

## CHAPTER 2

### Theory and Literature Review

#### 2.1 Theory

##### 2.1.1 Asset Return<sup>1</sup>

Campbell, Lo, and MacKinlay (1997) give reason for using returns that for average investors, return of an asset is a complete and scale-free summary of the investment opportunity.

Let  $P_t$  be the price of an asset at time index  $t$ . Assume for the moment that the asset pays no dividends.

One-Period Simple Return that holding the asset for one period from date  $t-1$  to date  $t$  would result in a simple gross return

$$1 + R_t = \frac{P_t}{P_{t-1}} \quad \text{or} \quad P_t = P_{t-1}(1 + R_t) \quad (1)$$

The corresponding one-period simple net return or simple return is

The arithmetic return is:

$$R_t = \frac{P_t}{P_{t-1}} - 1 = \frac{P_t - P_{t-1}}{P_{t-1}} \quad (2)$$

The logarithmic return is:

$$R_t = \ln\left(\frac{P_t}{P_{t-1}}\right) \quad (3)$$

##### 2.1.2 Time Series

Time series are data or observations which have been changing along times.

There might be either stationary or non-stationary changing for time series. If time series is able to explain or analyze the changes in the past, then those time series can be used as a tool to predict or estimate the future.

---

<sup>1</sup> Ruey S. Tsay. (1951). *Analysis of Financial Time Series*: 2<sup>nd</sup> edition, (University of Chicago: Graduate school of Business, 1951), p.41-42

“In regressing a time series variable on another time series variables, one often obtains a very high  $R^2$  (in excess of 0.9) even through there is no meaningful relationship between the two variables. Sometimes we expect no relationship between two variables, yet a regression of one on the other variable often shows a significant relationship”. (Gujarati, 2004 p.794), that means, the time series used should be stationary. There are ways to test whether time series are stationary or non-stationary. One of them is the Box-Jenkins method of time series analysis (Autocorrelation Coefficient Function: ACF). Another is Dickey-Fuller’s called unit root as would be study further.

**Stationary Stochastic Process:** “A stochastic process is said to be stationary if its mean and variance are constant over time and the value of the covariance between the two time periods depends only on the distance or gap or lag between the two time periods and not the actual time at which the covariance is computed”. (Gujarati, 2004 p.797)

Let  $X_t$  be stochastic time series with these properties:

$$\text{Mean: } E(X_t) = \mu$$

$$\text{Variance: } \text{var}(X_t) = E(X_t - \mu)^2 = \sigma^2$$

$$\text{Covariance: } \text{COV}(X_t, X_{t+k}) = E[(X_t - \mu)(X_{t+k} - \mu)] = \sigma_k - \mu$$

(Charemza and Deadman, 1992, pp.118)

### 2.1.3 Unit Roots Test

**Dickey-Fuller Test, Augmented Dickey-Fuller Test (1979).** Unit Root Test is test of stationary using an autoregressive model. The model follows as:

$$X_t = \rho X_{t-1} + \varepsilon_t \quad (4)$$

where  $X_t, X_{t-1}$  = Variable time series at time t and t-1

$\rho$  = Autocorrelation coefficient

$\varepsilon_t$  = Random error

If  $\rho = 1$  that is a unit root means that the model would be non-stationary in this case.

The hypotheses as:

$$H_0 : \rho = 1$$

$$H_a : |\rho| < 1 ; -1 < \rho < 1$$

If  $H_0$  is accepted, that means  $X_t$  has unit root and non-stationary.

From  $X_t = \rho X_{t-1} + \varepsilon_t$  can be manipulated by subtracting  $X_{t-1}$  on both sides:

$$\begin{aligned} X_t - X_{t-1} &= \rho X_{t-1} - X_{t-1} + \varepsilon_t \\ &= (\rho - 1)X_{t-1} + \varepsilon_t \end{aligned}$$

can be written as:

$$\Delta X_t = \delta X_{t-1} + \varepsilon_t \quad (5)$$

where  $\delta = \rho - 1$ . If  $\delta = 0$  then  $\rho = 1$  and  $H_0$  is accepted. In another way, if  $\delta = 1$  then  $\rho = 0$  and  $H_a$  is accepted that they are stationary. Then the hypotheses can be written as:

$$H_0 : \delta = 0$$

$$H_a : \delta < 0$$

If  $\delta = 0$  or  $\rho = 1$ , from equation (4);  $\Delta X_t = \delta X_{t-1} + \varepsilon_t$  we get:

$$\Delta X_t = \varepsilon_t \quad (6)$$

This is a version of Dickey-Fuller Test. There are three main versions of the test as:

1. Test for a random walk:  $\Delta X_t = \delta X_{t-1} + \varepsilon_t$  (7)

2. Test for a random walk with drift:  $\Delta X_t = \alpha_0 + \delta X_{t-1} + \varepsilon_t$  (8)

3. Test for a random walk with drift and deterministic time trend:

$$\Delta X_t = \alpha_0 + \alpha_1 t + \delta X_{t-1} + \varepsilon_t \quad (9)$$

where  $\alpha_0, \alpha_1, \delta$  = Parameters

$t$  = Trend

As mentioned, the one used in this study is an Augmented Dickey-Fuller Test (ADF-test) which a version for larger and more complicated set of time series model. This test is conducted by “augmenting” the preceding three equations by adding the lagged values of the dependent variable  $\Delta X_t$ .

There are three Augmented Dickey-Fuller Tests (ADF-test) are:

1. Test for a random walk:

$$\Delta X_t = \delta X_{t-1} + \sum_{i=1}^m \beta_i \Delta X_{t-i} + \varepsilon_t \quad (10)$$

2. Test for a random walk with drift:

$$\Delta X_t = \alpha_0 + \delta X_{t-1} + \sum_{i=1}^m \beta_i \Delta X_{t-i} + \varepsilon_t \quad (11)$$

3. Test for a random walk with drift and deterministic time trend:

$$\Delta X_t = \alpha_0 + \alpha_1 t + \delta X_{t-1} + \sum_{i=1}^m \beta_i \Delta X_{t-i} + \varepsilon_t \quad (12)$$

For the equations (10), (11) and (12) which terms added are  $\sum_{i=1}^m \beta_i \Delta X_{t-i}$ , they are lagged difference terms. It is test for Augmented Dickey-Fuller Test (ADF-Test) that has developed from Dickey-Fuller Test for solve problem of serial correlation. ADF-Test has asymptotic distribution same as DF-Test then it can be used the same critical values.

