An Analysis of Asynchronous Data

KIM CHATALL
NIKLAS JOHANSSON

Master of Science Thesis
Stockholm, Sweden 2013
An Analysis of Asynchronous Data

KIM CHATALL
NIKLAS JOHANSSON

Master’s Thesis in Mathematical Statistics (30 ECTS credits)
Master Programme in Mathematics (120 credits)
Supervisor at Handelsbanken was Fredrik Bohlin
Supervisor at KTH was Boualem Djehiche
Examiner was Boualem Djehiche

TRITA-MAT-E 2013:21
ISRN-KTH/MAT/E--13/21--SE

Royal Institute of Technology
School of Engineering Sciences
KTH SCI
SE-100 44 Stockholm, Sweden
URL: www.kth.se/sci
Abstract

Risk analysis and financial decision making requires true and appropriate estimates of correlations today and how they are expected to evolve in the future. If a portfolio consists of assets traded in markets with different trading hours, there could potentially occur an underestimation of the right correlation. This is due the asynchronous data - there exist an asynchronicity within the assets time series in the portfolio. The purpose of this paper is twofold. First, we suggest a modified synchronization model of Burns, Engle and Mezrich (1998) which replaces the first-order vector moving average with an first-order vector autoregressive process. Second, we study the time-varying dynamics along with forecasting the conditional variance-covariance and correlation through a DCC model. The performance of the DCC model is compared to the industrial standard RiskMetrics Exponentially Weighted Moving Averages (EWMA) model. The analysis shows that the covariance of the DCC model is slightly lower than of the RiskMetrics EWMA model. Our conclusion is that the DCC model is simple and powerful and therefore a promising tool. It provides good insight into how correlations are likely to evolve in the short-run time horizon.
Acknowledgements

We would like to express our deep gratitude to Professor Boualem Djehiche for his patient guidance, enthusiastic encouragement and useful critiques of this paper. Moreover, we also want to thank him for introducing stochastic calculus early in our engineering studies - this inspired us to go for advanced studies in mathematics. We would also like to thank Fredrik Bohlin, Quantitative Analyst at Handelsbanken Capital Markets, for his advice, assistance and thoughtful comments during the process. Our grateful thanks are also extended to Mattias Lundahl, Vice President at Goldman Sachs, for his assistance in finishing this paper. We would also like to extend our thanks to all the Quants at the Model Development Group at Handelsbanken Capital Markets for insightful discussions and for offering us work place during the entire process. Finally, we wish to thank our parents and girlfriends for their support and encouragement throughout our studies.
Chapter 1

Introduction

1.1 Background

The dynamics of daily correlations plays an important role in several applications in finance and economics. It will result either from correlations between risk premiums, dividend news events or expected returns. Riskmetrics use correlation to calculate Value-At-Risk at short horizons. Erb, Harvey and Viskanta (1994) present examples of the possibility that time varying correlation forecasts will influence optimal portfolio weights. Kroner and V.K. (1998) present how hedging ratios is affected by time varying covariance matrices. Burns, Engle, and Mezrich J. (1998) illustrate that a term structure of correlation is constructed from a multivariate GARCH model on a daily basis. The correlation term structure can be applied when pricing derivative products which have payoffs that depend on the values of more than one asset. During turbulent market conditions, one wish to value international portfolios in real-time. To calculate correct portfolio value, one need, among other things, correct correlation estimates. For example, standard portfolio theory claims that the tangency portfolio is the only efficient stock portfolio. On the other hand, it has been observed that an investment in the global minimum variance portfolio (GMVP) frequently yields better out-of-sample results than an investment in the tangency portfolio (Kempf and Memmel, 2006). The problem can be seen as a minimizations problem

\[
\min_{\omega_t} \omega_t' H_t \omega_t \\
\text{s.t.} \sum_{i=1}^{N} \omega_{i,t} = 1,
\]

(1.1.0.1)

where \(\omega\) is the vector of portfolio weights, \(H_t\) is the variance-covariance matrix of the assets. When the weights has been determined, the variance \(\sigma_t^2 = \omega_t' H_t \omega_t\) at time \(t\) can be computed. The most efficient property with the GMVP process is its uniqueness, which means that "the correct" covariance is associated with an improved performace. The portfolio with the most accurate covariance has the smallest variance at time \(t\) (Sheppard, 2003). This problem will be analyzed and put into perspective later in the paper.

The idea of modeling and forecasting volatilities and correlations through univariate time series was first
introduced by Engle (1982). Ever since the first paper, several attempts have been made to model multivariate GARCH models such as Engle et al. (1984), Bollerslev et al (1988;1994), Engle Mezrich (1996), Bauwens et al. (2006), Silvennoinen and Terasvirta (2008). Bollerslev (1990) introduced a class of multivariate GARCH models called constant conditional correlation (CCC). The main assumption in this model is that the conditional correlations between all assets are assumed to be time invariant. However, correlations tend to vary in time and the CCC model cannot incorporate this fact. Another attempt has been made by Alexander and Barbosa (2008) which is known as the Orthogonal GARCH (OGARCH). OGARCH assumes that every diagonal conditional variance is a univariate GARCH model. Engle (2002a) generalized the CCC model to the dynamic conditional correlation (DCC) model. It has the same structure as the CCC model, besides that it allows the correlations to vary over time instead of being constant.

If the portfolio consists of assets traded in markets with different trading hours, there could potentially occur an underestimation of the right correlation. This is due the asynchronous data - there exist an asynchronicity within the assets time series in the portfolio. Consequently, a common method to lessen the impact of asynchronous data is to use weekly or monthly data. However, weekly and monthly data are unable to capture daily correlation dynamics. Burns et al. (1998) and Riskmetrics (1996;2006) proposed a variety of approaches for treating the issue and calculating a synchronized correlation from a data set containing non-synchronization assets on a daily basis. Therefore Burns et al. (1998) and Riskmetrics (1996;2006) will form the foundation for the synchronization process presented in this paper. The purpose of this paper is twofold. First, we suggest a method for synchronize the returns. Second, we study the daily dynamics along with forecasting of the covariance, the conditional variance and correlation. The results are then tested in the global minimum variance portfolio problem.

1.2 Implication of Asynchronous Data

The difference in trading hours between the world’s stock exchanges plays a vital part when calculating correlations and asset prices. Typically prices are measured from one point in time to the same point 24 hours later. In some cases, stock exchanges in different markets are not open at the same time. Due to different trading hours, news that influences the prices of the assets in the open exchange will also affect the prices in the closed exchange. This is indicated in the opening price and therefore attributing to the following daily return. If returns are measured over distinguishable periods, the correlations may be understated due to asynchronous returns. If assets are traded in markets with diverse trading hours the correlation between those will be influenced. For instance, the correlation between the Japanese market and the U.S. market measured on daily closing prices is significant lower than when simultaneous returns are measured, the markets have a partial overlap during the day. Thus, news events influencing the Japanese market will influence the assets traded on the U.S market the day after (Burns et al, 1998), (Scholes and Williams, 1977)
and (Lo and McKinlay, 1990a). If the time then differ by several hours, the effects can have seriously impact. For instance, it is important for hedging strategies and value at risk measurers to have correct values, or estimates these, for any given point in time, to know the value of the assets. Thus, if prices are not measured at the same time for all assets in a portfolio, systematic errors can occur.

1.3 Fundamentals of Correlations

To achieve an understanding of correlations between assets and why they change, it is necessary to glance at the economics behind movements in asset prices. Investors hold assets in anticipation of payments to be made in the future. Thus, the value of an asset is related to forecasts of future payments; changes in prices are a function of changing forecasts of future payments. The changes in forecasts of future payments we simply call news. This is the foundation of the basic model for changes in asset prices (Samuelson, 1965). Hence, the return of an asset as well as its volatilities and correlation between other assets are dependent on news. The values of all assets are influenced by news to a greater or lesser extent. In equities, news tends to affect some equity prices greater than others because their lines of business are different. Thus, correlations in company’s returns tend to depend on their business. Naturally, if a company changes its business model, its correlations with other companies are likely to change. This is essential to why correlations change over time.

1.4 Fundamentals of Volatility

When observing correlations between assets, it is relevant to get a solid understanding of the expected volatility that might occur among them. Modeling and forecasting the volatility have attracted much attention in recent years, largely driven by its importance in asset-pricing models and risk measurements. There are certain patterns that financial time series exhibits which are essential for correct model specification, estimation and forecasting. Some of these patterns are briefly described below.

**Fat Tails** is when the distribution of asset returns exhibit fatter tails than those of a normal distribution. This is because they exhibit excess kurtosis.

**Volatility Clustering** is the clustering of periods of volatility where large movements are followed by further large movements. This is an indication of shock persistence. Corresponding Box-Ljung statistics show significant correlations which exist at extended lag lengths.

**Leverage Effects** is when volatility increases due to fall in asset prices.

**Long Memory** occurs in high frequency data, when volatility is highly persistent and there is evidence of near unit root behavior in the conditional variance process.
Co-Movements in Volatility has been observed in financial time series across different markets (currencies). Namely, that big movements in one currency is matched by big movements in another. This suggests the importance of multivariate models in modeling cross-correlation in different markets.

