etc.* The Markov-switching model is different from TAR model. TAR model uses a deterministic scheme to govern the model transition while MSA model uses a stochastic scheme. *Terasvirta and Anderson, 1992* developed Threshold model into Smooth Transition Autoregressive (STAR) model to be used in business cycles. They found that the nonlinearity is needed mainly to describe the responses of production to large negative shocks such as oil price shocks. Moreover, STAR model is different from Markov-switching model because we can define the probability function of choosing each regime while Markov-switching model cannot. The only thing that Markov-switching model can do is to predict the probability of using each regime and the probability will be constant over time. *Saechung, 2006* studied about Time varying excess returns on Thai Government Bond by using Vector STAR model. She claimed that Vector STAR model can be used to find the interaction between two regimes while VAR model cannot be used because VAR model cannot find the relationship of those regimes. Moreover, Vector STAR model is very flexible. We can apply this model to describe the behavior of macroeconomics variables and financial variables in both stock market and bond market.

## 1.2 Research Questions

1) Can macroeconomic factors explain stock returns from Thailand stock exchange significantly?

2) Do we have non-linear behavior in Thailand stock exchange returns from April 2004 to October 2016?

3) Which model has better forecasting performance among Autoregressive (AR) model, Threshold Autoregressive (TAR) model, and Smooth Transition Regression (STAR) model?

## 1.3 Contribution

1) Use Autoregressive (AR) model to find the linear relationships between macroeconomic variables and stock market returns in Thailand.

2) Use Threshold Autoregressive (TAR) model and Smooth Transition Autoregressive (STAR) model to find the non-linear relationships between macroeconomic variables and stock market returns in Thailand.

3) Comparing the forecasting performance among STAR model, TAR model, and AR model

# CHAPTER 2
# REVIEW OF LITERATURE

## 2.1 Linear relationship

*(Brahmasrene & Jiranyakul, 2007)* examine the relationship between stock market index and selected macroeconomic variables during the post-financial liberalization (pre-financial crisis) and post- financial crisis in Thailand by using the error correction mechanism (ECM) and vector autoregressive model (VAR). They find that the stock index is cointegrated with some macroeconomic variables in the pre-financial crisis, but not in the post-financial crisis. Moreover, they claim that monetary policy may be able to stimulate the stock market but only in short run. The error correction mechanism (ECM) and vector autoregressive model (VAR) also use to apply in several papers with different set of macroeconomic variables and different stock index with the same conclusion such as *Paisalyakit (2005)* and *Piantam (2009).*

*(Kanokwichitslip, 2008)* investigates causality the volatility among each country's macroeconomic, regional and global factors in South East Asia emerging markets as well as comparing their effects between before crisis and after crisis by using GARCH to model volatility and eight-variable VAR model, Granger Causality, and IRF investigates the relationship of all studied variables. The results show that the relationship of domestic macroeconomic, regional, and global variables is more related in post-crisis period. Moreover, based on IRF, the movement of stock return to its lag is significantly positive on the response of stock return volatility in four countries both pre-crisis and post-crisis period.

*(D. Gay, 2008)* investigate the time- series relationship between stock market index prices and the macroeconomic variables of exchange rate and oil price for Brazil, Russia, India, and China (BRIC) using the Box-Jenkins ARIMA model. He claims that there is no significant relationship was found between exchange rate and oil price on the stock market index. Also, there was no significant relationship found between present and past stock market returns.

(*Tangjitprom, 2012)* study the relationship between macroeconomic variables and Thailand stock market returns by using Vector Autoregressive model (VAR) and

Granger causality test. He finds macroeconomic variables have less relationship to forecast future stock return whereas stock return can be used to predict macroeconomic variables more. Additionally, the interest rate factor has the most relationship to explain the variance in stock return.

*(Hatipoglu et al., 2014)* finds the relationships between Borsa Istanbul Stock Exchange (BIST) stock prices and underlying macroeconomic shocks in the Turkish economy. They used Structural Vector Autoregressive (SVAR) models to examine the effects of fundamental shocks on stock price movements in Turkey. They found the relationship between stock prices and real activity variables is more than the relationship with key investments.

*(Joshi, 2015)* examines how fiscal fundamental macroeconomic variables affect the performance of the stock market in India. He used the Auto Regressive Distributed Lag (ARDL) bounds test and a Vector Error Correction Model (VECM) for short and long run dynamic relationships.

## 2.2 Non-linear relationship

*(McMillan, 2000)* test for a non-linear relationship between stock market returns and macroeconomic variables, and whether this non-linearity can be exploited for forecast improvements by using nonparametric regression technique, a smooth-transition threshold type model. The results suggest that a non-linear relationship exists between returns and interest rates.

*(Ali Siad, 2001)* test for nonlinear dynamics in The Stock Exchange of Thailand (SET) from 1975 to 1999. The study found the distribution of the daily return on the SET index is not normal distributed and leptokurtic. The results of the study suggest that stock market behavior is significant non-linear.

*(McMillan, 2002)* test for an exponential smooth transition threshold model can provide a better characteristic of UK stock market returns than a linear model or an alternate non-linear model.