#### 2.1.4 Phillips Perron Test (PP Test)<sup>2</sup>

Phillips and Perron (1988) propose nonparametric method for controlling of serial correlation when test a unit root. The PP method estimates the non-augmented DF test equation  $\Delta y_t = \alpha y_{t-1} + x_t' \delta + \varepsilon_t$ , and modifies the t-ratio of the  $\alpha$  coefficient

<sup>2</sup> Available: [http://www.er.uqam.ca/nobel/r20564/cours/unit\\_root\\_%20Ev4\\_1.pdf](http://www.er.uqam.ca/nobel/r20564/cours/unit_root_%20Ev4_1.pdf).

so that serial correlation does not affect the asymptotic distribution of the test statistic.

The PP test is based on the statistic:

$$\tilde{t}_\alpha = t_\alpha \left( \frac{\gamma_0}{f_0} \right)^{\frac{1}{2}} - \frac{T(f_0 - \gamma_0)(se(\hat{\alpha}))}{2f_0^{1/2}s} \quad (13)$$

where  $\hat{\alpha}$  is the estimate, and  $t_\alpha$  the  $\alpha$ -ratio,  $se(\hat{\alpha})$  is coefficient standard error, and  $s$  is the standard error of the test regression. In addition,  $\gamma_0$  is a consistent estimate of the error variance in  $\Delta y_t = \alpha y_{t-1} + x_t' \delta + \varepsilon_t$  (calculated as  $(T-k)s^2/T$ , where  $k$  is the number of regressors). The remaining term,  $f_0$ , is an estimator of the residual spectrum at frequency zero.

There are two choices you will have to make when performing the PP test. First, you must choose whether to include a constant, a constant and a linear time trend, or neither, in the test regression. Second, you will have to choose a method for estimating  $f_0$ .

### 2.1.5 Long Memory Test<sup>3</sup>

The presence of long memory may be defined in terms of the observed autocorrelations, which show high dependence between very distant observations. The autocorrelations of a long memory process is consistent with an essentially stationary process but takes far longer to decay than the exponential rate associated with the stationary ARMA class of processes. Defining the autocorrelation between observation at time  $t$  and observation at time  $t-j$  as  $\rho_j$ , long memory processes are characterized by the following property:

$$\lim_{t \rightarrow \infty} \sum_{j=-t}^t |\rho_j| = \infty \quad (14)$$

Fractionally integrated processes are long memory processes given the property in (14). It is possible to define the fractionally integrated process  $y_t$  as follows,

<sup>3</sup> Celso Brunetti. (1999). *Long Memory, The Taylor Effect and Intraday Volatility in Commodity Futures Markets*. Department of Business Studies, The University of Edinburgh.

$$(1-L)^d y_t = u_t \quad (15)$$

where  $d$  is not integer and represents the fractional order of integration and  $L$  is the lag operator. For values of  $d$  less than 0.5 and positive, the process  $y_t$  is long memory, and its autocorrelations are all positive and exhibit a hyperbolic rate of decay. For  $-0.5 < d < 0$  the process has short memory according to (14). It is possible to note from (15) that fractionally integrated processes are intermediate between I(0) and I(1) processes. Following Granger and Joyeux (1980) and Hosking (1981), rewrite (15) as follows,

$$(1-L)^d (y_t - \mu) = \varepsilon_t \quad (16)$$

where  $\mu$  is the mean of the process  $y_t$ ,  $E(\varepsilon_t) = 0$ ,  $E(\varepsilon_t^2) = \sigma^2$  and  $E(\varepsilon_t \varepsilon_s) = 0$  for  $s \neq t$ . Equation (16) defines a fractional white noise process. The fractional difference operator  $(1-L)^d$  is defined as follows,

$$(1-L)^d = \sum_{j=0}^{\infty} \frac{\Gamma(j-d)L^j}{\Gamma(j+1)\Gamma(-d)} \quad (17)$$

where  $\Gamma(\cdot)$  = the standard gamma function.

The autocorrelation function of (16) at lag  $j$  is equal to

$$\rho_j = \frac{\Gamma(j+d)\Gamma(1-d)}{\Gamma(j-d)\Gamma(d)} \quad (18)$$

The asymptotic approximation of Equation (18) is given by

$$\rho_j \approx c j^{2d-1}$$

where  $c = \frac{\Gamma(1-d)}{\Gamma(d)}$

Hence the autocorrelation coefficients exhibit slow hyperbolic decay for large  $j$ .



### The methods of Long memory test.

#### 1. R/S Analysis <sup>4</sup>

R/S test was used by Hurst (1951) to test for long memory in the pattern of flooding by the river Nile. It was modified by Mandelbrot and Wallis (1968).

The range  $R_n$  is defined as:

$$R_n = \left\{ \max_{1 \leq i \leq n} \sum_{t=1}^i (x_t - \bar{x}) - \min_{1 \leq i \leq n} \sum_{t=1}^i (x_t - \bar{x}) \right\} \quad \text{when } t = 1, 2, \dots, T \quad (19)$$

where  $\bar{x}$  is the sample mean given by:

$$\bar{x} = \frac{1}{T} \sum_{i=1}^n x_i \quad (20)$$

where  $s_T$  is the sample standard deviation given by:

$$s_T = \left[ \frac{1}{T} \sum_{t=1}^k (x_t - \bar{x})^2 \right]^{0.5} \quad (21)$$

The R/S statistics that the range rescaled ( $R$ ) divided by the sample standard deviation ( $S$ ) where  $C$  is constant and  $H$  is Hurst exponent follows as:

$$(R/S)_s = Ct^H \quad (22)$$

The Hurst exponent ( $H$ ) follows as:

$$\ln(R/S)_s = \ln(C) + H \ln(s). \quad \text{when } s = t_1, t_2, \dots, T \quad (23)$$

The value of Hurst exponent ( $H$ ) shows the different features that if  $H$  value is 0.5 means the Standard Brown Motion. If  $H$  value is ( $0 < H < 0.5$ ) means Fractal Brown Motion. If  $H$  value is ( $0.5 < H < 1$ ) means the process has long term memory and non-period cycle.