Investors are interested in modeling volatility in asset returns as volatility is essential in risk measurement and investors wants a premium for bearing risk. To illustrate this, the daily percentage change in the US stock market has periods of high and low volatility. High volatilities were observed during the financial crisis in 2008 and low volatilities in the middle of the 1990’s when there was a consolidation in the market. It has been observed that large changes in volatility tend to be followed by further large changes, and small changes tend to be followed by further small changes, and this is true for either sign. Consequently, there exists some sort of correlation between the magnitudes of the fluctuations. This phenomenon, when a series of data goes through periods of high and low volatility, is called volatility clustering. A more quantitative view of this fact is while asset returns themselves are uncorrelated, the absolute returns or their squares display a positive, significant and slowly decaying autocorrelation function. Due to the fact that the volatility appears in clusters, the variance of the daily returns can be forecasted even though the daily returns itself is difficult to forecast.
Chapter 2

Theory

2.1 The Univariate GARCH Model

Before we introduce the multivariate Dynamic Conditional Correlation model (DCC), we need to get a fundamental understanding of how the univariate GARCH model functions as it plays an essential role in the study of the DCC Model of Engle and Sheppard (2001). Because the DCC model is a linear combination of the individual GARCH models, and more interestingly, the correlation matrix from the DCC model originates from the GARCH model. Suppose that we have the following return process

\[ r_t = \mu_t + \xi_t. \]  

(2.1.0.1)

where the conditional expectation \( \mu_t = E[r_t|\mathcal{F}_{t-1}] \), the conditional error \( \xi_t \) and \( \mathcal{F}_{t-1} = \sigma(\{r_s : s \leq t-1\}) \) is the sigma field generated by the values of \( \{r_t\} \) up to time \( t-1 \). Furthermore, assume that the conditional error is the conditional standard deviation of the return times the I.I.D. normally distributed zero mean unit variance stochastic variable. Hence,

\[ \xi_t|\mathcal{F}_{t-1} = \sqrt{h_t}z_t \sim N(0, h_t), \text{ where } z_t \sim N(0, 1). \]  

(2.1.0.2)

Note that \( h_t, \xi_t \) are assumed to be independent of time \( t \). Finally, assume that \( \mu_t = 0 \), which gives us

\[ r_t = \sqrt{h_t}z_t \text{ and } r|\mathcal{F}_{t-1} \sim N(0, h_t). \]  

(2.1.0.3)

If \( \mu_t \neq 0 \) the process could be either ARMA filtered or demeaning. However, for \( \mu_t = 0 \) the variances of the returns and error coincides, therefore \( \xi_t \) is an innovation process. Bollerslev (1986) stated the GARCH(p,q) process which in general consists of three terms: the weighted long run variance \( \omega \), the autoregressive term \( \sum_{i=1}^{p} \gamma_i h_{t-i} \) (the sum of the previous lagged variances times the assigned weight for each lagged variance), and the moving average term \( \sum_{i=1}^{q} \delta_i \xi_{t-i}^2 \) (the sum of the previous lags of squared-innovations times the
assigned weight for each lagged square innovation). Hence, the process can be written as

\[ h_t = \omega + \sum_{i=1}^{p} \gamma_i h_{t-i} + \sum_{i=1}^{q} \delta_i \xi_{t-i}^2, \]  

(2.1.0.4)

where \( p \geq 0, \ q > 0, \ \omega \geq 0, \ \gamma_j, \delta_j \geq 0 \) for \( j=1,2,\ldots \).

One drawback with this model is that the innovations in the moving average term is raised to the power of two and does not assume asymmetry of the errors. Glosten, Jagannathan and Runkel (1993) developed an extensions of the GARCH model, GJR-GARCH, which embraces the asymmetry effect. It will not be analyzed further in this report. However, by definition, the variance process is non-negative, which implies that the process \( \{h_t\}_{t=0}^{\infty} \) must be non-negative valued. Further and more detailed constrains of the GARCH\((p,q)\) model can be found in Nelson and Cao (1992).

2.1.1 The Univariate GARCH(1,1) Model

One of the most common and popular applications of the generalized GARCH\((p,q)\) model is the simple GARCH\((1,1)\) model. This process has the following dynamics

\[ h_t = \omega + \delta \xi_{t-1}^2 + \gamma h_{t-1} \]  

(2.1.1.1)

The intuition behind this model is similar to the GARCH\((p,q)\) except that the squared innovation and the variance term only contributes to one lag each. If we successively apply backward recursion of \( h_t \) all the way up to time \( t - T \), we get the following expression

\[ h_t = \omega \left( 1 + \gamma + \gamma^2 + \cdots + \gamma^{T-1} \right) + \delta \sum_{k=1}^{T} \gamma^{k-1} \xi_{t-k}^2 + \gamma^T h_{t-T} \]  

(2.1.1.2)

Letting \( T \) approaching to infinity and since \( \gamma \in (0,1) \), we get the following limit convergence

\[ \lim_{T \to \infty} h_t = \frac{\omega}{1 - \gamma} + \delta \sum_{k=1}^{\infty} \gamma^{k-1} \xi_{t-k}^2 \]  

(2.1.1.3)

This shows that the current variance is an Exponential Weighted Moving Average (EMWA) of the past squared innovations. In spite of the fact that there are substantial differences between the GARCH\((1,1)\) and the EMWA model, in the GARCH process the parameters must be estimated. However, a mean-reverting approach has been incorporated in this model.
It is convenient to work with as few parameters as possible and Engle and Mezrich (1996) introduced
a method called “variance targeting” which makes the computation slightly easier. The idea behind "vari-
ance targeting" is as follows. Denote the unconditional variance $\bar{h}$, then equation (2.1.1.1) can be re-written
in terms of the unconditional variance

$$
h_t - \bar{h} = \omega - \bar{h} + \delta(\xi_{t-1}^2 - \bar{h}) + \gamma(h_{t-1} - \bar{h}) + \gamma \bar{h} + \delta \bar{h}
$$

(2.1.1.4)

If we let $\omega = (1 - \delta - \gamma)\bar{h}$ the above equation becomes

$$
h_t = (1 - \delta - \gamma)\bar{h} + \delta \xi_{t-1}^2 + \gamma h_{t-1}
$$

(2.1.1.5)

The advantages with this model is that the unconditional variance can be expressed as $\bar{h} = \frac{\omega}{1 - \delta - \gamma}$, which
makes the computation easier. Engle and Mezrich (1996) call this "variance targeting", as it forces the
variance matrix to take on a particular and plausible value. Such a moment conditions is particularly
attractive since it will be consistent regardless of whether the model (2.1.1.1) is correctly specified. Clearly,
this holds only under the assumptions that $\gamma + \delta < 1$ and if $\omega > 0$, $\delta > 0$, and $\delta > 0$. In order to ensure that
the conditional variance $h_t$ remains non-negative with probability one, the conditions $\omega \geq 0$, $\delta \geq 0$, and $\gamma \geq 0$
is sufficient in this case. However, another important feature that preserves the non-negativeness of the
conditional covariance $h_t$ is when the process is stationary. However, the GARCH(1,1) is weakly-stationary,
which means that neither the mean process nor the autocovariance of the process depends on the time $t$ with
expected value and covariance according to

$$
E(r_t) = 0
$$

$$
Cov(r_t, r_{t-s}) = \frac{\omega}{1 - \delta - \gamma}
$$

(2.1.1.6)

if and only if $\gamma + \delta < 1$

Hence, the inequality constrains for the GARCH(1,1) when using variance targeting is $\omega > 0, \delta > 0, \gamma > 0,$
and $\gamma + \delta < 1$, and is considered covariance stationary.

### 2.2 The Multivariate Dynamic Conditional Correlation

The Multivariate Dynamic Conditional Correlation (DCC) model is an extension from the univariate GARCH
model given in the previous section. Instead of the one dimensional case, suppose that we have a portfolio
consisting of $n$ assets. Let $r_t = (r_{1,t}, r_{2,t}, \ldots, r_{n,t})'$ be a $n$ dimensional column vector of asset returns at time
$t$, such that $r_t$ is normally distributed with $E[r_t|\mathcal{F}_{t-1}] = 0$ and covariance matrix $H_t = E[r_t r_t'|\mathcal{F}_{t-1}]$, where
\( \mathcal{F}_{t-1} \) is the complete set of asset returns up to time \( t-1 \). For example, \( r_t \) could be the returns of stocks in the S&P 500 equity index. Then we get that

\[
    r_t = H_t^{1/2} z_t, \\
    r_t | \mathcal{F}_{t-1} \sim N(0, H_t),
\]

(2.2.0.7)

where \( z_t = (z_{1,t}, z_{2,t}, \ldots, z_{n,t})' \sim N(0, I_n) \) and \( I_n \) is the identity matrix of order \( n \). One way to obtain \( H_t^{1/2} \) is by applying Cholesky decomposition of \( H_t \). Furthermore, the conditional covariance matrix in the DCC model is developed by Engle (2002) and is decomposed into a relation between the estimated univariate GARCH variances \( (D_t) \) and the conditional correlation matrix \( (R_t) \)

\[ H_t = D_t R_t D_t. \]

(2.2.0.8)

Clearly, \( H_t \) and \( R_t \) are positive definite when there are no linear dependencies in the returns. We need to be assure that all correlation and covariance matrices are positive definite. Thus, that all variances are non-negative. These matrices are in fact stochastic processes and need to be positive definite with probability one, so all past covariance matrices must also be positive definite. If not, there exist linear combinations of \( r_t \) and that gives negative or zero variances. Furthermore, \( D_t \) is a diagonal matrix of the estimated univariate GARCH variances, i.e.

\[
    D_t = \begin{pmatrix}
    \sqrt{h_{1,t}} & 0 & \cdots & 0 \\
    0 & \sqrt{h_{2,t}} & \cdots & 0 \\
    \vdots & \vdots & \ddots & \vdots \\
    0 & 0 & \cdots & \sqrt{h_{n,t}}
    \end{pmatrix}
\]

(2.2.0.9)

The elements in \( D_t \) are specified to go by calculate the quasi-correlation matrix \( Q_t \) by using a mean-reverting model \[ \text{(2.2.0.12)} \]. The \( n \) section (2.1). But it works for any GARCH\((p,q)\) process with normally distributed errors which fulfills the requirements to be a stationary process and the non-negative conditions. Moreover, \( R_t \) is defined as

\[
    R_t = \begin{pmatrix}
    1 & q_{12,t} & q_{13,t} & \cdots & q_{1n,t} \\
    q_{21,t} & 1 & q_{23,t} & \cdots & q_{2n,t} \\
    q_{31,t} & q_{32,t} & 1 & \cdots & q_{3n,t} \\
    \vdots & \vdots & \vdots & \ddots & \vdots \\
    q_{n1,t} & q_{n2,t} & q_{n3,t} & \cdots & 1
    \end{pmatrix}
\]

