Chapter 2

Theoretical Foundations

In this chapter, the multivariate volatility GARCH model is presented in Dynamic conditional and smooth transition correlation approach that has to test about time vary correlation before estimation. The fractional properties of variables are tested and are estimated by Semi Parametric ARFIMA model. The time vary CAPM are presented by State Space Estimator, Quantile regression and Bayesian estimator.

Time-Varying Correlation Models

The mean and volatility equations, with the following two subsections describing the Dynamic conditional and smooth transition correlation models are discussed in this section

The two-dimensional vector of oil price and each Asia stock market (y_t) has mean equation

$$y_t = E[y_t / \Im_{t-1}] + r_t, t = 1,2,...,T \tag{1}$$

where $\Im_{t-1}$ is all information available at time t-1, together with values of exogenous variables for time t. Since we are interested in the role oil prices in the evolution of y_t, $\Im_{t-1}$ together with lagged stock returns. Allowing oil prices changes affect stock returns for each country enables us to capture correlations. The mean equations in (1) are assumed linear. The conditional covariance follow

$$r_t | \Im_{t-1} \sim N(0, H_t) \tag{2}$$

where N denotes the bivariate normal distribution. From (2), each univariate error process can be written

$$r_{t,t} = h_{tt,t}^{1/2}\varepsilon_{t,t}, i = 1,2 \qquad (3)$$

Where $h_{tt,t} = E(r_{t,t}^2 / \Im_{t-1})$ and $\varepsilon_{t,t}$ is a sequence of independent random variables with mean zero and variance one. As common in empirical analyses, each conditional variance is assumed to follow a univariate GARCH (1,1) process

$$h_{tt,t} = \alpha_{i0} + \alpha_{i1} r_{i,t-1}^2 + \beta_{i1} h_{ii,t-1} \qquad (4)$$

with non-negativity and stationarity restrictions imposed.

Rather than modelling the off-diagonal elements of H directly, the definition

$$h_{12,t} = \rho_t (h_{11.t} h_{22,t})^{1/2} \qquad (5)$$

allows the focus to be placed on the time-varying correlations ρ_t. The Dynamic conditional and smooth transition models then differ in their definitions of ρ_t. The constant conditional correlation (CCC) model simply assumes that ρ_t is constant over time (McMaleer, 2005,Bauwens, L., S. Laurent and V.K. Rombouts, 2006).

Dynamic Conditional Correlation Model

Engle (2002) specifies the dynamic conditional correlation model through the GARCH(1,1)-type process

$$q_{i,j,t} = \bar{\rho}_{12}(1 - \alpha - \beta) + \alpha\varepsilon_{1,t-1}\varepsilon_{2,t-1} + \beta q_{i,j,t-1} \qquad (6)$$

Where $\bar{\rho}_{12}$ is the (assumed constant) unconditional correlation between $\varepsilon_{1,t}$ and $\varepsilon_{2,t}, \alpha$, is the news coefficient and β is the decay coefficient. In order to constrain the conditional correlation ρ_t to lie between -1 and +1, $q_{1,2,t}$ from (6) and the conditional correlation is obtained from

$$\rho_t = q_t/(q_{11}, q_{22,t})^{1/2} \quad (7)$$

The model is mean-reverting provided $\alpha + \beta < 1$, and when the sum is equal to 1 the conditional correlation process in (6) is integrated (Ling, S., and M. McAleer, 2003 a , Nektarios Aslanidis, 2007).

Smooth Transition Conditional Correlation Models

The smooth transition conditional correlation model considered by Silvennoinen and Terasvirta (2005) assumes the presence of two extreme states (or regimes) with state-specific constant correlations. These correlations are allowed to change smoothly between the two regimes as a function of an observable transition variable. The conditional correlation ρ_t follows

$$\rho_t = \rho_1(1 - G_t(S_t;\gamma,c)) + \rho_2 G_t(s_t;\gamma,c) \quad (8)$$

in which the transition function $G_t(s_t;\gamma,c)$ is assumed continuous and bounded by zero and unity, γ and c are parameters, whereas s_t is the transition variable. Since (8) implies $\rho_t = \rho_1$ when G = 0 and $\rho_t = \rho_2$ when G = 1, extreme values of the transition function identify the distinct correlations that apply in these regimes. A weighted mixture of these two correlations applies when $0 < G_t < 1$. A plausible and widely used specification for the transition function is the logistic function

$$(G_t(s_t;\gamma,c) = 1/1 + \exp[-\gamma(s_t - c)], \gamma > 0 \quad (9)$$

where the parameter c is the threshold between the two regimes. The slope parameter

$\gamma > 0$ determines the smoothness of the change in the value of the logistic function and thus the speed of the transition from one correlation state to the other. When $\gamma \to \infty, G_t(s_t;\gamma,c)$ becomes a step function $(G_t(s_t;\gamma,c) = 0$ if $s_t < c$ their transition variable can be deterministic or stochastic. $(G_t(s_t;\gamma,c) = 1 \cdot if \cdot s_t > c)$, and the transition between the two extreme correlation states becomes abrupt. In that case, the model approaches a threshold model in correlations. An important special case of the smooth transition conditional correlation model uses time as the transition, $s_t = t/T$, which gives rise to the time-varying conditional correlation (TVCC) model. The (smooth) change between correlation regimes, and as $\gamma \to \infty$ captures a structural break in the correlations (Bwo-Nung Huang, 2005 and Annastiina Silvennoinen, 2007). The Pooled AIC are used for selecting Smooth Transition Conditional Correlation Models (philip Hans Franses, 2004). The alternative AIC for 2-regime SETAR model as the sum of AICs for AR models in the two regimes, that is

$$AIC(p_1, p_2) = n_1 \ln \hat{\sigma}_1^2 + n_2 \ln \hat{\sigma}_2^2 + 2(p_1 + 1) + 2(p_2 + 1)$$

where n_j ,j=1,2, is the number of observations in the j_{th} regime,and $\hat{\sigma}_{\grave{e}}^2$, j=1,2,,is the variance of the residuals in the j_{th} regime. The BIC for a SETAR model can be defined analogously as

$$BIC(p_1, p_2) = n_1 \ln \hat{\sigma}_1^2 + n_2 \ln \hat{\sigma}_2^2 + (p_1 + 1)\ln n_1 + (p_2 + 1)\ln n_2$$