*(Saechung A. ,2006)* studies about the behavior of excess return in Thailand government bond by using Vector STAR model. The results show that the movement of excess return is non-linearity and varying by regime switching behavior. She

suggest that we should use Vector STAR model to describe the behavior of macroeconomic variables instead of VAR model.

*(Guidolin et al., 2009)* find the comparative predictive performance of a number of linear and non-linear models for stock and bond returns in the G7 countries by using Markov switching, threshold autoregressive (TAR), and smooth transition autoregressive (STAR) regime switching models. They estimate GARCH model to predict volatilities which maybe appear in the conditional mean function. They fail to find a consistent win ner/out-performer across all countries and asset markets. By the way, the results show that non-linear effects is the key to improve forecasting performance.

(Çil Yavuz, 2010) examine the dynamic relationship between stock returns and inflation in Turkey by using two-regime vector error- correction model with a single cointegrating vector and threshold effect in the error-correction term. The main findings suggest the presence of nonlinear cointegration between stock returns and inflation.

*(El Hedi AROURI, 2010)* examines the financial integration hypothesis for two emerging countries (the Philippines and Mexico) into the world stock market in a nonlinear framework. The results show the nonlinear financial integration However, this relationship is significantly more important for Mexico and that it has been powerfully stimulated after 1994.

*(Humpe & Macmillan, 2014)* use the logistic smooth transition regression model to examine the non-linear predictability of Japanese and US stock market returns by a set of macroeconomic variables. They find that the relationship between stock returns and macroeconomic variables is dependent upon the state of the market. Moreover, they also find some evidence of an inner and outer regime in the Japanese data but just outer regime for US.

# CHAPTER 3
# RESEARCH METHODOLOGY

## 3.1 Data

This study uses monthly data from April 2004 to October 2016. Consumer Price Index was collected from Bureau of trade. Other economic indicators and stock market returns were collected from The Stock Exchange of Thailand. Inflation is calculated from consumer price index. The exchange rate between Baht and US Dollar represents an external factor. The unemployment rate is used to represent the economic activity in Thailand because monthly GDP is not available. Five-year government bond yield represents the level of interest rate.

The results from Augmented Dickey-Fuller Test (Dickey and Fuller, 1979) and KPSS test (Kwiatkowski, Phillips, Schmidt, and Shin, 1992) shows that the log-return of stock market index, the inflation, and the exchange rate USD/THB are stationary. For unemployment rate and government bond 5Y, they are stationary after taking the first difference.

## 3.2 Research methodology

### 3.2.1 Descriptive statistics and stationary test

In order to model the time series data, we need the data to be stationary, otherwise we will face the spurious problem. A spurious problem is a relationship that two or more variables in the equation are not relevance but the statistical results also show that those variables are significant with the high value of R-square. Unit root test is the way to detect whether the series is stationary or not which included Dickey-Fuller test (Dickey and Fuller, 1979), ADF test and KPSS test (Kwiatkowski, Phillips, Schmidt, and Shin, 1992). Unit root test, such as Dickey-Fuller test, is the way to avoid the spurious problem.

Assume the AR(1) model,

$$y_t = \rho y_{t-1} + u_t; \qquad t = 1,2,3, \dots T$$

The conditions for the process to be stationary depends on the parameter $\rho$.

| | | | | |
|---|---|---|---|---|
| If | $\mid \rho \mid < 1$ | then | $y_t$ | is stationary |
| If | $\mid \rho \mid \geq 1$ | then | $y_t$ | is non-stationary |

We set the hypothesis as,

$H_0: \rho = 1$ (Non-Stationary)

$H_1: \rho < 1$ (Stationary)

If we cannot reject Null hypothesis, $y_t$ is non-stationary.

After we obtain the stationary data, we will find the descriptive statistics for preliminary information before we construct the model.

- Sample Mean: $\hat{\mu} = \frac{1}{T}\sum_{t=1}^{T} y_t$

Generally, Sample Mean $(\hat{\mu})$ is used to specify the location of the distribution.

- Sample Standard Deviation: $\hat{\sigma} = \sqrt{\frac{\sum_{t=1}^{T}(y_t - \hat{\mu})^2}{T-1}}$

Due to we study with the random variables, Sample Standard Deviation $(\hat{\sigma})$ will tell us about the magnitude of the distribution. The more $\hat{\sigma}$ is higher, the more fluctuations of the distribution.

- Sample Skewness: $\hat{S} = \frac{1}{\hat{\sigma}^3}\frac{\sum_{t=1}^{T}(y_t - \hat{\mu})^3}{T-1}$

Sample Skewness $(\hat{S})$ tell us about the asymmetric of the distribution. It can twist to the right if $\hat{S} > 0$, to the left if $\hat{S} < 0$ or even symmetry if $\hat{S} = 0$.

- Sample Kurtosis: $\hat{K} = \frac{1}{\hat{\sigma}^4}\frac{\sum_{t=1}^{T}(y_t - \hat{\mu})^4}{T-1}$

The Sample Kurtosis of the normal distribution is 3. If $\hat{K} > 3$, then we call it leptokurtic. If $\hat{K} < 3$, we will call it platykurtic.