(2.2.0.10)

and is the conditional correlation matrix of the standardized residuals \( \epsilon_t = D_t^{-1} r_t \sim N(0, R_t) \). Before we go further and analyze \( R_t \), we should take a step back and evaluate the covariance matrix \( H_t \). We know by definition that the covariance matrix is positive definite. Further we know from \[ \text{(2.2.0.8)} \] that \( H_t \) is...
on quadratic form based on \( R_t \). Then it follows that \( R_t \) must be positive definite in order to ensure that \( H_t \) is positive definite. Hence, by the definition of conditional correlation matrix, all elements in \( R_t \) must satisfy the requirement that they are less or equal to one. To guarantee that these requirements are met, \( R_t \) is decomposed to \( R_t = Q_t^{-1} Q_t Q_t^{-1} \) where \( Q_t^{-1} \) ensures that all elements in \( Q_t \) fulfills the requirement \( |q_{ij}| \leq 1 \). Note that \( Q_t \) is positive definite.

\[
Q_t^{-1} = \begin{pmatrix}
\frac{1}{\sqrt{q_{11,t}}} & 0 & \ldots & 0 \\
0 & \frac{1}{\sqrt{q_{11,t}}} & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & \frac{1}{\sqrt{q_{11,t}}}
\end{pmatrix}
\] (2.2.0.11)

Let us assume that \( Q_t \) follows the dynamics

\[
Q_t = \Omega + \alpha \epsilon_{t-1} \epsilon_{t-1} + \beta Q_{t-1}
\]
\[
\Omega = (1 - \alpha - \beta) \bar{R}
\]
\[
\bar{R} = \text{Cov}(\epsilon_t \epsilon_t^\prime) = E(\epsilon_t \epsilon_t^\prime)
\] (2.2.0.12)

where \( \alpha, \beta \) are scalars. The proposed structure of \( Q_t \) might be considered as complicated, but if we compare it with the structure derived from section (2.1.1), where the \( GARCH(1, 1) \) model is derived, things seems to make sense. Notice that the structure of \( Q_t \) is nearly identical to one of the \( GARCH(1, 1) \) case with variance targeting. In particular, this dynamical structure is called “mean reverting” an analogue of the "Scalar GARCH". However, one drawback in this model is that all correlations assume the same structure. Engle and Sheppard (2002) extended the simple model to a more general structure, called the \( DCC(P, Q) \) model. The correlation structure of this model is defined as,

\[
Q_t = (1 - \sum_{i=1}^P \alpha_i - \sum_{j=1}^Q \beta_j) \bar{R} + \sum_{i=1}^P \alpha_i \epsilon_{t-i} \epsilon_{t-i}^\prime + \sum_{j=1}^Q \beta_j Q_{t-j}
\] (2.2.0.13)

Further in this paper we are only going to consider the DCC(1,1) model. For more information about alternative procedures, see Engle and Sheppard (2002).

The specification and estimation of the DCC model contains three general steps. The first step is to "DE-GARCHING" the data, which means that the volatilities must be estimated to constructed standardized residuals (or volatility-adjusted returns). Secondly, we use these standardized residuals to estimate the quasi-correlation matrix \( Q_t \). The third step is to re-scale the quasi-correlation matrix so it becomes a valid correlation matrix, since the quasi-correlation is an approximation of the true correlation. Hence, some elements in the quasi-correlation matrix may not belong in the defined region of correlation \([-1, 1]\), which
in theory is not possible. Therefore, we need to adjust these mishaps after the first estimation is completed.

2.2.1 Step 1: DE-GARCHING

The first step is to construct the standardized residuals or the adjusted volatility-returns. Recall that for the DCC model, we have that

\[ H_t = D_t R_t D_t, \] (2.2.1.1)

\[ D_t^2 = \text{diag}[H_t]. \] (2.2.1.2)

We know from the previous section that the conditional correlation matrix is the covariance matrix of the standardized residuals, given by

\[ R_t = \text{Cov}(D_t^{-1} r_t) = E[\epsilon_t \epsilon_t'], \text{ given } \mathcal{F}_{t-1} \] (2.2.1.3)

All sufficient information that we need to estimate the conditional correlation is captured in these standardized residuals. But estimating \( H_t \) is difficult, so it is convenient to divided the estimation-procedure into two operations. First, we estimate the diagonal elements and then use these estimates to determine the elements not belonging to the diagonal. The diagonal elements of \( D_t \) are the expected standard deviation of each asset with respect to the complete set of information \( \mathcal{F}_{t-1} \). Hence,

\[ H_{i,i,t} = E[r_{i,t}^2] \text{ given } \mathcal{F}_{t-1} \] (2.2.1.4)

The issue that has gained a lot of attention over the years is to find an appropriate model to estimate the conditional variance. Bollerslev (1986) provides a short answer and argue that the variance of a random variable, conditioned on its past information, may be represented by a simple GARCH model. Therefore, we are considering the standard GARCH(1,1) in this case, defined as

\[ H_{i,i,t} = \omega_i + \alpha_i r_{i,t-1}^2 + \beta_i H_{i,i,t-1}. \] (2.2.1.5)

Thus, every univariate process in a multivariate portfolio of assets can be estimated using the above model to get its conditional covariance, so the standardized residuals are

\[ \epsilon_{i,t} = \frac{r_{i,t}}{\sqrt{H_{i,i,t}}} \] (2.2.1.6)

2.2.2 Step 2: Estimating the Quasi-Correlation

In this step we are GARCH(1,1) process embrace the assumption that most correlations changes temporary and are mean-reverting. This specification give us the dynamics of the quasi-correlation process in the
mean-reverting model between asset \(i, j\) and is specified as

\[
Q_{i,j,t} = \omega_{i,j} + \alpha \epsilon_{i,t-1} \epsilon_{j,t-1} + \beta Q_{i,j,t-1}.
\]  
(2.2.2.1)

In matrix notation, we can write the above process simply as

\[
Q_t = \Omega + \alpha \epsilon_{t-1} \epsilon_{t-1}' + \beta Q_{t-1}.
\]  
(2.2.2.2)

Correspondingly, there are two unknown parameters of the dynamical part \((\alpha, \beta)\), and \(\frac{1}{2} N(N-1)\) unknowns in the intercept matrix. However, there is a simple estimator available for the parameters in the intercept matrix that is called "correlation targeting" (compare with variance targeting). This simple estimator essentially amounts to using an estimate of the unconditional correlations among the volatility-adjusted returns (Engle, 2009). More explicitly, using

\[
\hat{\Omega} = (1 - \alpha - \beta) \bar{R},
\]  
(2.2.2.3)

where \(\bar{R} = \frac{1}{T} \sum_{t=1}^{T} \epsilon_t \epsilon_t'\), decreases the number of unknown parameters to two. This is something we need to consider and take into account when we are evaluating the properties of the estimator. Now, if we combine (2.2.2.3) and (2.2.2.2) will give us the dynamics for the mean-reverting DCC model:

\[
Q_t = \bar{R} + \alpha (\epsilon_{t-1} \epsilon_{t-1}' - \bar{R}) + \beta (Q_{t-1} - \bar{R}).
\]  
(2.2.2.4)

Accordingly we can assure that \(Q_t\) is positive definite (PD) if the initial value \(Q_1\) is PD and if

\[
\begin{align*}
\alpha > 0, & \quad \beta > 0 \\
\alpha + \beta < 1 & \\
(1 - \alpha - \beta) > 0
\end{align*}
\]  
(2.2.2.5)

An alternative way of seeing this is that each subsequent of \(Q_t\) is the sum of positive definite or positive semi definite matrices, so must \(Q_t\) be PD. How does this model behave? As we already know, the off-diagonal elements of \(Q_t\) evolves over time in response to new information in the returns. If the returns are moving in line at the same direction (either it goes up or down) the correlation will rise and remain over its average level for a while. However, as time goes by, the correlation will fall back to long-run levels as information will decay. Consequently, if the returns move in the opposite direction relative to each other, the correlation will (temporarily) fall below the unconditional value. Thus, this speed of adjustment is controlled by the parameters \((\alpha, \beta)\), which we need to estimate from the data set. Notice that this is a rather weak specification as only \(\alpha\) and \(\beta\) is used, without hardly considering the size of the system.
2.2.3 Step 3: Rescaling the Quasi-Correlation

The diagonal elements in the matrix $Q_t$ will be an approximation of the correlation matrix. But unfortunately not for every observation, as they may be outside the defined interval. Therefore, we cannot ensure that $Q_t$ is a correlation matrix. This problem can be solved through rescaling the matrix. We can simply estimate the correlation as

$$\rho_{i,j,t} = \frac{Q_{i,j,t}}{\sqrt{Q_{i,i,t}Q_{j,j,t}}}.$$  \hspace{1cm} (2.2.3.1)

While the expected value of $Q_{i,i,t}$ and $Q_{j,j,t}$ are one, they are not estimated to be 1 for every point in time. Denote this equation as rescaling and its matrix is

$$R_t = diag(Q_t^{-1/2})Q_t diag(Q_t^{-1/2})$$  \hspace{1cm} (2.2.3.2)

This will introduce nonlinearity into the estimator. In general, $Q_t$ will be linear in cross products and squares of the data. This implies that it is the case for $R_t$ as well. Thus, $R_t$ will not be an unbiased estimator of the correlation. Moreover, the forecasts are biased. This is true for all multivariate GARCH methods due to the obvious fact that correlations are bounded and the set of data are not.