For given upper bounds p_1^* and p_2^*, respectively, the selected lag orders in the two regimes are those for which the information criterion is minimized.

The SETAR model assumes that the threshold variable q_t is chosen to be a lagged value of the time series itself. The model is assumed in both regimes, a 2-regime SETAR model is given by

$$\rho_t = \begin{cases} \varphi_{0,1} + \varphi_{1,1}\rho_{t-1} + \varepsilon_t & If\ \rho_{t-1} \le c. \\ \varphi_{0,2} + \varphi_{1,2}\rho_{t-1} + \varepsilon_t & If\ \rho_{t-1} > c. \end{cases}$$

An alternative way to write the SETAR model is

$$\rho_t = \left(\varphi_{0,1} + \varphi_{1,1}\rho_{t-1}\right)\left(1 - I\left[\rho_{t-1} > c\right]\right) + \left(\varphi_{0,2} + \varphi_{1,2}\rho_{t-1}\right) I\left[\rho_{t-1} > c\right] + \varepsilon_t$$

where I [A] is an indicator function with I [A]=1 if the event A occurs and I [A] = 0 otherwise.

The SETAR model assume that the border between the two regime is given by a specific value of the threshold variable In particular, in the 2-regime SETAR model, $y_{Đ}$ will be estimated within the y_{t-1} (Philip Hans Franses, 2004 and Zivot., 2006).

Testing for constant correlations in a multivariate GARCH model

The constant-correlation hypothesis in a multivariate GARCH model is detected by Lagrange Multiplier (LM) (Tse.Y.K., 2000). The constant-correlation model set the conditional variances of y_{it} follow a GARCH process, while the correlations are constant. Denoting $\Gamma=\{\rho_{ij}\}$ as the correlation matrix, we have

$$\sigma_{it}^2 = \omega_i + \alpha_i \sigma_{i,t-1}^2 + \beta_i y_{i,t-1}^2, i = 1,..., K \qquad (10)$$

$$\sigma_{ijt} = \rho_{ij}\sigma_{it}\sigma_{jt}, 1 \le i < j \le K \qquad (11)$$

The assumption ω_i, α_i and β_i are nonnegative, $\alpha_i + \beta_i < 1$, for i=1,2,K and

Γ is positive definite. The LM test can then be applied to test for the restrictions. This approach only requires estimates under the constant-correlation model, and can thus conveniently exploit the computational simplicity of the model.

The equations allow time-varying correlations

$$\rho_{ijt} = \rho_{ij} + \delta_{ij} y_{i,t-1} y_{j,t-1}, \qquad (12)$$

The conditional covariances are given by

$$\sigma_{ijt} = \rho_{ijt}\sigma_{it}\sigma_{jt} \qquad (13)$$

Note that there are N=K^2+2K parameters in the extended model with time-varying correlations. The constant-correlation hypothesis can be tested by examining the hypothesis H_0: $\delta_{ij} = 0, \quad 1 \le i < j \le K$ Under H_0, there are M=K(K-1)/2 independent restrictions. The optimal properties under the null H_0 is the LM test .The model which is deified standardised residual as $\varepsilon_{it} = y_{it} / \sigma_{it}$ might be written as

$$\rho_{ijt} = \rho_{ij} + \delta'_{ij}\varepsilon_{i,t-1}\varepsilon_{j,t-1} \tag{14}$$

As ε_{it} depends on other parameters of the model through σ_{it} , analytic derivation of the LM statistic is intractable. The LM statistic of H_0 under the above framework which denote D_t as the diagonal matrix with diagonal elements given by σ_{it} , and $\Gamma=\{\rho_{ijt}\}$ as the time-varying correlation matrix. Hence the conditional-variance matrix of y_t is given by $\Omega_t = D_t \Gamma_t D_t$. Under the normality assumption the conditional log-likelihood of the observation at time t is given by (the constant term is ignored)

$$\ell_t = -\frac{1}{2}\ln|D_t\Gamma_t D_t| - \frac{1}{2}y'_t D_t^{-1}\Gamma_t^{-1}D_t^{-1}y_t$$

$$= -\frac{1}{2}\ln|\Gamma_t| - \frac{1}{2}\sum_{i=1}^{K}\ln\sigma_{it}^2 - \frac{1}{2}y'_t D_t^{-1}\Gamma_t^{-1}D_t^{-1}y_t,$$

for i=1,…,K and the log-likelihood function l is given by $l = \sum_{t=1}^{T} l_t$. For simplicity, the assumption of y_{i0} and σ_{i0}^2 are fixed and known. This assumption has no effects on the asymptotic distributions of the LM statistic. The derivatives of σ_{it}^2 with respect to ω_i, α_i and β_i for i =1... K