- Sample Correlation: $\hat{\rho}(x,y) = \frac{\hat{\sigma}_{xy}}{\sqrt{\hat{\sigma}_x^2 \hat{\sigma}_y^2}}$

The sample correlation is +1 in the case of a perfect increasing linear relationship. The sample correlation is -1 in the case of a perfect decreasing linear relationship.

### 3.2.2 Autoregressive (AR) model and estimation

An autoregressive model (AR) is representation of a type of random process. It specifies that the variable depends linearly on its own previous values and the stochastic term. Thus the model is in the form of a stochastic difference equation.

The autoregressive process of order p or AR(p) is defined

$$y_t = \sum_{j=1}^{p} \varphi_j y_{t-j} + w_t; \qquad w_t \sim N(0, \sigma^2)$$

$\varphi = (\varphi_1, \varphi_2, \dots, \varphi_p)$ is the vector of model coefficients and p is non-negative integer. For $p = 0, y_t = w_t$ and there is no autoregressive term.

The lag operator is denoted by B and used to express lagged values of the process

$$B^2 y_t = y_{t-2}, B^3 y_t = y_{t-3}, \dots, B^d y_t = y_{t-d}$$

If we define

$$\emptyset(B) = 1 - \sum_{j=1}^{p} \varphi_j B^j = 1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_p B^p$$

$\emptyset(B)$: Characteristic polynomial of the process and its roots determine when the process is stationary or not.

The AR(p) process is given by the equation

$$\emptyset(B) y_t = w_t; t = 1, \dots, n$$

The Breusch-Godfrey test is used to assess the validity of some of the modeling assumptions inherent in applying regression-like models to observed data series.

Because the test is based on the idea of Lagrange multiplier testing, it is sometimes referred to as LM test for serial correlation.

Consider:

$$Y = X\beta + u$$

In which we suspect that $u_t$ is first order serially correlated, i.e.

$$u_t = \rho_1 u_{t-1} + \varepsilon_t$$
$$\varepsilon_t \sim iid\ N(0, \sigma^2), t = 1,2, \dots, n$$

We need to test the null hypothesis:

$$H_0: \rho_1 = 0$$

$$H_a: \rho_1 \neq 0$$

Add and subtract $\rho_1 Y_{-1}$ at the right hand side of the main process:

$$Y = \rho_1 Y_{-1} - \rho_1 Y_{-1} + X\beta + u$$

However,

$$Y_{-1} = X_{-1}\beta + u_{-1}$$
$$Y = \rho_1 Y_{-1} - \rho_1 (X_{-1}\beta + u_{-1}) + X\beta + u$$
$$Y = \rho_1 Y_{-1} + (X - \rho_1 X_{-1})\beta + \varepsilon$$

Where, $\varepsilon = u - \rho_1 u_{-1}$

The restricted MLEs, $\hat{\beta}_0$, are the OLS estimates of $Y = X\beta + u$.

$$e_0 = Y - X\hat{\beta}_0$$

The LM test statistic is $LM = nR^2 \approx \chi_1^2$, where $R^2$ is obtained from the regression:

$e_{0,t} = \gamma x_t + \delta e_{0,t-1} + error$, if X contains lagged dependent variables.

$e_{0,t} = \delta e_{0,t-1} + error$, otherwise

Consider the transformed LM test statistic:

$$LM^* = \frac{n-k}{m}\frac{R^2}{1-R^2} \approx F(m, n-k)$$

m: number of restrictions

k: number of estimates coefficients

**3.2.3 Threshold Autoregressive (TAR) model and estimation**

Classic threshold models were developed by Schelling, Axelrod, and Granovetter to model collective behavior.

For the observation in regime $j = 0,1, \dots, m$ we have the linear regression specification

$$y_t = X_t'\beta + Z_t'\delta_j + \epsilon_t$$

$X$: The variables whose parameters do not vary across regimes

$Z$: The variables whose coefficients are regime-specific

Suppose that there is an observable threshold variable $q_t$ and strictly increasing threshold values $(\gamma_1 < \gamma_2 < \cdots < \gamma_m)$ such that we are in regime j if and only if:

$$\gamma_j \leq q_t \leq \gamma_{j-1}; \qquad \gamma_0 = -\infty \text{ and } \gamma_{m+1} = \infty$$

Thus, we are in regime $j$ if the value of the threshold variable is at least as large as the $j-th$ threshold, but not as large as the $(j+1)-th$ threshold.

Given the threshold variable and the regression specification in

$$y_t = X_t'\beta + Z_t'\delta_j + \epsilon_t$$

We wish to find the coefficients $\delta$ and $\beta$, and usually, the threshold values $\gamma$. We may also use model selection to identify the threshold variable $q_t$.