2.2.4 Estimation of the DCC Model

To estimate the DCC model, we make a assumption about the distribution of the data being used. Once this is done, the problem can be restated as a maximum likelihood problem. We will assume that the data has a multivariate normal distribution with given covariance structure and mean. Moreover, the estimator will be quasi maximum likelihood, due to the fact that it will be inefficient but consistent; the covariance and mean models can be accurate while the distribution assumption is inaccurate. Recall that if $r_t|F_{t-1} \sim N(0, D_tR_tD_t)$, then

$$\begin{align*}
D_t^2 &= diag(H_t) \\
H_{i,i,t} &= \omega_i + \alpha_i \epsilon_{i,t-1}^2 + \beta_i H_{i,i,t-1} \\
\epsilon_t &= D_t^{-1} r_t \\
R_t &= diag(Q_t^{-1/2})Q_t diag(Q_t^{-1/2}) \\
Q_t &= \Omega + \alpha \epsilon_{t-1}^2 + \beta Q_{t-1}
\end{align*}$$  \hspace{1cm} (2.2.4.1)

where $(\alpha_i, \beta_i)$ are positive $\forall i$ and has a sum less than the unity. In order to estimate $\theta = (\phi, \varphi) = (\omega_1, \delta_1, \gamma_1, \ldots, \omega_n, \delta_n, \gamma_n, \alpha, \beta)$ of $H_t$ for the data set $r_t = (r_{1,t}, \ldots, r_{n,t})$, we can set up the following log
likelihood equation (since the errors are assumed to be multivariate normally distributed)

\[
L(r_t, \theta) = -\frac{1}{2} \sum_{t=1}^{T} \left( n \log(2\pi) + \log |H_t| + r'_t H_t^{-1} r_t \right)
= -\frac{1}{2} \sum_{t=1}^{T} \left( n \log(2\pi) + \log |D_t R_t D_t| + r'_t D_t^{-1} R_t^{-1} D_t^{-1} r_t \right)
= -\frac{1}{2} \sum_{t=1}^{T} \left( n \log(2\pi) + 2 \log |D_t| + \log |R_t| + \epsilon'_t R_t^{-1} \epsilon_t \right)
= -\frac{1}{2} \sum_{t=1}^{T} \left( n \log(2\pi) + r'_t D_t^2 r_t + \epsilon'_t \epsilon_t + \log |R_t| + \epsilon'_t R_t^{-1} \epsilon_t \right).
\]

The above function can be maximized with respect to each parameter \( \theta \) in the model. In particular, the first three terms contain the returns and the variance parameters and the remaining parts contains the correlation parameters as well as the volatility adjusted returns. Hence, we can split the function into two separate parts, namely

\[
L(r_t, \theta) = L_1(r_t, \phi) + L_2(\epsilon, \varphi)
= \frac{(-1)}{2} \sum_{t=1}^{T} (n \log(2\pi) + 2 \log |D_t| + r'_t D_t^2 r_t) + \frac{(-1)}{2} \sum_{t=1}^{T} (\epsilon'_t \epsilon_t + \log |R_t| + \epsilon'_t R_t^{-1} \epsilon_t).
\]

Two-step estimation of the parameters

To estimate the parameters of the variance matrix \( H_t \) we use a two-step estimation method according to (Engle 2009). First, we maximize the variance part of the log-likelihood function where we treat each time series as independent and simply compute the univariate GARCH models for respectively time series. Hence, the \( R_t \) matrix is replaced with the identity matrix \( I_n \) and the variance part then becomes

\[
L_1(r_t, \phi) = -\frac{1}{2} \sum_{t=1}^{T} \left( n \log(2\pi) + 2 \log |D_t| + \log |I_n| + r'_t D_t^{-1} I_n D_t^{-1} r_t \right)
= -\frac{1}{2} \sum_{t=1}^{T} \left( n \log(2\pi) + 2 \log |D_t| + r'_t D_t^{-1} D_t^{-1} r_t \right)
= -\frac{1}{2} \sum_{t=1}^{T} \sum_{i=1}^{n} \left( \log(2\pi) + 2 \log(h_{i,t}) + \frac{r_{t,i}^2}{h_{i,t}} \right)
= -\frac{1}{2} \sum_{t=1}^{T} \left( T \log(2\pi) + \sum_{i=1}^{n} \left( 2 \log(h_{i,t}) + \frac{r_{t,i}^2}{h_{i,t}} \right) \right).
\]
This equation helps us to estimate the parameters \( \hat{\phi} = (\hat{\omega}_1, \hat{\delta}_1, \hat{\gamma}_1, \ldots, \hat{\omega}_n, \hat{\delta}_n, \hat{\gamma}_n) \) for each univariate GARCH process of \( r_t \). Further as \( h_{i,t} \) is estimated for \( t \in [1, T] \), so all elements in \( D_t \) are estimated over the same period. The second step clearly involves to estimate the parameters of the correlation part, i.e. \( \varphi = (\alpha, \beta) \) conditioned on the previously estimated parameters \( \hat{\phi} \) from the first step. We have from the log-likelihood equation

\[
L_2(r_t, \varphi|\hat{\phi}) = -\frac{1}{2} \sum_{t=1}^{T} \left( n \log(2\pi) + 2 \log |D_t| + \log |R_t| + \epsilon_t^\prime R_t^{-1} \epsilon_t \right) 
\]

\[
\approx -\frac{1}{2} \sum_{t=1}^{T} \left( \log |R_t| + \epsilon_t^\prime R_t^{-1} \epsilon_t \right). 
\]

The approximation evolves because the first two terms \( n \log(2\pi) + 2 \log |D_t| \) are constant and we are only interested to optimize the remaining parts which includes the \( R_t \) matrix. The residuals \( \epsilon_t \) are calculated according to (2.2.4.1) and the covariance matrix is then estimated by \( R = \frac{1}{T} \sum_{t=1}^{T} \epsilon_t \epsilon_t^\prime \). Recall the specification of the correlation matrix below. With the imposed restrictions of \( \Omega \) gives the dynamics of the correlation

\[
Q_t = \Omega + \alpha \epsilon_{t-1} \epsilon_{t-1}^\prime + \beta Q_{t-1} 
\]

\[
\Omega = \bar{R}(1 - \alpha - \beta). 
\]
Chapter 3

Synchronization

3.1 The Data Set

The data set that we consider in this report includes equity indices returns from global markets around the world. Particularly, we intend to consider daily returns on a basis of a logarithmic approach as it captures the daily compounding and new information available from the market (A.0.4.1). The goal is then to study and enlighten the asynchronous properties of these global equity indices, especially when there are no common trading hours. However, there are some markets with partial overlaps, such as the U.K. and the U.S., when both markets are open at the same time. For now on, let us consider a global portfolio \( G \) of six equity indices, namely the S&P500 (U.S.), OMXS30 (Sweden), DAX (Germany), HSCE (Hong Kong) and NIKKEI (Japan) for the period February 2006 to March 2013. They were selected due to their size and trading hours which are given in the table below.

<table>
<thead>
<tr>
<th>Stock Exchange(Index)</th>
<th>Open(UTC)</th>
<th>Close(UTC)</th>
<th>Lunch</th>
</tr>
</thead>
<tbody>
<tr>
<td>Japan (NIKKEI 225)</td>
<td>00:00</td>
<td>06:00</td>
<td>02:30-03:50</td>
</tr>
<tr>
<td>Hong Kong (HSCE)</td>
<td>01:20</td>
<td>08:00</td>
<td>04:00-05:00</td>
</tr>
<tr>
<td>Germany (DAX30)</td>
<td>07:00</td>
<td>21:00</td>
<td>No</td>
</tr>
<tr>
<td>UK (FTSE100)</td>
<td>08:00</td>
<td>16:30</td>
<td>No</td>
</tr>
<tr>
<td>Sweden (OMXS30)</td>
<td>08:00</td>
<td>16:30</td>
<td>No</td>
</tr>
<tr>
<td>U.S. (S&amp;P500)</td>
<td>14:30</td>
<td>21:00</td>
<td>No</td>
</tr>
</tbody>
</table>

One concern that needs to be addressed before the evaluation of the portfolio \( G \) is the treatment of missing data points. The portfolio may contain gaps where no closing price (no returns) has been observed. The reason for this is mostly due to bank holidays which create an asymmetry between the other markets. Recently, the case has also been that something catastrophic has occurred in a nation which causes the markets to temporarily close. To solve this gap-issue and fill the missing data points in our time-series, we apply a classic linear interpolation method Meijering, (2002). Naturally, there are other methods to fill the time series gap issue, such as the 'spline' method (smooth polynomial function that is piecewise-defined),
which are not considered in this paper. As mentioned above the log-returns captures the daily information available among the markets, but it also exhibit another beneficial characteristic. By plotting the time series for respectively index, we see that it exhibits the sought stationarity property as it fluctuate around a common mean (Brockwell and Davis, 1991). It is often desired to deal with a stationary data set, primarily because the trait allows us to assume that models are independent of a particular starting point. The time series and log returns are shown in Appendix C.

3.2 Synchronization of the Data

The synchronization technique presented in this section is influenced by the approach proposed by Burns, Engle and Mezrich (1998). For example, consider a sub-portfolio $G$ containing only the U.S. market and the U.K. market. To measure the value of this portfolio when the market is closed in the U.S., we need to estimate the value of the corresponding shares traded in the U.K. Assume that the market in the U.S. falls by 5 percent after the market close in the U.K., then it is inappropriate to value the portfolio at the U.S. closing time. More general, for the portfolio $G$, let us denote $S_{t,j}$, as the price measured continuously of an equity index $j$, such that $S_{t,j} \in G$. Let $t$ for $t \in \mathbb{N}$ be the daily local time in the U.S. so that $S_{1,1}$ is the price of the U.S. equity index at 16:00 on the first day of trading. The closing time in the U.S. at 16:00 corresponds to 22:00 in the U.K. as the U.K. closes 4 hours before the U.S. Denote the observed closing price of the equity index in the U.K. on the first trading day as $S^{0.83}_{0,2}$ (Burns, Engle and Mezrich, 1998). The observed data are captured from the closing times of the markets and has the structure $S_{t,j,j}$ where $t_j = t_1 - c_j$ $(0 \leq c_j \leq 1)$, $j = (1, \ldots, 6)$. We have to synchronize to $t_1$, the closing time in the U.S. of equity index $j = 1$, where $t_1 \in \{1,2,\ldots,T\}$. Our intention is to produce synchronized prices $S^s_{t_j,j}$, where $t \in \{1,2,\ldots,T\}$, for all equity indices. The synchronized prices $S^s_{t_j,j}$ is defined as

$$
\log(S^s_{t_j,j}) = E[\log(S_{t,j})|\mathcal{F}_t], \ where \ \mathcal{F}_t = \{S_{t_j,j}; t_j \leq t, j = 1,\ldots,6\}, \ (3.2.0.1)
$$

where the logarithms are consistent with continuously compounded returns. Thus, the best predicted log prices at $t$ are the synchronized log prices given the complete information $\mathcal{F}_t$, where $\mathcal{F}_t$ contains all recorded prices up to time $t$. Moreover, the complete information contains only the prices $S_{t_j,j}$ with $t_j t$. A strict relationship $t_j < t$ is often observed when equity index $j$ is trading in a market with a different closing time than the first equity in the U.S. For the equity index in the U.S., the closing price $S_t$ that is observed at time $t \in \mathbb{N}$ has the conditional expectation, given $\mathcal{F}_t$, the observed price. In the case when a market closes before $t$, its past closing prices and closing prices from other markets can be useful for predicting $S_t$ at time $t$. For simplicity, assume that given $\mathcal{F}_t$ the predicted log prices at $t$ and at the following closing time $t_j + 1$ are the same. Thus, future changes from predictions at time $t$ to predicted returns at $t_j + 1$ are not predictable.