$$d_{it} = \partial\sigma_{it}^2 / \partial\omega_i, e_{it} = \partial\sigma_{it}^2 / \partial\alpha_i, f_{it} = \partial\sigma_{it}^2 / \partial\beta_i$$

$$\begin{aligned} d_{it} &= 1 + \alpha_i d_{i,t-1}, \\ e_{it} &= \sigma_{i,t-1}^2 + \alpha_i e_{i,t-1}, \\ f_{it} &= \alpha_i f_{i,t-1} + y_{i,t-1}^2, \end{aligned} \tag{15}$$

where the starting values are given by $d_{i1} = 1, e_{i1} = \sigma_{i0}^2$ and $f_{i1} = y_{i0}^2$ The first partial derivatives of l_t with respect to the model parameters are given by

$$\begin{aligned}
\frac{\partial \ell_t}{\partial \omega_i} &= \frac{\left(\varepsilon_{it}^* \varepsilon_{it} - 1\right) d_{it}}{2\sigma_{it}^2}, \\
\frac{\partial \ell_t}{\partial \alpha_i} &= \frac{\left(\varepsilon_{it}^* \varepsilon_{it} - 1\right) e_{it}}{2\sigma_{it}^2}, \\
\frac{\partial \ell_t}{\partial \beta_i} &= \frac{\left(\varepsilon_{it}^* \varepsilon_{it} - 1\right) f_{it}}{2\sigma_{it}^2}, \\
\frac{\partial \ell_t}{\partial \delta_{ij}} &= \left(\varepsilon_{it}^* \varepsilon_{jt}^* - \rho_t^{ij}\right) y_{i,t-1} y_{j,t-1},
\end{aligned} \tag{16}$$

Where $\varepsilon_t^* = (\varepsilon_{1t}^*, ..., \varepsilon_{Kt}^*)' = \Gamma_t^{-1}\varepsilon_t$ and $\Gamma_t^{-1} = \{\rho_t^{ij}\}$. thus, if we denote the parameters of the model as these analytic derivatives can facilitate the evaluation of the MLE of the $\theta = \left(\omega_1.\alpha_t, \beta_1, \omega_2, ..., \beta_K, \rho_{12}, \rho_{13}, ..., \rho_{K-1,K}, \delta_{12}, ..., \delta_{K-1,K}\right)'$, extended model if desired. Note that on $H_0 : \Gamma_t = \Gamma$ for all t, so that $\varepsilon_t^* = \Gamma^{-1}\varepsilon_t$ and $\rho_t^{ij} = \rho^{ij}$. In this case, ε_t are just the standardized residuals calculated from the algorithm suggested by We shall denote $\hat{\theta}$ as the MLE of θ under H_0

The N-element score vector given by $s = \partial l / \partial \theta$ and V as the N*N information matrix given by $V = E\left(-\partial^2 l / \partial\theta\partial\theta'\right)$, where E(.) denotes the expectation operator, the LM statistic for H_0 is given by $\hat{s}'\hat{V}^{-1}\hat{s}$, where the hats denote evaluation at $\hat{\theta}$. V may be replaced by the (negative of the) Hessian matrix or the sum of the cross products of the first derivatives of l_t. The S is denoted the T*N $\frac{\partial l_t}{\partial \theta'}$ matrix the rows of which are the partial derivatives for t=1,…T . the LM statistic for H_0 can be calculated using the following formula

$$LMC = \hat{s}'\left(\hat{S}'\hat{S}\right)^{-1}\hat{s} \tag{17}$$

$$= l'\hat{S}\left(\hat{S}'\hat{S}\right)^{-1}\hat{S}'l, \tag{18}$$

where l is the T*1 column vector of ones and $\hat{S}$ is S evaluated $\hat{\theta}$. Under the usual regularity conditions LMC is asymptotically distributed as $\chi 2_M$. Eq. (18) shows that LMC can be interpreted as T times R^2, where R^2 is the uncentered coefficient of determination of the regression of l on $\hat{S}$. It is well-known that other forms of the LM statistic are available. For example, further simplification can be obtained by making use of the fact that in $\hat{S}' l$ the elements corresponding to the unrestricted parameters is zero. Eq. (18) is a convenient form.

$$\sigma_{it}^2 = \omega_i + \sum_{k=1}^{P} \alpha_{ih}\sigma_{i,t-1}^2 + \sum_{k=1}^{q} \beta_{ik} y_{i,t-k}^2, i = 1,..., K$$
$$d_{it} = 1 + \sum_{h}^{P} \alpha_{ih} d_{i,t-h}$$
$$e_{iht} = \sigma_{i,t-h}^2 + \sum_{h'=1}^{P} \alpha_{ih'} e_{ih,t-h'},$$
$$f_{ikt} = \sum_{h=1}^{P} \alpha_{ih} f_{ik,t-h} + y_{i,t-k}^2$$

The first partial derivatives of l_t with respect to ω_i, α_{ih} and β_{ik} (p+q+1 derivatives altogether) can be calculated using (16), with e_{iht} and f_{ikt} replacing e_{it} and f_{it}, respectively.