Nonlinear least squares is a natural approach for estimating the parameters of the model. If we define the sum-of-squares objective function.

$$S(\delta,\beta,\gamma)=\sum_{t=1}^{T}(y_t-X_t'\beta-\sum_{j=0}^{m}1_j(q_t,\gamma)\cdot Z_t'\delta_j)^2$$

And we may obtain threshold regression estimates by minimizing $S(\delta,\beta,\gamma)$ with respect to the parameters.

Taking advantage of the fact that for a given $\gamma$, say $\tilde{\gamma}$, minimization of the concentrated objective $S(\delta,\beta,\gamma)$ is a simple least squares problem, we can view estimation as finding the set of thresholds and corresponding OLS coefficient estimates that minimize the sum-of-squares across all possible sets of $m-$ threshold partitions.

**3.2.4 Smooth transition autoregressive (STAR) model and estimation**

*Teräsvirta and Anderson (1992)* developed Smooth Transition Autoregressive (STAR) model which is one of the regime switching model but difference from Markov switching model. The difference between two models is STAR model has observed threshold variable but Markov switching model has unobserved variables that drive the changes. The Smooth Transition Autoregressive (STAR) model which is the combination of the average from each regime has two main parts those we have to consider, the equation for each regime and the weighted of each function.

The smooth transition autoregressive model (STAR) of order $p$ is defined as

$$y_t=\theta' w_t+(\pi' w_t)F(y_{t-d};\gamma,c)+u_t \text{ ; Where } u_t{\sim}nid(0,\sigma_u^2)$$

$$\theta=\left(\theta_1,\theta_2\ldots,\theta_p\right)',\pi=\left(\pi_1,\pi_2,\ldots\pi_p\right)', and\ w_t=(1,y_{t-1},\ldots,y_{t-p})'$$

$y_{t-d}$: Transition Variable; the transition might depend on the past values of the variables, or exogenous variables.

$\gamma$: The parameter in the transition function that can tell you about the speed of changing from regime to another regime.

$c$: The parameter in the transition function.

The transition function: $F(y_{t-d};\gamma,c)$ that is normally used had two forms. The first one is in the form of Logistic function and the second is the Exponential function.

- Logistic STAR (LSTAR):

$$F(y_{t-d};\gamma,c) = (1+\exp\{-\gamma(y_{t-d}-c)\})^{-1}; \qquad \gamma > 0$$

- Exponential STAR (ESTAR):

$$F(y_{t-d};\gamma,c) = 1-\exp\{-\gamma(y_{t-d}-c)^2\}; \qquad \gamma > 0$$

Linearity test: LM test

- From STAR model, regress $y_t$ on $w_t$ by OLS and compute the residuals

$$\hat{v}_t = y_t - \hat{\pi} w_t \text{ and their } SSR_0.$$

- Compute the auxiliary regression by OLS and the SSR.
- Calculate the statistic $LM = \frac{T(SSR_0 - SSR)}{SSR_0}$.

Transition variable significance

To test if a variable is statistically significant enough to be the transition variable the auxiliary regression must be calculated:

$$\hat{v}_t = \beta_0' w_t + \beta_1' w_t w_{td} + \beta_2' w_t w_{td}^2 + \beta_3' w_t w_{td}^3 + \eta_t$$

Where $\hat{v}_t$ are the residuals of the model. The null hypothesis is

$$H_0: \beta_1 = \beta_2 = \beta_3 = 0$$

If the null hypothesis is rejected, the several choices of $x_{td}$ one must select the one with the smallest p-value or the biggest statistical.

<u>Test between STAR structures</u>

This test decides between LSTAR or ESTAR structure. From the auxiliary regression, the sequence of nested test showed below.

$$H_{04}: \beta_3 = 0$$
$$H_{03}: \beta_2 = 0 | \beta_3 = 0$$
$$H_{02}: \beta_1 = 0 | \beta_3 = \beta_2 = 0$$

The selection rule stated that if the rejection of the $H_{03}$ is the strongest, H is the strongest, Teräsvirta (1998) advises choosing the ESTAR model. In case of the strongest rejection of the hypothesizes $H_{04}$ or $H_{02}$, LSTAR is chosen as the appropriate model. This heuristic decision rule is based on expressing the parameter vectors $\beta_1, \beta_2$, and $\beta_3$ from auxiliary regression as functions of the parameters $\gamma, c$ and $\theta$.

**3.2.5 Forecasting Performance**

The forecast error is simply $e_t = y_t - \hat{y}_t$. The accuracy based on $e_t$ cannot be used to make comparisons between different scales of the series.

The method to forecast error is Mean Squared Error. To measure the ability to forecast from each model, the model that has lower Mean Squared Error should have better ability to forecast. The two most commonly used methods are based on the squared errors and absolute errors.

Mean Absolute Error (MAE) = $\frac{\sum_{t=1}^{T} |e_t|}{T}$

Root Mean Squared Error (RMSE) = $\sqrt{\frac{\sum_{t=1}^{T} e_t^2}{T}}$

For time series data, the training set consists only of observations that occurred prior to the observation that forms the test set. Thus, no future observations can be used in constructing the forecast.

Suppose $k$ observations are required to forecast. Thus, the process works as follows.

