KASSI, AYMARD N’ZI. Essays in Modeling of Daily Returns and Realized Volatility. (Under the direction of Denis Pelletier.)

The dissertation presents three essays in the joint modeling of returns and realized volatility. The common characteristic of the three essays is the use of measurements of realized volatility, computed with high frequency data, to improve the fit and forecast of volatility models for daily returns. The relatively easy access to high frequency data makes this approach interesting as intraday data carry much more information about the dynamics of short term variance than low frequency (daily) data does.

The first essay introduces a new multivariate conditional volatility model for returns that utilizes realized covariance matrices. The model decomposes the conditional and realized covariance matrices into standard deviation and correlation matrices. On a first level, the univariate variances are estimated by a modified Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model that exploits intraday information. On the second level the conditional correlation matrices follow a regime switching Markov process. The inference about the regimes and the regime-switching correlations exploits the information contained in the realized correlation. An empirical application shows the ease of estimation and a forecasting exercise shows superior predictive ability when high-frequency information is incorporated.

The second essay puts forward a multivariate extension of a modified GARCH model that incorporates realized measure of covariance. The modification of the multivariate GARCH consists in replacing the outer product of past returns by the past realized covariance matrix. The curse of dimensionality is alleviated by a decentralization of the estimation procedure. The estimation process is computationally easier because there is no need to invert large covariance matrices and can therefore be applied to large data set.

The previous two chapters establish the gain in statistical accuracy of including intraday information in the modeling and forecasting of conditional covariance. The last essay presents empirical evidence on the economic gain to an investor in a situation of portfolio allocation. An empirical exercise presents the economic value of the access to intraday information for an agent who faces a simple asset allocation problem.

As an introduction to the dissertation, chapter one presents the methods used throughout the dissertation to calculate the realized covariance matrices and to compare the different predictions made by the models proposed.
Essays in Modeling of Daily Returns and Realized Volatility

by

Aymard N’Zi Kassi

A dissertation submitted to the Graduate Faculty of North Carolina State University in partial fulfillment of the requirements for the Degree of Doctor of Philosophy

Economics

Raleigh, North Carolina

2015

APPROVED BY:

Mehmet Caner

Walter Thurman

Barry Goodwin

Denis Pelletier

Chair of Advisory Committee
DEDICATION

A toi Maman...
BIOGRAPHY

Aymard N’Zi Kassi was born in the Ivory Coast. He moved to Raleigh in August of 2010 to pursue a Ph.D. in Economics.
ACKNOWLEDGEMENTS

I would like to thank my advisor Dr. Denis Pelletier for his advice. I also thank my committee members, Dr. Walter Thurman, Dr. Mehmet Caner and Dr. Barry Goodwin, for their helpful comments and advice.
## TABLE OF CONTENTS

### LIST OF TABLES

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Introduction and Preliminaries</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Preliminaries</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Multivariate Realized Kernel</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Statistical Forecast Evaluation</td>
<td>5</td>
</tr>
</tbody>
</table>

### LIST OF FIGURES

- viii

### Chapter 1 Introduction and Preliminaries

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>1</td>
</tr>
<tr>
<td>1.2</td>
<td>3</td>
</tr>
<tr>
<td>1.2.1</td>
<td>3</td>
</tr>
<tr>
<td>1.2.2</td>
<td>5</td>
</tr>
</tbody>
</table>

### Chapter 2 The Realized RSDC Model

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>7</td>
</tr>
<tr>
<td>2.2</td>
<td>8</td>
</tr>
<tr>
<td>2.2.1</td>
<td>9</td>
</tr>
<tr>
<td>2.2.2</td>
<td>9</td>
</tr>
<tr>
<td>2.2.3</td>
<td>11</td>
</tr>
<tr>
<td>2.3</td>
<td>12</td>
</tr>
<tr>
<td>2.3.1</td>
<td>13</td>
</tr>
<tr>
<td>2.3.2</td>
<td>13</td>
</tr>
<tr>
<td>2.3.3</td>
<td>15</td>
</tr>
<tr>
<td>2.3.4</td>
<td>16</td>
</tr>
<tr>
<td>2.4</td>
<td>17</td>
</tr>
<tr>
<td>2.4.1</td>
<td>17</td>
</tr>
<tr>
<td>2.4.2</td>
<td>18</td>
</tr>
<tr>
<td>2.5</td>
<td>20</td>
</tr>
<tr>
<td>2.6</td>
<td>22</td>
</tr>
<tr>
<td>2.7</td>
<td>23</td>
</tr>
</tbody>
</table>

### Chapter 3 The Multivariate GARCHX

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>39</td>
</tr>
<tr>
<td>3.2</td>
<td>40</td>
</tr>
<tr>
<td>3.2.1</td>
<td>40</td>
</tr>
<tr>
<td>3.3</td>
<td>42</td>
</tr>
<tr>
<td>3.3.1</td>
<td>42</td>
</tr>
<tr>
<td>3.3.2</td>
<td>44</td>
</tr>
<tr>
<td>3.3.3</td>
<td>45</td>
</tr>
<tr>
<td>3.3.4</td>
<td>45</td>
</tr>
<tr>
<td>3.4</td>
<td>45</td>
</tr>
<tr>
<td>3.4.1</td>
<td>46</td>
</tr>
<tr>
<td>3.4.2</td>
<td>47</td>
</tr>
<tr>
<td>3.4.3</td>
<td>47</td>
</tr>
<tr>
<td>3.5</td>
<td>48</td>
</tr>
<tr>
<td>3.5.1</td>
<td>48</td>
</tr>
</tbody>
</table>
### Chapter 4 The Economic Value of Intraday Data in the Modeling of Conditional Covariance

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1 Introduction</td>
<td>62</td>
</tr>
<tr>
<td>4.2 Methodology</td>
<td>63</td>
</tr>
<tr>
<td>4.2.1 The General Problem</td>
<td>63</td>
</tr>
<tr>
<td>4.2.2 The Expected Return</td>
<td>63</td>
</tr>
<tr>
<td>4.2.3 The Conditional Covariance Matrices</td>
<td>64</td>
</tr>
<tr>
<td>4.2.4 The Economic Measure</td>
<td>64</td>
</tr>
<tr>
<td>4.3 Empirical Results</td>
<td>65</td>
</tr>
<tr>
<td>4.3.1 Sharpe Ratio and Portfolio Volatility</td>
<td>66</td>
</tr>
<tr>
<td>4.3.2 The Value of High-Frequency Information</td>
<td>68</td>
</tr>
<tr>
<td>4.4 Conclusion</td>
<td>68</td>
</tr>
<tr>
<td>4.5 Tables and Figures</td>
<td>70</td>
</tr>
</tbody>
</table>

### Chapter 5 Conclusion

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>81</td>
</tr>
</tbody>
</table>

### BIBLIOGRAPHY

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>83</td>
</tr>
</tbody>
</table>

### APPENDIX

<table>
<thead>
<tr>
<th>Appendix</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Appendix</td>
<td>88</td>
</tr>
<tr>
<td>A.1 Proof of Eq. 2.17 and 2.18</td>
<td>88</td>
</tr>
<tr>
<td>A.2 Proof of Eq. 2.19 and 2.20</td>
<td>90</td>
</tr>
<tr>
<td>A.2.1 Notations</td>
<td>90</td>
</tr>
<tr>
<td>A.2.2 Proof of Eq. 2.19</td>
<td>90</td>
</tr>
<tr>
<td>A.2.3 Proof of Eq. 2.20</td>
<td>92</td>
</tr>
<tr>
<td>A.3 Details of the Probabilities Computation</td>
<td>93</td>
</tr>
<tr>
<td>A.4 Proof of Eq. 2.28</td>
<td>94</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

Table 2.1 : Summary Statistics for Refresh Sampling: January 03, 2006 to April 30, 2012 .................................................. 23
Table 2.2 : Summary Statistics for the Daily Log Returns and the Annualized Realized Volatilities: January 03, 2006 to April 30, 2012 .......................................................... 23
Table 2.3 : Summary Statistics for the Realized Correlation Coefficients: January 03, 2006 to April 30, 2012 .......... 24
Table 2.4 : Parameter Estimates for the Univariate Realized GARCH Model .................................................. 25
Table 2.5 : Persistence of the Conditional Variances Implied by the R-RSDC model .................................................. 25
Table 2.6 : Correlation Estimates Regime 1 .................................................. 26
Table 2.7 : Correlation Estimates Regime 2 .................................................. 26
Table 2.8 : Percentage Change in Estimated Correlation Coefficients Between Regime 1 and 2 .................................................. 27
Table 2.9 : Transition Probabilities Estimates .................................................. 27
Table 2.10 : Statistical Test of Predictive Ability: Forecast Horizons = 1, 5, 10 and 22 days .................................................. 27

Table 3.1 : Summary Statistics for Refresh Sampling: January 03, 2006 to April 30, 2012 .................................................. 51
Table 3.2 : Summary Statistics for the Daily Log Returns and the Annualized Realized Volatilities: January 03, 2006 to April 30, 2012 .......................................................... 52
Table 3.3 : Average Values of the Parameters Estimates for the MGARCHX Model .................................................. 52
Table 3.4 : Statistical Test of Predictive Ability: Forecast Horizons = 1, 2, 3, 5, 10 and 22 days .................................................. 53
Table 3.5 : VaR Back-Testing Results for the MGARCHX and MHEAVY Models .................................................. 53
Table 3.6 : Parameter Estimates of the “A” Matrix for the MGARCHX Model .................................................. 57
Table 3.7 : Parameter Estimates of the “B” Matrix for the MGARCHX Model .................................................. 58
Table 3.8 : Parameter Estimates of the “C” Matrix for the MGARCHX Model .................................................. 59
Table 3.9 : Parameter Estimates of the “M” Matrix for the MGARCHX Model .................................................. 60
Table 3.10 : Parameter Estimates of the “QQ’ ” Matrix for the MGARCHX Model .................................................. 61

Table 4.1 : Summary Statistics for the Daily Log Returns : In-Sample and Out-of-Sample Period .................................................. 70
Table 4.2 : Median Contribution to Total Portfolio Risk of a Single Risky Asset .................................................. 71
LIST OF FIGURES

Figure 2.1 : Data Retention Percentages .................................................. 28
Figure 2.2 : Annualized Realized Volatility for AXP, BAC, COF and GS - January
03, 2006 to April 30, 2012 ........................................................................ 29
Figure 2.3 : Annualized Realized Volatility for JPM, MS and USB - January 03, 2006
to April 30, 2012 ...................................................................................... 30
Figure 2.4 : Annualized Realized Volatility for AXP, BAC, COF and GS - September
1, 2008 to June 1, 2009 ............................................................................. 31
Figure 2.5 : Annualized Realized Volatility for JPM, MS and USB - September 1,
2008 to June 1, 2009 .................................................................................. 32
Figure 2.6 Smoothed Probability of Occurrence of Regime 1 or 2 - January 03, 2006
to April 30, 2012 ....................................................................................... 33
Figure 2.7 : Filtered Probability of Occurrence of Regime 1 or 2 - January 03, 2006
to April 30, 2012 ......................................................................................... 34
Figure 2.8 Smoothed Probability of Occurrence of Regime 1 or 2 - September 1, 2008
to June 1, 2009 ........................................................................................... 35
Figure 2.9 : Filtered Probability of Occurrence of Regime 1 or 2 : September 1, 2008
to June 1, 2009 ........................................................................................... 36
Figure 2.10 : Realized and Conditional covariance and correlation JPM and BAC -
January 03, 2006 to April 30, 2012 ............................................................. 37
Figure 2.11 : Realized and Conditional covariance and correlation JPM and BAC -
September 1, 2008 to June 1, 2009 .............................................................. 38

Figure 3.1 : Data Retention Percentages ...................................................... 54
Figure 3.2 : 1% Portfolio Value-at-Risk Standardized Exceedences ................. 55
Figure 3.3 : Actual Portfolio Returns and 1% Value-at-Risk from the MGARCHX
and MHEAVY Models ................................................................................ 55
Figure 3.4 : 5% Portfolio Value-at-Risk Standardized Exceedences ................. 56
Figure 3.5 : Actual Portfolio Returns and 5% Value-at-Risk from the MGARCHX
and MHEAVY Models ................................................................................ 56

Figure 4.1 : Time Evolution of the Bayes-Stein Estimator of the Expected Return in
% : AXP and BAC ....................................................................................... 71
Figure 4.2 : Time Evolution of the Bayes-Stein Estimator of the Expected Return in
% : CVX and KO ........................................................................................ 72
Figure 4.3 : Time Evolution of the Bayes-Stein Estimator of the Expected Return in
% : DD and GE ........................................................................................... 72
Figure 4.4 : Time Evolution of the Bayes-Stein Estimator of the Expected Return in
% : IBM and JPM ........................................................................................ 73
Figure 4.5 : Time Evolution of the Bayes-Stein Estimator of the Expected Return in
% : MSFT and XOM .................................................................................... 73
Figure 4.6 : Optimal Portfolio Weights in % Construced with the MGARCHX Model:
AXP, BAC and the Risk-Free Rate Asset ..................................................... 74
Figure 4.7: Optimal Portfolio Weights in % Constructed with the MGARCHX Model: CVX, KO and the Risk-Free Rate Asset
Figure 4.8: Optimal Portfolio Weights in % Constructed with the MGARCHX Model: DD, GE and the Risk-Free Rate Asset
Figure 4.9: Optimal Portfolio Weights in % Constructed with the MGARCHX Model: IBM, JPM and the Risk-Free Rate Asset
Figure 4.10: Optimal Portfolio Weights in % Constructed with the MGARCHX Model: MSFT, XOM and the Risk-Free Rate Asset
Figure 4.11: Forecasted Portfolio Annualized Standard Deviation in %
Figure 4.12: Density of the Forecasted Portfolio Annualized Standard Deviation
Figure 4.13: Ex-post Portfolio Annualized Standard Deviation in % Calculated with Outer Product of Vector of Actual Returns
Figure 4.14: Density of the Ex-post Portfolio Annualized Standard Deviation Calculated with Outer Product of Vector of Actual Returns
Figure 4.15: Ex-post Portfolio Annualized Standard Deviation in % Calculated with the Realized Kernel
Figure 4.16: Density of the Ex-post Portfolio Annualized Standard Deviation Calculated with the Realized Kernel
Figure 4.17: Portfolio Expected Mean and Variances
Figure 4.18: Sensitivity of Cost to Target Return and Relative Risk Aversion
Chapter 1

Introduction and Preliminaries

1.1 Introduction

Easy access to intraday data in financial econometrics has created the need for models that can exploit efficiently large high-frequency information. One potential advantage of higher frequency data in finance is a better understanding and a better forecasting of time-varying covariance matrices. Understanding and predicting covariances between assets or between markets have important implications in areas such as portfolio allocation, asset pricing and risk management. Since the important work of Engle [1982] and Bollerslev [1986], extensive research in the area of univariate and multivariate volatility models for low-frequency (daily) returns has been undertaken.

In recent years however, starting with Andersen and Bollerslev [1998], the focus of volatility modeling has shifted from returns computed with low-frequency observations to intraday returns. Under some general assumptions, fairly precise measurement of daily volatility, called realized volatility, can be estimated by non-parametric methods.

There exists a considerable amount of work on issues related to these estimators. The asymptotic properties [e.g., Barndorff-Nielsen and Shephard [2002]], the optimal sampling frequency [e.g., Bandi and Russell [2008]], the robustness to jumps [e.g., Barndorff-Nielsen and Shephard [2004]] and the robustness to market microstructure noise [e.g., Hansen and Lunde [2006]] have all been extensively studied. However, the focus has mostly been on the measurement of volatility, less so on its modeling and forecasting.

Recently and on the univariate level, Shephard and Sheppard [2010] and Hansen et al. [2012] have introduced frameworks that take advantage of the information contained in intraday data. In both studies, the conditional variance is specified as a modified version of the Generalized
1.1. INTRODUCTION

Autoregressive Conditional Heteroskedasticity (GARCH) model of Bollerslev [1986] with the modification consisting in replacing the lagged square return by the lagged value of the realized variance. They have shown that the use of high-frequency information improves the reaction time of the forecasted variances to rapid changes in financial markets. The improvement results from the fact that the realized volatility is a better proxy for the true and unobserved second moment of the returns. In both specifications, the evolution of realized variance is also modeled with the objective of forecasting daily conditional volatility at horizon greater than one period ahead. While Shephard and Sheppard [2010] use the Multiplicative Error Model (MEM) of Engle [2002a] to describe the dynamics of the conditional mean of the realized measures, Hansen et al. [2012] supplement the conditional variance specification with a measurement equation that links linearly the realized and the conditional volatility.

High-frequency information has also been applied in a multivariate setting. Gourieroux et al. [2009] introduce the Wishart autoregressive (WAR) model in which they assume that the realized covariance follows a non-central Wishart distribution. The non-centrality parameter accommodates temporal dependence in the scale matrix. Jin and Maheu [2012], however propose to jointly model the returns and the realized covariances and assume that the realized covariance matrix follows a Wishart distribution with a time varying scale matrix. Noureldin and Sheppard [2012] extend Shephard and Sheppard [2010] to a vector of asset returns using multivariate MEM. Once again, the assumption of the Wishart distribution is made for the conditional and the realized covariance matrices. Finally, Hansen et al. [2014] propose a multivariate extension of Hansen et al. [2012]. The marginal model for each asset is the same as in their univariate specification and the correlation between assets is assumed to be determined exclusively by the correlation between each asset and an observed index factor (the S&P 500 for instance).

This dissertation has two objectives. The first one is the introduction of two frameworks to incorporate intraday information into the forecasting of daily conditional covariance matrix. The second is to measure the economic benefit of high-frequency information in portfolio allocation.

The first goal is challenging because of two constraints. On the one hand, any multivariate model of volatility is required to produce positive definite matrices of covariance while on the other, the number of the parameters to estimate increases rapidly with the number of assets included in the analysis. Various solutions to these problems have been proposed. These include parameter restrictions in straightforward multivariate extensions of the GARCH model [e.g. Bollerslev et al. [1988], Engle and Kroner [1995]]. It is also possible to model just a few factors believed to be the main drivers of the covariance dynamic [e.g Engle and Ng [1990]]. Finally, some authors have decomposed the covariance matrix allowing thus to separately model the dynamic of the conditional variances and the correlations [e.g. Bollerslev [1990], Engle [2002b],
1.2. PRELIMINARIES

The first model, proposed in Chapter 2, makes use of the latter solution. On a first level, the univariate variances are estimated by a modified GARCH specification that exploits intraday information. On the second level the conditional correlation matrices follows a regime switching Markov process. The inference about the regimes and the regime-switching correlations takes advantage of the information contained in the realized correlation.