Long Memory Time Series

A stationary process y_t has long memory, or long range dependence, if its autocorrelation function behaves like

$$\rho(k) \rightarrow C_\rho k^{-\alpha} \ as \ k \rightarrow \infty \tag{19}$$

where C_ρ is a positive constant, and α is a real number between 0 and 1. Thus the autocorrelation function of a long memory process decays slowly at a hyperbolic rate. In fact, it decays so slowly that the autocorrelations are not summable:

$$\sum_{k=-\infty}^{\infty} \rho(k) = \infty$$

For a stationary process, the autocorrelation function contains the same information as its spectral density. In particular, the spectral density is defined as:

$$f(\omega) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \rho(k) e^{ik\omega}$$

Where ω is the Fourier frequency (Halmilton, 1994). From (1) it can be shown that

$$f\left(\omega\right) \rightarrow C_f \omega^{\alpha-1} \text{ as } \omega \rightarrow 0 \quad (20)$$

Where C_f is a positive constant. So for a long memory process, its spectral density tends to infinity at zero frequency. Instead of using α , in practice use

$$\text{H}=1-\alpha/2 \in (0.5,1), \quad (21)$$

Which is known as the Hurst coefficient (Hurst, 1951) to measure the long memory in y_t. The larger H is the longer memory the stationary process has.

Based on the scalling property in (19) and the frequency domain property in (20), Hosking (Hosking, 1981) independently showed that a long memory process y_t can also be modeled parametrically by extending an integrated process to a fractionally integrated process. In particular, allow for fractional integration in a time series y_t as follow:

$$(1-L)^d (y_t - \mu) = u_t \quad (22)$$

where L denotes the lag operater , d is the fractional integration or fractional difference parameter , μ_t is a stationary short-memory disturbance with zero mean.

The time series is highly persistent or appears to be non-stationary, let d = 1 and difference the time series once to achieve stationarity. However, for some highly persistent economic and financial time series, it appear that an integer difference may be too much, which is indicated by the fact that spectral density vanishes at the zero frequency for the differenced time series. To allow for long memory and avoid taking an integer of y_t, allow d to be fractional. The fractional difference filter is defined as follows, for any real d > -1:

$$(1-L)^d = \sum_{k=0}^{\infty} \binom{d}{k} (-1)^k L^k \quad (23)$$

With binomial coefficients: $\binom{d}{k} = \frac{d!}{k!(d-k)!} = \frac{\Gamma(d+1)}{\Gamma(k+1)\Gamma(d-k+1)}$

Notice that the fractional difference filter can be equivalent treated as an infinite order autoregressive filter. It can be show that when $|d| > 1/2$, y_t is stationary and has short memory , and is sometimes refer to as anti-persistent.

When a fractionally integrated series y_t has long memory, it can also be shown that

$$d = H - 1/2 \quad (24)$$

and thus d and H can be used interchangeably as the measure of long memory. Hosking(1981) showed that the scaling property in (19) and the frequency domain property in (20) are satisfied when $0 < d < 1/2$.

ARFIMA models

The traditional approach to modeling an I(0) time series y_t is to use the ARIMA model:

$$\varphi(B)(1-B)^d \{y_t - \mu\} = \theta(B) \in_t \quad (25)$$

Where $\varphi(B)$ and $\theta(B)$ are lag polynomials

$$\varphi(B) = 1 - \sum_{i=1}^{p} \phi_i B^i$$

$$\theta(B) = 1 - \sum_{j=1}^{q} \theta_j B^{ji}$$

With root out side the unit circle, and $\in_t$ is assumed to be an i.i.d normal random variable .this is usually referred to as the ARMA (p,d,q) model. By allow d to be the real number instead of a positive integer, the ARIMA model becomes the Autoregressive fractionally integrated moving average (ARFIMA) model. The stationary FARIMA model is -1/2<d<1/2,(Sowel, 1992). The ARFIMA or FARIMA was extended by Beran (Beran j. , 1995).

$$\varphi(B)(1-B)^{\delta}\left\{(1-B)^{m} y_t - \mu\right\} = \theta(B)\in_t \qquad (26)$$

where δ , -1/2< δ<1/2 and m is the number of times that y_t must be differenced to achieve stationarity. The difference parameter is given by d = δ+m. The restriction of m is either 0 or 1 , when m=0, μ is the expectation of y_t; in contrast, when m=1, μ is the slop of the linear trend component in y_t.

SEMIFAR models

Many observed time series exhibit apparent trends. Forecasts will differ greatly, depending on how these trends are modelled. A trend may be deterministic, i.e. defined by a deterministic function and purely stochastic or mixture of both.

SEMIFAR models are define by (Beran J. A., 1999): A *Gaussian process* Y_i *is called* a semiparametric fractional autoregressive model (*or* SEMIFAR *model*) *or order p , if there exists a smallest integer* $m \in \{0,1\}$ such that

$$\varphi(B)(1-B)^{\delta}\left\{(1-B)^{m} Y_i - g(t_i)\right\} = \in_i \qquad (27)$$

where $\delta \in (-0.5, 0.5)$.

Estimation for SEMIFAR model Let

$\theta^{o} = \left(\sigma^{2}_{\in,o}, d^{o}, \phi^{o}_{1}, \ldots, \phi^{o}_{p}\right)^{t} = \left(\sigma^{2}_{\in,o}, \eta^{o}\right)^{t}$ be the true unknown parameter vector in (26) where $d^{0} = m^{0} + \delta^{0}, -1/2 < \delta^{0} < 1/2 \; and \; m^{0} \in \{0,1\}$. Combining maximum likelihood with kernel estimation , the following method for estimating θ^{0} and the trend function g is obtained in (Beran J. A., 1999): Let K be a symmetric polynomial kernel define by

$$K(x) = \sum_{l=0}^{r} \alpha_{l} x^{2l}, |x| \le 1, \; and \; K(x) = 0 \; if \; |x| > 1, r \in \{0,1,2,\ldots\}$$,