1. Select the observation at time $k+i$ for the testing set, and use the observations at times $1,2,\ldots,k+i-1$ to estimate the forecasting model.
2. Compute the error on the forecast at time $k+i$.
3. Repeat the above step for $i=1,2,\ldots,T-k$ where $T$ is the total number of observations.
4. Compute the forecasting performance measures based on the errors obtained.

# CHAPTER 4
# RESULTS AND DISCUSSION

## 4.1 Descriptive statistics and stationary test

Log return of stock market index, exchange rate USD/THB, and consumer price index (Inflation) is stationary at level while unemployment rate and government bond 5Y is stationary at first difference. Table 4.1 shows the results of ADF-statistic of every variable is larger than absolute MacKinnon value at 99% 95% and 90% confidence level which means we can reject the null hypothesis. Therefore, all variables are stationary.

Table 4.1: The results of stationary by using ADF test

| Augmented Dickey-Fuller (ADF) Test | | | | | |
|---|---|---|---|---|---|
| Variables | ADF | 1% MacKinnon | 5% MacKinnon | 10% MacKinnon | Results |
| LNRETURN | -5.450539 | -4.021691 | -3.440681 | -3.144830 | Stationary |
| LNFX | -11.96132 | -4.020822 | -3.440263 | -3.144585 | Stationary |
| LNCPI | -10.30959 | -4.020822 | -3.440263 | -3.144585 | Stationary |
| FIRSTDIFUNEM | -4.825829 | -4.025924 | -3.442712 | -3.146022 | Stationary |
| FIRSTDIFGB | -11.0934 | -4.020822 | -3.440263 | -3.144585 | Stationary |

Table 4.2 shows that the results of KPSS-Statistic of every variable is smaller than KPSS value at 99% 95% and 90% confidence level which means we cannot reject the null hypothesis and it implies that the result is stationary.

Table 4.2: The results of stationary by using KPSS test

| Kwiatkowski-Phillips-Schmidt-Shin (KPSS) Test | | | | | |
|---|---|---|---|---|---|
| Variables | KPSS | 1% KPSS | 5% KPSS | 10% KPSS | Results |
| LNRETURN | 0.051737 | 0.216000 | 0.146000 | 0.119000 | Stationary |
| LNFX | 0.045734 | 0.216000 | 0.146000 | 0.119000 | Stationary |
| LNCPI | 0.034450 | 0.216000 | 0.146000 | 0.119000 | Stationary |
| FIRSTDIFUNEM | 0.022739 | 0.216000 | 0.146000 | 0.119000 | Stationary |
| FIRSTDIFGB | 0.038808 | 0.216000 | 0.146000 | 0.119000 | Stationary |

Table 4.3 shows the descriptive statistics of the variables. Most of the variable has a positive skew except the exchange rate USD/THB and every variable has the

kurtosis value more than 3 which means they are fat tail. Table 4.4 shows the correlations between each variable in the system.

Table 4.3: Descriptive Statistics

| | LNRETURN | LNFX | LNCPI | FIRSTDIFUNEM | FIRSTDIFGB |
|---|---|---|---|---|---|
| Mean | -0.0056 | 0.0009 | 0.0024 | -0.0001 | -0.0001 |
| Median | -0.0137 | 0.0027 | 0.0018 | -0.0004 | -0.0001 |
| Maximum | 0.3592 | 0.0568 | 0.0634 | 0.0182 | 0.0110 |
| Minimum | -0.1308 | -0.0828 | -0.0306 | -0.0099 | -0.0112 |
| Std. Dev. | 0.0591 | 0.0196 | 0.0075 | 0.0033 | 0.0030 |
| Skewness | 1.7850 | -0.3425 | 3.1067 | 1.2351 | 0.0063 |
| Kurtosis | 11.6653 | 4.8711 | 33.8920 | 9.5210 | 6.1943 |
| | | | | | |
| Jarque-Bera | 548.9432 | 24.8143 | 6205.7510 | 303.9089 | 63.7736 |
| Probability | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | | | | | |
| Sum | -0.8362 | 0.1339 | 0.3579 | -0.0158 | -0.0144 |
| Sum Sq. Dev. | 0.5196 | 0.0574 | 0.0083 | 0.0016 | 0.0013 |
| | | | | | |
| Observations | 150 | 150 | 150 | 150 | 150 |

Table 4.4: Correlation

| | LNRETURN | LNFX | LNCPI | FIRSTDIFUNEM | FIRSTDIFGB |
|---|---|---|---|---|---|
| LNRETURN | 1.0000 | -0.5530 | -0.1323 | 0.0107 | -0.0919 |
| LNFX | -0.5530 | 1.0000 | 0.1567 | 0.0665 | -0.0149 |
| LNCPI | -0.1323 | 0.1567 | 1.0000 | 0.1484 | 0.2363 |
| FIRSTDIFUNEM | 0.0107 | 0.0665 | 0.1484 | 1.0000 | -0.1003 |
| FIRSTDIFGB | -0.0919 | -0.0149 | 0.2363 | -0.1003 | 1.0000 |

## 4.2 The Autoregressive (AR) Model and estimation

Before we can estimate the Autoregressive Model (AR), we need to select the lag length for the AR(p) model by using AIC and SIC statistic.