$$
\log(S^s_{t_j,j}) = E[\log(S_{t,j})|\mathcal{F}_t] = E[\log(S_{t_j+1,j})|\mathcal{F}_t], \ t_j \leq t < t_j + 1 \ (t \in \mathbb{N}) \quad (3.2.0.2)
$$
By definition in (3.2.0.1), the first equality holds. The approximation above will be the foundation for the synchronization formula in (3.2.0.2). Now, denote $R_t$ as the vector of log-returns McNeil and Frey (2000) of the 6 markets in our portfolio $G$ at different time points

$$R_t = \begin{pmatrix} r_t^{U.S.} \\ r_t^{SWE} \\ r_t^{UK} \\ r_t^{GER} \\ r_t^{HKG} \\ r_t^{JPY} \end{pmatrix} = 100 \begin{pmatrix} \log \left( \frac{S_{t,1}}{S_{t-1,1}} \right) \\ \vdots \\ \log \left( \frac{S_{t,6}}{S_{t-1,6}} \right) \end{pmatrix} = 100(\log(S_t) - \log(S_{t-1})) \tag{3.2.0.3}$$

where $t = (t_1, t_2, \ldots, t_6)$. The synchronized returns can now be written in term of the synchronized prices

$$R_s^t = \begin{pmatrix} r_t^{U.S.} \\ r_t^{SWE} \\ r_t^{UK} \\ r_t^{GER} \\ r_t^{HKG} \\ r_t^{JPY} \end{pmatrix} = 100 \begin{pmatrix} \log \left( \frac{S_{t,1}}{S_{t-1,1}} \right) \\ \vdots \\ \log \left( \frac{S_{t,6}}{S_{t-1,6}} \right) \end{pmatrix} = 100(\log(S_t^s) - \log(S_{t-1}^s)), \ t \in \mathbb{N} \tag{3.2.0.4}$$

The synchronized returns have to be modeled as it depend on unknown conditional expectations. We assume a multivariate AR(1) process to fit these returns

$$R_t = AR_{t-1} + \epsilon_t \tag{3.2.0.5}$$

where $\epsilon_t$ is the error term and $E[\epsilon_t | \mathcal{F}_{t-1}] = 0$ and $A$ is a $6 \times 6$ matrix. Burns, Engle and Mezrich (1998) suggest an first-order vector moving average while (Audriano, F. and Bühlmann, P. 2004) argue that an first-order vector autoregressive process will simplify the synchronization, due to the Markovian structure, i.e. the expected returns $E[R_t | \mathcal{F}_{t-1}]$ only depends on the previous $R_{t-1}$. Moreover, (Audriano, F. and Bühlmann, P. 2004) provide empirical evidence that their model is superior to Burns, Engle and Mezrich (1998) in predictive performance. In (3.2.0.6) it will be shown that the synchronized returns achieved with (3.2.0.5) are functions of $R_t$ and $R_{t-1}$ only. We then obtain the synchronized returns by substituting (3.2.0.2) into (3.2.0.4)

$$R_s^t = 100(\log(S_t^s) - \log(S_{t-1}^s)) \tag{3.2.0.6}$$

The synchronized returns can now be written in terms of asynchronous returns. This is done by adding and subtracting $E[\log(S_t) | \mathcal{F}_{t-1}] = \log(S_t)$ for $t$ and $t^{-1}$ on the right-hand side of (3.2.0.6). Substituting (3.2.0.3)
and (3.2.0.5) into (3.2.0.6)

\[ R_s^t = 100(E[\log(S_{t+1}) - \log(S_t)]|F_t] - E[\log(S_t) - \log(S_{t-1})|F_{t-1}] + \log\left(\frac{S_t}{S_{t-1}}\right)). \quad (3.2.0.7) \]

This gives us that,

\[ R_s^t = E[R_{t+1}|F_t] - E[R_t|F_{t-1}] + R_t \]
\[ = R_t + AR_t - AR_{t-1} \]
\[ = R_t + A(R_t - R_{t-1}) \]

(3.2.0.8)

The synchronized returns are equivalent to the asynchronous returns and a correction term. The correction term consists of linear combinations of increments from \( t - 1 \) to \( t \). If \( A \) is the zero matrix, it is clear that \( R_s^t = R_t \) since the data is synchronized. Since we are benchmarking the synchronization against the U.S. market the corresponding row in \( A \) is a zero row, as the data are already synchronized. A method for estimating \( A \) is presented in the next section.

### 3.2.1 Estimation of the A-matrix

In the previous section, we proposed a synchronous DCC GARCH model which includes the \( A \) matrix with \( 6^2 \) unknown parameters. To reduce the number of parameters we set some elements in \( A \) equal to zero if they are statistically insignificant. This is helpful when working with high-dimensional portfolios, i.e. portfolios with multivariate-assets returns. We will apply a two step approach to estimate the \( A \) matrix.

**Step 1**

First we apply the Yule-Walker estimator from Brockwell and Davis (1991) to estimate the \( 6^2 \) parameters of \( A \) and the corresponding \( \Sigma \) matrix. For a multivariate \( AR(1) \) process, the Yule-Walker covariance relation is written as

\[ \gamma(0) = \gamma(-1)A' + \Sigma = \gamma(1)A' + \Sigma \]
\[ \gamma(1) = \gamma(0)A', \text{ where } \gamma(k) = E[R_{t-k}R_t'] \]

(3.2.1.1)

Secondly, calculate the standard errors of the estimated elements of the matrix \( A \), i.e. find standard errors of \( \hat{A} \).

\[ \text{s.e. } (\hat{A}_{i,j}) = \sqrt{\text{Var}(\hat{A}_{i,j})} \]
\[ = \sqrt{\frac{\hat{\sigma}_{i,i}}{T-1} \gamma(0)^{-1}_{j,j}} \]

(3.2.1.2)

where \( \gamma(0) = \frac{1}{T} \sum_{t=1}^{T} (R_t - \bar{R})(R_t - \bar{R})' \) with \( \bar{R} = \frac{1}{T} \sum_{t=1}^{T} R_t \) and \( \hat{\sigma}_{i,i} \) is the estimate of the \( i \)th diagonal element of the covariance matrix \( \Sigma \), and \( T \) is the sample size.
Step 2

The second step involves adopting the elements which has a significant impact on the model. We set $A_{ij} = 0$ if the $t$-statistics

$$t_{ij} = \left| \frac{\hat{A}_{ij}}{\text{s.e.}(\hat{A}_{ij})} \right| \leq 1.96 \ (5\% \ significans \ level) \quad (3.2.1.3)$$

and $A_{ij} = 0$ for all $j$ with $i$ corresponding to the U.S. stocks (in this case for all $i = 1$).
Chapter 4

Results

4.1 Synchronization

4.1.1 Estimating the A matrix

The estimated $\hat{A}$ matrix is obtained by using the procedure described in section 3.2.1. The variables in $\hat{A}$ are ordered as U.S, SWE, U.K., GER, HK and JPY, where the first column corresponds to the U.S. and therefore has the highest coefficients. Recall that the first row is the zero row as returns for U.S. is already synchronized. There is a significant predictability of the other indices from the U.S. the day before. Not only is the U.S. market a major player and determining aspect in the global markets, but the U.S. is the last to close so the observed patterns comes quite natural. Moreover, there seem to be some kind of predictability from the stock exchange in the U.K. with the rest. The negative signs can be viewed as some kind of correction impulse for the markets in Europe. This could described by some kind of joint effect from the U.S. and the U.K. and could be the reason for the large impact of the U.S. on the U.K. In addition, the U.K., HK and JPY seem to be autocorrelated. SWE and the U.K. close simultaneously and have coefficients equal or close to zero.

$$\hat{A} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0.2646 & 0 & 0 & -0.0668 & 0 & 0 \\ 0.3114 & -0.0489 & -0.1283 & 0 & 0.0283 & 0 \\ 0.2452 & 0 & -0.1303 & 0 & 0 & 0 \\ 0.3882 & 0 & 0.2642 & 0 & -0.0507 & -0.0792 \\ 0.3872 & 0 & 0.1862 & 0.0609 & 0 & -0.0849 \end{pmatrix}$$

The figure below indicates that synchronization has primarily an effect on returns of small or medium size.
4.2 Numerical Evaluation of the DCC Model