and $K(x) = 0 \; if \; |x| > 1, r \in \{0,1,2,\ldots\}$ and the coefficient α_{l} such that $\int_{-1}^{1} K(x)dx = 1$. Let $b_{n}(n \in N)$ be a sequence of positive bandwidths such that $b_{n} \to 0$ and $nb_{n} \to \infty$ and define $\hat{g}(t_{i}) = \hat{g}(t_{i}; m)$ by

$$\hat{g}\left(t_{i}; m\right) = \frac{1}{nb_{n}} \sum_{j=1}^{n} K\left(\frac{t_{i} - t_{j}}{b_{n}}\right) \tilde{Y}_{j} \tag{28}$$

where $\tilde{Y}_{j} = (1-B)^{m} Y_{j}$ (for $m = 1, set \; \tilde{Y}_{1} = 0$). Using equations (9) and (10), define approximate residuals

$$e_{i}(\eta) = \sum_{j=0}^{i-m-1} a_{j}(\eta)[c_{j}(\eta) Y_{i-j} - \hat{g}(t_{i-j}; m)], \tag{29}$$

With coefficient a_{j} and c_{j} obtained from (9), and denote by $r_{i}(\theta) = e_{i}(\eta)/\sqrt{\theta_{1}}$ the standardized residual as a function of a trial value $\theta = \left(\sigma^{2}_{\in}, m+\delta, \phi_{1}, \ldots, \phi_{p}\right)^{t}$. Then $\hat{\theta}$ is defined by maximizing the approximate log-likelihood

$$l(Y_{1}, \ldots, Y_{n}; \theta) = -\frac{n}{2} \log 2\pi - \frac{n}{2} \log \sigma^{2}_{\in} - \frac{1}{2} n^{-1} \sum_{i=m+2}^{n} r_{i}^{2} \tag{30}$$

with respect to θ and $\hat{g}(t_i)$ is set equal to $\hat{g}(t_i;\hat{m})$.

The asymptotic behavior of $\hat{g}$ and $\hat{\theta}$ is derived in Beran(1999). As $n \rightarrow \infty, \hat{g}$ converges in probability to g, the optimal mean squared error of $\hat{g}$ is proportional to $n^{(4\delta-2)/(5-2\delta)}$ and $\sqrt{n}(\hat{\theta}-\theta)$ converges in distribution to a zero mean normal vector with covariance matrix $V = 2D^{-1}$ where

$$D_{ij} = (2\pi)^{-1}[\int_{-\pi}^{\pi} \frac{\partial}{\partial\theta_i} \log f(x) \frac{\partial}{\partial\theta_j} \log f(x)dx] \Big| \theta = \theta_*^o \tag{31}$$

with $\theta_*^o = \left(\sigma_{\in,o}^2, \eta_*^o\right)^T = \left(\sigma_{\in,o}^2, \delta^o, \eta_2^o, \ldots, \eta_{p+1}^o\right)^T$. The same result hold if a consent model choice criterion is used for the estimation of the autoregressive order *p*. It should be emphasized , in particular , that here both , the integer differencing parameter $m^o = \left[d^o + 0.5\right]$ and the fractional differencing parameter $\delta^o = d^o - m^o$ are estimated from the data. Also, the same central limit theorem holds if the innovation $\in_i$ are not normal, and satisfy suitable moment conditions. Finally note that the asymptotic covariance matrix does not depend on m^o

R/S Statistic

The R/S statistic is the range of partial sum of deviation of a time series from its mean, rescaled by its standard deviation. Specifically, consider a time series series y_t for $t = 1, \ldots, T$. The R/S statistic is defined as:

$$Q_T = \frac{1}{s_T}\left[\max_{1\le k\le T}\sum_{j=1}^{k}(y_j - \bar{y}) - \min_{1\le k\le T}\sum_{j=1}^{k}(y_i - \bar{y})\right] \tag{32}$$

Where $\bar{y} = 1/T\Sigma_{i=1}^{T} y_i$ and $s_T = \sqrt{1/T\sum_{i=1}^{T}(y_i - \bar{y})^2}$. If y_t's are i.i.d. normal random variables, then

$$\frac{1}{\sqrt{T}} Q_T \Rightarrow V$$

where $\Rightarrow$ denote weak convergence and V is range of a Brownian bridge on the unit interval. Lo(1991) gives selected quantiles of V

Lo(1991) pointed out that the R/S statistic is not robust to short range dependence. In particular, if y_t is autocorrelated(has short memory) then the limiting distribution of $Q_T / \sqrt{T}$ is V scaled by the square root of the long run variance of y_t. To allow for short range dependence in y_t ,Lo(1991) modified the R/S statistic as follow:

$$\tilde{Q}_T = \frac{1}{\hat{\sigma}_T(q)} \left[\max_{1 \le k \le T} \sum_{j=1}^{k} (y_j - \bar{y}) - \min_{1 \le k \le T} \sum_{j=1}^{k} (y_j - \bar{y}) \right] \tag{33}$$

Where the sample standard deviation is replaced by the square root of the Newey-West estimate of the long run variance with bandwidth q^2. Lo(1991) showed that if there is short memory but no long memory in y_t, $\tilde{Q}_T$ also converges to V, the range of a Brownian bridge.