The third Chapter puts forward a second model that also incorporates a realized measures of covariance. The difference with the first model is that we do not decompose the covariance matrix and the univariate modified GARCH is simply extended to a portfolio of assets. In this case, the modification of the multivariate GARCH consists in replacing the outer product of past returns by the lagged realized covariance matrix. The curse of dimensionality is mitigated by a decentralization of the estimation procedure. The estimation process is computationally easier because there is no need to invert large covariance matrices and can therefore be applied to a large data set.

The previous two chapters established the gain in statistical accuracy of including intraday information in the modeling and forecasting of conditional covariance. The fourth chapter presents empirical evidence on the economic gain for an investor in a situation of portfolio allocation.

The common characteristic of the three chapters is the use of measurements of realized volatility computed from high-frequency data. Thus, for ease of exposition, the methodology adopted to calculate them is presented in the next section. The statistical method of forecast evaluation generated by the different models, which is identical in the next two chapters, is also explained in the next section.

1.2 Preliminaries

1.2.1 Multivariate Realized Kernel

Notation and Assumptions

Consider the n-dimensional vector of efficient intraday log prices, modeled as a Brownian semimartingale on a filtered probability space \((\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})\):

\[
P(t) = \int_0^t a(s)ds + \int_0^t \sigma(s)dB(s),
\]

(1.1)
where \( a(t) \) and \( \sigma(t) \) are the predictable processes for the drift and the instantaneous volatility. \( B(t) \) is a vector of independent Brownian motions. The integrated covariance process is then

\[
[P](t) = \int_0^t \sigma(s)\sigma(s)'ds.
\] (1.2)

Consider the quadratic covariation, \( v_n(t) = \sum_{i=0}^{n-1} (P(t^n_{i+1}) - P(t^n_i))(P(t^n_{i+1}) - P(t^n_i))' \) for a partition \( \{t^n_i\} \) of \([0,t]\). It can be shown that \( v_n(t) \) converges in probability to \([P](t)\). See for example Klebaner [2005, Section 8.5] and Protter [2004, Chapter 2, Section 6]. Therefore by increasing the sampling frequency we can approximate the integrated covariance by the quadratic covariation. However, microstructure frictions render imprecise the estimation of the realized covariance when the sampling frequency increases. If we denote the noise by \( MN \), the observed price at a given \( \tau_i \) is:

\[
\tilde{P}(\tau_i) = P(\tau_i) + MN(\tau_i),
\] (1.3)

Thus, the observed price \( \tilde{P} \), is the noisy measurement of the efficient price and the quadratic covariation is no longer a consistent estimator for the latent integrated covariances. Examples of microstructure noise include bid-ask bounces and discreteness of the price. Hansen and Lunde [2006] have found that the market microstructure noise is:

1. Correlated with the efficient price
2. Time-dependent
3. Small for DJIA stocks

In a multivariate framework, the measurement of realized covariance is complicated further because the observation are irregularly spaced and non-synchronuous causing a bias toward zero of the covariances, the so called Epps effect [see Epps [1979]].

**The Realized Kernel**

Barndorff-Nielsen et al. [2011] introduce a framework to construct matrices of realized covariances that are robust to the asynchronous observations and the presence of noise. The multivariate realized kernel is defined as:

\[
X_t = \gamma_0 + \sum_{q=1}^{Q} K(q/Q) (\gamma_q + \gamma'_q)
\] (1.4)
1.2. PRELIMINARIES

Where, $K(q/Q)$ is the Parzen kernel function$^1$, $Q$ is the bandwidth and $\gamma_q$ the autocovariance at $q$ laggeds of the high-frequency synchronized returns.

The computation of the multivariate realized kernel starts by the synchronization of the high-frequency prices for each day in the sample. The sampling method, called the “refresh time sampling,” defines the time clock for the $i$th trade as the time at which the last asset in the sample is traded for that $i$th trade. More precisely, if $\tau_i$ is the refresh time then: $\tau_i = \max(t_1^{(1)}, t_1^{(2)}, ..., t_1^{(n)})$ and subsequent refresh times are $\tau_{j+1} = \max(t_{N_{r_j}+1}^{(1)}, t_{N_{r_j}+1}^{(2)}, ..., t_{N_{r_j}+1}^{(3)})$, where $N_{r_j}$ is a count of the transactions up to refresh time $\tau_j$. A measure of the data retention in this case is $p = \frac{nN}{\sum_{i=1}^{n} n^i}$, where $n$ is the number of assets, $n^i$ and $N$ are respectively the initial and synchronized number of transactions for the $i$th asset.

An important input of the multivariate realized kernel estimation is the bandwidth. The bandwidth parameter for each asset $Q^i$ is optimized with respect to the mean squared error criterion by setting $Q^i = c^*\xi^{4/5}N^{3/5}$, where $c^* = 3.5134$, $\xi^2 = \omega^2/\sqrt{IQ}$ denotes the noise-to-signal ratio, $\omega^2$ is a measure of microstructure noise variance and $IQ$ is the integrated quarticity as defined in Barndorff-Nielsen et al. [2011]. On a single day, the bandwidth $Q$, selected for the kernel estimation is the median of the bandwidths estimated for each asset. Notice that the bandwidth also corresponds to the number of autocovariances included in the computation of the realized matrices because the Parzen kernel is the weight function.

Often high-frequency data contain a considerable amount of errors in recording, therefore data cleaning is an important preliminary step of any volatility estimation based on intraday data.

Procedure for Cleaning the High-frequency Data

The data used in the dissertation consists of stock prices from the S&P 100. The analysis is restricted to the regular trading hours, that is from 9:30am to 4:00pm EST. Furthermore, entries with corrected trades are deleted. The transactions with a sale condition equal to $Z$ are also excluded because they are reported to the tape at a later time than their occurrences. Multiple transactions with the same time stamp are aggregated and the median price is used as the aggregated price.

1.2.2 Statistical Forecast Evaluation

The Diebold and Mariano [2002] and West [1996] test (DMW) given a loss function is used for the pair-wise comparison of the volatility forecasts. The quasi-likelihood (QLIK) is the chosen

\[ K(x) = \begin{cases} 1 - 6x^2 + 6x^3 & \text{if } 0 \leq x \leq 0.5, \\ = 2(1-x)^3 & \text{if } 0.5 \leq x \leq 1, \\ = 0 & \text{if } x > 1. \]

---

$^1$K(x) = 1 − 6x² + 6x³ if 0 ≤ x ≤ 0.5, = 2(1 − x)³ if 0.5 ≤ x ≤ 1, = 0 if x > 1.
loss function:

\[ L_{t,L}(H_0^{t+L}, H_i^{t+L|t}) = \ln |H_i^{t+L|t}| + \text{tr} \left( (H_i^{t+L|t})^{-1} H_0^{t+L} \right), \quad \forall \quad i = 1, 2, \quad (1.5) \]

where the index \( i = 1, 2 \) represents the models to compare. \( H_0^{t+L} \) is the true and unobserved covariance matrix while \( H_i^{t+L|t} \) is the \( L \)-step-ahead forecast generated by the \( i \)th model. The QLIK function is robust in the sense of Patton and Sheppard [2009] and Laurent et al. [2013] because it preserves statistical ordering regardless of the proxy of the latent covariance matrix.

With a proxy for the latent covariance matrix and the forecasted conditional covariance matrices, the loss is calculated for each model and the loss differential is computed as:

\[ d_{t,L} = L_{t,L}(H_0^{t+L}, H_1^{t+L|t}) - L_{t,L}(H_0^{t+L}, H_2^{t+L|t}) \]

The direction of the difference implies that the forecast generated by the model 2 is more accurate than the one generated by the model 1 if \( d_{t,L} \) is positive. With the time series of \( d_{t,L} \), we test for all \( L \):

\[ H_0 : E[L_{t,L}(H_0^{t+L}, H_1^{t+L|t})] = E[L_{t,L}(H_0^{t+L}, H_2^{t+L|t})], \]

\[ \text{vs.} \quad H_1 : E[L_{t,L}(H_0^{t+L}, H_1^{t+L|t})] > E[L_{t,L}(H_0^{t+L}, H_2^{t+L|t})], \]

\[ H_2 : E[L_{t,L}(H_0^{t+L}, H_1^{t+L|t})] < E[L_{t,L}(H_0^{t+L}, H_2^{t+L|t})]. \]

Therefore the DMW tests the null of equal predictive accuracy against the alternative that one of the model performs better. The test is computed using a standard t-test after accounting for serial correlation. Under the null, Diebold and Mariano [2002] show that the test is asymptotically Normally distributed. Practically, the differential is regressed on a constant and a t-test is constructed with heteroskedasticity and autocorrelation consistent estimate of the standard errors.

Comparison of forecasts in a multivariate context has the additional issue of interpretation. The reason is the difficulty to separate the effect of the forecasted variances and covariances on the predictive ability of the model. Noureldin and Sheppard [2012] propose to decompose the QLIK in Eq. 1.5 into marginal and copula-style contribution. Thus, the marginal DMW gives the contribution of the forecast of the variance of an asset while the copula-style DMW gives the contribution of the correlation forecast.
Chapter 2

The Realized RSDC Model

2.1 Introduction

This chapter introduces a new and parsimonious way to generalize to multivariate data the Realized GARCH framework of Hansen et al. [2012]. The model is inspired by the decomposition of the covariance matrix adopted successfully for low frequency returns in the Constant Conditional Correlation (CCC) model of Bollerslev [1990] and the Dynamic Conditional Correlation (DCC) models of Engle [2002b] and Tse and Tsui [2002]. As indicated by their names, the CCC model assumes that the conditional correlations between assets are constant while the DCC describes a dynamic correlation matrix with an autoregressive process. The Regime Switching Dynamic Correlations (RSDC) model of Pelletier [2006] is an intermediary step because it assumes a Markov switching process with correlations that are different across regimes but constant within a particular regime. In the latter model, the transition probability matrix governs the persistence of the correlations. The model, the Realized RSDC (R-RSDC), builds on Pelletier [2006] by assuming regime-dependent conditional correlations matrices. Furthermore, the realized covariance matrix is assumed to follow a mixture of Wishart distributions.

Our model presents three main benefits. First the information set for inference of the regime switching correlations is richer than in Pelletier [2006], potentially allowing for a better description of the correlations and covariances. Second, the method of estimation, the Expectation-Maximization algorithm introduced by Dempster et al. [1977], breaks the curse of dimensionality in that the correlation matrices can be estimated through simple weighted sums. Third, the idea of describing the market through different regimes allows us to give economic interpretation to the overall levels of correlation. It is possible for instance to tie a general increase of correlations between assets to particular events in the market such as a financial crisis.
2.2. MODEL

This chapter is organized as follows. Section 2 introduces the R-RSDC model. The estimation procedure is discussed in Section 3. Section 4 presents an empirical application with transaction-level data for seven stocks. Section 5 evaluates the forecasting ability of the model and the concluding remarks are given in Section 6.

2.2 Model

Let $Y_t$ be the $n \times 1$ vector of daily demeaned log-returns such that:

$$
Y_t = H_t^{1/2}U_t, \\
E(U_t) = 0, \\
Var(U_t) = I_n,
$$

(2.1)

where $H_t^{1/2}$ is a $n \times n$ positive definite matrix, $I_n$ is the identity matrix and $0$ is a $n \times 1$ vector of zeros. Conditional on $\mathcal{F}_{t-1}$, the $t-1$ information set, the variance of $Y_t$ is:

$$
Var(Y_t|\mathcal{F}_{t-1}) = Var_{t-1}(Y_t) = H_t.
$$

(2.2)

The conditional covariance matrix $H_t$ can be decomposed into standard deviation and correlation matrices, that is:

$$
H_t = D_t \Gamma_t D_t,
$$

(2.3)

where $D_t = diag(\sqrt{h_{1,t}}, \sqrt{h_{2,t}}, ..., \sqrt{h_{n,t}})$ and $\Gamma_t$ is the correlation matrix. Therefore the sequence of daily conditional variances for the $i^{th}$ asset is $\{h_{i,t}\}_{t=1}^T$, where $T$ is the time of the last observation.

Contrary to the basic GARCH model where the information set is composed exclusively of the daily log-returns, the R-RSDC model also exploits the information from the corresponding daily realized covariance matrices $\{X_t\}_{t=1}^T$ computed with intraday returns.

In what follows, the R-RSDC model proposes a dynamic for $\{H_t\}_{t=1}^T$ and $\{X_t\}_{t=1}^T$ assuming multivariate Gaussian distribution for the innovation, that is $U_t \sim i.i.d.N(0, I_n)$ and Wishart distribution for $X_t$. 
2.2. MODEL

2.2.1 The Conditional Variances

With the set up above, the dependence of the squared elements of $D_t$ conditioning on the past information of the daily returns and realized variances is specified as the following modified GARCH:

$$h_{i,t} = \omega_i + \beta_i h_{i,t-1} + \alpha_i x_{i,t-1}. \quad (2.4)$$

As Eq. 2.4 shows, $x_{i,t-1}$, the lagged realized variance replaces the lagged square return from the usual GARCH specification. The objective of the substitution is to improve the dynamic of the conditional variance. This can be seen by solving forward for $h_{i,t}$:

$$h_{i,t} = \frac{\omega_i}{1 - \beta_i} + \sum_{j=0}^{\infty} [\alpha_i \beta_i^j] x_{i,t-1-j}. \quad (2.5)$$

The past values of the realized variance drive the dynamic of the conditional variance. The former, being a better proxy for the latent variance (Lopez [2001]), improves the dynamic of the latter. This is the underlying idea of the recent univariate models proposed by Shephard and Sheppard [2010] and Hansen et al. [2012]. There is however an issue if one only specifies the dynamic of the conditional variance without explicitly modeling the evolution of the daily realized variances. That is, it is not possible to forecast at an horizon greater than one-period ahead. Therefore, as in Shephard and Sheppard [2010] and Hansen et al. [2012], we also propose a specification for the daily measures of realized volatility.

2.2.2 The Realized Variances

First, the realized covariances are assumed to follow a Wishart distribution with degree of freedom, $\nu$ and time varying scale matrix $\Sigma_t/\nu$. The implication of the Wishart assumption is that $E(X_t | \mathcal{F}_{t-1}) = \Sigma_t$. Next, let us decompose the scale matrix into standard deviation and correlation matrices. Finally, we extend the idea of the Regime Switching Dynamic Correlation of Pelletier [2006] to the correlation matrix implied by $\Sigma_t$. 

9
More precisely we have the relations:

\[ X_t \sim W_n(\nu, \Sigma_t / \nu), \]
\[ \Sigma_t = \tilde{D}_t \Gamma_t \tilde{D}_t, \]
\[ \Gamma_t = \sum_{i=1}^{M} \Gamma^{(i)} \mathbb{1}_{\{s_t = i\}}, \quad \Gamma^{(i)} \neq \Gamma^{(j)} \quad \forall \quad i \neq j, \quad (2.6) \]

where \( \tilde{D}_t \) is diagonal and holds the square root of the conditional mean of the realized variances. The correlation matrix \( \Gamma_t \) follows a regime switching dynamic with M regimes. The unobserved variable \( s_t \) represents the state at time \( t \) and the indicator function \( \mathbb{1}_{\{s_t = i\}} \) is equal 1 when the condition \( \{s_t = i\} \) is true and 0 otherwise. Let us denote by \( \Pi \) the associated transition probability matrix and by \( \pi_{i,j} \) the element on row \( j \) and column \( i \) of \( \Pi \). Then, \( \pi_{i,j} \) is the probability of transition from regime \( i \) in period \( t \) to regime \( j \) in period \( t + 1 \) and \( \pi_i \) is the limiting probability of regime \( i \). The standard assumptions on the Markov chain (aperiodicity, irreducibility and ergodicity) are made, see Ross [2006, Chapter 4].

Notice that the conditional correlation matrix provides a link at time \( t \) between the realized and conditional covariances. The conditional mean of the realized variance is modeled as a function of the conditional variance which is \( \mathcal{F}_{t-1} \) measurable:

\[ x_{i,t} = \tilde{h}_{i,t} \epsilon_{i,t}, \]
\[ \tilde{h}_{i,t} = \gamma_{i}^{(0)} + \gamma_{i}^{(1)} h_{i,t}, \quad (2.7) \]

where \( \tilde{h}_{i,t} \) is the conditional mean of the realized variance. The multiplicative error term, \( \epsilon_{i,t} \) is \( Ga(\nu/2, 2/\nu) \) and has a mean of 1. The Gamma distribution is a consequence of the Wishart distribution assumption made for the realized covariance matrix. Given Eq. 2.7, let \( \tilde{D}_t = \text{diag}(\sqrt{\tilde{h}_{1,t}}, \sqrt{\tilde{h}_{2,t}}, ..., \sqrt{\tilde{h}_{n,t}}) \) and the following decomposition is obtained:

\[ X_t = \tilde{D}_t \tilde{S}_t \tilde{D}_t, \quad (2.8) \]
\[ \tilde{S}_t \sim W_n(\nu, \Gamma_t / \nu). \quad (2.9) \]

Thus, by the properties of the Wishart distribution, the \( i^{th} \) diagonal element of \( X_t \) has a Gamma marginal distribution with time varying scale parameter equal to \( \tilde{h}_{i,t} \). The realized variance \( x_{i,t} \), is equal to \( \tilde{h}_{i,t} \) times the element \( (i, i) \) of a random matrix \( \tilde{S}_t \). Finally, \( E[\tilde{S}_t | \Gamma_t] = \Gamma_t \).
2.2.3 Comparison With Existing Models

On a univariate level, relationships between existing models and ours can be established. In the Realized GARCH model of Hansen et al. [2012], the contemporaneous link between the realized variance $x_{i,t}$ and the latent conditional variance $h_{i,t}$ of the daily return (neglecting a leverage effect term) is $x_{i,t} = \gamma_i^{(0)} + \gamma_i^{(1)} h_{i,t} + \epsilon_{i,t}$. Therefore as opposed to an additive error term, the error is multiplicative in the R-RSDC. Even though Shephard and Sheppard [2010] also implement a MEM, our distributional assumption is different.