<sup>4</sup> Svetlana Danilenko. (2009). *Long-Term Memory effect in stock prices analysis*. Economics and Management, 14. ISSN. 1822-6515.

## 2. Modified R/S Test <sup>5</sup>

Lo (1991) modified the classical R/S statistic which was proposed by Hurst (1951). Lo was showed that its statistical behavior is invariant over a general class of short memory processes, but deviates for long memory ones.

As a result, Lo modified the classical R/S method, mainly according to

$$s_T = \left[ \frac{1}{T} \sum_{t=1}^k (x_t - \bar{x})^2 \right]^{0.5} \text{ and the modified function is } (R/S)_s = \frac{1}{A} \sum_{n=1}^A \frac{R_n}{\sigma_n(q)}$$

where

$$R_n = \left\{ \max_{1 \leq i \leq n} \sum_{t=1}^i (x_t - \bar{x}) - \min_{1 \leq i \leq n} \sum_{t=1}^i (x_t - \bar{x}) \right\} \text{ when } t = 1, 2, \dots, T$$

The modified R/S statistic is defined as:

$$\tilde{Q}_n = \frac{1}{\hat{\sigma}_n(q)} \left[ \max_{1 \leq i \leq n} \sum_{t=1}^i (x_t - \bar{x}) - \min_{1 \leq i \leq n} \sum_{t=1}^i (x_t - \bar{x}) \right] \quad (24)$$

where

$$\hat{\sigma}_n^2(q) = \frac{1}{T} \sum_{t=1}^k (x_t - \bar{x})^2 + \frac{2}{T} \sum_{t=1}^k (\omega_j(q) \left[ \sum_{i=t+1}^T (x_i - \bar{x})(x_{i-t} - \bar{x}) \right]) \quad (25)$$

$$= \sigma_n^2 + 2 \sum_{t=1}^k (\omega_j(q) \gamma_j) \quad (26)$$

where  $\omega_j(q) = 1 - \frac{i}{q+1}$  when  $q < T$ ,  $\sigma_n^2$  and  $\gamma_j$  are the sample variance and autocovariance of  $x$ .

## 3. GPH Test <sup>6</sup>

Geweke and Porter-Hudak (1983) proposed to use information from the periodogram to estimate the spectral density at very low frequencies.

The spectral regression follows as:

$$\ln \left[ \hat{I}(\lambda_j) \right] = c + d \ln \left[ 2 \sin \left( \frac{\lambda_j}{2} \right) \right]^2 + e_j \quad \text{when } j = 1, 2, \dots, m \quad (27)$$

<sup>5</sup> Huang Fei-xue, Jin Jian-dong and Li Yan-xi (2010). *Comparable Analysis of Long-Term Memory of EUR/USD Based on Non-parametrical Statistics*. Management science and engineering, Vol. 4, No. 3, pp.117-127.

<sup>6</sup> Koong, C.S., Tsui, A.K.C., and Chan, W.S. *On tests for Long Memory in Pacific Basin Stock Return*. National University of Singapore. pp. 89-94.



where  $\hat{I}(\lambda_j) = \frac{1}{2\pi} \left[ \sum_{i=-(T-1)}^{T-1} \hat{\gamma}_i \cos(i\lambda_j) \right]$  is the periodogram at harmonic frequency.

$\lambda_j = \frac{2\pi j}{T}$  are the Fourier frequencies for  $j = 1, 2, \dots, m$  with  $m = \sqrt{T}$  and  $\hat{\gamma}_i$  is the usual lag- $i$  sample autocovariance. The null hypothesis of no memory, the slope of the regression  $d$  equal zero and the usual t-statistic can be employed to perform the test.

### 2.1.6 AR Model <sup>7</sup>

The Autoregressive Model (AR model) is the simplest form of ARIMA model which is the similar to a linear regression model. The model can be follows as

$$Y_t = C + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \dots + \alpha_p Y_{t-p} + d_t \quad (28)$$

where  $C$  is the constant level,  $Y_{t-1}, Y_{t-2}, Y_{t-p}$  are part series values (lags),  $\alpha_r$  are coefficients to be estimated,  $d_t$  is random variable (assumed to be independent and represent random shocks). If  $Y_{t-p}$  is the further lag with a nonzero coefficient then AR model is to be order  $p$  (AR ( $p$ )).

### 2.1.7 MA Model <sup>7</sup>

The Moving Average Model (MA Model) is an extreme autoregressive model which is the current observations depends on all past observations. The model can be follows as

$$Y_t = -\phi_1 Y_{t-1} - \phi^2 Y_{t-2} - \phi^3 Y_{t-3} - \dots + d_t \quad (\text{if } \alpha_i = -\phi^i) \quad (29)$$

assume that  $C$  is zero without loss of generality

$$d_t = Y_t + \phi_1 Y_{t-1} + \phi^2 Y_{t-2} + \phi^3 Y_{t-3} + \dots \quad (30)$$

Multiply the expression for  $d_{t-1}$  by  $\phi$

$$\phi d_{t-1} = \phi_1 Y_{t-1} + \phi^2 Y_{t-2} + \phi^3 Y_{t-3} + \dots \quad (31)$$

so by subtraction

$$d_t - \phi d_{t-1} = Y_t \quad (32)$$

<sup>7</sup> Box-Jenkins Time Series Analysis. pp. 469-471. Available: <http://www.statistical-solutions-software.com/BMDP-documents/BMDP-2T.pdf>