GPH Test

Based on the fractionally integated process representation of a long memory time series, (Geweke, 1983) proposed a semi-nonparametric approach to testing for long memory. In particular, the spectral density of the fractionally integrated process y_t is given by:

$$f(\omega) = \left[4 \sin^2 (\frac{\omega}{2}) \right]^{-d} f_u(\omega) \tag{34}$$

where ω is the Fourier frequency, and $f_u(\omega)$ is the spectral density corresponding to u_t. Note that the fractional difference parameter d can be estimated by the following regression:

$$\ln f(\omega_j) = \beta - d \ln[4\sin^2(\frac{\omega_j}{2})] + e_j \tag{35}$$

for $j = 1,2,...,n_f(T)$. Geweke and Porter – Hudak (Geweke, 1983) showed that using a *periodogram* estimate of $f(\omega_j)$, the least square estimate $\hat{d}$ using the above regression is normally integrated in large samples if $n_f(T) = T^{\alpha}$ with $0 < \alpha < 1$

$$\hat{d} \sim N(d, \frac{\pi^2}{6\sum_{j=1}^{n_f}(U_j - \bar{U})^2})$$

where

$$U_j = \ln[4\sin^2(\frac{\omega_j}{2})]$$

and $\bar{U}$ is the sample mean of U_j, $j = 1,...,n_f$. Under the null hypothesis of no long memory $(d = 0)$, the t-statistic

$$t_{d=0} = \hat{d}.(\frac{\pi^2}{6\sum_{j=1}^{n_f}(U_j - \bar{U})^2})^{-1/2}$$

(18)

It has a limiting standard normal distribution.

Model overview of Long Memory Time Series

The raw daily data, Thai and Asia stock index, N225 (Nikkei Stock Average 225) Tokyo Stock Exchange, KLSE (KLSE Composite Index), Malaysian stock market, TWSE (Taiwan's composite index) Taiwan Stock Exchange, SETI (SET Composite Index) the Stock Exchange of Thailand, SSEC (Shanghai Composite Index) Shanghai Stock Exchange, The BSESN (Bombay SE Sensitive Index) Bombay Stock Exchange, JKSE (Jakarta Composite) Indonesia Jakarta Composite, PSI (PSE Composite Index) Philippine Stock Exchange, KS11 (KOSPI Index) Korean Stock Exchange are collected from Reuters for the period November 10, 1998 to November 10, 2008.

A stationary process y_t is the set of log daily price Asia Indexes. Based on the scalling property in (19) and the frequency domain property in (20) showed that a long memory process y_t can also be modeled parametrically by extending an integrated process to a fractionally integrated process. The fractional integration in a time series y_t as follow:

$$(1-B)^d (y_t - \mu) = u_t$$

where B denotes the lag operater , d is the fractional integration or fractional difference parameter , μ_t is a stationary short-memory disturbance with zero mean.

SEMIFAR models are define by (Beran J. A., 1999) such that

$$\phi(B)(1-B)^{\delta}\left\{(1-B)^m Y_i - g(t_i)\right\} = \in_i \qquad (36)$$

The SEMIFAR model extended by δ, m which -1/2< δ<1/2 for any d > -1/2. The number of times is m that y_t must be differenced to achieve stationary (Beran j. , 1995). The difference parameter is given by d = δ+m. The restriction of m is either 0 or 1 , when m=0, μ is the expectation of y_t; in contrast, when m=1, μ is the slop of the linear trend component in y_t

To allow for a possible deterministic trend in a time series, in addition to a stochastic trend, long memory and short memory component. The SEMIFAR model is based on the following extension to the FARIMA (p,d,0) model. The constant term μ is replaced by g(i_t), a smooth trend function on [0.1], with i=t/T. Using BIC choose autoregressive order p which is proposed by (Beran J. A., 1999)

Financial and Econometric Model Base with Time Varying

In this section, we are going to give a brief summary about the time varying models, such us State Space CAMP, Bayesian CAMP and Quantile regression CAMP that economists often used. State Space modeling in macroeconomics and finance has become widespread over the last decade. Many dynamic time series models in economics and finance may be represented in State Space form, as the system of equation. The work of (Zellner and Chetty 1965) shows the optimal Bayesian portfolio problem by Bayes' rule, the posterior density $p\left(r|\theta\right)$ is proportional to the product of the sampling density (the likelihood function) and the prior density, $f\left(r|\theta\right)p\left(\theta\right)$. The Koenker and Bassett (1978) developed the median (quantile) regression estimator to minimize the symmetrically weighted sum of absolute errors (where the weight is equal to 0.5) to estimate the conditional median (quantile) function. Then we will present an integrated procedure to construct an appropriate model for the stock data.

State Space CAPM

Typically, the State Space models can be found in most books (cf. Durbin & Koopman (2001), and Chan (2002)). The State Space model equation can be compactly expressed as

$$\begin{pmatrix} \alpha_{t+1} \\ y_t \end{pmatrix} = \underset{(m\times N)\times 1}{\delta_t} + \underset{(m+N)\times m}{\Phi_t} \, . \underset{m\times 1}{\alpha_t} + \underset{(m\times N)\times 1}{\mu_t} \qquad (37)$$

where $\alpha_t \sim$ N(**a**, **P**), $\mu_t \sim$ WN (0,1)

and $\delta_t \begin{pmatrix} d_t \\ c_t \end{pmatrix}, \Phi_t = \begin{pmatrix} T_t \\ Z_t \end{pmatrix}, \mu_t = \begin{pmatrix} H_t \mu_t \\ G_t \varepsilon_t \end{pmatrix}, \Omega_t = \begin{pmatrix} H_t H_t{}' & 0 \\ 0 & G_t G_t{}' \end{pmatrix}$.

The initial value parameters are summarized in the (m + 1) × m matrix $\Sigma = \begin{pmatrix} P \\ a' \end{pmatrix}$.