Compared to the regular GARCH specification, the univariate level of the R-RSDC adds the intraday information to the lower frequency (daily) information set. It also adds variation in the persistence of the conditional variance. To see this, Eq. 2.4 and 2.7 can be combined to find an alternative expression for the conditional volatility. Replacing $x_{i,t-1}$ in Eq. 2.4 by its expression yields:

$$h_{i,t} = \left[ \omega_i + \alpha_i \gamma_i^{(0)} \epsilon_{i,t-1} \right] + \left[ \beta_i + \alpha_i \gamma_i^{(1)} \epsilon_{i,t-1} \right] h_{i,t-1}. \quad (2.10)$$

Thus, the measure of persistence of the conditional variance, $\beta_i + \alpha_i \gamma_i^{(1)} \epsilon_{i,t-1}$, varies with the innovations $\epsilon_{i,t-1}$. This flexibility allows the conditional variance to react faster than in a typical GARCH to a market shock. Given that $E(\epsilon_{i,t}) = 1$, the long run measure of persistence is $\beta_i + \alpha_i \gamma_i^{(1)}$ and the unconditional mean of the $i^{th}$ conditional and realized variances are:

$$E[h_i] = \frac{\omega_i + \alpha_i \gamma_i^{(0)}}{1 - [\beta_i + \alpha_i \gamma_i^{(1)}]} \quad \text{and} \quad (2.11)$$

$$E[x_i] = \gamma_i^{(0)} + \frac{\gamma_i^{(1)} (\omega_i + \alpha_i \gamma_i^{(0)})}{1 - [\beta_i + \alpha_i \gamma_i^{(1)}]}. \quad (2.12)$$

To our knowledge, the R-RSDC is unique in the literature of multivariate joint modeling of returns and realized covariances because no other study models the correlations as regime dependent. However, we borrow from the multivariate stochastic volatility literature by assuming a Wishart distribution for the realized covariance matrices but we estimate the parameters by maximum likelihood instead of the Bayesian methodology.
2.3 Estimation

The estimation of the parameters by full Maximum Likelihood (ML) would require to evaluate the log-likelihood function given the observations \( \{Y_t, X_t\}_{t=1}^T \):

\[
\mathcal{L}(\theta; Y, X) = \sum_{t=1}^T \ln f(Y_t, X_t | \mathcal{F}_{t-1}),
\]

(2.13)

where \( Y = \{Y_t\}_{t=1}^T \), \( X = \{X_t\}_{t=1}^T \) and \( \theta \) is the vector of parameters. When the covariance matrices to estimate are very large, maximum likelihood is not feasible. Our estimation strategy instead takes advantage of the decomposition of the conditional and realized covariance matrices in Eq. 2.3 and 2.8. The method is of course less efficient but yields consistent estimates. The joint distribution of returns and conditional variances can be expressed as:

\[
f(Y_t, X_t | \mathcal{F}_{t-1}) = f(Y_t | \mathcal{F}_{t-1}) f(X_t | Y_t, \mathcal{F}_{t-1}),
\]

(2.14)

\[
f(Y_t | \mathcal{F}_{t-1}) = (2\pi)^{-n/2} |H_t|^{-1/2} \exp \left( -\frac{1}{2} Y_t' H_t^{-1} Y_t \right),
\]

(2.15)

\[
f(X_t | Y_t, \mathcal{F}_{t-1}) = \left| X_t \right|^{-\nu-1}_2 \nu/2 \Gamma_{\nu/2}(\nu/2) \exp \left( -\frac{1}{2} X_t \nu^{-1}_2 \Sigma_t \right).
\]

(2.16)

\( \Gamma_{\nu/2}(\nu/2) \) is the multivariate gamma function evaluated at \( \nu/2 \) and \( \text{etr} \) represents the exponential of the trace function. The probability distribution functions in Eq. 2.15 and 2.16 are used to derive the log-likelihood functions. With \( \theta_1 = \{\omega, \beta, \alpha, \gamma^{(0)}_i, \gamma^{(1)}_i, \nu\} \) and \( \theta_2 = \{\Gamma^{(1)}, ..., \Gamma^{(M)}, \pi_{1,1}, \pi_{1,2}, ..., \pi_{M,M}\} \) and after replacing the matrices \( X_t, H_t \) and \( \Sigma_t \) by their decompositions, we can split the log likelihood function into two parts (proof in appendix A):

\[
\mathcal{L}_{\text{step1}}(\theta_1) = \sum_{t=1}^T \sum_{i=1}^n \left\{ \left[ -\frac{1}{2} \ln (2\pi) - \frac{1}{2} \ln h_{i,t} - \frac{1}{2} \frac{y_{i,t}^2}{h_{i,t}} \right] + \left[ \frac{\nu}{2} \ln (\nu/2) - \ln \Gamma(\nu/2) - \frac{\nu}{2} \ln (\bar{h}_{i,t}) + \frac{\nu-2}{2} \ln x_{i,t} - \frac{\nu x_{i,t}}{2 \bar{h}_{i,t}} \right] \right\},
\]

(2.17)
2.3. ESTIMATION

and

\[ L_{\text{step2}}(\theta_2) = \sum_{t=1}^{T} \left\{ \left[ \frac{1}{2} \hat{U}_t' \hat{U}_t - \frac{1}{2} \ln |\Gamma_t| - \frac{1}{2} \hat{U}_t' \Gamma_t^{-1} \hat{U}_t \right] + \left[ \frac{1}{2} \frac{1-n}{\nu} \ln |V_t^n| - \ln |\Gamma_t| + n \ln \Gamma(\nu/2) - \ln \Gamma_n(\nu/2) + \frac{\nu}{2} \text{tr}(\hat{S}_t[I_n - \Gamma_t^{-1}]) \right\} \],

(2.18)

where, \( y_{i,t} \) is the \( i^{th} \) element of \( Y_t \) in Eq. 2.1 and \( \hat{U}_t = D_t^{-1}Y_t \). In the first step, the estimation of the parameters of the conditional and realized variances is achieved for each asset by maximizing the log-likelihood function 2.17. In the second step, and conditional on the first step estimates, the states, the transition probability matrix \( \Pi \) and the correlation matrices \( \Gamma^{(j)}, \forall j = 1, 2, ..., M \) are estimated by maximizing 2.18.

2.3.1 Univariate Model Estimation

The univariate parameters are estimated by maximizing 2.17. With the estimated parameters, the sequence of matrices and vectors \( \{\tilde{S}_t\}_{t=1}^{T} \) and \( \{\hat{U}_t\}_{t=1}^{T} \) are constructed.

Notice that in this first step, \( \nu \) has no effect on the estimation of the parameters because the univariate likelihood (\( L_{\text{step1}}(\theta_1) \)) is a combination of Normal and Gamma distribution. The Gamma part, where \( \nu \) appears, can be written as:

\[ C_{x,\nu} - \frac{\nu}{2} \sum_{t=1}^{T} \left\{ \ln \tilde{h}_{i,t} + \frac{x_{i,t}}{\tilde{h}_{i,t}} \right\}, \]

where \( C_{x,\nu} \) is function of \( \nu \) and \( \{x_{i,t}\}_{t=1}^{T} \) only. So the first order conditions involved to find the parameters of \( \tilde{h}_{i,t} \) do not depend on \( \nu \). This is, however not the case in the second step as we see it below.

2.3.2 Regime Switching Model Estimation

The direct maximization of 2.18 may not be feasible in portfolio analysis with large number of assets. For instance, if we assume a two-state Markov process for \( n = 10 \) assets, not only do we need to infer the unobserved states at each time \( t \) but also the model implies that there are \( 2^{n(n-1)/2} = 90 \) unique correlations coefficients to estimate. A solution to this problem is the Expectation-Maximization (EM) algorithm introduced by Dempster et al. [1977]. In the following, we show that by using the EM framework to maximize the likelihood we can break
2.3. ESTIMATION

the curse of dimensionality in that the correlation matrices $\Gamma^{(1)}, \ldots, \Gamma^{(M)}$ can be estimated through simple weighted sums.

The central concept of the EM algorithm is the auxiliary function $Q(\bullet|\bullet)$ called the intermediate quantity of EM. It is the expected value of the log-likelihood with respect to the unknown states given the data and the current guess of the parameters value. The algorithm consists in building iteratively a sequence $\{\theta_2^{(k)}\}_{k \geq 1}$ of parameter estimates given an initial guess $\theta_2^{(0)}$. For the $k^{th}$ iteration, the steps of the EM algorithm are:

- **Expectation:** Compute $Q(\theta|\theta_2^{(k)})$.
- **Maximization:** Find $\theta_2^{(k+1)}$ such that $Q(\theta|\theta_2^{(k)})$ is maximized.

Given regularity assumptions (Cappé et al. [2005]), the convergence of the iteration achieves local maximization of the log-likelihood.

Applied to the problem at hand, the EM algorithm consists of the a set of simple updating equations (proof in appendix B). For states, $i,j = 1, \ldots, M$ there are:

\[
\hat{\pi}_{ij} = \frac{\sum_{t=1}^{T} P(s_t = j, s_{t-1} = i|\hat{U}, \tilde{S}, \theta_2^{(k)})}{\sum_{t=1}^{T} P(s_{t-1} = i|U, \tilde{S}, \theta_2^{(k)})}, \tag{2.19}
\]

\[
\hat{\Gamma}^{(j)} = \frac{\sum_{t=1}^{T} C_t P(s_t = j|\hat{U}, \tilde{S}, \theta_2^{(k)})}{\sum_{t=1}^{T} P(s_t = j|U, \tilde{S}, \theta_2^{(k)})}, \tag{2.20}
\]

\[
C_t = \frac{1}{\nu + 1} \{\hat{U}_t \hat{U}_t' + \nu \tilde{S}_t\}, \tag{2.21}
\]

where $\tilde{S} = \{\tilde{S}_t\}_{t=1}^{T}$ and $\hat{U} = \{\hat{U}_t\}_{t=1}^{T}$. When the only source of information is the time series of log returns, the correlation model collapses to the RSDC of Pelletier [2006].

\[
\hat{\pi}_{ij} = \frac{\sum_{t=1}^{T} P(s_t = j, s_{t-1} = i|\hat{U}, \theta_2^{(k)})}{\sum_{t=1}^{T} P(s_{t-1} = i|U, \theta_2^{(k)})}, \tag{2.22}
\]

\[
\hat{\Gamma}^{(j)} = \frac{\sum_{t=1}^{T} \hat{U}_t \hat{U}_t' P(s_t = j|\hat{U}, \theta_2^{(k)})}{\sum_{t=1}^{T} P(s_t = j|\hat{U}, \theta_2^{(k)})}. \tag{2.23}
\]

Finally, if the only available dataset is a time series of realized covariances we have the following
2.3. ESTIMATION

updating equations:

\[
\hat{\pi}_{ij} = \frac{\sum_{t=1}^{T} P(s_t = j, s_{t-1} = i | \tilde{S}, \theta^{(k)})}{\sum_{t=1}^{T} P(s_{t-1} = i | \tilde{S}, \theta^{(k)})},
\]

(2.24)

\[
\hat{\Gamma}^{(j)} = \frac{\sum_{t=1}^{T} \tilde{S}_t P(s_t = j | \tilde{S}, \theta^{(k)})}{\sum_{t=1}^{T} P(s_t = j | \tilde{S}, \theta^{(k)})},
\]

(2.25)

where an appropriate model for the conditional mean of the realized variance is estimated in a first step to construct the sequence of \( \tilde{S}_t \) matrices.

The smoothed, joint and conditional probabilities required in the updating equations are computed by forward-backward inference algorithm as explained in Hamilton [1990] or Kim [1994] and pioneered by Baum et al. [1970] (see appendix C for the details).

The impact of the degree of freedom \( \nu \) on the estimation when the number of assets increases can be seen in the updating Eq. 2.21. \( C_t \) is a weighted average of \( \hat{U}_t \) and \( \tilde{S}_t \) with the weights depending on \( \nu \). Given that the non-singular Wishart distribution requires that \( \nu > n - 1 \), on the one hand, the contribution of the standardized returns in the estimation will decrease as the ratio \( \frac{1}{\nu+1} \) get smaller. On the other hand, most of the multivariate information would be provided by the scaled realized correlation \( \tilde{S}_t \) as \( \frac{\nu}{\nu+1} \approx 1 \).

2.3.3 Asymptotic Distribution of the Two-Step MLE

The derivation of the asymptotic distribution of the two-step MLE can be found in Murphy and Topel [2002]. Let us denote by \( \frac{1}{T} \mathcal{V}_1 \) the asymptotic covariance matrix of the first step estimators. Furthermore, let \( \frac{1}{T} \mathcal{V}_2 \) be the asymptotic covariance matrix of the second step estimators without accounting for the first step. \( \frac{1}{T} \mathcal{V}_2 \) must be corrected for the fact that an estimate of \( \theta_1 \) is used in the estimation of \( \theta_2 \) instead of the true \( \theta_1 \). The corrected form is \( \frac{1}{T} \mathcal{V}_2^* \) where \( \mathcal{V}_2^* \) is defined as:

\[
\mathcal{V}_2^* = \mathcal{V}_2 + \mathcal{V}_2 (\mathcal{M} \mathcal{V}_1 \mathcal{M}' - \mathcal{R} \mathcal{V}_1 \mathcal{M}' - \mathcal{M} \mathcal{V}_1 \mathcal{R}) \mathcal{V}_2,
\]

(2.26)

with,

\[
\mathcal{M} = E \left[ \left( \frac{\partial L_{\text{step}2}}{\partial \theta_2} \right) \left( \frac{\partial L_{\text{step}2}}{\partial \theta_1} \right) \right], \quad \mathcal{R} = E \left[ \left( \frac{\partial L_{\text{step}2}}{\partial \theta_2} \right) \left( \frac{\partial L_{\text{step}1}}{\partial \theta_1} \right) \right].
\]
2.3. ESTIMATION

In finite sample, the estimates of \( V_1, V_2, M \) and \( R \) are:

\[
\hat{V}_1 = \left[ \frac{1}{T} \sum_{t=1}^{T} \left( \frac{\partial L_{\text{step}1}}{\partial \theta_1} \right) \left( \frac{\partial L_{\text{step}1}}{\partial \theta_1'} \right) \right]^{-1}, \quad \hat{V}_2 = \left[ \frac{1}{T} \sum_{t=1}^{T} \left( \frac{\partial L_{\text{step}2}}{\partial \theta_2} \right) \left( \frac{\partial L_{\text{step}2}}{\partial \theta_2'} \right) \right]^{-1},
\]

\[
\hat{M} = \frac{1}{T} \sum_{t=1}^{T} \left( \frac{\partial L_{\text{step}2}}{\partial \theta_2} \right) \left( \frac{\partial L_{\text{step}2}}{\partial \theta_1'} \right) \quad \text{and} \quad \hat{R} = \frac{1}{T} \sum_{t=1}^{T} \left( \frac{\partial L_{\text{step}2}}{\partial \theta_2} \right) \left( \frac{\partial L_{\text{step}1}}{\partial \theta_1'} \right).
\]

2.3.4 Multistep Ahead Forecast

Unfortunately, as in the DCC of Engle [2002b], the non-linearity of our specification precludes from the derivation of a closed form expression for the multistep ahead forecast. However, the assumption of Markov regime switching allows to forecast the correlation matrices independently of the variances. We also can easily forecast the individual conditional and realized variances. So the following forecasting scheme is proposed:

\[
E_T[H_{T+L}] = H_{T+L|T} = D_{T+L|T} \Gamma_{T+L|T} D_{T+L|T}, \quad \forall \ L \geq 1 \quad (2.27)
\]

where \( D_{T|T+L} \) holds the square root of the \( L \)-step ahead forecasts of the conditional variances and \( \Gamma_{T|T+L} \) is the \( L \)-step ahead forecast of the correlation matrix.

2.3.4.1 Conditional and Realized Variances Forecasts

Given the specification in Eq. 2.4 and 2.7, the L-step ahead forecast of the conditional variance is (See appendix 4 for proof):

\[
E_T[h_{i,T+L}] = h_{i,T+L|T} = \sum_{j=0}^{L-1} \omega_j \beta_i^2 + \beta_i^T h_i, T + \sum_{k=0}^{L-1} \alpha_i \beta_i x_{i,T+L-1-k|T}, \quad \forall \ L \geq 1 \quad (2.28)
\]

where \( x_{i,T+L-1-k|T} = E_T x_{i,T+L-1-k} = \gamma_0 + \gamma_1 E_T h_{i,T+L-1-k|T} \). The forecast is therefore recursive and depends non trivially on the forecast of the realized variance. Typically, it is expected that \( \alpha_i \beta_i < 1 \) so the effect of past forecast of the realized variance should decrease with the number of periods. Similarly, \( \beta_i < 1 \) and the longer the horizon is, the smaller the effect of the conditional variance at time \( T \) will be. Finally, notice that it is also possible to generate \( E_T[\sqrt{h_{i,T+L}}] \) by simulation to avoid the Jensen’s inequality \( E_T[\sqrt{h_{i,T+L}}] \neq \sqrt{E_T[h_{i,T+L}]} \).
2.4. EMPIRICAL ILLUSTRATION

2.3.4.2 Correlations Forecast

Given the properties of the Markov chain and using the notation introduced in appendix (D), the \( L \)-step ahead forecast of the correlation matrix \( \Gamma \) is:

\[
E_T[\Gamma_{T+L}] = \Gamma_{T+L} = \sum_{s_{T+L}=i}^{M} \hat{\xi}_{s_{T+L}|T} \Gamma^{(i)}_{T+L}, \quad \forall i = 1, ..., M
\]

(2.29)

where \( \hat{\xi}_{s_{T+L}|T} \) is the \( i \)th row of:

\[
\hat{\xi}_{s_{T+L}|T} = \Pi^L \hat{\xi}_{T|T}
\]

(2.30)

2.4 Empirical Illustration

The model is applied to high-frequency prices for seven stocks from the financial sector of the S&P 100 \(^1\). The main focus of the present section is to illustrate the ability to model the correlations and the conditional variances. The sample period runs from January 3, 2006 to April 30, 2012 corresponding to a sample size of 1,593. However, days for which the daily transaction period is less than the normal 6.5 hours are removed\(^2\), leaving a final sample size of 1,581.