To examine the specifications of the DCC model, we intend to look at its relative performance to the industry standard RiskMetrics exponential smoother. Our aim is to study the models performances. We use a method to test the variance of returns for a portfolio against the predicted variance. The portfolio we are considering consists of six equity index from different markets, see section 3.1. To calculate the weights we use a method called the global minimum variance portfolio (GMVP). This is an interesting method to use as all weights in the portfolio are calculated by the estimated variance-covariance matrix $H_t$ derived from the DCC model. If the conditional covariance is properly specified, we should expect the variance of the portfolio to be specified as $\sigma_t^2 = \omega_t' H_t \omega_t$, where $\omega_t$ is the weight at time $t$. The GMVP also possess the unique property that the most correct conditional covariance leads to improved performance. Thus, the portfolio with the best estimated covariance tends to have the smallest variance. We calculate the time-varying weights of the portfolio by using the following structure,

$$\omega_t = \frac{H_t^{-1}}{C_t} \quad (4.2.0.1)$$
where $C_t = \gamma H_t^{-1}$ and $H_t$ is the (one step ahead) conditional covariance up to time $t - 1$ and $1$ is a $k$ by 1 dimensional vector of ones. If the variance of each portfolio is extremely small (relative to the predicted variance) it is an indication of excess correlation, while the opposite would imply an underestimation of the correlation. The performance of the DCC model is compared to the industrial standard RiskMetrics Exponentially Weighted Moving Averages (EWMA) model, which is an alternative to the classic EMWA model (Riskmetrics, 1996). The core function of this model is that it allows for more weight on recent information, therefore a very popularized method when estimating volatility. The structure of RiskMetrics is defined as

$$H_t = (1 - \lambda)\epsilon_t\epsilon_t^\prime + \lambda H_{t-1}, \text{ where } \lambda \in (0, 1)$$

(4.2.0.2)

where $\lambda$ is the smoothing parameter and the initial covariance, $H_0$, can be set to the sample covariance matrix or any other suitable selection of presampled data. It is worth noting that the smoothing parameter $\lambda$ already is determined in this model. The single parameter is usually set to 0.94 for daily data and 0.97 for monthly data based on the recommendations from RiskMetrics (RiskMetrics, 1996). This is obviously an advantage as it simplifies the estimation of the model. But it also provides a drawback as it forces all assets in the portfolio to assume the same smoothing coefficient. However, the simplifications make it extremely popular and is used by many applicants in the financial industry, especially those within risk management and Value-At-Risk measures.

The comparison of GMVP is considered for two different cases. The first involves the non-synchronous returns $R_t$, i.e. without taking into account the estimated $\hat{A}$-matrix from section 3.2.1. Secondly, we use the synchronized returns $R_t^s$, i.e. using equation (3.2.0.9) where the $\hat{A}$-matrix are used for both the DCC(1,1) and RiskMetrics EMWA model. Figure 2 shows the weights for GMVP in the Japanese market with non-synchronous. At a first glance, the daily weights look quite similar, despite the slightly different covariance structure. However, Table 4.1 provide us with the calculated (annualized) standard deviations 13.45 and 14.21 for DCC(1,1) and RiskMetrics, respectively. Consequently, the DCC model tends to "beat" the RiskMetrics model, since the realized variance is slightly smaller. Further investigation of Table 4.1 provides us with information that it seems to hold for all the other markets aswell. In particular, the Hong Kong index shows a remarkably difference of 5.471 in terms of annualized standard deviation, which confirms the superiority of DCC model. In this case, it is quite obvious that the DCC model proves to have the smallest variance.
4.2.1 Estimating the Quasi-Correlations

Using the $A$ matrix from section 3.2.1 and the synchronization formula (3.2.0.8), we obtain the synchronized returns $R_s$. The asynchronous correlations and the synchronized correlations are shown below. At a first glance at the correlations we claim that synchronized data frequently exhibit bigger instantaneous correlations among different returns from indices from the same day. We spot that the correlations between asynchronous indices, particularly the ones for the U.S. and HK/JPY, are significant lower than the synchronized returns. It should be clear that synchronized data does not always have to yield higher correlations. The result is similar and consistent to the analysis made by Burns, Engle and Mezrich (1998). Similarly, the same reasoning seems to hold for the quasi correlations for $t + 1$. In this case, the asynchronous correlations tend
to decrease, except for U.S. and HK/JPY which increase. All synchronized correlations decrease for $t + 1$ except U.S./HK which increase. Thus, correlations tend to decrease when predicting.

### Table 4.2: Asynchronous Quasi-Correlations

<table>
<thead>
<tr>
<th>Stock Exchange</th>
<th>U.S.</th>
<th>SWE</th>
<th>UK</th>
<th>GER</th>
<th>HK</th>
<th>JPY</th>
</tr>
</thead>
<tbody>
<tr>
<td>U.S.</td>
<td>1</td>
<td>0.50933</td>
<td>0.57652</td>
<td>0.50065</td>
<td>0.074304</td>
<td>-0.064002</td>
</tr>
<tr>
<td>SWE</td>
<td>0.50933</td>
<td>1</td>
<td>0.86144</td>
<td>0.79769</td>
<td>0.33705</td>
<td>0.22857</td>
</tr>
<tr>
<td>UK</td>
<td>0.57652</td>
<td>0.86144</td>
<td>1</td>
<td>0.97232</td>
<td>0.42014</td>
<td>0.31744</td>
</tr>
<tr>
<td>GER</td>
<td>0.50065</td>
<td>0.79769</td>
<td>0.97232</td>
<td>1</td>
<td>0.34142</td>
<td>0.3002</td>
</tr>
<tr>
<td>HK</td>
<td>0.074304</td>
<td>0.33705</td>
<td>0.42014</td>
<td>0.34142</td>
<td>1</td>
<td>0.54565</td>
</tr>
<tr>
<td>JPY</td>
<td>-0.064002</td>
<td>0.22857</td>
<td>0.31744</td>
<td>0.3002</td>
<td>0.54565</td>
<td>1</td>
</tr>
</tbody>
</table>

### Table 4.3: Synchronous Quasi-Correlations

<table>
<thead>
<tr>
<th>Stock Exchange</th>
<th>U.S.</th>
<th>SWE</th>
<th>UK</th>
<th>GER</th>
<th>HK</th>
<th>JPY</th>
</tr>
</thead>
<tbody>
<tr>
<td>U.S.</td>
<td>1</td>
<td>0.64239</td>
<td>0.70013</td>
<td>0.61956</td>
<td>0.40803</td>
<td>0.32849</td>
</tr>
<tr>
<td>SWE</td>
<td>0.64239</td>
<td>1</td>
<td>0.72291</td>
<td>0.71291</td>
<td>0.45912</td>
<td>0.39812</td>
</tr>
<tr>
<td>UK</td>
<td>0.70013</td>
<td>0.72291</td>
<td>1</td>
<td>0.78361</td>
<td>0.53904</td>
<td>0.4951</td>
</tr>
<tr>
<td>GER</td>
<td>0.61956</td>
<td>0.71291</td>
<td>0.78361</td>
<td>1</td>
<td>0.46218</td>
<td>0.48034</td>
</tr>
<tr>
<td>HK</td>
<td>0.40803</td>
<td>0.45912</td>
<td>0.53904</td>
<td>0.46218</td>
<td>1</td>
<td>0.42832</td>
</tr>
<tr>
<td>JPY</td>
<td>0.32849</td>
<td>0.39812</td>
<td>0.4951</td>
<td>0.48034</td>
<td>0.42832</td>
<td>1</td>
</tr>
</tbody>
</table>

### Table 4.4: Asynchronous Quasi-Correlations (t+1)

<table>
<thead>
<tr>
<th>Stock Exchange</th>
<th>U.S.</th>
<th>SWE</th>
<th>UK</th>
<th>GER</th>
<th>HK</th>
<th>JPY</th>
</tr>
</thead>
<tbody>
<tr>
<td>U.S.</td>
<td>1</td>
<td>0.51518</td>
<td>0.58267</td>
<td>0.50396</td>
<td>0.071449</td>
<td>-0.072</td>
</tr>
<tr>
<td>SWE</td>
<td>0.51518</td>
<td>1</td>
<td>0.8754</td>
<td>0.8092</td>
<td>0.34194</td>
<td>0.23123</td>
</tr>
<tr>
<td>UK</td>
<td>0.58267</td>
<td>0.8754</td>
<td>1</td>
<td>0.98988</td>
<td>0.42829</td>
<td>0.32337</td>
</tr>
<tr>
<td>GER</td>
<td>0.50396</td>
<td>0.8092</td>
<td>0.98988</td>
<td>1</td>
<td>0.34558</td>
<td>0.3002</td>
</tr>
<tr>
<td>HK</td>
<td>0.074304</td>
<td>0.33705</td>
<td>0.42014</td>
<td>0.34142</td>
<td>1</td>
<td>0.54565</td>
</tr>
<tr>
<td>JPY</td>
<td>-0.072</td>
<td>0.23123</td>
<td>0.32337</td>
<td>0.30624</td>
<td>0.55541</td>
<td>1</td>
</tr>
</tbody>
</table>

### 4.2.2 Forecasting the Quasi-Correlations
Table 4.4: This table represents the forecasted quasi-correlation between different markets around when asynchronized returns $R_{t,i}$ have been applied in the Multivariate DCC-GARCH(1,1)

<table>
<thead>
<tr>
<th>Stock Exchange</th>
<th>U.S.</th>
<th>SWE</th>
<th>UK</th>
<th>GER</th>
<th>HK</th>
<th>JPY</th>
</tr>
</thead>
<tbody>
<tr>
<td>U.S.</td>
<td>1</td>
<td>0.6497</td>
<td>0.7074</td>
<td>0.62518</td>
<td>0.4123</td>
<td>0.3283</td>
</tr>
<tr>
<td>SWE</td>
<td>0.6497</td>
<td>1</td>
<td>0.73036</td>
<td>0.71959</td>
<td>0.46326</td>
<td>0.39996</td>
</tr>
<tr>
<td>UK</td>
<td>0.7074</td>
<td>0.73036</td>
<td>1</td>
<td>0.79234</td>
<td>0.54491</td>
<td>0.49856</td>
</tr>
<tr>
<td>GER</td>
<td>0.62518</td>
<td>0.71959</td>
<td>0.79234</td>
<td>1</td>
<td>0.46559</td>
<td>0.48446</td>
</tr>
<tr>
<td>HK</td>
<td>0.4123</td>
<td>0.46326</td>
<td>0.54491</td>
<td>0.46559</td>
<td>1</td>
<td>0.43117</td>
</tr>
<tr>
<td>JPY</td>
<td>0.3283</td>
<td>0.39996</td>
<td>0.49856</td>
<td>0.48446</td>
<td>0.43117</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 4.5: This table represents the forecasted quasi-correlation between different markets when synchronized returns $R_{t,i}^s$ have been applied in the Multivariate DCC-GARCH(1,1)