The smoothed estimate of the response y_t and its variance are computed using

$$\hat{y}_t = c_t + Z_t \hat{\alpha}_t \qquad \mathrm{var}(\alpha_t | Y_n) = Z_t \ \mathrm{var}(\alpha_t | Y_n) Z_t' \tag{38}$$

The smoothed disturbance estimates are the estimates of the measurement equations innovations ε_t and transition equation innovations η_t based on all available information Y_n, and are denoted $\hat{\varepsilon}_t = E[\alpha_t | Y_n]$ (or $\varepsilon_{t|n}$) and $\hat{\eta}_t = E[\eta_t | Y_n]$ (or $\eta_{t|n}$), respectively. The computation of $\hat{\varepsilon}_t$ and $\hat{\eta}_t$ from the Kalman smoother algorithm is described in Durbin & Koopman (2001). These smoothed disturbance estimates can be useful for parameter estimation by maximum likelihood and for diagnostic checking. The vector of prediction errors v_t and prediction error variance matrices F_t are computed from the Kalman filtered recursions.

State Space representation of a time varying parameter regression model consider a Capital Asset Pricing Model (CAPM) with time varying intercept and slope

$$\begin{aligned} y_t &= \alpha_t + \beta_{M,t} x_{M,t} + v_t, && v_t \sim \mathrm{WN}, \\ \alpha_{t+1} &= \alpha_t + \xi_t, && \xi_t \sim \mathrm{WN}, \\ \beta_{M,t+1} &= \beta_{M,t} + \varsigma_t && \varsigma_t \sim \mathrm{WN}, \end{aligned} \tag{39}$$

where y_t denotes the return on an asset in excess of the risk free rate, and $x_{M,t}$ denotes the excess return on a market index. In this model, both the abnormal excess return α_t and asset risk $\beta_{M,t}$ are allowed to vary over time following a random walk specification. Let $\alpha_t = (\alpha_t, \beta_{M,t})'$, $x_t = (1_t, x_{M,t})'$, $H_t = diagccc(\sigma_{\xi_t}, \sigma_{\zeta})'$ and $G_t = \sigma_v$. Then the State Space form equation (37) of equation (39) is $\begin{pmatrix} \alpha_{t+1} \\ y_t \end{pmatrix} = \begin{pmatrix} I_2 \\ x_t' \end{pmatrix} \alpha_t + \begin{pmatrix} H_t \eta_t \\ G_t \varepsilon_t \end{pmatrix}$ and has parameters

$$\Phi_t = \begin{pmatrix} I_2 \\ x_t' \end{pmatrix}, \Omega = \begin{pmatrix} \sigma_\xi^2 & 0 & 0 \\ 0 & \sigma_\zeta^2 & 0 \\ 0 & 0 & \sigma_{v\xi}^2 \end{pmatrix} \tag{40}$$

Since α_t is I(1) the initial state vector α_1 requires an infinite variance so it is customary to set a = 0 and P = $k\mathrm{I}_2$ with $k \to \infty$. Using equation (23), the initial variance is specified with $p_* = 0$ and $p_\infty = \mathrm{I}_2$. Therefore, the initial state matrix $\Sigma = \begin{pmatrix} P \\ a' \end{pmatrix}$ for the time varying CAPM has the form $\Sigma = \begin{pmatrix} -1 & 0 \\ 0 & -1 \\ 0 & 0 \end{pmatrix}$. The State Space parameter matrix Φ_t in equation (40) has a time varying system element $Z_t = x_t'$. The specification of the State Space form for the time varying CAPM requires values for the variances σ_ξ^2 σ_ζ^2 and σ_v^2 as well as a data matrix X whose rows correspond with $Z_t = x_t' = (1, r_{M,t})$. The values Φ_t associated with x_t' in the third row are set to zero. In the general State Space model equation (37), it is possible that all of the system matrices δ_t, Φ_t and Ω_t have time varying elements.

The typical CAPM regression model is, $y_t = \alpha + \beta_M x_{M,t} + \xi_t$, ξ_t~WN. The y_t denotes the return on an asset in excess of the risk free rate, and $x_{M,t}$ is the excess return on a market index. The State Space representation is given by $\begin{pmatrix} \alpha_{t+1} \\ y_t \end{pmatrix} = \begin{pmatrix} I_k \\ x_t' \end{pmatrix} \alpha_t + \begin{pmatrix} 0 \\ \sigma_\xi \xi_t \end{pmatrix}$ with $\mathrm{x}_t = (1, x_{M,t})'$ and the state vector satisfies $\alpha_{t+1} = \alpha_t = \beta = (\alpha, \beta_M)'$. The State Space system matrices are $T_t = I_\kappa$, $Z_t = x_t'$, $G_t = \sigma_\xi$ and $H_t = 0$. Estimating the CAPM with time varying coefficients in equation (22) subject to random walk evolution are showed in data. Neumann (2002) surveys several estimation strategies for time varying parameter models and concludes that the State Space model with random walk specifications for the evolution of the time varying parameters generally performs very well. The log-likelihood is parameterized using $\varphi = \left(\ln(\sigma_\xi^2), \ln(\sigma_\varsigma^2), \ln(\sigma_v^2)\right)'$ so that $\sigma^2 = (\exp(\varphi_1), \exp(\varphi_2), \exp(\varphi_3))'$. The maximum likelihood estimates for φ which estimates of $\varphi = \left(\ln(\sigma_\xi^2), \ln(\sigma_\varsigma^2), \ln(\sigma_v^2)\right)'$. These methods estimated the standard deviations σ_ξ σ_ς and σ_v as well as estimated standard errors.