2.4.1 Data Description

2.4.1.1 Realized Volatility Description

The multivariate realized kernel estimator of Barndorff-Nielsen et al. [2011] is used as measure of realized covariance. The data retention ratio introduced in section 2.1 of chapter 1 is \( p = \frac{nN}{\sum_{i=1}^{n} n^I} \), where \( n \) is the number of assets, \( n^I \) and \( N \) are respectively the initial and synchronized number of transactions for the \( i \)th asset. Fig. 2.1 presents the evolution of the retention percentage for the sample period. On average 36.98% of the daily data is retained. The minimum retention is 22.48% while the maximum is 79.93%. It is worthwhile to notice that the maximum retention percentage occurs on September 18, 2008\(^3\), when all the stocks in the sample were actively traded. Table 2.1 shows the statistics for the percentage of retention per asset in the sample.

---

\(^1\)The ticker symbols are AXP, BAC, COF, GS, JPM, MS and USB

\(^2\)Independence Day, the days before Thanksgiving and Christmas

\(^3\)That day, in all the major financial markets, swap lines were expanded by $180 Billion enabling central banks to lend dollars into their domestic banking system
The percentages are computed as the ratio of the number of daily transactions selected over the total of daily transactions for a particular asset. The sampling percentage can be as low as 9.66% and as high as 88%.

As for the bandwidth selection, the minimum, maximum and mean of the bandwidths are respectively 29, 141 and 60. The bandwidth also corresponds to the number of autocovariances included in the computation of the realized matrices.

The summary statistics for the annualized realized volatilities are reported in the second part of Table 2.2. The annualized realized volatility is the square root of 252 times the realized kernel estimates. Morgan Stanley (MS) is on average the most volatile of all the assets with the highest average (42.38%) and the highest standard deviation (43.64%). Fig. 2.2 and 2.3 show the behavior of the annualized realized volatility for the entire sample period. The financial crisis of 2008-2009 is noticeable midway in the sample as there is increased volatility. The period spanning from September 1, 2008 to June 1, 2009 is depicted in Fig. 2.5 and 2.4. Notice that the maximum values of the annualized realized volatility for all the assets occur in this specific time frame.

The averages and standard deviations of the realized correlations of the pairs of assets are presented in Table 2.3. The average of the daily realized correlations oscillates between 0.46 and 0.62 while the standard deviations are between 0.12 and 0.19. The last two columns of Table 2.3 present the maximum and minimum realized correlation coefficients along with their day of occurrence. The majority of the lowest correlations dates (17 out of 21) are before the financial crises of 2008. While, except from the correlation between AXP and GS, the highest values of realized correlations are recorded between 2008 and 2011.

2.4.1.2 Daily Log-Return Description

The statistics for the low-frequency data, the daily log close-to-close returns, are presented in the first part of Table 2.2. As expected, the averages are in the neighborhood of 0% except for BAC. It appears also that MS is the most volatile stock in the sample even at the daily frequency. The range of the return is large, for instance, the minimum daily return for BAC is −33.79% and the maximum is 31.66%.

2.4.2 Estimation Results

This subsection presents the results of the estimation as introduced in section 3. For illustration purpose, it is assumed that there are two regimes ($M = 2$) with respectively high and low

---

4 There were 17,790 total transactions for BAC on 12/27/2010, only 1,718.6 transactions were selected.
correlations. With a convergence criteria for the second step set at $10^{-8}$, the local maximum is attained after 33 iterations. The individual estimation of the univariate Realized GARCH models, which are based on a combination of a Gaussian and a Gamma term, are presented in Tables 2.4 and 2.5. The estimates of the coefficients of correlation for the different regimes and the transition probabilities are presented in Tables 2.6, 2.7, 2.8 and 2.9.

### 2.4.2.1 Step 1: Univariate Results

The autoregressive parameters in the conditional variance specification, $\beta_i$, is considerably less than it would be under a regular GARCH model. The estimated coefficients are in the range of 0.36 to 0.50. The typical value in a GARCH specification is between 0.95 and 0.99 which is indicative of the strong persistence of the conditional variance. However a direct interpretation of $\beta_i$ as a persistence parameter is not correct. As shown in section 2.3, the appropriate R-RSDC model time-varying measure of persistence is $\beta_i + \alpha_i \gamma_i^{(1)} \epsilon_{i,t-1}$. Given that $E(\epsilon_{i,t-1}) = 1$, we present the averages in Table 2.5 and we can see that the values are similar to what a typical GARCH model would predict.

The effect of the lagged realized volatility is in the range of 0.65 to 0.92. Hansen et al. [2012] have found a value for the SPY index in that range (0.87). The estimates of $\gamma_i^{(1)}$ are less than one, suggesting that the variance measured while the market is open corresponds to between 56% and 78% of the daily variance.

As for the estimates of the shape parameter $\nu$, they vary between 6.35 and 8.87 with a median value of 8.39. A non-singular Wishart distribution requires that the degree of freedom be greater than the number of assets minus one. In our case, the degree of freedom must be greater than 6. For the second step therefore, the unique shape parameter is set to 8.39, the median first step estimates.

### 2.4.2.2 Step 2: Multivariate Results

The sequence of scaled covariance matrices $\tilde{S}_t$ and the standardized return vectors $\tilde{U}_t$ are constructed with the first step parameters estimates. As evidenced in Table 2.8, the difference in correlation between the two regimes is in the range of 15.85% to 44.13%. Table 2.9 shows that the regime of high correlation is slightly more persistent. Fig. 2.6 and 2.7 show the evolution of the probabilities (smooth and filtered) of the occurrence of regime 1 and 2 for the entire sample period. Except for brief periods, the regime constantly switches between regime 1 and 2. One of these exceptions is the financial crisis of 2008-2009 where the level of correlation appears higher (regime 2). Fig. 2.8 and 2.9 reproduce the probability dynamic for the period of
September 1, 2008 to June 1, 2009. We can see that large portions in the month of October 2008 to April 2009 belong to the regime of high correlation. This suggests that during periods of high volatility, the financial assets in the sample tend to behave uniformly. As an illustration, the evolution of the annualized realized volatility along with the realized, conditional correlations and covariances for BAC and JPM are presented in Fig. 2.10 and 2.11. It is noticeable that in Fig. 2.10, the period of the financial crisis is characterized by high volatilities and covariances. The picture is clear when we zoom in on the period in Fig. 2.11. The realized volatilities for JPM and BAC behave almost identically from September to December 2008. The conditional covariance dynamic implied by the R-RSDC specification follows closely the erratic movement of the realized covariance.

2.4.2.3 Model Fit

To assess the fit of the model, the multivariate Ljung-Box portmanteau (or HM) test of Hosking [1980] is applied to the standardized residuals. The test is used to detect any residual ARCH effects and has the form:

\[ HM(W) = T^2 \sum_{j=1}^{W} (T - j)^{-1} \text{tr}\{A_{\hat{z}_t}^{-1}(0)A_{\hat{z}_t}(j)A_{\hat{z}_t}^{-1}(0)A_{\hat{z}_t}(j)\}, \]

where \( \hat{z}_t = \text{vech}(H_t^{-1/2}Y_tY_t'H_t^{-1/2}) \) and \( A_{\hat{z}_t}(j) \) is the associated sample autocovariance matrix of order \( j \). Under the null hypothesis of no ARCH effects, the test is distributed asymptotically as Chi-square with degree of freedom \( W \times n^2 \).

With \( W = 20 \) lags, the HM statistics is 998.3220. Given that the 5% critical value is 1053.9, there is no sign of ARCH effects in the residual.

2.5 Out-of-sample Model Evaluation

The goal of this section is to show the advantage of the high-frequency information in the forecasting of conditional covariance matrices. Therefore, our model is compared to a multivariate specification that uses only daily information. Notice that a benchmark specification that also incorporates high frequency information would be the HEAVY of Noureldin and Sheppard [2012].

To conduct the forecast analysis, the sample is split into two parts. The in-sample for the initial estimation runs from January 4, 2006 to April 29, 2011 and has a size of 1329. With a

\footnote{A 20 lags HM test on \( Y_t \) was initially performed and the Test statistics was 3,240.3}
2.5. OUT-OF-SAMPLE MODEL EVALUATION

size of 251 (one year), the out-of-sample spans from May 2, 2011 to April 30, 2012.

The R-RSDC model is close in its design to DCC model of Engle [2002b] because it splits
the covariance matrix into standard deviation and correlation matrices. Therefore to assess
the gain from the proposed model, the DCC is the competing model. The parameters of both
models are updated daily and the horizon considered for the predictions are 1, 5, 10 and 20.
These correspond respectively to one day, one week, two weeks and one month ahead forecast.

We implement the DMW test explained in section 2.2 of Chapter 1. Given that \( H^0_{t+L} \)
is unobserved, a proxy is constructed by scaling the multivariate realized kernel. The scaling
is necessary because the realized kernel measures the volatility while the market is open (6.5 hours)
and the forecast objective is the daily variance and covariance. For each univariate forecast of
the conditional variance, the scale factor is:

\[
\zeta_i = \frac{\sum_{t=1}^{\tilde{T}} y^2_{i,t}}{\sum_{t=1}^{\tilde{T}} x_{i,t}}, \forall \quad i = 1, \ldots, n
\]

where \( \tilde{T} \) is the size of the estimation sample. One effect of the scaling is that the sum of the
scaled daily realized variance is equal to the sum of the square daily returns. The overall realized
covariance matrix is constructed with the scaled realized variance and the realized correlation.
The realized correlation are not transformed because it is assumed that the overnight effect on
the correlation is negligible.

With a proxy for the latent covariance matrix and the forecasted conditional covariance
matrices, the loss is calculated for each model and the loss differential is computed as:

\[
d_{t,L} = L_{t,L}(H^0_{t+L}, H^{\text{DCC}}_{t+L|t}) - L_{t,L}(H^0_{t+L}, H^{R-RSDC}_{t+L|t})
\]

The direction of the difference implies that the forecast generated by the R-RSDC model is
more accurate than the one generated by the DCC model if \( d_{t,L} \) is positive. With the time
series of \( d_{t,L} \), we test for all \( L \):

\[
H_0 : E[L_{t,L}(H^0_{t+L}, H^{\text{DCC}}_{t+L|t})] = E[L_{t,L}(H^0_{t+L}, H^{R-RSDC}_{t+L|t})],
\]

\[
vs. \ H_1 : E[L_{t,L}(H^0_{t+L}, H^{\text{DCC}}_{t+L|t})] > E[L_{t,L}(H^0_{t+L}, H^{R-RSDC}_{t+L|t})],
\]

\[
H_2 : E[L_{t,L}(H^0_{t+L}, H^{\text{DCC}}_{t+L|t})] < E[L_{t,L}(H^0_{t+L}, H^{R-RSDC}_{t+L|t})].
\]

The results, in Table 2.10, show that the R-RSDC model outperforms the DCC model. The
univariate variance model has a clear superior predictive ability, especially in shorter horizon. Except for BAC, all the tests in the one-step ahead case reveal that the R-RSDC model is more accurate than the DCC model with a 95% confidence level or more. The DMW test for the correlation forecasts shows a clear advantage over the DCC model even at longer horizon. As a result, with a confidence level of 99%, the multivariate forecast is more accurate in all the horizons, even though on the margin the gain from the univariate model is mixed in longer horizon.

2.6 Conclusion

The chapter has introduced the Realized Regime Switching for Dynamic Correlations (R-RSDC) model for vectors of daily returns and corresponding realized measures. This model builds on the univariate Realized GARCH of Hansen et al. [2012], the DCC of Engle [2002b] and the RSDC of Pelletier [2006].

The Realized GARCH framework describes the dynamic of the conditional variance with a modified GARCH equation. The decomposition of the realized and conditional covariance matrices permits the specification of an univariate measurement equation for each asset. The inference of regime switching correlation matrices exploits the daily information and the high frequency realized correlation matrices.

The curse of dimensionality is mitigated by the DCC-like decomposition of the matrices of covariance. The realized covariance matrices considered are the multivariate realized kernel of Barndorff-Nielsen et al. [2011]. The estimation is broken into two steps where the first step consists in estimating the parameters of univariate models. The second step uses the EM algorithm introduced by Dempster et al. [1977].

In an illustrative empirical application, we have estimated the model for a dataset of seven assets and have inferred with accuracy the state of the market. We have presented evidence that there is an increase in correlation between assets during periods of crisis. In an out-of-sample comparison with the DCC model, the proposed model improves the forecast accuracy because of the additional high frequency information. The marginal (asset by asset) accuracy level deteriorates slightly with the increase of the forecast horizon. However, the correlation forecast contribution is considerable even at our largest horizon of prediction (twenty-two days).
2.7. TABLES AND FIGURES

2.7 Tables and Figures

Table 2.1: Summary Statistics for Refresh Sampling: January 03, 2006 to April 30, 2012

<table>
<thead>
<tr>
<th>Asset</th>
<th>Mean</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>AXP</td>
<td>46.25</td>
<td>14.63</td>
<td>87.94</td>
</tr>
<tr>
<td>BAC</td>
<td>26.42</td>
<td>9.66</td>
<td>76.03</td>
</tr>
<tr>
<td>COF</td>
<td>54.20</td>
<td>18.42</td>
<td>88.06</td>
</tr>
<tr>
<td>GS</td>
<td>35.76</td>
<td>14.83</td>
<td>75.64</td>
</tr>
<tr>
<td>JPM</td>
<td>30.33</td>
<td>15.72</td>
<td>77.05</td>
</tr>
<tr>
<td>MS</td>
<td>38.99</td>
<td>18.31</td>
<td>75.68</td>
</tr>
<tr>
<td>USB</td>
<td>48.16</td>
<td>23.58</td>
<td>82.95</td>
</tr>
</tbody>
</table>

The numbers are the statistics for the percentages of retention after synchronization. The retention percentage is computed as \( \frac{n_i^t}{N_t} \times 100 \), where \( n_i^t \) is the initial number of transactions for asset \( i \) at time \( t \) and \( N_t \) is the number of transactions selected after synchronization.

Table 2.2: Summary Statistics for the Daily Log Returns and the Annualized Realized Volatilities: January 03, 2006 to April 30, 2012

<table>
<thead>
<tr>
<th>Stocks</th>
<th>Log returns</th>
<th>Realized volatility</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>St. dev</td>
</tr>
<tr>
<td>AXP</td>
<td>0.01</td>
<td>2.93</td>
</tr>
<tr>
<td>BAC</td>
<td>-0.11</td>
<td>4.32</td>
</tr>
<tr>
<td>COF</td>
<td>-0.03</td>
<td>3.88</td>
</tr>
<tr>
<td>GS</td>
<td>-0.01</td>
<td>2.95</td>
</tr>
<tr>
<td>JPM</td>
<td>0.00</td>
<td>3.21</td>
</tr>
<tr>
<td>MS</td>
<td>-0.08</td>
<td>4.34</td>
</tr>
<tr>
<td>USB</td>
<td>0.00</td>
<td>2.91</td>
</tr>
</tbody>
</table>

The log returns are in percentage point. The annualized realized volatility is the square root of 252 times the realized kernel.
### Table 2.3: Summary Statistics for the Realized Correlation Coefficients: January 03, 2006 to April 30, 2012

<table>
<thead>
<tr>
<th>Pairs</th>
<th>Mean</th>
<th>St.dev</th>
<th>Max</th>
<th>Date Max</th>
<th>Min</th>
<th>Date Min</th>
</tr>
</thead>
<tbody>
<tr>
<td>AXP-BAC</td>
<td>0.49</td>
<td>0.16</td>
<td>0.88</td>
<td>14-Mar-2008</td>
<td>-0.07</td>
<td>13-Mar-2006</td>
</tr>
<tr>
<td>AXP-COF</td>
<td>0.50</td>
<td>0.19</td>
<td>0.91</td>
<td>08-Jun-2010</td>
<td>-0.17</td>
<td>23-Jan-2007</td>
</tr>
<tr>
<td>AXP-GS</td>
<td>0.47</td>
<td>0.16</td>
<td>0.84</td>
<td>30-Mar-2007</td>
<td>-0.13</td>
<td>21-Dec-2010</td>
</tr>
<tr>
<td>AXP-JPM</td>
<td>0.51</td>
<td>0.17</td>
<td>0.89</td>
<td>12-Aug-2011</td>
<td>-0.14</td>
<td>02-Nov-2010</td>
</tr>
<tr>
<td>AXP-MS</td>
<td>0.47</td>
<td>0.17</td>
<td>0.84</td>
<td>14-Sep-2011</td>
<td>-0.27</td>
<td>21-Apr-2011</td>
</tr>
<tr>
<td>AXP-USB</td>
<td>0.48</td>
<td>0.19</td>
<td>0.91</td>
<td>14-Mar-2008</td>
<td>-0.40</td>
<td>13-Jul-2007</td>
</tr>
<tr>
<td>BAC-COF</td>
<td>0.46</td>
<td>0.17</td>
<td>0.87</td>
<td>25-Feb-2008</td>
<td>-0.17</td>
<td>19-Jan-2007</td>
</tr>
<tr>
<td>BAC-GS</td>
<td>0.49</td>
<td>0.15</td>
<td>0.85</td>
<td>01-Nov-2011</td>
<td>-0.11</td>
<td>19-Sep-2006</td>
</tr>
<tr>
<td>BAC-JPM</td>
<td>0.59</td>
<td>0.15</td>
<td>0.91</td>
<td>13-Mar-2008</td>
<td>-0.18</td>
<td>25-Jan-2007</td>
</tr>
<tr>
<td>BAC-MS</td>
<td>0.50</td>
<td>0.15</td>
<td>0.86</td>
<td>04-Mar-2008</td>
<td>-0.20</td>
<td>25-Apr-2011</td>
</tr>
<tr>
<td>BAC-USB</td>
<td>0.54</td>
<td>0.17</td>
<td>0.93</td>
<td>25-Feb-2008</td>
<td>-0.11</td>
<td>17-Oct-2006</td>
</tr>
<tr>
<td>COF-GS</td>
<td>0.44</td>
<td>0.17</td>
<td>0.81</td>
<td>28-Sep-2011</td>
<td>-0.12</td>
<td>13-Mar-2006</td>
</tr>
<tr>
<td>COF-JPM</td>
<td>0.48</td>
<td>0.18</td>
<td>0.90</td>
<td>26-Aug-2011</td>
<td>-0.20</td>
<td>17-Mar-2006</td>
</tr>
<tr>
<td>COF-MS</td>
<td>0.44</td>
<td>0.18</td>
<td>0.85</td>
<td>01-Sep-2011</td>
<td>-0.31</td>
<td>19-Mar-2007</td>
</tr>
<tr>
<td>COF-USB</td>
<td>0.46</td>
<td>0.19</td>
<td>0.88</td>
<td>26-Aug-2011</td>
<td>-0.34</td>
<td>16-Feb-2007</td>
</tr>
<tr>
<td>GS-JPM</td>
<td>0.54</td>
<td>0.15</td>
<td>0.85</td>
<td>12-Aug-2011</td>
<td>-0.05</td>
<td>09-May-2006</td>
</tr>
<tr>
<td>GS-MS</td>
<td>0.62</td>
<td>0.12</td>
<td>0.87</td>
<td>01-Jul-2008</td>
<td>0.10</td>
<td>15-Mar-2006</td>
</tr>
<tr>
<td>GS-USB</td>
<td>0.46</td>
<td>0.17</td>
<td>0.83</td>
<td>13-Mar-2008</td>
<td>-0.22</td>
<td>27-Dec-2006</td>
</tr>
<tr>
<td>JPM-MS</td>
<td>0.54</td>
<td>0.15</td>
<td>0.87</td>
<td>12-Jul-2010</td>
<td>0.01</td>
<td>17-Mar-2006</td>
</tr>
<tr>
<td>JPM-USB</td>
<td>0.56</td>
<td>0.18</td>
<td>0.92</td>
<td>12-Aug-2011</td>
<td>-0.35</td>
<td>20-Apr-2007</td>
</tr>
<tr>
<td>MS-USB</td>
<td>0.47</td>
<td>0.18</td>
<td>0.89</td>
<td>26-Aug-2011</td>
<td>-0.32</td>
<td>27-Dec-2006</td>
</tr>
</tbody>
</table>