Table 4.5 shows the result of AIC and SIC statistics for each AR(p) model. From the results, the AIC suggests the AR(4) model while SIC suggests the AR(1) model. In order to obtain the parsimonious model, AR(1) model is selected.

Table 4.5: AIC and SIC statistics among AR(p)

| Model | AIC | SIC |
|---|---|---|
| AR(1) | -2.8064 | -2.6854 |
| AR(2) | -2.7795 | -2.5567 |
| AR(3) | -2.7948 | -2.4693 |
| AR(4) | -2.8157 | -2.3866 |
| AR(5) | -2.8001 | -2.2664 |
| AR(6) | -2.7473 | -2.1080 |

Table 4.6 shows the result of the AR(1) model with exogenous variables. LNRETURN is the log-return of the stock market index which is the dependent variable of this model and the first lag of LNRETURN is used to explain itself. The first lag of independent variables are used, Log-return of exchange rate USD/THB (LNFX), Inflation (LNCPI), the first different of unemployment rate (FIRSTDIFUNEM), and the first different of 5Y-government bond (FIRSTDIFGB), to be the exogenous variables in this model.

Only the first lag of itself LNRETURN(-1) is significant at 95% confidence level. However, we still include the exogenous variables in the model because they still have the relationship with the dependent variable and it will cause serious problem if we drop some variable in the model.

Table 4.6: AR(1) estimation

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|---|---|---|---|---|
| LNRETURN(-1) | 0.2650 | 0.0981 | 2.7019 | 0.0077 |
| LNFX(-1) | 0.3373 | 0.2967 | 1.1370 | 0.2575 |
| LNCPI(-1) | 0.1388 | 0.6784 | 0.2045 | 0.8382 |
| FIRSTDIFUNEM(-1) | -1.6605 | 1.5056 | -1.1029 | 0.2719 |
| FIRSTDIFGB(-1) | 2.0420 | 1.6883 | 1.2095 | 0.2285 |
| C | -0.0049 | 0.0051 | -0.9609 | 0.3382 |

### 4.3 Serial Correlation problem test

Table 4.7 shows the result of serial correlation LM test. The null hypothesis is there is no serial correlation. From the result of AR(1) model, p-value of the F-statistic is 0.07 which is insignificant at 95% confident level thus we cannot reject the null hypothesis which means this model has no serial correlation problem. Then, we can use this model to obtain forecast values.

Table 4.7: Serial correlation LM test statistics

| Breusch-Godfrey Serial Correlation LM Test | | | |
|---|---|---|---|
| F-statistic | 3.3337 | Prob. F(1,142) | 0.0700 |
| Obs*R-squared | 3.4178 | Prob. Chi-Square(1) | 0.0645 |

**4.4 Threshold Autoregressive (TAR) Model and estimation**

Table 4.8 shows that the threshold model with the log-return of stock market index as threshold variable. We also add other macroeconomic variables as exogenous variables. In this model we estimate a two-regime threshold regression model with the first lag of LNRETURN as the threshold variable. The threshold value is 0.04540432 which tell us that if the threshold variable is less than 0.0450432 then we are in first regime and when the threshold variable is greater or equal than 0.0450432 then we are in second regime.

Table 4.8: TAR(1) model estimation

| First Regime (125 obs) | | | | |
|---|---|---|---|---|
| Threshold value < 0.04540432 | | | | |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
| LNRETURN(-1) | -0.0010 | 0.1097 | -0.0094 | 0.9925 |
| LNFX | -1.3703 | 0.2048 | -6.6922 | 0.0000 |
| LNCPI | 0.0985 | 0.5893 | 0.1672 | 0.8675 |
| FIRSTDIFUNEM | -0.0028 | 1.2470 | -0.0023 | 0.9982 |
| FIRSTDIFGB | 0.7986 | 1.4554 | 0.5487 | 0.5841 |
| Second Regime (24 obs) | | | | |
| Threshold value >= 0.04540432 | | | | |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
| LNRETURN(-1) | 0.0383 | 0.1110 | 0.3452 | 0.7305 |
| LNFX | -3.8846 | 0.6072 | -6.3972 | 0.0000 |
| LNCPI | 2.2324 | 1.3969 | 1.5980 | 0.1123 |
| FIRSTDIFUNEM | 5.2324 | 3.3203 | 1.5759 | 0.1173 |
| FIRSTDIFGB | -12.7291 | 3.3354 | -3.8163 | 0.0002 |
| Non-Threshold Variables | | | | |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
| C | -0.0090 | 0.0048 | -1.8634 | 0.0645 |

**4.5 Smooth Transition Autoregressive (STAR) model test**

Table 4.9 shows the results of Linearity test for the STAR model. The null hypothesis of the test is the series is Linear which means if we can reject the null hypothesis then the model is non-linear distribution. From the results, the p-value

(0.014) is less than the significant level (0.05) which means that the model has non-linear distribution.