### 4.2.3 Parameters

#### Synchronous Univariate-GARCH(1, 1)

<table>
<thead>
<tr>
<th>Stock Exchange</th>
<th>$\alpha_{async}$</th>
<th>$\alpha_{sync}$</th>
<th>$\beta_{async}$</th>
<th>$\beta_{sync}$</th>
<th>$\omega_{async}$</th>
<th>$\omega_{sync}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>U.S.</td>
<td>0.061969</td>
<td>0.061969</td>
<td>0.92936</td>
<td>0.92936</td>
<td>0.0091177</td>
<td>0.0091177</td>
</tr>
<tr>
<td>SWE</td>
<td>0.053889</td>
<td>0.05489</td>
<td>0.93766</td>
<td>0.93546</td>
<td>0.013504</td>
<td>0.017189</td>
</tr>
<tr>
<td>UK</td>
<td>0.067475</td>
<td>0.059809</td>
<td>0.93252</td>
<td>0.93285</td>
<td>0.000002</td>
<td>0.0093723</td>
</tr>
<tr>
<td>GER</td>
<td>0.059492</td>
<td>0.056456</td>
<td>0.92909</td>
<td>0.93338</td>
<td>0.01604</td>
<td>0.014793</td>
</tr>
<tr>
<td>HK</td>
<td>0.050069</td>
<td>0.048255</td>
<td>0.94372</td>
<td>0.94447</td>
<td>0.018588</td>
<td>0.021981</td>
</tr>
<tr>
<td>JPY</td>
<td>0.069818</td>
<td>0.079187</td>
<td>0.91567</td>
<td>0.9039</td>
<td>0.02174</td>
<td>0.02643</td>
</tr>
</tbody>
</table>
4.2.4 Volatility and Correlation over time

Sweden

Figure 4.3: Sweden synchronized returns - volatility and correlation over time

Figure 4.4: Sweden asynchronized returns - volatility and correlation over time
Germany

Figure 4.5: Germany synchronized returns - volatility and correlation over time

Figure 4.6: Germany asynchronized returns - volatility and correlation over time
United Kingdom

Figure 4.7: United Kingdom synchronized returns - volatility and correlation over time

Figure 4.8: United Kingdom asynchronized returns - volatility and correlation over time
Hong Kong

Figure 4.9: Hong Kong synchronized returns - volatility and correlation over time

Figure 4.10: Hong Kong asynchronized returns - volatility and correlation over time
Japan

Figure 4.11: Japan synchronized returns - volatility and correlation over time

Figure 4.12: Japan asynchronous returns - volatility and correlation over time
Chapter 5

Conclusion

Risk analysis and financial decision making requires true and appropriate estimates of correlations today and how they are expected to evolve in the future. We have presented methods for analyzing correlations and synchronization of data. The DCC model is simple and powerful and therefore a promising tool. It provides good insight into how correlations are likely to evolve in the future. The method for synchronization data is an extension of Burns et al. (1998) where a first-order vector autoregressive process is applied, which will simplify the synchronization, due to the Markovian structure.

When synchronizing the returns, all correlations against the United States increase. In particular, the results exhibits that for markets with non-partial overlaps, the correlation increases the most. For example, the asynchronous correlation between the United States and Japan is -0.064 while the synchronous correlation switches sign and becomes 0.3285. This is, indeed, a significant result which confirms our claim that correlations between asynchronous markets are underestimated.

Further, we look into the volatility for the asynchronous and synchronous returns, by numerically evaluate the DCC model against the RiskMetrics EMWA model. As we just concluded, the correlation seems to increase for synchronized returns, which in turn leads to higher volatility. This is noticed in Table 4.1 and the GVMP, where the volatility increases for all markets despite which model we use. However, the DCC model provides lower volatility for both the asynchronous and synchronous returns. It is clear from the results that the DCC model is superior to RiskMetrics EMWA model and the conclusion is based on the volatility of the portfolio.

The financial markets are complex and continuously evolving processes. The DCC model developed in this paper has the potential to adjust to unexpected changes in the financial markets and therefore provide a dynamic picture of volatilities and correlations. This model is suited for the short-run, focusing on what can occur in the near future. Risk managers must essentially be concerned with the longer run. The interest reader can learn about such versions of the DCC model in Engle (2009). One of these is the Factor DCC,
which models the determinants of factor volatilities. These are main determinants of longer-term correlations and volatilities. This paper is presented with the wish that with this type of sane modeling, investors will in the future be given the opportunity to take simply the risks they desire to capture and leave the rest for others with diverse appetites. Certainly, the foundation of our theories of finance is the preference between return and risk for all investors. This involves ever-changing judgments of risk in a world of such enormous and remarkable complexity.
Appendix A

An Economic Model of Correlations

We can illustrate how correlations in returns are based on correlations in the news by a mathematical derivation. Denote $r$ as the continuously compounded returns

$$r_{t+1} = \log(P_{t+1} + D_{t+1}) - \log(P_t)$$  \hspace{1cm} (A.0.4.1)

where $P$ is the share price and $D$ the dividend per share. If we apply the log-linearization on the returns, proposed by Campbell and Shiller (1988), the above relation can be approximated as

$$r_{t+1} \approx k + \rho p_{t+1} + (1 - \rho)d_{t+1} - d_t$$  \hspace{1cm} (A.0.4.2)

where $k$ is a constant of linearization and lowercase letters refer to logs. This approximation is desired when the ratio of the components is relatively constant and small (Engle, 2009). These conditions are often satisfied for equity prices. The parameter $\rho$ is the discount rate and assumingly close to one. Furthermore, if we assume that stock prices do not diverge to $\infty$, we can solve for $p_t$ according to

$$p_t = \frac{k}{1 - \rho} + (1 - \rho) \sum_{j=0}^{\infty} \rho^j d_{t+1+j} - \sum_{j=0}^{\infty} \rho^j r_{t+j}.$$  \hspace{1cm} (A.0.4.3)

Take the expectations with respect to the news at time $t$

$$p_t = \frac{k}{1 - \rho} + (1 - \rho) \sum_{j=0}^{\infty} \rho^j E(d_{t+1+j}) - \sum_{j=0}^{\infty} \rho^j E(r_{t+1+j})$$  \hspace{1cm} (A.0.4.4)

Similarly, take expectations with respect to the news at time $t - 1$, which gives the one step ahead predictor of prices. Take the difference between the log of the price excepted at $t$ and the expected at $t - 1$ to achieve the surprise in returns. Thus,

$$r_t - E_{t-1}(r_t) = E_t(p_t) - E_{t-1}(p_t)$$  \hspace{1cm} (A.0.4.5)
and

$$r_t - E_{t-1}(r_t) = (1 - \rho) \sum_{j=0}^{\infty} \rho^j (E_t - E_{t-1})(d_{t+1+j}) - \sum_{j=0}^{\infty} \rho^j (E_t - E_{t-1})(r_{t+1+j})$$  \hspace{1cm} (A.0.4.6)$$

We notice that there are two components of unexpected returns; surprises in future expected returns and future dividends. It is convenient to express this by the relation

$$r_t - E_{t-1}(r_t) = \eta_t^d - \eta_t^r$$  \hspace{1cm} (A.0.4.7)$$

These two innovations comprise the news. The news is used to forecast the expected returns and discounted future dividends. These innovations are the shocks to a weighted average of expected returns and future dividends. Therefore, each is a martingale difference sequence. It is clear from (A.0.4.6) that a small price of news observed during time $t$ could have a great effect on stock prices if it influence expected dividends for numerous periods into the future. Trivially, it could have a petite effect if it affects dividends for a short period. From (A.0.4.7) the conditional variance of an asset return is given

$$V_{t-1}(r_t) = V_{t-1}(\eta_t^d) + V_{t-1}(\eta_t^r) - 2Cov(\eta_t^d, \eta_t^r)$$  \hspace{1cm} (A.0.4.8)$$

Clearly, each term measures the significance in today’s news in forecasting expected returns or future dividends. If $d$ is an $\infty$-order $MA$, potentially with weights that do not converge

$$d_t = \sum_{i=1}^{\infty} \theta_i \epsilon_{t-1}^d$$  \hspace{1cm} (A.0.4.9)$$

then

$$\eta_t^d = \epsilon_t^d (1 - \rho) \sum_{j=0}^{\infty} \rho^j \theta_{j+1}$$  \hspace{1cm} (A.0.4.10)$$

and

$$V_{t-1}(\eta_t^d) = V_{t-1}(\epsilon_t^d) - \left[ \sum_{j=0}^{\infty} \theta_{j+1} \rho^j (1 - \rho) \right]^2$$  \hspace{1cm} (A.0.4.11)$$

We have that, time variation arises simply from substituting volatility in the innovation for dividends. This is the conditional variance of returns when there is no predictability in expected returns. When the memory of the dividend process is long, the effect is essential and the volatility is greater. We can express the conditional covariance between two asset returns in exactly the same terms

$$Cov_{t-1}(r_1^1, r_1^2) = Cov_{t-1}(\eta_t^{d1}, \eta_t^{d2}) + Cov_{t-1}(\eta_t^{r1}, \eta_t^{r2}) - Cov_{t-1}(\eta_t^{d1}, \eta_t^{r2}) - Cov_{t-1}(\eta_t^{r1}, \eta_t^{d1})$$  \hspace{1cm} (A.0.4.12)$$
If the dividends are fixed-weight MA, as in (A.0.4.9), and expected returns are constant, we denote the parameters for each asset as $(\rho^1, \theta^1)$ and $(\rho^2, \theta^2)$

$$ov_{t-1}(r^1_t, r^2_t) = Cov_{t-1}(\epsilon^{d1}_t, \epsilon^{d2}_t)(1 - \rho^1)(1 - \rho^2)[\sum_{j=0}^{\infty} \theta^{j1}_1(\rho^1)^j][\sum_{j=0}^{\infty} \theta^{j2}_2(\rho^2)^j]$$  \(\text{(A.0.4.13)}\)