Column 2 presents the mean and standard deviation of daily realized correlation coefficients. Column 3 presents the values of the maximum daily realized correlation coefficients and their dates of occurrence. Column 4 presents the values of the minimum daily realized correlation coefficients and their dates of occurrence.
### Table 2.4: Parameter Estimates for the Univariate Realized GARCH Model

<table>
<thead>
<tr>
<th></th>
<th>AXP</th>
<th>BAC</th>
<th>COF</th>
<th>GS</th>
<th>JPM</th>
<th>MS</th>
<th>USB</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\omega}$</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>(0.0000)</td>
<td>(0.0000)</td>
<td>(0.0000)</td>
<td>(0.0000)</td>
<td>(0.0000)</td>
<td>(0.0000)</td>
<td>(0.0000)</td>
</tr>
<tr>
<td>$\hat{\beta}$</td>
<td>0.3853</td>
<td>0.4345</td>
<td>0.3968</td>
<td>0.5090</td>
<td>0.4043</td>
<td>0.3637</td>
<td>0.3750</td>
</tr>
<tr>
<td></td>
<td>(0.0196)</td>
<td>(0.0149)</td>
<td>(0.0193)</td>
<td>(0.0145)</td>
<td>(0.0196)</td>
<td>(0.0201)</td>
<td>(0.0171)</td>
</tr>
<tr>
<td>$\hat{\alpha}$</td>
<td>0.8658</td>
<td>0.9649</td>
<td>0.8584</td>
<td>0.6537</td>
<td>0.8720</td>
<td>0.9259</td>
<td>0.7784</td>
</tr>
<tr>
<td></td>
<td>(0.0481)</td>
<td>(0.0460)</td>
<td>(0.0468)</td>
<td>(0.0450)</td>
<td>(0.0520)</td>
<td>(0.0542)</td>
<td>(0.0405)</td>
</tr>
<tr>
<td>$\hat{\gamma}_0$</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>(0.0000)</td>
<td>(0.0000)</td>
<td>(0.0000)</td>
<td>(0.0000)</td>
<td>(0.0000)</td>
<td>(0.0000)</td>
<td>(0.0000)</td>
</tr>
<tr>
<td>$\hat{\gamma}_1$</td>
<td>0.6857</td>
<td>0.5686</td>
<td>0.6711</td>
<td>0.6613</td>
<td>0.6546</td>
<td>0.6314</td>
<td>0.7807</td>
</tr>
<tr>
<td></td>
<td>(0.0326)</td>
<td>(0.0214)</td>
<td>(0.0337)</td>
<td>(0.0453)</td>
<td>(0.0341)</td>
<td>(0.0322)</td>
<td>(0.0324)</td>
</tr>
<tr>
<td>$\hat{v}$</td>
<td>8.8691</td>
<td>6.1575</td>
<td>8.7315</td>
<td>6.3525</td>
<td>8.3953</td>
<td>8.3977</td>
<td>7.2103</td>
</tr>
<tr>
<td></td>
<td>(0.2404)</td>
<td>(0.1299)</td>
<td>(0.2660)</td>
<td>(0.1627)</td>
<td>(0.2238)</td>
<td>(0.2276)</td>
<td>(0.1575)</td>
</tr>
</tbody>
</table>

The Standard errors are in parenthesis. The table reports the estimates from the model in Eq. 2.4 and 2.7.

### Table 2.5: Persistence of the Conditional Variances Implied by the R-RSDC model

<table>
<thead>
<tr>
<th></th>
<th>AXP</th>
<th>BAC</th>
<th>COF</th>
<th>GS</th>
<th>JPM</th>
<th>MS</th>
<th>USB</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_i + \alpha_i\gamma_i^{(1)}$</td>
<td>0.9796</td>
<td>0.9831</td>
<td>0.9729</td>
<td>0.9413</td>
<td>0.9751</td>
<td>0.9484</td>
<td>0.9826</td>
</tr>
</tbody>
</table>

The table reports the average persistence estimated by $\beta_i + \alpha_i\gamma_i^{(1)}$ since $E[\epsilon] = 1$.
2.7. TABLES AND FIGURES

Table 2.6 : Correlation Estimates Regime 1

<table>
<thead>
<tr>
<th></th>
<th>AXP</th>
<th>BAC</th>
<th>COF</th>
<th>GS</th>
<th>JPM</th>
<th>MS</th>
<th>USB</th>
</tr>
</thead>
<tbody>
<tr>
<td>AXP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAC</td>
<td>0.4307(0.0136)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COF</td>
<td>0.4284(0.0107)</td>
<td>0.4102(0.0137)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GS</td>
<td>0.4189(0.0160)</td>
<td>0.4501(0.0126)</td>
<td>0.3771(0.0173)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JPM</td>
<td>0.4545(0.0131)</td>
<td>0.5680(0.0229)</td>
<td>0.4263(0.0211)</td>
<td>0.5068(0.0123)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MS</td>
<td>0.4181(0.0176)</td>
<td>0.4598(0.0207)</td>
<td>0.3789(0.0132)</td>
<td>0.5815(0.0118)</td>
<td>0.4988(0.0225)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>USB</td>
<td>0.4124(0.0211)</td>
<td>0.4809(0.0140)</td>
<td>0.3978(0.0260)</td>
<td>0.4090(0.0202)</td>
<td>0.5118(0.0206)</td>
<td>0.4078(0.0224)</td>
<td></td>
</tr>
</tbody>
</table>

The standard errors are in parenthesis. The table reports the estimates of the correlation coefficient when the regime is in state 1. The estimates are the results of the implementation of the EM algorithms Eq. 2.19, 2.20 and 2.21.

Table 2.7 : Correlation Estimates Regime 2

<table>
<thead>
<tr>
<th></th>
<th>AXP</th>
<th>BAC</th>
<th>COF</th>
<th>GS</th>
<th>JPM</th>
<th>MS</th>
<th>USB</th>
</tr>
</thead>
<tbody>
<tr>
<td>AXP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAC</td>
<td>0.5796(0.0162)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COF</td>
<td>0.6123(0.0126)</td>
<td>0.5699(0.0138)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GS</td>
<td>0.5523(0.0242)</td>
<td>0.5745(0.0202)</td>
<td>0.5333(0.0218)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JPM</td>
<td>0.6092(0.0158)</td>
<td>0.6820(0.0272)</td>
<td>0.5936(0.0279)</td>
<td>0.6288(0.0136)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MS</td>
<td>0.5632(0.0233)</td>
<td>0.5928(0.0272)</td>
<td>0.5451(0.025)</td>
<td>0.6736(0.0219)</td>
<td>0.6348(0.0269)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>USB</td>
<td>0.5954(0.0285)</td>
<td>0.6382(0.0156)</td>
<td>0.5782(0.0296)</td>
<td>0.5559(0.0258)</td>
<td>0.6701(0.0263)</td>
<td>0.5726(0.0285)</td>
<td></td>
</tr>
</tbody>
</table>

The standard errors are in parenthesis. The table reports the estimates of the correlation coefficient when the regime is in state 2. The estimates are the results of the implementation of the EM algorithms Eq. 2.19, 2.20 and 2.21.
2.7. TABLES AND FIGURES

Table 2.8 : Percentage Change in Estimated Correlation Coefficients Between Regime 1 and 2

<table>
<thead>
<tr>
<th></th>
<th>AXP</th>
<th>BAC</th>
<th>COF</th>
<th>GS</th>
<th>JPM</th>
<th>MS</th>
<th>USB</th>
</tr>
</thead>
<tbody>
<tr>
<td>AXP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAC</td>
<td>34.55</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COF</td>
<td>42.92</td>
<td>38.92</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GS</td>
<td>31.84</td>
<td>27.65</td>
<td>41.43</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JPM</td>
<td>34.06</td>
<td>20.08</td>
<td>39.23</td>
<td>24.08</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MS</td>
<td>34.72</td>
<td>28.93</td>
<td>43.87</td>
<td>15.83</td>
<td>27.28</td>
<td></td>
<td></td>
</tr>
<tr>
<td>USB</td>
<td>44.39</td>
<td>32.69</td>
<td>45.36</td>
<td>35.93</td>
<td>30.92</td>
<td>40.41</td>
<td></td>
</tr>
</tbody>
</table>

The table reports the change in percentage between coefficients of correlation from Table 2.6 and Table 2.7.

Table 2.9 : Transition Probabilities Estimates

<table>
<thead>
<tr>
<th></th>
<th>Probabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\hat{\pi}_{11})</td>
<td>0.85 (0.02)</td>
</tr>
<tr>
<td>(\hat{\pi}_{22})</td>
<td>0.87 (0.02)</td>
</tr>
</tbody>
</table>

The standard errors are in parenthesis. The table reports the estimates of the transition probabilities. The estimates are the results of the implementation of the EM algorithms Eq. 2.19, 2.20 and 2.21. \(\hat{\pi}_{ii}\) is the estimated probability of staying in state \(i\) when the initial state is \(i\).

Table 2.10 : Statistical Test of Predictive Ability: Forecast Horizons = 1, 5, 10 and 22 days

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(5)</th>
<th>(10)</th>
<th>(22)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Margin 1 (AXP)</td>
<td>2.17</td>
<td>1.50</td>
<td>1.78</td>
<td>2.32</td>
</tr>
<tr>
<td>Margin 1 (BAC)</td>
<td>1.79</td>
<td>0.89</td>
<td>-0.23</td>
<td>0.43</td>
</tr>
<tr>
<td>Margin 1 (COF)</td>
<td>5.37</td>
<td>3.43</td>
<td>0.77</td>
<td>2.64</td>
</tr>
<tr>
<td>Margin 1 (GS)</td>
<td>2.40</td>
<td>1.72</td>
<td>2.49</td>
<td>2.30</td>
</tr>
<tr>
<td>Margin 1 (JPM)</td>
<td>2.35</td>
<td>1.80</td>
<td>1.83</td>
<td>3.23</td>
</tr>
<tr>
<td>Margin 1 (MS)</td>
<td>6.08</td>
<td>1.57</td>
<td>0.42</td>
<td>0.42</td>
</tr>
<tr>
<td>Margin 1 (USB)</td>
<td>2.88</td>
<td>0.35</td>
<td>-0.64</td>
<td>0.03</td>
</tr>
<tr>
<td>Correlation Contribution</td>
<td>3.08</td>
<td>2.72</td>
<td>3.60</td>
<td>2.54</td>
</tr>
<tr>
<td>Multivariate test</td>
<td>3.14</td>
<td>2.66</td>
<td>4.54</td>
<td>2.92</td>
</tr>
</tbody>
</table>

Test of predictive ability: Loss(DCC) - Loss(RSDDC). A positive value means the RRSDC has better predictive ability. The critical values are: 2.5758(1 percent), 1.96 (5 percent), 1.6449 (10 percent)
Figure 2.1: Data Retention Percentages

The graph depicts the evolution of the measure of data retention when the "refresh sampling" methodology is used to synchronize the daily intraday returns. The measure is $p = \frac{\sum_{i=1}^{N} n_i}{n} \times 100$, where $n$ is the number of assets, $n_i$ the initial number of transactions for the $i^{th}$ asset and $N$ is the common number of transactions selected after the synchronization.
Figure 2.2: Annualized Realized Volatility for AXP, BAC, COF and GS - January 03, 2006 to April 30, 2012
2.7. TABLES AND FIGURES

Figure 2.3: Annualized Realized Volatility for JPM, MS and USB - January 03, 2006 to April 30, 2012
Figure 2.4: Annualized Realized Volatility for AXP, BAC, COF and GS - September 1, 2008 to June 1, 2009
2.7. TABLES AND FIGURES

Figure 2.5: Annualized Realized Volatility for JPM, MS and USB - September 1, 2008 to June 1, 2009
Figure 2.6 Smoothed Probability of Occurrence of Regime 1 or 2 - January 03, 2006 to April 30, 2012
Figure 2.7: Filtered Probability of Occurrence of Regime 1 or 2 - January 03, 2006 to April 30, 2012
Figure 2.8 Smoothed Probability of Occurrence of Regime 1 or 2 September 1, 2008 to June 1, 2009
Figure 2.9: Filtered Probability of Occurrence of Regime 1 or 2: September 1, 2008 to June 1, 2009
Figure 2.10: Realized and Conditional covariance and correlation JPM and BAC - January 03, 2006 to April 30, 2012
Figure 2.11: Realized and Conditional covariance and correlation JPM and BAC - September 1, 2008 to June 1, 2009
Chapter 3

The Multivariate GARCHX

3.1 Introduction

Chapter 3 introduces a methodology for incorporating high frequency information in Multivariate GARCH modeling. Inspired by the Flexible GARCH method of Ledoit et al. [2003], the framework presented is potentially applicable to large-scale problems. The proposed methodology is a straightforward extension of the univariate modified GARCH model with the modification consisting in the substitution of the outer product of the lagged returns for a realized measure of volatility. The model is called MGARCHX for Multivariate GARCH-X, where the X stands for the realized covariance. The computational burden is alleviated by a decentralization of the estimation method. The MGARCHX also provides a multivariate relationship between conditional and realized variances allowing multistep ahead forecasting of both conditional and realized covariances. In an empirical application, the MGARCHX model is used to estimate the conditional covariances of the daily returns for ten assets. The forecasts generated by the proposed model are compared to those generated by the scalar Multivariate HEAVY (MHEAVY) of Noureldin and Sheppard [2012] using the Diebold and Mariano [2002] and West [1996] test. Our proposed model shows a statistical improvement of the forecasting accuracy. However, when the comparison criteria is the one-period ahead forecasted Value-at-Risk, the performance of the MGARCHX model is comparable to that of the scalar MHEAVY model.

Section 2 below, presents the details of the model. The estimation and forecasting procedures are discussed in Section 3. Section 4 presents an empirical application with transaction-level data for ten stocks. Section 5 evaluates the forecasting ability of the model in comparison to the MHEAVY model and the conclusion follows in section 6.
3.2 The Multivariate GARCHX

Let $Y_t$ be a $n \times 1$ vector of daily demeaned log-returns such that:

$$Y_t = H_t^{1/2}U_t,$$

$$E(U_t) = 0,$$

$$\text{var}(U_t) = I_n.$$  

(3.1)

$H_t^{1/2}$ is a $n \times n$ positive definite matrix, $I_n$ is the identity matrix and $0$ is a $n \times 1$ vector of zeros. Conditional on $\mathcal{F}_{t-1}$, the $t-1$ information set, the variance of $Y_t$ is:

$$\text{var}(Y_t|\mathcal{F}_{t-1}) = H_t.$$  

(3.2)

Thus, $H_t^{1/2}$ could be the Cholesky factor of the positive definite matrix $H_t$ which is measurable with respect to $\mathcal{F}_{t-1}$.

Given, the observed vectors of daily returns and the calculated realized covariance matrices $\{X_t\}_{t=1}^T$, the proposed multivariate model describes the dynamic of $\{H_t\}_{t=1}^T$ and $\{X_t\}_{t=1}^T$.

3.2.1 Model Description

The formulation proposed follows the general GARCH specification of Bollerslev et al. [1988]. However, the matrix of outer product of past error vector is replaced by a matrix of realized covariance and the dynamic is:

$$H_t = C + B \circ H_{t-1} + A \circ X_{t-1},$$  

(3.3)

where $\circ$ is the Hadamard multiplication sign. The symmetric nature of all the matrices in Eq. 3.3 allows a ‘vech’ transformation.

$$\text{vech } H_t = \text{vech } C + \text{vech } B \circ \text{vech } H_{t-1} + \text{vech } A \circ \text{vech } X_{t-1},$$

where $\text{vech}(\bullet)$ is the operator that stacks the lower triangular elements of a $n \times n$ matrix in a $\frac{n(n+1)}{2} \times 1$ vector. It is therefore implied that the conditional (co)variance evolves as in the

$\text{vech } H_t$
3.2. THE MULTIVARIATE GARCH

following:

\[ h_{ij,t} = c_{ij} + b_{ij} h_{ij,t-1} + a_{ij} x_{ij,t-1}, \quad i, j = 1, 2, ..., n, \]  

(3.4)

where \( h_{ij,t} \) and \( x_{ij,t} \) are the conditional and realized covariances. \( c_{ij}, a_{ij} \) and \( b_{ij} \) are the elements of the matrices \( C, A \) and \( B^2 \). The conditional covariance is therefore a function of its own lagged value and the lag value of the realized covariance.