This equation shows that  $Y_t$  is a linear function of all the past lags is equivalent to  $Y_t$  as only a few past shocks of linear function. The MA model of order  $q$  or MA( $q$ ) can be written as

$$Y_t = d_t - \phi_1 d_{t-1} - \phi_2 d_{t-2} - \phi_q Y_{t-q} \quad (33)$$

### 2.1.8 ARMA Model <sup>7</sup>

The ARMA model combined from the AR and MA models for stationary series for both past values and past shocks which the model can called ARMA ( $p, q$ ) model with  $p$  order AR terms and  $q$  order MA terms. The model can be follows as

$$Y_t = C + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \alpha_p Y_{t-p} + d_t - \phi_1 d_{t-1} - \phi_2 d_{t-2} - \phi_q Y_{t-q} \quad (34)$$

where  $\alpha_p$  is coefficient of  $Y_{t-p}$

### 2.1.9 ARIMA Model <sup>7</sup>

The autoregressive integrated moving-average (ARIMA) model is reference by Box and Jenkins (1976). The order of an ARIMA model is usually denoted by the notation ARIMA ( $p, d, q$ ), where  $p$  is the order of the autoregressive part,  $d$  is the order of the differencing and  $q$  is the order of the moving-average process.

The ARIMA ( $p, d, q$ ) has the general multiplicative seasonal form as:

$$\phi(\beta)\varphi(\beta^s)(1-\beta^s)(1-\beta)^d Z_t = \theta(\beta)\rho(\beta^s)d_t \quad (35)$$

where  $\phi(\beta) = (1 - \phi_1\beta - \dots - \phi_p\beta^p)$  is non-seasonal autoregressive operator

$\varphi(\beta^s) = (1 - \phi_1\beta^s - \dots - \phi_p\beta^{ps})$  is the seasonal autoregressive operator of order  $p$

$\rho(\beta^s) = (1 + \rho_1\beta^s + \dots + \rho_q\beta^{qs})$  is the seasonal moving average operator of order  $q$

$\theta(\beta) = (1 - \theta_1\beta - \dots - \theta_q\beta^q)$  is non-seasonal moving average operator

$(1 - \beta^s)(1 - \beta)^d$  is seasonal differencing and non-seasonal differencing of

order  $d$

$\beta$  is the backshift operator that  $\beta(Y_t) = Y_{t-1}$ .

### 2.1.10 ARFIMA Model <sup>8</sup>

The autoregressive fractionally integrated moving-average (ARFIMA) model was proposed by Granger and Joyeux (1980) and Hosking (1981) proposed the processes are capable long term persistence of model.

The model can be written the general form as:

$$\varphi_p(\beta)(1-\beta)^d Y_t = \omega + \theta_q(\beta)\varepsilon_t \quad (36)$$

Where  $d$  is the fractional difference parameter,  $\beta$  is the backshift operator that  $\beta(Y_t) = Y_{t-1}$ ,  $(1-\beta)^d$  is differencing operator of order  $d$ ,  $\omega$  is constant term,  $\varphi_p(\beta)$  is the autoregressive operator of order  $p$ ,  $\theta_q(\beta)$  is the moving average operator of order  $q$ ,  $\varepsilon_t$  is error term. When  $0 < d \leq 0.5$  means that the process is long memory. For  $-0.5 < d < 0$  the process is called intermediate memory. The process exhibits nonstationary if  $0.5 < d < 1$ . A series exhibits a stationary and invertible ARMA process with geometrically bounded autocorrelations if  $-0.5 < d < 0.5$ .

### 2.1.11 ARFIMAX Model<sup>9</sup>

Ebens (1999) proposed the ARFIMAX model and estimated the realized volatilities of Dow Jones Industrial Average (DJIA) portfolio using this model. His original model is as follow,

$$(1-L)^d (1-\phi(L_p)) \ln(h_t^2) = \omega_0 + \omega_1 r_{t-1} I^- + \omega_2 r_{t-1} I^+ + (1+\theta(L_p))\varepsilon_t \quad (37)$$

where  $\varepsilon_t \sim i.i.d.N(0, \sigma_\varepsilon^2)$ ,  $L$  is the back shift operator,  $\phi(L_q) = \sum_{i=1}^q \phi_i L^i$  and

$\theta(L_q) = \sum_{i=1}^q \theta_i L^i$ . Realized volatilities are denoted by  $h_t^2$ , the indicator  $I^- (I^+)$  takes value of one when return  $r_{t-1} < 0 (r_{t-1} \geq 0)$  and is zero otherwise. This model was

<sup>8</sup>Olanrewaju. I. Shittu and Olaoluwa Simon Yaya. (2009) *Measuring Forecast Performance of ARMA and ARFIMA Models: An Application to US Dollar/UK Pound Foreign Exchange Rate*. Department of Statistics, University of Ibadan Nigeria. European Journal of Scientific Research, ISSN 1450-216X Vol.32 No.2 (2009), pp.167-176

<sup>9</sup>Jia Geng. (2009) *Two Essays on Financial Econometrics*. The Graduate School of Clemson University. pp. 20-23.

generated based on the classical ARMA (p, q) model where the ARMA coefficients are  $\omega_0$ ,  $\phi(L_p)$  and  $\theta(L_p)$ . The new items in the model are a fractional integration parameter ( $d$ ) to capture the slow hyperbolic decay in the sample autocorrelation function; lagged negative ( $\omega_1$ ) and positive ( $\omega_2$ ) returns to capture the leverage effect in the distribution of  $\ln(h_t^2)$ .