Bayesian CAPM

The predictive density function reflects estimation risk explicitly since it integrates over the posterior distribution, which summarizes the uncertainty about the model parameters, updated with the information contained in the observed data. The Bayes' rule, the posterior density $p\left(r|\theta\right)$ is proportional to the product of the sampling density (the likelihood function) and the prior density, $f\left(r|\theta\right)p\left(\theta\right)$.

The decision-making under uncertainty are represented portfolio choice problem. Let r_{T+1} denote the vector $\left(N\times 1\right)$ of next-period returns and W current wealth. The next-period wealth is $W_{T+1}=W\left(1+\omega' r_{T+1}\right)$ in the absence of a risk-free asset. The next-period wealth $W_{T+1}=W\left(1+r_f+\omega' r_{T+1}\right)$ is a risk-free asset with return r_f is present. Let ω denote the vector of asset allocations. The optimal portfolio decision consists of choosing ω that maximizes the expected utility of next-period's wealth, $\max_{\omega} E\left(U\left(W_{T+1}\right)\right)=\max_{\omega}\int U\left(W_{T+1}\right)p\left(r|\theta\right)dr,$ subject to feasibility constraints, where θ is the parameter vector of the return distribution and U is a utility function generally characterized by a quadratic or a negative exponential functional form. The distribution of returns is $p\left(r|\theta\right),$. The $\max_{\omega} E\left(U\left(W_{T+1}\right)\right)=\max_{\omega}\int U\left(W_{T+1}\right)p\left(r|\theta\right)dr,$ is conditional on the unknown parameter vector θ, which are set θ equal to its estimate $\widehat{\theta}\left(r\right)$ based on some estimator of the data r (often the maximum likelihood estimator). Then, the optimal allocation given by $\omega^{*}=\arg\max_{\omega} E\left(U\left(\omega' r\right)\middle|\theta=\widehat{\theta}\left(r\right)\right)$.

The return generating process for the stock's excess return is $r_i=\alpha+\beta' f_t+\varepsilon_t, t=1,\ldots,T,$. The f_t is denoted a (K x 1) vector of factor returns (returns to benchmark portfolios), and ε_t is a mean-zero disturbance term. Then, the slopes of the regression in $r_i=\alpha+\beta' f_t+\varepsilon_t$ are stock's sensitivities (betas). In a single factor model such as the CAPM, the benchmark portfolio is the market portfolio. The implications for portfolio selection of varying prior beliefs about a pricing model are expressed, the prior mean of α, α_0, is set equal to zero. It could have a non-zero value. The prior variance σ_{α} of α reflects the investor's degree of confidence in the prior mean a zero value of

σ_α represents dogmatic belief in the validity of the model; $\sigma_\alpha = \infty$ suggests complete lack of confidence in its pricing power.

Quantile Regression CAPM

The other conditional quantile functions are estimated by minimizing an asymmetrically weighted sum of absolute errors, where the weights are functions of the quantile of interest. Thus, Quantile regression is robust to the presence of outliers. Engle and Manganelli (1999) and Morillo (2000) used in financial applications. The general Quantile regression model, as described by Buchinsky (1998), is $y_i = x_i'\beta_\theta + \mu_{\theta i}$ or, alternatively, $\theta = \int_{-\infty}^{x_i'\beta_\theta} f_y\left(s \middle| x_i\right) ds$, where β_θ is an unknown k × 1 vector of regression parameters associated with the θ_{th} percentile x_i is a k × 1 vector of independent variables, y_i is the dependent variable of interest, and $\mu_{\theta i}$ is an unknown error term. The θ_{th} conditional quantile of y given x is $Quant_\theta\left(\mu_{\theta i} \middle| x_i\right) = x_i'\beta_\theta$.Its estimate is given by $x_i'\hat{\beta}_\theta$. As θ increases continuously, the conditional distribution of y given x is traced out. Although many of the empirical Quantile regression papers assume that the errors are independently and identically distributed (i.i.d.), the only necessary assumption concerning $\mu_{\theta i}$ is $Quant_\theta\left(\mu_{\theta i} \middle| x_i\right) = 0$, That is, the conditional θ_{th} quantile of the error term is equal to zero. Thus, the Quantile regression method involves allowing the marginal effects to change for firms at different points in the conditional distribution by estimating β_θ using several different values of θ, $\theta \in (0,1)$ It is in this way that Quantile regression allows for parameter heterogeneity across different types of assets. Thus, the Quantile regression estimator can be found as the solution to the following minimization problem: $\widehat{\beta}_\theta = \arg_\beta \min\left(\sum_{i:y_i \succ x_i'\beta} \theta\left|y_j - x_i'\beta\right| + \sum_{i:y_i \prec x_i'\beta} (1-\theta)\left|y_i - x_i'\beta\right|\right)$ By minimizing a weighted sum of the absolute errors, the weights are symmetric for the median regression case ($\theta = 0.5$) and asymmetric otherwise. The former implies that the method is computationally straightforward while the latter implies that $\sqrt{n}\left(\widehat{\beta}_\theta - \beta_\theta\right) \xrightarrow{d} N\left(0, \Omega_\theta\right)$, The CAPM presents $E_t\left(R_{i,t+1}\right) = \gamma_{1,t+1}\beta_{i,\tau}$ as τ < t + 1.

The beta-risk is determined over moving samples. The $\beta_{i,\tau}$ is the beta-risk obtained from a time series regression. $R_{i,\tau} = \alpha_i + \beta_{i,\tau} R_{m,\tau} + \mu_{i,\tau}$. The $R_{i,\tau}$ and $R_{m,\tau}$ are the excess return on the asset and the market portfolio, respectively.