Table 4.9: Linearity test (LM Test)

| Tests on LNRETURN | | |
|---|---|---|
| Transition variable: LNRETURN(-1) | | |
| H0: Linearity | | |
| | LM statistic | P-value |
| Transition variable | 14.2580 | 0.0140 |

Table 4.10 shows the test of the transition variables. The null hypothesis of the test is the transition variable is not significant which means if we can reject the null hypothesis, then the transition variable is appropriated to use. The transition value of the model is the first lag of the log-return of stock market index (LNRETURN(-1)). The results show that it is significant at 95% confidence level. Therefore, we can conclude that the log return of stock market index is fit to the model.

Table 4.10: Transition variables significance

| Transition variable tests on: LNRETURN(-1) | | |
|---|---|---|
| H0: The transition variable is not significant | | |
| Transition variable (STR) | F-statistic | P-value |
| LNRETURN(-1) | 130.4440 | 0.0000 |

Table 4.11 shows the results of the STAR structures test between ESTAR and LSTAR. The results show that we should use the ESTAR structure to estimate the model.

Table 4.11: STAR structures test

| Structure tests on: LNRETURN(-1) | |
|---|---|
| Choice: LSTR or ESTR at 5% of significance | |
| Transition variable (STR) | Structure |
| LNRETURN(-1) | ESTR |

### 4.6 Smooth Transition Autoregressive (STAR) model and estimation

Table 4.12 shows the estimated ESTAR model. The estimation of threshold value c (Coef(14)) is 0.118716 and the speed of changing from regime to another regime ($\gamma$) is 3.409552 (Coef(13)). The rest of the coefficients are the coefficient of each variable in the model.

From the estimation, the value of $\gamma$ = 3.409552, if LNRETURN(-1) $\geq$ 0.118716 then the process is described by the coefficients in the table with F(LNRETURN(-1)) =1 as the first regime and F(LNRETURN(-1)) = 0 otherwise.

Table 4.12: STAR(1) model estimation

LNRETURN= @coef(1) + @coef(2) * LNRETURN (-1) + @coef(3) * LNFX+ @coef(4) * LNCPI + @coef(5) * FIRSTDIFUNEM+ @coef(6) * FIRSTDIFGB + (@coef(7) + @coef(8) * LNRETURN (-1) + @coef(9) * LNFX + @coef(10) * LNCPI + @coef(11) * FIRSTDIFUNEM + @coef(12) * FIRSTDIFGB) / (1 - @exp((-@coef(13) * (lnreturn(-1) - @coef(14))^2) / 0.00348698447684603))

| | Coefficient | Prob. |
|---|---|---|
| Coef(1) | -0.2764990 | 0.7848 |
| Coef(2) | 3.0805250 | 0.7829 |
| Coef(3) | 0.6841960 | 0.9192 |
| Coef(4) | -1.2470560 | 0.7075 |
| Coef(5) | -9.1217330 | 0.7036 |
| Coef(6) | 31.4132700 | 0.7354 |
| Coef(7) | 0.2696400 | 0.7902 |
| Coef(8) | -3.0263800 | 0.7863 |
| Coef(9) | -2.1267750 | 0.7527 |
| Coef(10) | 1.2536450 | 0.6953 |
| Coef(11) | 9.2130820 | 0.6991 |
| Coef(12) | -30.3541600 | 0.7443 |
| Coef(13) | 3.4095520 | 0.5998 |
| Coef(14) | 0.1187160 | 0.0000 |

### 4.7 Forecasting performance

After we estimate all of the models correctly, we can forecast the log-return of the stock market index (LNRETURN) for each model. After that, we compare the forecasting performance.

Table 4.13 shows the RMSE and MAE results. We forecast the results for ten month. The forecast period is on January, 2016 to October, 2016. We start form the data at April, 2004 to December, 2015 to forecast the log-return of the stock market index in January, 2016 and roll the data from April, 2004 to January, 2016 to forecast February, 2016 and so on. Both of them imply the forecasting performance of each model, the lower of the statistics value, the better of the forecasting performance. The results show that AR(1) has the best ability to forecast at RMSE = 0.000048493 and MAE = 0.001346753 follow by TAR(1) at RMSE = 0.000055512 and MAE = 0.001626202 and STAR(1) at RMSE = 0.000056085 and MAE = 0.00161745.

Table 4.13: Mean Squared Error

| | AR(1) | TAR(1) | STAR(1) |
|---|---|---|---|
| RMSE | 0.000048493 | 0.000055512 | 0.000056085 |
| MAE | 0.001346753 | 0.001626202 | 0.001617450 |

# CHAPTER 5
# CONCLUSIONS AND RECOMMENDATIONS

## 5.1 Conclusions

This study's purpose is to find the non-linear relationship and the ability to forecast between stock market return and macroeconomic factors by using the Autoregressive (AR) model, Threshold Autoregressive (TAR) model and Smooth transition autoregressive (STAR) model, which one is the best. All of the model is estimated at first lag, AR(1), TAR(1) and ESTAR(1). Applying the TAR and ESTAR model allows me to separate the relationship between stock market return and macroeconomic factors into an upper and lower regime.