Jia Geng (2009) modified ARFIMAX model which is given below,

$$(1-L)^d(1-\phi(L_p))(\ln h_t^2 - K - k_1 r_{t-1} I^- - k_2 r_{t-1} I^+) = (1+\theta(L_p))\varepsilon_t \quad (38)$$

Compare a model with the original ARFIMAX model. The general form of modified model can be written as:

$$(1-L)^d(1-\phi(L_p))(y - XB) = (1+\theta(L_p))\varepsilon_t \quad (39)$$

Ebens (1999) used conditional sum-of-squares maximum likelihood (SSML) estimator (advocated by Hosking 1984) to estimate the coefficients of the model. Jia Geng (2009) use modified profile likelihood method (MPL) to estimate the model. Bloomfield (1993), and Doornik and Ooms (1999) proved, based on Monte Carlo simulation, that MPL will eliminate the negative bias commonly found in SSML. The likelihood function follows as:

$$\log L(d, \phi, \theta, \beta, \sigma_\varepsilon^2) = -\frac{T}{2} \log(2\pi) - \frac{1}{2} \log |\Sigma| - \frac{1}{2} \log z' \Sigma^{-1} z \quad (40)$$

where  $z = y - XB$ ,  $\Sigma$  is the auto covariance matrix of  $y = (y_1, \dots, y_T)'$ . The auto

correlation matrix  $R = \sum \frac{1}{\sigma_\varepsilon^2}$ , then rewrite (40) into

$$\log L(d, \phi, \theta, \beta) = -\frac{T}{2} \log(2\pi) - \frac{1}{2} \log |R| - \frac{T}{2} \log \sigma_\varepsilon^2 - \frac{1}{2\sigma_\varepsilon^2} (z' R^{-1} z) \quad (41)$$

Take the derivative of (41) with respect to  $\sigma_\varepsilon^2$  and let it equal to zero, then

$$\log L(d, \phi, \theta, \beta) = -\frac{T}{2} \log(2\pi) - \frac{1}{2} \log |R| - \frac{T}{2} - \frac{T}{2} \log(T^{-1} z' R^{-1} z) \quad (42)$$

Take the derivative with respect to  $\beta$  and get,

$$\log L(d, \phi, \theta) = -\frac{T}{2} \log(2\pi) - \frac{1}{2} \log |R| - \frac{T}{2} - \frac{T}{2} \log(T^{-1} \hat{z}' R^{-1} \hat{z}) \quad (43)$$

Then modified profile likelihood for ARFIMAX ( $p, d, q$ ) as follows equation

$$\log L(d, \phi, \theta) = -\frac{T}{2} (1 + \log(2\pi)) - \left(\frac{1}{2} - \frac{1}{T}\right) \log |R| - \frac{T-k-2}{2} \log(T^{-1} \hat{z}' R^{-1} \hat{z}) - \frac{1}{2} \log |X' R X| \quad (44)$$

where  $k$  is the degree of freedom

Cheung and Diebold (1994) found that most of the errors in fractional-integrated estimation are from the mean. If the sample is not very large, we can use the average of the sample to replace the mean in the likelihood function and get a better estimation. Following this approach:

$$(1-L)^d (1-\phi(L_p)) (\ln h_t^2 - \hat{\mu} - k_1 r_{t-1} I^+) = (1+\theta(L_p)) \varepsilon_t \quad (45)$$

where  $\hat{\mu}$  is the average of  $\ln h_t^2$

Ebens (1999) only estimated the ARFIMAX model without autoregression term, or FIMAX model. Jia Geng (2009) estimate model using the likelihood function. ARFIMAX (1,d,1,X) is the full model where “1” is the first order autoregression term, “d” is the fractional integration parameter, “1” is the first order moving average term, and “X” means there are exogenous variables in the model.

### 2.1.12 ARCH Model <sup>10</sup>

The autoregressive conditional heteroscedastic model were introduced by Engle (1982). The model designed to capture the volatility of financial returns.

The model can be written as:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_1 \varepsilon_{t-2}^2 + \dots + \alpha_i \varepsilon_{t-p}^2 \quad (46)$$

or 
$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 \quad \text{where } \alpha_0 > 0 \text{ and } \alpha_i \geq 0, i > 0 \quad (47)$$

<sup>10</sup> The free encyclopedia. (2011), ARCH Model Available:  
[http://en.wikipedia.org/wiki/Autoregressive\\_conditional\\_heteroskedasticity](http://en.wikipedia.org/wiki/Autoregressive_conditional_heteroskedasticity)

### 2.1.13 ARFIMAX-GARCH Model <sup>11</sup>

The univariate GARCH specification allows to define dynamics for the conditional mean from the general ARFIMAX model with the addition of ARCH-in-mean effects introduced in Engle (1987). The ARFIMAX-ARCH-in-mean specification may be formally defined as,

$$\phi(L)(1-L)^d (y_t - \mu_t) = \theta(L)\varepsilon_t \quad \text{when } \varepsilon_t \sim i.i.d.N(0, \sigma_\varepsilon^2) \quad (48)$$

with the left hand side denoting the Fractional AR specification on the demeaned data and the left hand side the MA specification on the residuals. ( $L$ ) is the lag operator,  $(1-L)^d$  is the long memory fractional process with  $0 < d < 1$ , and equivalent to the Hurst Exponent  $H - 0.5$ , and  $\mu_t$  defined as,

$$\mu_t = \mu + \sum_{i=1}^m \delta_i x_{i,t} + \xi \sigma_t^k \quad (49)$$

where we allow for  $m$  external regressors  $x$  and ARCH-in-mean on either the conditional standard deviation,  $k = 1$  or conditional variance  $k = 2$ .