To understand the dynamic of the model, consider the following univariate GARCH of order 1:

\[ h_{ij,t} = \phi_0 + \phi_1 h_{ij,t-1} + \phi_2 v_{t-1}, \quad i, j = 1, 2, ..., n, \]  

(3.5)

where \( v_{t-1} \) is a martingale difference sequence (MDS) equal to \( y_{ij,t-1}^2 - h_{ij,t-1} \). The past square return \( y_{ij,t-1}^2 \) is used as proxy for the past latent variance. We can replace \( y_{ij,t-1}^2 \) in \( v_{t-1} \) by a realized measure of variance, \( x_{ij,t-1} \) as it has been shown to be a better proxy for the true variance (Lopez [2001]). However, \( x_{ij,t-1} - h_{ij,t-1} \) is no longer a MDS because in general \( x_{ij,t-1} < h_{ij,t-1} \). That is, for stock returns, the variance measured during open period is expected to be less than the daily variance because non-trivial overnight activities are not accounted. A simple solution to this problem is to define the realized variance as a linear function of the daily conditional variance. We can express \( x_{ij,t} \) as approximately equal to \( \phi_3 + \phi_4 h_{ij,t} \), with the expectation that \( \phi_4 \leq 1 \). Eq. 3.5 can therefore be modified as such:

\[ h_{ij,t} = \phi_0 + \phi_1 h_{ij,t-1} + \phi_2 (x_{ij,t-1} - \phi_3 - \phi_4 h_{ij,t-1}) \quad \text{or,} \]
\[ h_{ij,t} = (\phi_0 - \phi_2 \phi_3) + (\phi_1 - \phi_2 \phi_4) h_{ij,t-1} + \phi_2 x_{ij,t-1}. \]  

(3.6)

Comparing Eq. 3.4 and 3.6, we can see that \( c_{ij} = \phi_0 - \phi_2 \phi_3, b_{ij} = \phi_1 - \phi_2 \phi_4 \) and \( a_{ij} = \phi_2 \). A second dynamic for \( x_{ij,t} \) is therefore required, not only to completely identify all the parameters but also to allow forecasting at more than one period ahead.

The natural multivariate extension of the MDS condition is:

\[ E(X_t|\mathcal{F}_{t-1}) = QQ' + MH_tM, \]  

(3.7)

where the \( n \times n \) matrices \( Q \) and \( M \) are respectively lower triangular and diagonal. The form \( QQ' \) constrains the constant matrix to be positive definite while the diagonal assumption of \( M \) allows a simplification of the estimation procedure. Given the properties of the Haddamard

\[ \text{when } i = j, \text{ we have the conditional and realized variances} \]
multiplication, $MH_tM = mm' \odot H_t$, with $m$ being the vector of the diagonal elements of $M$. Thus, we have for each realized (co)variance at time $t$:

$$x_{ij,t} = q_{ij} + m_i m_j h_{ij,t}, \quad i, j = 1, 2, ..., n,$$

where $q_{ij}$ is the element on the $i^{th}$ row and $j^{th}$ column element of $QQ'$. The model as proposed in Eq. 3.3 and 3.7 faces the two issues in multivariate modeling of covariances. First, there are $4n(n+1)/2 + n$ parameters to estimate. Furthermore, the coefficients matrices $A$, $B$ and $C$ must meet some necessary conditions in order for the conditional covariance matrix to be positive definite. The method, explained below, not only ensures that estimated $A$, $B$ and $C$ yield positive definite covariance matrix but also allows the easy estimation of the model in high-dimensional case.

3.3 Estimation, Inference and Forecasting

Let us denote by $\Theta^1$ and $\Theta^2$ the respective parameters vectors in Eq. 3.3 and 3.7 and let us assume that $\Theta^1 \perp \Theta^2$. The parameters $\Theta^1 = \{\text{vech } A, \text{vech } B, \text{vech } C\}$ are estimated by a modification of the flexible GARCH method introduced by Ledoit et al. [2003], while $\Theta^2 = \{\text{vech } Q, m\}$ are estimated by the method of moments to avoid any distributional assumption on the realized covariance matrices.

3.3.1 Estimation of the Multivariate GARCHX

3.3.1.1 The Flexible GARCH

The Flexible GARCH model of Ledoit et al. [2003] reduces the computational burden of large-scale inference by a decentralization of the estimation. The idea consists in estimating the parameters $\{c_{ij}, b_{ij}, a_{ij}\}$ for each $(i, j)$ separately in the equation,

$$H_t = C + B \odot H_{t-1} + A \odot Y_{t-1} Y'_{t-1}. \quad (3.8)$$

Therefore, the problem reduces to the estimation of one-dimensional (variances) or two-dimensional (covariances) models. Using Maximum Likelihood (ML) or Quasi Maximum Likelihood (QML) there are $\frac{n(n+1)}{2}$ optimization routines which, given modern computer capabilities, is feasible. Once all the parameters are estimated, they are collected in $\hat{A}, \hat{B}$ and $\hat{C}$. However at this point, these estimated matrices will not necessarily imply positive definite conditional covariance matrices. To obtain a sequence, $\{H_t\}_{t=1}^T$ of positive definite matrices, $A$, $B$ and $C$ must...
meet certain conditions. It is easy to see that they must be symmetric. Furthermore, Ledoit et al. [2003] show that $A, B$ and $D = C \odot (ii - B)$ must be positive definite\(^3\). New coefficients matrices $\tilde{A}, \tilde{B}$ and $\tilde{C}$, which satisfy the conditions above are chosen to be the closest in norm to $\hat{A}, \hat{B}$ and $\hat{C}$. Formally, the following problems are solved:

\[
\min_{\tilde{D}} \| \tilde{D} - \hat{D} \|
\]

s.t. $\tilde{D}$ is positive-definite and $\tilde{d}_{ii} = \hat{d}_{ii}, \forall i = 1,.., n,$

\[
\min_{\tilde{A}} \| \tilde{A} - \hat{A} \|
\]

s.t. $\tilde{A}$ is positive-definite and $\tilde{a}_{ii} = \hat{a}_{ii}, \forall i = 1,.., n,$

\[
\min_{\tilde{B}} \| \tilde{B} - \hat{B} \|
\]

s.t. $\tilde{B}$ is positive-definite and $\tilde{b}_{ii} = \hat{b}_{ii}, \forall i = 1,.., n.$

Notice that the diagonal elements of the new matrices, $\tilde{A}, \tilde{B}$ and $\tilde{C}$, are equal to the diagonal elements of $\hat{A}, \hat{B}$ and $\hat{C}$.

**3.3.1.2 Estimation of the Conditional Covariance Equation**

Application of the Flexible GARCH method to our model yields estimates of $A, B$ and $C$ in Eq. 3.3. The steps for the estimation of $\Theta^1$ are the following:

1. **Diagonal Coefficients**: The diagonal elements of $A, B$ and $C$ are consistently estimated by QML. For each $i = 1,.., n$, the problem is:

\[
\max_{c_{ii}, b_{ii}, a_{ii}} - \frac{1}{2} \sum_{t=1}^{T} \left\{ \ln 2\pi + \ln h_{ii,t} + \frac{y_{it}^2}{h_{ii,t}} \right\}
\]

s.t. $h_{ii,t} = c_{ii} + b_{ii} h_{ii,t-1} + a_{ii} x_{ii,t-1},$

$c_{ii}, b_{ii}, a_{ii} > 0; \quad b_{ii} < 1, \quad \forall i = 1,.., n.$ \(3.9\)

2. **Off-Diagonal Coefficients**: With the diagonal parameter estimates the sequences of conditional variances, $\{h_{i,1}, h_{i,2}, ..., h_{i,T}\}_{i=1}^{n}$ are constructed. These sequences are used to estimate the off-diagonal elements by QML. For each $i = 1,.., n$ and $j = i + 1,.., n$, the

---

\(^3\)The term $\odot$ is the element by element division sign and $ii$ is a matrix of the same order as $B$ composed of 1s.
maximization problem is:

$$
\max_{c_{ij}, b_{ij}, a_{ij}} \left\{ 2 \ln 2\pi + \ln |H_{ij,t}| + Y_{ij,t}H_{ij,t}^{-1}Y_{ij,t} \right\}
$$

s.t. \( h_{ij,t} = c_{ij} + b_{ij} h_{ij,t-1} + a_{ij} x_{ij,t-1} \),

\( |c_{ij}| \leq (\hat{c}_{ii}\hat{c}_{jj})^{1/2} \), \( 0 \leq a_{ii} \leq (\hat{a}_{ii}\hat{a}_{jj})^{1/2} \) and \( |b_{ij}| \leq (\hat{b}_{ii}\hat{b}_{jj})^{1/2} \). \tag{3.10}

\( H_{ij,t} \) is the \( 2 \times 2 \) covariance matrix of assets \( i \) and \( j \). That is:

$$
H_{ij,t} = \begin{pmatrix}
    h_{ii,t} & h_{ij,t} \\
    h_{ji,t} & h_{jj,t}
\end{pmatrix}.
$$

\( Y_{ij,t} \) is the 2-dimensional vector of returns \( y_{i,t} \) and \( y_{j,t} \). The constraints on the parameters \( c_{ij}, a_{ij} \) and \( b_{ij} \) are to ensure that the sub-covariance matrix \( H_{ij,t} \) is positive definite.

3. Non-linear Transformation of Parameters Matrix Estimates: The resulting estimates \( \hat{C}, \hat{B} \) and \( \hat{A} \) are transformed as explained above into matrices \( \tilde{C}, \tilde{B} \) and \( \tilde{A} \). The latter estimates are used to calculate the conditional covariance matrices of the system.

3.3.2 Estimation of the Multivariate Measurement Equation

The estimation of \( \Theta^2 = \{ \text{vech } Q, m \} \) is greatly simplified because of the moment conditions and the assumption of diagonality of the matrix \( M \):

$$
E(X_t | \mathcal{F}_{t-1}) = QQ' + MH_t M,
$$

\( x_{ij,t} = q_{ij} + m_i m_j h_{ij,t} \) \( i, j = 1, 2, ..., n. \)

Therefore we can use the non-linear least squares method of Jennrich [1969] and Malinvaud [1970]. That is, \( \text{vech } Q \) and \( m \) are estimated by minimizing the prediction errors for the unique elements of \( X_t \):

$$
\min_{M} \sum_{t=1}^{T} \| \text{vech } X_t - \text{vech } E(X_t | \mathcal{F}_{t-1}) \|^2 \quad \text{or,}
$$

$$
\min_{M} \sum_{t=1}^{T} \left\{ \sum_{i=1}^{n} \sum_{j \geq i}^{n} \left( x_{ij,t} - q_{ij} - m_i m_j h_{ij,t} \right)^2 \right\}. \tag{3.11}
$$

or,

$$
\min_{M} \sum_{t=1}^{T} \left\{ \sum_{i=1}^{n} \sum_{j \geq i}^{n} \left( x_{ij,t} - q_{ij} - m_i m_j h_{ij,t} \right)^2 \right\}. \tag{3.12}
$$
3.3.3 Bootstrap for Standard Errors of Parameter Estimates

The disadvantage of the estimation procedure is that the standard errors are not straightforward to estimate. This is the case because the matrices \( \tilde{C}, \tilde{B} \) and \( \tilde{A} \) are non-linear transformations of \( \hat{C}, \hat{B} \) and \( \hat{A} \). With no form for the derivative of the transformation, the delta method cannot be applied here. Much as in Ledoit et al. [2003], the standard errors can be calculated by bootstrap. However, unlike in the Flexible GARCH model, there is the need to sample from the “observed” realized covariances. The complete steps are the following:

1. For \( k = 1, \ldots, K \), generate the data \( \{(Y_{k,1}^*, X_{k,1}^*), (Y_{k,1}^*, X_{k,1}^*), \ldots, (Y_{k,T}^*, X_{k,T}^*)\} \) by resampling blocks of data. The block bootstrap method ensures that the dependence in the data is preserved.

2. Compute the estimates \( \tilde{C}_k, \tilde{B}_k, \tilde{A}_k \) and \( \tilde{M}_k \).

3. The sample standard deviations of the elements of the matrices are the bootstrap standard errors.

Ledoit et al. [2003] proposes to use \( K \geq 100 \) for the algorithm above. This extra step comes at a computational cost when the number of assets is large. However, as noticed by Jondeau et al. [2007], the need for standard errors for large-scale multivariate GARCH models is not clear.

3.3.4 Multi-step Ahead Forecasting

Forecasts of the conditional covariance for more than one step ahead is achieved by iteration. For horizon \( L \geq 1 \), we have:

\[
E_T[H_{T+L}] = C + B \odot E_T[H_{T+L-1}] + A \odot E_T[X_{T+L-1}],
\]

\[
E_T[X_{T+L}] = QQ' + ME_T[H_{T+L}]M = QQ' + mm \odot E_T[H_{T+L}]. \tag{3.13}
\]

Thus, to find \( E_T[H_{T+L}] \), one iterates 3.13 for \( l = 1, \ldots, L \). Notice also that, the realized covariance forecast is a by-product of the model.

3.4 Empirical Application

The model is applied to high-frequency prices for ten assets from the S&P 100\(^4\). The sample period runs from January 3, 2006 to April 30, 2012 corresponding to a sample size of 1,593.

\(^4\)The ticker symbols of the assets are AXP, BAC, CVX, KO, DD, GE, IBM, JPM, MSFT and XOM
3.4. EMPIRICAL APPLICATION

However, days for which the daily transaction period is less than the normal 6.5 hours are removed\(^5\). The final sample size is 1,581 days.

3.4.1 Data Description

3.4.1.1 Realized Volatility Description

The multivariate realized kernel estimator of Barndorff-Nielsen et al. [2011] is used as measure of realized covariance. The data retention ratio introduced in section 2.1 of chapter 1 is

\[ p = \frac{n N}{\sum_{i=1}^{n_i} n_i}, \]

where \( n \) is the number of assets, \( n_i \) and \( N \) are respectively the initial and synchronized number of transactions for the \( i^{th} \) asset. On average 32.9\% of the daily data is retained. The minimum retention is 20.7\% while the maximum is 70.5\%. Table 3.1 shows the statistics for the percentage of retention per asset in the sample. These percentages are computed as the ratio of the number of daily transactions selected over the total of daily transactions for a particular asset. One drawback of the realized kernel estimation is the increasing loss of data when the number of assets considered is large. This can be seen here as the sampling percentage can be as low as 9.62\% and as high as 82.66\%.

As for the bandwidth selection, the minimum, maximum and mean of the bandwidths are respectively 24, 72 and 51. The bandwidth also corresponds to the number of autocovariances included in the computation of the realized matrices.

The summary statistics for the annualized realized volatilities are reported in the first part of Table 2.2\(^6\). The assets from the financial sector are the most volatile in the sample. Bank of America (BAC) is the most volatile with an average of 37.18\% while the less volatile asset is the Coca-Cola Company (KO) with 15.94\%.

3.4.1.2 Daily Log-Return Description

The summary statistics for the daily log close-to-close returns are presented in the second part of Table 3.2. As expected, the averages are in the neighborhood of 0\% except for BAC. The standard deviations of the daily return confirm that the assets from the financial sector are the most volatile. The variation of returns is the largest for BAC and the lowest for KO. The ranges of the return are large, for instance the minimum daily return for BAC is \(-33.79\%\) and the maximum is 31.66\%.

---

\(^5\)Independence Day, the days before Thanksgiving and Christmas

\(^6\)The annualized realized volatility is the square root of 252 times the realized kernel estimates.
3.4. EMPIRICAL APPLICATION

3.4.2 Estimation Results

This subsection presents the results of the estimation as introduced in section 3. For clarity of exposition the estimates along with their bootstrap standard errors are presented at the very end of the chapter. However, a summary of the estimates are shown in Table 3.3.

The first five rows of Table 3.3 present the average of the parameters estimates as in Eq. 3.7 and 3.3. A first look at these results confirms that the effects of the past realized variances and covariances ($a_{ij}$) are non-trivial as they average around 0.58 and 0.56. The persistence of the daily conditional covariances and variances is calculated (Eq. 3.6) and the averages are presented in the last row of the table. The autoregressive parameters of the conditional variances is equal to 1 on average. This result is consistent with the existing literature and explains why some researchers often model the daily conditional variance as an Integrated GARCH. As seen in Eq. 3.6, the parameter $a_{ij}$ represents the effect of the MDS, $v_{t-1}$. The substitution of the function of the realized variance for the square return allows the model to correct the dynamic of the conditional variance when $v_{t-1} \neq 0$. The strength of this correction is proportional to $a_{ij}$. Finally, we also observe that although persistent, the covariances are less so than the variances as the average of the estimates is 0.879. For the parameters of the multivariate measurement equation, the estimates of $m_{ij}$ are on average less than one, suggesting that the variance measured while the market is open corresponds to about 66% of the daily variance.

3.4.3 Model Fit

To assess the fit of the model, the multivariate Ljung-Box portmanteau (or HM) test of Hosking [1980] is applied to the standardized residuals to detect any residual ARCH effects. It has the form:

$$HM(W) = T^2 \sum_{j=1}^{W} (T-j)^{-1} \text{tr}\{A_{\hat{z}_t}^{-1}(0)A_{\hat{z}_t}(j)A_{\hat{z}_t}^{-1}(0)A'_{\hat{z}_t}(j)\},$$

where $\hat{z}_t = \text{vech}\left[H_t^{-1/2}Y_tY_t'\right]$ and $A_{\hat{z}_t}(j)$ is the sample autocovariance matrix of order $j$. Under the null hypothesis of no ARCH effects, the test is distributed asymptotically as Chi-square with degree of freedom $W \times n^2$.

With $W = 20$ lags, the HM statistics is 2,094. Given that the 5% critical value is 2,105, there is no sign of ARCH effects in the standardized residuals.

A 20 lags HM test on $Y_t$ was initially performed and the Test statistics was 4,614.