The results from this study show that the stock market return has a non-linear distribution where the behavior of itself changing by regime switching. The ability to forecast what is measured by Mean Squared Error including Root Mean Squared Error (RMSE) and Absolute Mean Squared Error shows that AR(1) has the best ability to forecast the outputs from January, 2016 to October, 2016.

## 5.2 Recommendations

1. Even if the Smooth transition autoregressive (STAR) model is commonly used for non-linear distributed model, but the results in this study show that Autoregressive (AR) model and Threshold Autoregressive (TAR) model has better forecasting performance. Thus, we can only use the Autoregressive (AR) model to obtain forecast values. However, we need to compare forecasting performance for each model.

2. You can apply this study by using more advance model likes Vector Smooth Transition Autoregressive (Vector STAR) model to find the non-linear behavior and compare ability to forecast with another models.

3. You can focus on the other market such as bond market to find the non-linear behavior and forecasting performance for each model.

# REFERENCES

1. Mungmanee, M. (2009). “COINTEGRATION AND CAUSALITY BETWEEN THAI STOCK MARKET AND MACROECONOMIC VARIABLES”, Journal of Economic Payap University, Vol.6.

2. Khunacharoensuk, C. (2006). “LONG-TERM EQUILIBRIUM BETWEEN MACROECONOMIC VARIABLES AND SECURITY MAKET INDICES”, Ramkhamhaeng University, 10.14457/RU.the.2006.289.

3. น้ำริน ผลไสว (2007). “Volatility analysis of the rate of returns of price index in the stock exchange of Thailand, Singapore, Malaysia, Indonesia and Philippines”, Chiang Mai University, 10.14457/CMU.the.2007.243.

4. Brahmasrene, T. and Jiranyakul, K. (2007). “COINTEGRATION AND CAUSALITY BETWEEN STOCK INDEX AND MACROECONOMIC VARIABLES IN AN EMERGING MARKET”, Academy of Accounting and Financial Studies Journal, Vol.11, No.3.

5. Tanjitprom, N. (2012). “Macroeconomic Factors of Emerging Stock Market: The Evidence from Thailand”, International Journal of Financial Research, Vol.3, No.2, doi:10.5430/ijfr.v3n2p105. http://dx.doi.org/10.5430/ijfr.v3n2p105.

6. Engle, R. et al. (2013). “Stock Market Volatility and Macroeconomic Fundamentals”, The Review of Economics and Statistics, Vol.95, No.3, Pages 776-797, doi:10.1162/REST_a_00300.

7. Emenike Kalu O. and Odili Okwuchukwu (2014). “Stock Market Return Volatility and Macroeconomic Variables in Nigeria”, International Journal of Empirical Finance, Vol.2, No.2, 75-82.

8. Pooja Joshi and A.K. Giri (2015). “Cointegration and Causality between Macroeconomic variables and Stock Prices: Empirical Analysis from Indian Economy”, Business and Economic Research, ISSN 2162-4860, Vol.5, No.2.

9. Robert D. Gay, Jr (2008). “Effect of Macroeconomic Variables on Stock Market Returns for Four Emerging Economies: Brazil, Russia, India, and China”, International Business & Economics Research Journal, Vol. 7, No.3.

10. Hatipoglu et al. (2014). “The Impact of Fundamental Shocks on Stock Prices: Evidence from Turkey”, Research in Applied Economics, ISSN 1948-5433, Vol.6, No.1, doi:10.5296/rae.v6i1.4494. http://dx.doi.org/10.5296/rae.v6i1.4494

11. Humpe A. and Macmillan P. (2014). “Non-linear predictability of stock market returns: comparative evidence from Japan and the US”, Investment Management and Financial Innovations, Vol. 11, Issues 4.

12. Mohamed El Hedi AROURI and Fredj JAWADI (2010). “Nonlinear Stock Market Integration in Emerging Countries”, International Journal of Economics and Finance, Vol. 2, No. 5.

13. Ismail Ali Siad (2001). “Testing for Nonlinear Dynamic in The Stock Exchange of Thailand (SET)”, ABAC Journal, Vol. 21, No.1.

14. Tanjitprom, N. (2012). “The Review of Macroeconomic Factors and Stock Returns”, International Business Research, ISSN 1913-9004, E-ISSN 1913-9012, Vol.5, No.8.

15. Kavkler A. et al. (2008). “NONLINEAR ECONOMETRIC MODELS: THE SMOOTH TRANSITION REGRESSION APPROACH”, CERGE-EI Foundation under a program of the Global Development Network.

16. Saechung A. (2006). “Time varying excess returns on Thai Government Bond: A STAR Model” , Thammasat Economic Journal, HG5750.55.ก4 อ473

# BIOGRAPHY

| | |
|---|---|
| Name | Mr. Sarun Panmanotham |
| Date of Birth | July 4, 1991 |
| Educational Attainment | 2012: Bachelor degree of Economics |
| | Thammasat University |