### The standard GARCH model <sup>11</sup>

The standard GARCH model (Bollerslev (1986)) can be written as:

$$\sigma_t^2 = \left( \omega + \sum_{j=1}^m \zeta_j \nu_{jt} \right) + \sum_{j=1}^p \alpha_j \varepsilon_{t-j}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2 \quad (50)$$

with  $\sigma_t^2$  denoting the conditional variance,  $\omega$  the intercept and  $\varepsilon_t^2$  the residuals from the mean filtration process discussed previously. The GARCH order is defined by  $(p, q)$ , with possibly  $m$  external regressors  $\nu_j$  which are passed pre-lagged. If variance targeting is used, then  $\omega$  is replaced by,

$$\bar{\sigma}^2 (1 - \hat{P}) \quad (51)$$

<sup>11</sup> Alexios Ghalanos. (2011) *A package for flexible GARCH modelling in R*. Introduction to the rugarch package. pp.3-5. Available: [http://cran.r-project.org/web/packages/rugarch/vignettes/Introduction\\_to\\_the\\_rugarch\\_package.pdf](http://cran.r-project.org/web/packages/rugarch/vignettes/Introduction_to_the_rugarch_package.pdf)



where  $\bar{\sigma}^2$  is the unconditional variance of  $\varepsilon^2$  which is consistently estimated by its sample counterpart at every iteration of the solver following the mean equation filtration, and  $\bar{v}_j$  represents the sample mean of the  $j^{\text{th}}$  external regressors in the variance equation (assuming its stationarity), and  $\hat{P}$  is the persistence and defined below. One of the key features of the observed behavior of financial data which GARCH models capture is volatility clustering which may be quantified in the persistence parameter  $\hat{P}$ .

For the GARCH model this may be calculated as,

$$\hat{P} = \sum_{j=1}^p \alpha_j + \sum_{j=1}^q \beta_j \quad (52)$$

Related to this measure is the 'half-life' (call it  $h2l$ ) defined as the number of days it takes for half of the expected reversion back towards  $E(\sigma^2)$  to occur,

$$h2l = \frac{-\log_e 2}{\log_e \hat{P}} \quad (53)$$

Finally, the unconditional variance of the model  $\hat{\sigma}^2$ , and related to its persistence, is,

$$\hat{\sigma}^2 = \frac{\hat{\omega}}{1 - \hat{P}} \quad (54)$$

where  $\hat{\omega}$  is the estimated value of the intercept from the GARCH model.

#### 2.1.14 Information Criteria: Akaike Information Criteria (AIC), Bayesian Information Criterion (BIC) and Hannan-Quinn Information Criterion (HQC)<sup>12</sup>

The Akaike (1974, 1976) and Schwarz (1978) information criteria for selecting the most parsimonious correct model are respectively.

<sup>12</sup> Herman J. Bierens. (2006). *Information Criteria and Model Selection*. Pennsylvania State University. Available: <http://econ.la.psu.edu/~hbierens/INFORMATIONCRIT.PDF>.

$$\text{Akaike: } c_n(k) = \frac{-2 \cdot \ln(L_n(k))}{n} + \frac{2k}{n} \quad (55)$$

$$\text{Hannan-Quinn: } c_n(k) = \frac{-2 \cdot \ln(L_n(k))}{n} + \frac{2k \cdot \ln(\ln(n))}{n} \quad (56)$$

$$\text{Schwarz: } c_n(k) = \frac{-2 \cdot \ln(L_n(k))}{n} + \frac{k \cdot \ln(n)}{n} \quad (57)$$

where  $L_n(k)$  = the maximized value of the likelihood function for the estimated model.

$k$  = number of parameters used.

$n$  = the sample size.

Since the Schwarz information criterion is derived using Bayesian arguments, this criterion is also known as the Bayesian Information Criterion (BIC).

These criteria take the general form as:

$$c_n(k) = \frac{-2 \cdot \ln(L_n(k))}{n} + \frac{k \cdot \varphi(n)}{n} \quad (58)$$

where  $\varphi(n)=2$  in the Akaike case and  $\varphi(n) = \ln(n)$  in the Schwarz case. Using these criteria, the model is selected that corresponds to  $\hat{k} = \arg \min_{k \leq m} c_n(k)$

### 2.1.15 The Mean Absolute Error (MAE)<sup>13</sup>

The Mean Absolute Error (MAE) is a quantity used to measure how close forecasts or predictions are to the eventual outcomes. In form of the mean absolute error (MAE) follows as:

$$MAE = \frac{1}{n} \sum_{i=1}^n |f_i - y_i| = \frac{1}{n} \sum_{i=1}^n |e_i| \quad (59)$$

<sup>13</sup>The free encyclopedia. (2011). *The Mean Absolute Error (MAE)*. Available: [http://en.wikipedia.org/wiki/Mean\\_absolute\\_error](http://en.wikipedia.org/wiki/Mean_absolute_error).

As the name suggests, the mean absolute error is an average of the absolute errors  $e_i = f_i - y_i$ , where  $f_i$  is the prediction and  $y_i$  is the true value. The alternative formulations may include relative frequencies as weight factors.

### 2.1.16 The Mean Absolute Percentage Error (MAPE) <sup>14</sup>

The Mean Absolute Error (MAE) is measure of accuracy in a fitted time series value in statistics. It usually has been expressed accuracy by a percentage and the formula of MAPE follows as:

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{A_t - F_t}{A_t} \right| \quad (60)$$

where  $A_t$  = the actual value  
 $F_t$  = the forecast value

Rule of the mean absolute percentage error (MAPE) can interpretation follow as:

- If the MAPE value is less than 10%, it is highly accurate forecast.
- If the MAPE value is between 10% to 20%, it is good forecast.
- If the MAPE value is between 20% to 50%, it is reasonable forecast.
- If the MAPE value is greater than 50%, it is inaccurate forecast.

### 2.1.17 The Root Mean Squared Error (RMSE) <sup>15</sup>

The Root Mean Squared Error (RMSE) is a frequently used measure of the differences between values predicted by a model and the values actually observed. The formula of RMSE follows as:

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (\hat{y}_t - y_t)^2}{n}} \quad (61)$$

where  $y_t$  = the actual value  
 $\hat{y}_t$  = the forecast value

<sup>14</sup>The free encyclopedia. (2011). *The Mean Absolute Percentage Error (MAPE)*. Available: [http://en.wikipedia.org/wiki/Mean\\_absolute\\_percentage\\_error](http://en.wikipedia.org/wiki/Mean_absolute_percentage_error).