7A 20 lags HM test on $Y_t$ was initially performed and the Test statistics was 4,614.
3.5 Out-of-sample Forecast Evaluation

In this section, the forecasts generated by the MGARCHX model is compared to those generated by the scalar MHEAVY model of Noureldin and Sheppard [2012]. The full MHEAVY model has the following form:

\[
\begin{align*}
H_t &= C_H + A_H H_{t-1} A_H + B_H X_{t-1} B_H, \\
M_t &= C_M + A_M M_{t-1} A_M + B_M X_{t-1} B_M,
\end{align*}
\]

where \(C_H, A_H, B_H, C_M, A_M\) and \(B_M\) are \(n \times n\) matrices. The full MHEAVY model is adequate for very low-dimensional use as it suffers from the curse of dimensionality. A simplification consisting in assuming scalar coefficients \(a_H, b_H, a_M\) and \(b_M\) is also introduced in Noureldin and Sheppard [2012]. Along with covariance targeting, the scalar coefficients assumption reduces the computational burden of the MHEAVY model. For the data used in this chapter, it takes 11 seconds to estimate the parameters of the model (without calculating standard errors). One advantage of the MGARCHX model however, is that only 35 seconds is needed to estimate the full model (without calculating standard errors) for 10 assets.

To conduct the forecast analysis, the sample is split into two parts. The in-sample for the initial estimation runs from January 4, 2006 to April 30, 2010 and has a size of 1079. With a size of 502 (approximately two years), the out-of-sample spans from May 3, 2010 to April 30, 2012. The models are estimated using a rolling window of 1,079 observations and the parameter estimates are used to forecast the conditional and realized covariances at horizons 1, 2, 3, 5, 10 and 22 days.

3.5.1 Statistical Forecast Evaluation

We implement the DMW test explained in section 2.2 of Chapter 1 and the proxy for the true covariance is taken to be the realized kernel (as in Noureldin and Sheppard [2012]). The loss function is evaluated for each model and the loss differential is computed as:

\[
d_{t,L} = L_{t,L}(H_t^0, H_{t+L}^{MHEAVY}) - L_{t,L}(H_t^0, H_{t+L}^{MGARCHX}).
\]

The direction of the difference implies that the forecast generated by the MGARCHX model is more accurate than the one generated by the MHEAVY model if \(d_{t,L}\) is positive. With the
time series of $d_{t,L}$, we test for all $L$:

$$H_0 : E[L_{t,L}(H_{t+L}^0, H_{t+L}^{MHEAVY})] = E[L_{t,L}(H_{t+L}^0, H_{t+L}^{MGARCHX})],$$

vs.

$$H_1 : E[L_{t,L}(H_{t+L}^0, H_{t+L}^{MHEAVY})] > E[L_{t,L}(H_{t+L}^0, H_{t+L}^{MGARCHX})],$$

$$H_2 : E[L_{t,L}(H_{t+L}^0, H_{t+L}^{MHEAVY})] < E[L_{t,L}(H_{t+L}^0, H_{t+L}^{MGARCHX})].$$

The outcomes of the test and its decomposition are reported in Table 3.4. The results show superior predictive ability of the MGARCHX model for all the horizons considered. The differences in the loss function are all statistically significant for the joint distribution. We can however notice that the gain in accuracy diminishes with the horizon but stays significant at the 22 days forecast. The margin effect also shows superior predictive power, especially in short horizon (one, two and three days). The tendency is reversed however for IBM, MSFT and KO starting at the five-day horizon. As for the correlation forecast, the MGARCHX model is statistically more accurate than the MHEAVY model. One reason for these results could be attributed to the fact that the scalar MHEAVY model assumes scalar coefficients so that the dynamic of the covariance matrix is not fully described.

**3.5.2 Economic Forecast Evaluation**

In this subsection, the forecasting ability of the two models are evaluated by testing the performance of the Value-at-Risk (VaR) forecasts based on the conditional covariance predictions. The VaR, which is defined as the $\alpha$-quantile of the conditional distribution of the return of a portfolio, is calculated using filtered historical simulation. That is, a portfolio with equal weight composed of the ten assets is created. Then, the conditional covariances from each model are used to compute the standardized residuals. Finally, the one and five % quantile of these standardized residuals are calculated using 252 observations. Thus, the VaR for $t + 1$ is equal to:

$$VaR_{\alpha,t+1} = \sigma_{p,t+1} F_{p,t}^{-1}(\alpha), \quad \alpha = 1, 5\%,$$

where $\sigma_{p,t+1}$ is the forecasted portfolio standard error and $F_{p,t}^{-1}(\alpha)$ is the $\alpha$-quantile. The backtesting procedure of Christoffersen [2012] are applied to the VaRs obtained. The unconditional coverage test assesses whether or not the fraction of violations obtained from a model is significantly different from the one implied by $\alpha$. The second test, assesses the independence of the sequence of violations. Finally, the conditional coverage test checks simultaneously that the
3.6. CONCLUSION

violations are independent and that the average number of violations is correct. The results of the three tests are presented in Table 3.5.

For the 1% VaR, the two methods perform fairly well. They all pass the three tests however, the number of violations for the MGARCHX model exceeds the expected number of violations by 2 while the MHEAVY model exceeds them by 3. As evidenced in Fig. 3.2 and 3.3, the 1% VaRs appears to predict well the minimum loss of the portfolio. However, the violations for the MHEAVY model are clustered in the period of August 2011.

The results for the 5% VaR is different. The two methods only pass the unconditional coverage test, and have the same number of violations (28). Fig. 3.4 and 3.5 show that the majority of the violations in both models happens in the period of July to October 2011.

In summary both model appears to perform equally when compared in a VaR framework. They have a similar number of violations even though there is more variation in the VaR predicted by the MGARCHX model.

3.6 Conclusion

This chapter has introduced the Multivariate GARCHX (MGARCHX) model for vectors of daily returns and corresponding realized measures. The estimation method is inspired by the Flexible GARCH model of Ledoit et al. [2003] where the computational burden of large-scale covariance matrices inference is simplified by a decentralized estimation procedure. The conditional covariance has a GARCH-type structure but with the outer product of the returns vector replaced by the realized covariance. To ensure positive definitiveness of the resulting covariance matrices, a non-linear transformation of the parameters matrices is carried. Furthermore, a multivariate relation between realized and conditional covariance is derived. The latter allows the forecasting of the conditional covariance at horizon greater than one-period ahead. Compared to the scalar HEAVY model of Noureldin and Sheppard [2012], in a forecasting exercise for data composed of daily returns and realized covariances for ten assets, the MGARCHX has a better statistical predictive power. However, using a Value-at-Risk criteria to compare the models, we find that they perform equally.

The main drawback of the model is that there is no straightforward way to generate standard errors for the parameters estimates and they are computed by bootstrap in this chapter.
3.7 Tables and Figures

Table 3.1: Summary Statistics for Refresh Sampling: January 03, 2006 to April 30, 2012

<table>
<thead>
<tr>
<th>Asset</th>
<th>Mean</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>AXP</td>
<td>44.66</td>
<td>19.54</td>
<td>73.04</td>
</tr>
<tr>
<td>BAC</td>
<td>25.28</td>
<td>10.02</td>
<td>65.38</td>
</tr>
<tr>
<td>CVX</td>
<td>35.27</td>
<td>13.04</td>
<td>68.91</td>
</tr>
<tr>
<td>KO</td>
<td>44.62</td>
<td>17.49</td>
<td>75.37</td>
</tr>
<tr>
<td>DD</td>
<td>51.23</td>
<td>24.82</td>
<td>82.66</td>
</tr>
<tr>
<td>GE</td>
<td>26.82</td>
<td>11.99</td>
<td>65.26</td>
</tr>
<tr>
<td>IBM</td>
<td>42.83</td>
<td>15.54</td>
<td>81.34</td>
</tr>
<tr>
<td>JPM</td>
<td>29.07</td>
<td>15.20</td>
<td>65.90</td>
</tr>
<tr>
<td>MSFT</td>
<td>29.04</td>
<td>9.62</td>
<td>67.27</td>
</tr>
<tr>
<td>XOM</td>
<td>28.07</td>
<td>11.33</td>
<td>65.52</td>
</tr>
</tbody>
</table>

The numbers are the statistics for the percentages of retention after synchronization. The retention percentage is computed as \( \frac{n_t}{n_i} \times 100 \), where \( n_i \) is the initial number of transactions for asset \( i \) at time \( t \) and \( N_t \) is the number of transactions selected after synchronization.
### 3.7. TABLES AND FIGURES

#### Table 3.2: Summary Statistics for the Daily Log Returns and the Annualized Realized Volatilities: January 03, 2006 to April 30, 2012

<table>
<thead>
<tr>
<th>Stocks</th>
<th>Realized volatility</th>
<th>Log returns</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>St. dev</td>
</tr>
<tr>
<td>AXP</td>
<td>31.26</td>
<td>24.65</td>
</tr>
<tr>
<td>BAC</td>
<td>37.18</td>
<td>35.88</td>
</tr>
<tr>
<td>CVX</td>
<td>23.12</td>
<td>14.54</td>
</tr>
<tr>
<td>KO</td>
<td>15.94</td>
<td>10.29</td>
</tr>
<tr>
<td>DD</td>
<td>25.34</td>
<td>14.68</td>
</tr>
<tr>
<td>GE</td>
<td>26.00</td>
<td>20.97</td>
</tr>
<tr>
<td>IBM</td>
<td>19.09</td>
<td>12.91</td>
</tr>
<tr>
<td>JPM</td>
<td>32.67</td>
<td>26.20</td>
</tr>
<tr>
<td>MSFT</td>
<td>22.11</td>
<td>12.47</td>
</tr>
<tr>
<td>XOM</td>
<td>21.31</td>
<td>13.72</td>
</tr>
</tbody>
</table>

The log returns are in percentage point. The annualized realized volatility is the square root of 252 times the realized kernel.

#### Table 3.3: Average Values of the Parameters Estimates for the MGARCHX Model

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Variances ((i = j))</th>
<th>Covariances ((i \neq j))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a_{ij})</td>
<td>0.584</td>
<td>0.564</td>
</tr>
<tr>
<td>(b_{ij})</td>
<td>0.529</td>
<td>0.510</td>
</tr>
<tr>
<td>(c_{ij})</td>
<td>0.085</td>
<td>0.054</td>
</tr>
<tr>
<td>(m_{ij})</td>
<td>0.661</td>
<td>0.658</td>
</tr>
<tr>
<td>(q_{ij})</td>
<td>0.549</td>
<td>0.213</td>
</tr>
<tr>
<td>(\phi_1)</td>
<td>1.000</td>
<td>0.879</td>
</tr>
</tbody>
</table>

There are 10 variances and 45 covariances. The numbers reported are the mean across variances and covariances.
3.7. TABLES AND FIGURES

Table 3.4: Statistical Test of Predictive Ability: Forecast Horizons = 1, 2, 3, 5, 10 and 22 days

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(5)</th>
<th>(10)</th>
<th>(22)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Margin 1 (AXP)</td>
<td>14.08</td>
<td>13.71</td>
<td>13.15</td>
<td>12.28</td>
<td>0.92</td>
<td>-14.50</td>
</tr>
<tr>
<td>Margin 1 (BAC)</td>
<td>13.79</td>
<td>12.03</td>
<td>10.87</td>
<td>9.47</td>
<td>8.67</td>
<td>5.20</td>
</tr>
<tr>
<td>Margin 1 (CVX)</td>
<td>9.68</td>
<td>8.73</td>
<td>8.03</td>
<td>8.21</td>
<td>6.92</td>
<td>10.54</td>
</tr>
<tr>
<td>Margin 1 (KO)</td>
<td>9.58</td>
<td>8.59</td>
<td>5.04</td>
<td>-8.21</td>
<td>-7.29</td>
<td>-7.80</td>
</tr>
<tr>
<td>Margin 1 (DD)</td>
<td>7.99</td>
<td>7.41</td>
<td>6.91</td>
<td>8.26</td>
<td>1.10</td>
<td>-5.47</td>
</tr>
<tr>
<td>Margin 1 (GE)</td>
<td>9.55</td>
<td>9.32</td>
<td>7.82</td>
<td>-2.52</td>
<td>-13.36</td>
<td>-12.50</td>
</tr>
<tr>
<td>Margin 1 (IBM)</td>
<td>8.15</td>
<td>7.14</td>
<td>4.01</td>
<td>-7.11</td>
<td>-7.93</td>
<td>-8.44</td>
</tr>
<tr>
<td>Margin 1 (JPM)</td>
<td>14.17</td>
<td>13.45</td>
<td>12.92</td>
<td>13.10</td>
<td>15.10</td>
<td>10.32</td>
</tr>
<tr>
<td>Margin 1 (MSFT)</td>
<td>10.41</td>
<td>9.84</td>
<td>9.57</td>
<td>9.29</td>
<td>6.04</td>
<td>0.86</td>
</tr>
<tr>
<td>Margin 1 (XOM)</td>
<td>9.74</td>
<td>8.93</td>
<td>8.22</td>
<td>8.26</td>
<td>1.85</td>
<td>-3.22</td>
</tr>
<tr>
<td>Copula</td>
<td>4.34</td>
<td>5.91</td>
<td>6.05</td>
<td>5.64</td>
<td>3.71</td>
<td>3.44</td>
</tr>
<tr>
<td>Joint distribution</td>
<td>18.28</td>
<td>14.90</td>
<td>11.88</td>
<td>8.45</td>
<td>3.96</td>
<td>2.79</td>
</tr>
</tbody>
</table>

Test of predictive ability: Loss(DCC) - Loss(MGARCHX). A positive value means the MGARCHX has better predictive ability. The critical values are: 2.5758 (1 percent), 1.96 (5 percent), 1.6449 (10 percent).

Table 3.5: VaR Back-Testing Results for the MGARCHX and MHEAVY Models

<table>
<thead>
<tr>
<th></th>
<th>Number of exceed.</th>
<th>Uncond. coverage</th>
<th>Indep. of hits</th>
<th>Cond. coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 % VaR</td>
<td>exp(5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MGARCHX</td>
<td>7</td>
<td>0.71</td>
<td>0.23</td>
<td>0.94</td>
</tr>
<tr>
<td>MHEAVY</td>
<td>8</td>
<td>1.53</td>
<td>0.29</td>
<td>1.82</td>
</tr>
<tr>
<td>5 % VaR</td>
<td>exp(25)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MGARCHX</td>
<td>28</td>
<td>0.35</td>
<td>5.96</td>
<td>6.32</td>
</tr>
<tr>
<td>MHEAVY</td>
<td>28</td>
<td>0.35</td>
<td>7.12</td>
<td>7.47</td>
</tr>
</tbody>
</table>

The LR tests are presented in the table. The VaR are computed using Filtered Historical Simulation (FHS) rolling window of 252 days. The critical values are: 2.7055 (10 percent), 3.8415 (5 percent). The expected number of exceedences are 5 and 25 for the 1 and 5 % VaR.

53
3.7. TABLES AND FIGURES

Table 3.1: Data Retention Percentages

<table>
<thead>
<tr>
<th>Period</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan07</td>
<td>20</td>
</tr>
<tr>
<td>Jan08</td>
<td>30</td>
</tr>
<tr>
<td>Jan09</td>
<td>40</td>
</tr>
<tr>
<td>Jan10</td>
<td>50</td>
</tr>
<tr>
<td>Jan11</td>
<td>60</td>
</tr>
<tr>
<td>Jan12</td>
<td>70</td>
</tr>
</tbody>
</table>

Figure 3.1: Data Retention Percentages

The graph depicts the evolution of the measure of data retention when the “refresh sampling” methodology is used to synchronize the daily intraday returns. The measure is $p = \frac{n \cdot N}{\sum_{i=1}^{n} n^i} \times 100$, where $n$ is the number of assets, $n^i$ the initial number of transactions for the $i^{th}$ asset and $N$ is the common number of transactions selected after the synchronization.
3.7. TABLES AND FIGURES

Figure 3.2: 1% Portfolio Value-at-Risk Standardized Exceedences

The standardized portfolio exceedences are computed as \( \min \{ r_{t+1} - (\text{VaR}_{t+1}), 0 \} / \sigma_{p,t+1} \), where \( r_{t+1} \) is the realized portfolio return at \( t + 1 \).