Clearly, by combining (A.0.4.11) and (A.0.4.13) we state the conditional correlation

$$corr_{t-1}(r^1_t, r^2_t) = corr_{t-1}(\epsilon^{d1}_t, \epsilon^{d2}_t)$$  \(\text{(A.0.4.14)}\)

Thus, returns are correlated due to the fact that news is correlated. In the simple case above, they are equal. We form a more general expression, from the relation (A.0.4.7), for the covariance matrix of returns. Denote $\eta$ as the vector of innovation, due to expected returns or dividend, and $r$ as the vector of asset returns

$$r_t - E_{t-1}(r_t) = \eta^d_t - \eta^r_t$$  \(\text{(A.0.4.15)}\)

The covariance matrix is now given by

$$Cov_{t-1}(r_t) = Cov_{t-1}(\eta^d_t) + Cov_{t-1}(\eta^r_t) - Cov_{t-1}(\eta^d_t, \eta^r_t) - Cov_{t-1}(\eta^r_t, \eta^d_t)$$  \(\text{(A.0.4.16)}\)

Hence, correlation will result either from correlations between risk premiums, correlation between dividend news events or expected returns. For monthly data, the largest part of unconditional variance is the expected return for stock returns (Campbell and Ammar 1993). Additionally, this is also the largest part of the correlation between US stocks and UK stocks (Ammer and Mei 1996).
Appendix B

Alternative Models for Forecasting

B.1 Forecasting and Modeling Volatility

The idea of modeling and forecasting volatilities through univariate econometric time series was first introduced by Engle (1982). Ever since the first paper, several attempts have been made to model multivariate GARCH models Engle et al. (1984), Bollerslev et al (1988;1994), Engle Mezrich (1996), Bauwens et al. (2006), Silvennoinen Terasvirta (2008). However, before these breakthroughs, there have been a range of models developed for time varying correlations based on exponential smoothing and historical data windows. Numerous of these models are used today and present essential insights into the key features required for a correlation model. Practitioners in this field tend to look for parameterizations for the covariance matrix of a set of data containing random variables. It is conditional on a set of observed state variables. These are in general taken to be the precedent filtration of the dependent variable. One wants to estimate and parameterize

\[ H_t = V_{t-1}(r_t) \]  

where \( r \) is the vector of returns for any given asset. Thus, the conditional correlation between asset \( i \) and \( j \) is

\[ \rho_{i,j,t} = \frac{E_{t-1}((r_{i,t} - E_{t-1}(r_{i,t}))(r_{j,t} - E_{t-1}(r_{j,t})))}{\sqrt{V_{t-1}(r_{i,t})V_{t-1}(r_{j,t})}} \]

\[ = \frac{H_{i,j,t}}{\sqrt{H_{i,i,t}H_{j,j,t}}} \]  

The conditional correlation matrix and the variance matrix are

\[ R_t = D_t^{-1}H_tD_t^{-1} \text{ and } D_t^2 = diag[H_t] \]  

These relations imply the well-known covariance matrix

\[ H_t = D_tR_tD_t \]
Clearly, $H$ and $R$ are positive definite when there are no linear dependencies in the random variable. During the parameterization for this problem, one should assure that all correlation and covariance matrices are positive definite. Thus, that all volatilities are non-negative. These matrices are in fact stochastic processes and need therefore simply to be positive definite with probability one. Hence, all past covariance matrices should be positive definite. If not, there exist linear combinations of $r$ that have negative or zero variances.

## B.2 Constant Conditional Correlation

Bollerslev (1990) introduced a class of multivariate GARCH models called constant conditional correlation (CCC). The main assumption in this model is that the conditional correlations between all assets are assumed to be time invariant. Hence, the covariance matrix is

$$H_{i,j,t} = \rho_{i,j} \sqrt{H_{i,i,t} H_{j,j,t}}$$

(B.2.0.5)

matrix notation gives us

$$H_t = D_t RD_t, D_t = \sqrt{\text{diag} H_t}$$

(B.2.0.6)

The conditional covariances are not allowed to move so much that the correlations between each pair of assets change. However, the variances of the returns, $y$, may follow any type of process. Unfortunately, one cannot make analytical multistep forecasts of the covariance, nevertheless approximations are available. Likewise, the unconditional covariance cannot be calculated accurately. This model is suited for large systems; the defined likelihood function is simply estimated with a two step approach. The approach is that, one estimates the univariate models followed by computing of the sample correlation among the standardized residuals. As mentioned earlier in this paper, correlations tend to vary in time and the CCC model cannot incorporate this fact.

## B.3 Orthogonal GARCH

Assume that the nonsingular linear-combination of the variables has a CCC structure and that the correlation matrix is the identity

$$R = I$$

(B.3.0.7)

Let $P$ be a matrix such that

$$V_{t-1}(Pr_t) = D_t RD_t$$

(B.3.0.8)

$$V_{t-1}(r_t) = P^{-1} D_t RD_t P^{-1}$$

(B.3.0.9)
Pr is a CCC model although r is not. The conditional correlation matrix is

\[
corr_{t-1}(r_t) = \frac{P^{-1}D_tRD_tP^{-1}}{\sqrt{\text{diag}(P^{-1}D_tRD_tP^{-1})\text{diag}(P^{-1}D_tRD_tP^{-1})}} \tag{B.3.0.10}
\]

the conditional correlation is time varying. However, if the matrices D and P commute, the matrix will be independent on D_t, which make it time invariant. Moreover, D and P will only commute when P is itself diagonal. Thus, \((B.3.0.8)\) is in fact a generalization, where the parameters of P potentially estimable. It is common to specify that P is triangular. P can be chosen such that the unconditional covariance matrix of the linear combinations is diagonal, without loss of generality. This is called Cholesky factorization of the covariance matrix. This can be computed with least squares regressions. Namely, regress r_2 on r_1 and r_3 on r_2 and r_1, and so on. By construction, the residuals are uncorrelated across the equations. This model can now be expressed as

\[
V(Pr_t) = \Lambda, \; \Lambda \sim \text{diagonal} \tag{B.3.0.11}
\]

with the underlying assumption that the conditional covariance matrix is diagonal with univariate GARCH models for every series. Equivalently, this is the same thing as saying that the linear combination of the variables has a CCC structure. Due to the crucial fact that the unconditional covariance matrix is diagonal, we have that R = I. Every residual series computed in \((B.3.0.11)\) is taken to be a GARCH process. Its conditional variance is estimated as

\[
V_{t-1}(Pr_t) = G_t^2 \tag{B.3.0.12}
\]

\[
G_t = \text{diag}(\sqrt{h_{1,t}}, \sqrt{h_{2,t}}, \ldots, \sqrt{h_{n,t}}) \tag{B.3.0.13}
\]

\[
h_{i,t} \sim \text{GARCH}, \; i = (1, \cdots, n) \tag{B.3.0.13}
\]

Then we have that

\[
V_{t-1}(r_t) = P^{-1}G_t^2P^{-1} \tag{B.3.0.14}
\]

One version of this model Alexander (2002), Alexander and Barbosa (2008) is known as the Orthogonal GARCH (OGARCH). OGARCH assumes that every diagonal conditional variance is a univariate GARCH model. This model has a different choice of P, where P^{-1} is the matrix of eigenvectors of the unconditional covariance matrix. Now we call the random variables Pr principal components of r

\[
V(r_t) = \Sigma = P^{-1}\Lambda P^{-1}, \; \Lambda \sim \text{diagonal} \tag{B.3.0.15}
\]

We assume that the conditional covariance matrix is

\[
V_{t-1}(Pr_t) = G_t^2 \tag{B.3.0.16}
\]

\[
V_{t-1}(r_t) = P^{-1}G_t^2P^{-1} \tag{B.3.0.17}
\]
where every component is following a univariate GARCH. Naturally, the unconditional covariance matrix is

$$V(r_t) = P^{-1} E(G_t^2) P^{-1} = P^{-1} \Lambda P^{-1} = \Sigma$$  \hspace{1cm} (B.3.0.18)

We can, as Alexander proposes, transform \( r_t \) so that it has a unit variance. This results in that the principal components and eigenvectors are computed from the unconditional correlation matrix. We can find and express the eigenvector of the correlation matrix as \((P^{-1})\). Thus,

$$\tilde{D} \equiv \sqrt{\text{diag} \Sigma}$$  \hspace{1cm} (B.3.0.19)

$$V(D^{-1}r_t) \equiv \tilde{R} = \tilde{P}^{-1} \tilde{\Lambda} \tilde{P}^{-1}$$  \hspace{1cm} (B.3.0.20)

It is then assumed that,

$$V_{t-1}(\tilde{P} \tilde{D}^{-1}r_t) = \tilde{G}_t^2 \text{ where } \tilde{G}_t^2 \sim \text{diagonal}$$  \hspace{1cm} (B.3.0.21)

and have components that are univariate GARCH models. Thus, the conditional covariance matrix

$$V_{t-1}(r_t) = \tilde{D} \tilde{P}^{-1} \tilde{G}_t^2 \tilde{P}^{-1} \tilde{D}$$  \hspace{1cm} (B.3.0.22)

\( P \) or \( \tilde{P} \) as well as \( \tilde{D} \) are computed from the data. According to (B.3.0.16) or (B.3.0.21), one can estimate the variance of each of the principal components. As in CCC, this is a two step process. Simply extract the principal components from \( S \) followed by estimating univariate models for each of these.

(B.3.0.23)
Appendix C

Asynchronous and Synchronous log-returns

Figure C.1: U.S. SP 500
Figure C.2: Sweden OMXS30

Figure C.3: Hong Kong HSCE
Figure C.4: Japan NIKKEI225

Figure C.5: United Kingdom FTSE100
Figure C.6: Germany DAX30
Bibliography


46


[41] Sheppard, K. *Multi-Step estimation of Multivariate GARCH models.*


47