<sup>15</sup>The free encyclopedia. (2011). *The Root Mean Squared Error (RMSE)*. Available: [http://en.wikipedia.org/wiki/Root-mean-square\\_deviation#Normalized\\_root-mean-square\\_deviation](http://en.wikipedia.org/wiki/Root-mean-square_deviation#Normalized_root-mean-square_deviation).

## 2.2 Literature review

### Literature review about gold

**Narisa Samutsakhon (2004)** was studied how to forecast the gold price. The data used the selling price of gold bars and ornamental gold based on 120 monthly from 1994 to 2003. The model is ARIMA model using Box-Jenkins follow as 4 step that identification, parameter estimation, diagnostics checking and forecasting.

The result of the study found that the AR(2), MA(2), MA(5) have lowest value of root mean square error and Theil's inequality coefficient so it can be the best model to forecast the price of gold bars and ornamental gold. For the value of root mean square error of the AR(2), MA(2), MA(5) is 0.020343 and Theil's inequality coefficient of the AR(2), MA(2), MA(5) is 0.001139. For monthly prices of gold bars from January to April 2004 using the AR(2), MA(2), MA(5) model be 7,692.72, 7,715.80, 7,755.11 and 7,761.17 baht per Thai gold weight unit. For monthly prices of ornamental gold at the same period using the AR(2), MA(2), MA(5) model be 7,871.89, 7,893.76, 7,915.98 and 7,917.87 baht per Thai gold weight unit.

**Nuchara Gaysornpratoom (2007)**, this study forecasts the gold price using Neural Networks Model compare with ARIMA and GARCH – M Models. The data used the daily gold bars from January 2, 2004 to March 28, 2006 totally 550 observations.

The results of the study found that the best model for forecasting the gold price by using ARIMA that be ARIMA (6, 2, 0) which means absolute percentage error is 0.66. For the best GARCH-M that be GARCH-M (1, 3) which means absolute percentage error is 0.76. For forecasting the gold price on the Neural Networks model which yielded the lowest mean absolute percentage error of 2.05. This study concludes that the model for forecasting the gold price upon the value of mean absolute percentage error so ARIMA model is the best model because it used the error term to explain variable.

**Kanokwan Kampiew (2010)**, the aims of this study is to examine the relationship between world's gold spot price and the U.S. exchange rate using cointegration. The data used daily data of world's gold spot price, the U.S. exchange

rate against the Euro and the U.S. exchange rate against Yen from 4<sup>th</sup> January 1999 to 11<sup>th</sup> May 2010 totally 2,937 observations. The model was used cointegration method.

The results of this study for ADF-test found that the world's gold spot price, the U.S. exchange rate against the Euro and the U.S. exchange rate against Yen has unit root with an I(1) process. For the cointegration test and the error correction mechanism that world's gold spot price, the U.S. exchange rate against the Euro and the U.S. exchange rate against Yen has relationship in the short run and the long run. For the result of Granger Causality test that the U.S. exchange rate against the Euro has a relationship in the form of bi-directional but in case of the U.S. exchange rate against Yen has a relationship in form of one directional.

#### **Literature review about Stock**

**Pierre Giot and S'ebastien Laurent (2001)**, this study is how to compute a daily VaR measure for two stock indexes. The data used the French CAC40 stock index for the period from 1995 to 1999, totally 1249 daily observations and the SP500 futures contracts on the Chicago Mercantile Exchange for the period from January 1989 to December 2000, totally 3241 daily observations using the one-day-ahead forecast of the daily realized volatility. The model used the skewed Student APARCH model for daily data and ARFIMAX model.

The result of the study found that all the P-value of the CAC40 data are larger than 0.05 for both for long and short VaR. For all the P-value of the SP500 data are larger than 0.05 except short VaR at level  $\alpha = 1\%$  and  $\alpha = 0.25\%$ . However, VaR based on ARFIMAX (0,d,1)-skewed student model for daily realized volatility provides adequate one-day-ahead VaR forecast.

**Adnan Kasman and Erdost Torun (2007)**, this paper examines the Long Memory in the Turkish stock market return and volatility using the ARFIMA-FIGARCH model. The data are daily returns of Istanbul Stock Exchange (ISE) for the period 1988 to 2007.

The results show that the ARFIMA-FIGARCH model has long memory in return series. The long memory in the Turkish stock returns implies that the stock prices consist of the shocks in the past. However, the evidence of long memory in



volatility shows that risk is an important determinant of the behavior for the stock data in the Turkish stock market.

**Hussein Ali Al-Zeaud (2011)**, the aim of this study was to predict the volatility for banks sector using Autoregressive Integrated Moving Average (ARIMA) models. The data is the Amman Stock Exchange (ASE) in the period from 1/1/2005 to 1/4/2010.

The results of the study shows that the financial time series on some period had high fluctuations of the banks sector because the affect from the decision makers in the investment portfolios, asset allocation decisions, risk management conditions, selling or buying in the financial markets, information available to the investors. In this test the best model from ARIMA test was chosen based on the value of MSE that the best model for banks sector is ARIMA (2, 0, 2) because this model gives the minimum mean square error which is 0.0001003.

**Pattararat Puapanpattana (2011)**, the aims of this study was to forecast the rate of return on the Stock Exchange of Thailand using the ARFIMAX method. The data was used from the closing prices of five stocks such as PTT, PTTEP, SCC, KBANK and CRALL based on daily data and covering the period from June 1, 2009 to May 30, 2010 totally 225 days. This study used the unit root test such as ADF test and PP test, the result shows that the data has stationary at  $I(0)$ . For Long Memory Test by using R/S test, modified R/S test and GPH test, the result shows that the rate of return on stock has long memory.

The results of the study found that the best model for forecasting the rate of return on the PTT stock is ARFIMAX (0, -0.150750, 0, 3.64425, 5.68629, 5.51736), for the rate of return on the PTTEP stock is ARFIMAX (0, -0.165075, 0, 2.74490, 2.16771, 1.32123), for the rate of return on the SCC stock is ARFIMAX (1,-0.230285, 0, 3.53497, 4.58694, 1.07939), for the rate of return on the KBANK stock is ARFIMAX (0, -0.0270916, 0, 5.11903, -7.24994, -1.55422) and the rate of return on the CPALL stock is ARFIMAX (0, -0.0207286, 0, -0.0106983, 0.391607, 0.000135854). However, the MAPE value for the rate of return from the data has accurate result.



### **Literature review about exchange rate**

**Abdol S. Soofi, Shouyang Wang and Yuqin Zhang (2006)**, this paper tested long memory in the Asian Foreign Exchange Rates. The data used was 12 Asian daily dollar exchange rates series such as Chinese Renminbi (CNY), Hong Kong Dollar (HKD), Indonesia Rupiah (IDR), Indian Rupees (INR), Japanese Yen (JPY), South Korean Won (KRW), Malaysian Ringgit (MYR), Philippines Pesos (PHP), Pakistani Rupees (PKR), Singapore Dollar (SGD), Thai Baht (THB) and Taiwan Dollar (TWD). The method for this research based on long memory method such as rescaled range analysis (R/S analysis), the GPH analysis and the Whittle method.

The results of this study found that the first difference of log of the remminbi per dollar exchange rate series at the level has unit root and follow a random walk process. For the log of the return in Korean won is  $A=0.20$  and Singapore dollar is  $A=0.20, 0.25$  that the confidence interval contains zero. For the log of the return in Malaysian ringgit is  $A=0.25, 0.30$  that the confidence interval does not contain zero. However, Japanese yen and Malaysian ringgit found that the evidence of long memory and the return of the other series cannot reject the hypothesis because they are generated from short memory.

**Huang Fei-Xue, Jin Jian-dong and Li Yan-Xi (2010)**, this study was a comparable analysis of Long-Term Memory of EUR/USD based on non-parametrical statistics. The data used daily exchange closing price series of EUR to USD and covers from 4<sup>th</sup> January 1999 to 29<sup>th</sup> November 2008. This research based on long memory methods such as classical R/S method, modified R/S method and V/S method and takes the method of non-parametrical statistics.

The results of this study found that the histogram test on distribution of returns between EUR and USD shows the skewness is not 0 and kurtosis is large than 3. For Jarque-Bera (JB) test is the null hypothesis of normal distribution has rejected because the value of JB test is larger than threshold value at 1% and 5%. For the result of the classical R/S test that H value is 0.6124, the recycle length is 160 days and the correlation scale is 1.3432. For the result of the modified R/S test and classical R/S test that V has 2% bias statistics. This study concludes that the daily return rate of EUR/USD has long memory.

### Literature review about ARFIMAX model

**Chin Wen Cheong (2007)**, this study proposed to combine a few long memory models in a financial market volatility model using daily, range and high frequency data. The data used the Kuala Lumpur stock exchange (KLSE) index for the period from 1<sup>st</sup> January 2003 to 15<sup>th</sup> January 2006 totally 745 and 266710 observations. The model used in combination with ARFIMAX, HAR and ARCH-type generalized models.

The result of the study found that the HARFIMAX-GARCH has the highest log-likelihood value that it is -104.78 and the AIC value is 0.3491 to compare with others. The BIC value of the HARFIMAX-GARCH is 0.4299 that it is suffered greater in comparison with the others. For AR-GARCH and ARFIMA-FIGARCH shows similar results in log-likelihood value. In the ARCH-type model shows the smallest results that MSE is 0.0072, ME is -0.0017 and MAE is 0.0658. However, the ARCH-type model cannot be used to predict future volatility.

**Stavros Degiannakis (2008)**, this study was applied in estimating and forecasting the intra-day realized volatility using ARFIMAX and ARFIMAX-TARCH model. The data is CAC40 and DAX30 for the period from July 1995 to December 2003.

The results of the study found that the in-sample evaluation for the ARFIMAX (1,  $d$ , 1)-TARCH (1, 1) model was estimated CAC40 is 1.411473 and DAX30 is 1.291441 that achieves the minimize value of Schwarz's Bayesian Criterion (SBC). The Mean Square error (MSE) for CAC40 is 28.032 and DAX30 is 7.329 that the ARFIMAX (1,  $d$ , 1)-TARCH (1, 1) model has lowest value. For out-of-sample evaluation, the ARFIMAX (1,  $d$ , 0) -TARCH (1, 1) model in case of DAX30 is superior to the ARFIMAX (2,  $d$ , 2) model. In case of CAC40, the ARFIMAX (1,  $d$ , 2) -TARCH (2, 1) model does not achieve the lowest value in the predicted mean squared error loss function. However, case of the in-sample evaluation that the ARFIMAX-TARCH is superior to the ARFIMAX and for DAX30, in case of the out-of-sample evaluation that the ARFIMAX-TARCH is superior to the ARFIMAX.

**Toshiaki Watanabe, Masato Ubukata (2009)**, this research analyzed daily realized volatility and ARCH Type model. The data was chosen that the Nikkei 225 stock index options and the period from May 29, 1996 to September 27, 2007. The model has been used to estimate that the ARFIMA(X) model, ARCH model and find the performance of option pricing that used the MAE (Mean Absolute Error), RMSE (root mean square error), MAPE (Mean Absolute Percentage Error) and RMSPE (Root Mean Square Percentage Error).

The results of the study found that GARCH or EGARCH model perform best in DITM (deep-in-the-money). For ARFIMAX model performs best for the root mean square percentage error and the mean absolute percentage error in DOTM (deep-out-of-the-money). For OTM (out-of-the-money) that the ARFIMAX model performs best for the root mean square error. For ATM (at-the-money) and ITM (in-the-money) that the ARFIMAX model performs best no matter which loss function is used. However, the ARFIMAX model performs best with daily realized volatility.