Figure 3.3: Actual Portfolio Returns and 1% Value-at-Risk from the MGARCHX and MHEAVY Models

The vertical axis represents the rate of returns (\%)
3.7. TABLES AND FIGURES

![MGARCHX - Exceedence 5% VaR](image1)

![MHEAVY - Exceedence 5% VaR](image2)

Figure 3.4: 5% Portfolio Value-at-Risk Standardized Exceedences

The standardized portfolio exceedences are computed as
\[
\min \left\{ \frac{r_{t+1} - (VaR_{t+1}, 0)}{\sigma_{p,t+1}} \right\}, \text{ where } r_{t+1} \text{ is the realized portfolio return at } t + 1.
\]

![Actual returns](image3)

Figure 3.5: Actual Portfolio Returns and 5% Value-at-Risk from the MGARCHX and MHEAVY Models

The vertical axis represents the rate of returns (%)
### Table 3.6: Parameter Estimates of the “A” Matrix for the MGARCHX Model

<table>
<thead>
<tr>
<th></th>
<th>AXP</th>
<th>BAC</th>
<th>CVX</th>
<th>KO</th>
<th>DD</th>
<th>GE</th>
<th>IBM</th>
<th>JPM</th>
<th>MSFT</th>
<th>XOM</th>
</tr>
</thead>
<tbody>
<tr>
<td>AXP</td>
<td>0.7319</td>
<td>(0.0740)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAC</td>
<td>0.6449</td>
<td>(0.0573)</td>
<td>0.5714</td>
<td>(0.0725)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CVX</td>
<td>0.4489</td>
<td>(0.0440)</td>
<td>0.3966</td>
<td>(0.0432)</td>
<td>0.2773</td>
<td>(0.0444)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KO</td>
<td>0.6683</td>
<td>(0.0629)</td>
<td>0.5901</td>
<td>(0.0589)</td>
<td>0.4110</td>
<td>(0.0442)</td>
<td>0.6135</td>
<td>(0.0835)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DD</td>
<td>0.5560</td>
<td>(0.0547)</td>
<td>0.4912</td>
<td>(0.0508)</td>
<td>0.3419</td>
<td>(0.0414)</td>
<td>0.5090</td>
<td>(0.0630)</td>
<td>0.4248</td>
<td></td>
</tr>
<tr>
<td>GE</td>
<td>0.7194</td>
<td>(0.0725)</td>
<td>0.6356</td>
<td>(0.0649)</td>
<td>0.4424</td>
<td>(0.0516)</td>
<td>0.6586</td>
<td>(0.0750)</td>
<td>0.5479</td>
<td>(0.0631)</td>
</tr>
<tr>
<td>IBM</td>
<td>0.7087</td>
<td>(0.0654)</td>
<td>0.6261</td>
<td>(0.0655)</td>
<td>0.4358</td>
<td>(0.0489)</td>
<td>0.6488</td>
<td>(0.0684)</td>
<td>0.5397</td>
<td>(0.0571)</td>
</tr>
<tr>
<td>JPM</td>
<td>0.5334</td>
<td>(0.0475)</td>
<td>0.4712</td>
<td>(0.0494)</td>
<td>0.3280</td>
<td>(0.0354)</td>
<td>0.4883</td>
<td>(0.0472)</td>
<td>0.4062</td>
<td>(0.0421)</td>
</tr>
<tr>
<td>MSFT</td>
<td>0.8419</td>
<td>(0.0764)</td>
<td>0.7438</td>
<td>(0.0696)</td>
<td>0.5177</td>
<td>(0.0549)</td>
<td>0.7707</td>
<td>(0.0830)</td>
<td>0.6412</td>
<td>(0.0652)</td>
</tr>
<tr>
<td>XOM</td>
<td>0.5762</td>
<td>(0.0547)</td>
<td>0.5091</td>
<td>(0.0513)</td>
<td>0.3543</td>
<td>(0.0464)</td>
<td>0.5275</td>
<td>(0.0567)</td>
<td>0.4388</td>
<td>(0.0491)</td>
</tr>
</tbody>
</table>

The Standard errors are in parenthesis. The table reports the estimates from the model in Eq. 3.3
### Table 3.7: Parameter Estimates of the “B” Matrix for the MGARCHX Model

<table>
<thead>
<tr>
<th></th>
<th>AXP</th>
<th>BAC</th>
<th>CVX</th>
<th>KO</th>
<th>DD</th>
<th>GE</th>
<th>IBM</th>
<th>JPM</th>
<th>MSFT</th>
<th>XOM</th>
</tr>
</thead>
<tbody>
<tr>
<td>AXP</td>
<td>0.4849</td>
<td>(0.0373)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAC</td>
<td>0.5608</td>
<td>(0.0293)</td>
<td>0.6524</td>
<td>(0.0351)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CVX</td>
<td>0.5991</td>
<td>(0.0280)</td>
<td>0.6952</td>
<td>(0.0263)</td>
<td>0.7443</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KO</td>
<td>0.4381</td>
<td>(0.0341)</td>
<td>0.5081</td>
<td>(0.0352)</td>
<td>0.5431</td>
<td>0.3984</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DD</td>
<td>0.5597</td>
<td>(0.0299)</td>
<td>0.6494</td>
<td>(0.0276)</td>
<td>0.6938</td>
<td>0.5073</td>
<td>0.6495</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GE</td>
<td>0.4708</td>
<td>(0.0348)</td>
<td>0.5463</td>
<td>(0.0354)</td>
<td>0.5836</td>
<td>0.4268</td>
<td>0.5451</td>
<td>0.4597</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IBM</td>
<td>0.4243</td>
<td>(0.0356)</td>
<td>0.4924</td>
<td>(0.0392)</td>
<td>0.5260</td>
<td>0.3847</td>
<td>0.4914</td>
<td>0.4133</td>
<td>0.3736</td>
<td></td>
</tr>
<tr>
<td>JPM</td>
<td>0.5944</td>
<td>(0.0277)</td>
<td>0.6896</td>
<td>(0.0271)</td>
<td>0.7368</td>
<td>0.5388</td>
<td>0.6882</td>
<td>0.5799</td>
<td>0.5218</td>
<td>0.7320</td>
</tr>
<tr>
<td>MSFT</td>
<td>0.2952</td>
<td>(0.0580)</td>
<td>0.3426</td>
<td>(0.0665)</td>
<td>0.3660</td>
<td>0.2676</td>
<td>0.3419</td>
<td>0.2875</td>
<td>0.2592</td>
<td>0.3630</td>
</tr>
<tr>
<td>XOM</td>
<td>0.5462</td>
<td>(0.0313)</td>
<td>0.6337</td>
<td>(0.0288)</td>
<td>0.6770</td>
<td>0.4951</td>
<td>0.6324</td>
<td>0.5319</td>
<td>0.4795</td>
<td>0.6715</td>
</tr>
</tbody>
</table>

The Standard errors are in parenthesis. The table reports the estimates from the model in Eq. 3.3.
3.7. TABLES AND FIGURES

Table 3.8: Parameter Estimates of the “C” Matrix for the MGARCHX Model

<table>
<thead>
<tr>
<th></th>
<th>AXP</th>
<th>BAC</th>
<th>CVX</th>
<th>KO</th>
<th>DD</th>
<th>GE</th>
<th>IBM</th>
<th>JPM</th>
<th>MSFT</th>
<th>XOM</th>
</tr>
</thead>
<tbody>
<tr>
<td>AXP</td>
<td>0.0622</td>
<td>(0.1241)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAC</td>
<td>0.0000</td>
<td>(0.0927)</td>
<td>0.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CVX</td>
<td>0.0674</td>
<td>(0.0749)</td>
<td>0.0000</td>
<td>0.0725</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KO</td>
<td>0.0699</td>
<td>(0.0632)</td>
<td>0.0000</td>
<td>0.0755</td>
<td>0.0782</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DD</td>
<td>0.0395</td>
<td>(0.0761)</td>
<td>0.0000</td>
<td>0.0427</td>
<td>0.0442</td>
<td>0.0249</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GE</td>
<td>0.0665</td>
<td>(0.0920)</td>
<td>0.0000</td>
<td>0.0718</td>
<td>0.0745</td>
<td>0.0421</td>
<td>0.0707</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IBM</td>
<td>0.0908</td>
<td>(0.0975)</td>
<td>0.0000</td>
<td>0.0981</td>
<td>0.1018</td>
<td>0.0574</td>
<td>0.0968</td>
<td>0.1320</td>
<td></td>
<td></td>
</tr>
<tr>
<td>JPM</td>
<td>0.0000</td>
<td>(0.0701)</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td>MSFT</td>
<td>0.1521</td>
<td>(0.1352)</td>
<td>0.0001</td>
<td>0.1642</td>
<td>0.1705</td>
<td>0.0962</td>
<td>0.1621</td>
<td>0.2215</td>
<td>0.0001</td>
<td>0.3707</td>
</tr>
<tr>
<td>XOM</td>
<td>0.0503</td>
<td>(0.0726)</td>
<td>0.0544</td>
<td>0.0564</td>
<td>0.0318</td>
<td>0.0536</td>
<td>0.0732</td>
<td>0.1226</td>
<td>0.0404</td>
<td></td>
</tr>
</tbody>
</table>

The Standard errors are in parenthesis. The table reports the estimates from the model in Eq. 3.3.
3.7. TABLES AND FIGURES

Table 3.9: Parameter Estimates of the “M” Matrix for the MGARCHX Model

<table>
<thead>
<tr>
<th></th>
<th>AXP</th>
<th>BAC</th>
<th>CVX</th>
<th>KO</th>
<th>DD</th>
<th>GE</th>
<th>IBM</th>
<th>JPM</th>
<th>MSFT</th>
<th>XOM</th>
</tr>
</thead>
<tbody>
<tr>
<td>AXP</td>
<td>0.7827</td>
<td>0.7270</td>
<td>0.8871</td>
<td>0.8343</td>
<td>0.8522</td>
<td>0.7955</td>
<td>0.8563</td>
<td>0.7869</td>
<td>0.8000</td>
<td>0.7940</td>
</tr>
<tr>
<td></td>
<td>(0.0930)</td>
<td>(0.0955)</td>
<td>(0.1158)</td>
<td>(0.0872)</td>
<td>(0.0984)</td>
<td>(0.0887)</td>
<td>(0.0918)</td>
<td>(0.1068)</td>
<td>(0.0774)</td>
<td>(0.0958)</td>
</tr>
</tbody>
</table>

The Standard errors are in parenthesis. The table reports the estimates from the model in Eq. 3.7.
### Table 3.10: Parameter Estimates of the “QQ” Matrix for the MGARCHX Model

<table>
<thead>
<tr>
<th></th>
<th>AXP</th>
<th>BAC</th>
<th>CVX</th>
<th>KO</th>
<th>DD</th>
<th>GE</th>
<th>IBM</th>
<th>JPM</th>
<th>MSFT</th>
<th>XOM</th>
</tr>
</thead>
<tbody>
<tr>
<td>AXP</td>
<td>0.7365</td>
<td>(0.3979)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAC</td>
<td>0.4845</td>
<td>(0.2753)</td>
<td>1.3752</td>
<td>(0.8422)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CVX</td>
<td>0.1350</td>
<td>(0.1431)</td>
<td>0.2571</td>
<td>(0.1763)</td>
<td>0.2102</td>
<td>(0.2528)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KO</td>
<td>0.1150</td>
<td>(0.0713)</td>
<td>0.2032</td>
<td>(0.0805)</td>
<td>0.0902</td>
<td>(0.0815)</td>
<td>0.3252</td>
<td>(0.0990)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DD</td>
<td>0.2143</td>
<td>(0.1300)</td>
<td>0.2917</td>
<td>(0.1654)</td>
<td>0.0911</td>
<td>(0.1362)</td>
<td>0.1229</td>
<td>(0.0692)</td>
<td>0.3525</td>
<td>(0.1907)</td>
</tr>
<tr>
<td>GE</td>
<td>0.2676</td>
<td>(0.1500)</td>
<td>0.3812</td>
<td>(0.2065)</td>
<td>0.1464</td>
<td>(0.1246)</td>
<td>0.1585</td>
<td>(0.0678)</td>
<td>0.2277</td>
<td>(0.1117)</td>
</tr>
<tr>
<td>IBM</td>
<td>0.1509</td>
<td>(0.1039)</td>
<td>0.3194</td>
<td>(0.1270)</td>
<td>0.0641</td>
<td>(0.1062)</td>
<td>0.0637</td>
<td>(0.0589)</td>
<td>0.1482</td>
<td>(0.0934)</td>
</tr>
<tr>
<td>JPM</td>
<td>0.4167</td>
<td>(0.2405)</td>
<td>0.5744</td>
<td>(0.3860)</td>
<td>0.2252</td>
<td>(0.1671)</td>
<td>0.2247</td>
<td>(0.0835)</td>
<td>0.2929</td>
<td>(0.1552)</td>
</tr>
<tr>
<td>MSFT</td>
<td>0.1702</td>
<td>(0.1007)</td>
<td>0.3914</td>
<td>(0.1227)</td>
<td>0.0757</td>
<td>(0.1003)</td>
<td>0.0381</td>
<td>(0.0547)</td>
<td>0.1542</td>
<td>(0.0876)</td>
</tr>
<tr>
<td>XOM</td>
<td>0.1816</td>
<td>(0.1259)</td>
<td>0.2864</td>
<td>(0.1526)</td>
<td>0.2306</td>
<td>(0.1869)</td>
<td>0.1458</td>
<td>(0.0759)</td>
<td>0.1495</td>
<td>(0.1186)</td>
</tr>
</tbody>
</table>

The Standard errors are in parenthesis. The table reports the estimates from the model in Eq. 3.7.
Chapter 4

The Economic Value of Intraday Data in the Modeling of Conditional Covariance

4.1 Introduction

The chapter presents an economic evaluation of the addition of high-frequency information in the modeling of the conditional covariance. More precisely, the value of high-frequency data to an agent facing a portfolio decision is determined. Using the method of Fleming [2003], chapter 4 evaluates the maximum cost than an agent who invests a fixed amount each day, is willing to pay to have access to high-frequency data. This maximum cost is the fee that makes the agent indifferent to whether or not she/he has access to high-frequency information. When there is access to intraday data, the forecasts of the conditional covariance matrices necessary for the portfolio allocation, are provided by the Multivariate GARCHX (MGARCHX) model of chapter 3. When the only available information are the daily returns, the Dynamic Conditional Correlation (DCC) model of Engle [2002b] is used.

With a dataset composed of the returns of ten stocks and one risk-free asset, we find that the cost is increasing both with the target portfolio return and the relative risk aversion. We also find that it is sizable, for example it is estimated at about 174 annual basis points if the target portfolio return is set to 10% and the relative risk aversion is set to 2 in a power utility function. Furthermore, the Sharpe ratio of the strategy which incorporates the intraday data is higher than the one that does not: 0.14 for the latter and 0.17 for the former.

Section 2 presents the details of the methodology. The empirical results are presented in section 3 and the conclusion follows in section 4.
4.2 Methodology

4.2.1 The General Problem

Let us consider two investors with the same level of risk aversion who allocate a fixed amount \( W \) each day, in \( n \) risky assets and a risk-free rate asset. At the end of the day the agents consume the proceed from the investment. Only one agent has access to intraday data at an additional fixed cost \( c \). The portfolio weights are calculated by minimizing the portfolio variance subject to a target return \( \mu_p \). That is:

\[
\begin{align*}
\min_{w_t} & \quad w_t' \Sigma_{t+1|t} w_t, \\
\text{s.t.} & \quad w_t' r_{p,t+1|t} + (1 - w_t' i) r_f = \mu_p,
\end{align*}
\]

where \( w_t \) is the \( n \times 1 \) vector of weights for the risky assets and \( i \) is a \( n \times 1 \) vector of ones. \( r_f \) is the risk-free rate, \( \Sigma_{t+1|t} \) is the \( t+1 \) forecast of the conditional covariance and \( r_{p,t+1|t} \) is the \( n \times 1 \) vector of expected returns for the risky assets. The solution of the problem is:

\[
w_t = \frac{\mu_p - r_f}{(r_{p,t+1|t} - ir_f)' \Sigma_{t+1|t}' (r_{p,t+1|t} - ir_f)} \Sigma_{t+1|t}^{-1} (r_{p,t+1|t} - ir_f).
\]

Given the target return and the risk-free rate, the input in the solution 4.2.2 are the forecasts of the asset returns and the conditional covariance.

4.2.2 The Expected Return

Stein [1956] shows that the classical sample average is a poor estimator of the normal mean when the number of assets is greater than 2. To correct for this in our study, the Bayes-Stein estimator of Jorion [1986] is used as the expected return, \( r_{p,t+1|t} \). It is computed as the following:

\[
E^{BS}[r] = (1 - \omega) \bar{r} + \omega r_0, \quad \text{where:}
\]

\[
r_0 = \frac{S^{-1}\bar{r}}{iS^{-1}i},
\]

\[
\omega = \lambda/(\lambda + T) \quad \text{and}
\]

\[
\lambda = \frac{(n + 2)(T - 1)}{(\bar{r} - r_0 i)' S^{-1} (\bar{r} - r_0 i)(T - N - 2)}.
\]

S is the usual sample covariance matrix and \( T \) is the sample size. Notice that each sample average is shrunk toward a common value \( r_0 \), which happens to be the mean return of the minimum-variance portfolio. Jorion [1986] shows in a simulation analysis that the shrinkage
estimator 4.2.3 provides significant gains in portfolio selection problems.

### 4.2.3 The Conditional Covariance Matrices

The one-period ahead forecasts of the conditional covariance matrices are the results of two different models. The agent with access to intraday data uses the MGARCHX model introduced in chapter 3 while the other estimates and forecasts the conditional covariance with the DCC model of Engle [2002b] which has the form:

\[
H_t = D_t \Gamma_t D_t,
\]

\[
h_{i,t} = \omega_i + \beta_i h_{i,t-1} + \alpha_i y_{i,t-1}^2,
\]

\[
Q_t = \bar{Q}(1 - a - b) + bQ_{t-1} + a(u_{t-1}u'_{t-1}). \tag{4.2.4}
\]

The conditional covariance matrix at time \( t \) is \( H_t \) and its \( i^{th} \) diagonal element is \( h_{i,t} \). The daily return for asset \( i \) is \( y_{i,t} \). \( D_t \) and \( \Gamma_t \) are the matrices of standard deviations and correlations. Finally, \( \Gamma_t = diag(Q_t)^{-1/2} \times Q_t \times diag(Q_t)^{-1/2}, u_{t-1} = \{ y_{i,t-1}/h_{i,t-1} \}_{i=1,...,n} \) and \( \bar{Q} \) is the unconditional covariance matrix of \( u_t \). Each agent generates the forecast of the next day covariance matrix on a rolling window basis that is updated daily.

### 4.2.4 The Economic Measure

The cost \( c \) is the measure of the economic gain of using high-frequency. That is, we find the maximum cost \( c \) that the agent with access to intraday data is willing to pay.

More precisely, we assume that the portfolio returns are approximately normal and that both agents have a power utility with the coefficient of relative risk aversion (RRA). If the ranking function is the power utility and if we assume that the returns are approximately normally distributed, the portfolio expected return derived from an utility maximization problem behaves as in a mean-variance portfolio setting. One advantage of the power utility is that we can fix the coefficient of RRA for the two agents and compare their utility levels, which is problematic in the quadratic utility case.

The agents have the following value function at time \( t \) period:

\[
V_t = U(W_t) + \beta E_t[U(W_{t+1})], \quad \text{with}
\]

\[
U(W_t) = \frac{1}{1 - \gamma} (WR_{p,t})^{1-\gamma}. \tag{4.2.5}
\]

\( R_{P,t} \) is the actual gross portfolio return at time \( t \) and the parameter \( \beta \) is the discounting factor (set to 0.99 in the empirical application). Given the concavity of the power utility function, \( E_t[U(W_{t+1})] \leq U(E_t[W_{t+1}]) \) or, \( E_t[U(W_{t+1})] = U(E_t[W_{t+1}]) - \delta \), with \( \delta \geq 0 \). Therefore, we can
4.3. EMPIRICAL RESULTS

rewrite Eq. 4.2.5 as:

\[ V_t = \tilde{V}_t - \beta \delta, \quad \text{with} \]
\[ \tilde{V}_t = U(\bar{W}R_{p,t}) + \beta U(\bar{W}(1 + \mu_p)). \]  

(4.2.6)

Eq. 4.2.6 uses the fact that the expected portfolio return is the target return, \( \mu_p \), in the minimization problem 4.2.1.

Each day, the agents rebalance their portfolios and have the daily value function 4.2.6. The sum of these value functions over the entire forecasting sample \( T \) is:

\[ V = -T \beta \delta + \sum_{t=1}^{T} \tilde{V}_t \]  

(4.2.7)

To measure the economic benefit of incorporating high-frequency information in the construction of the portfolios, the maximum fixed cost \( c \) that the agent with access to high-frequency information is willing to pay is calculated. This maximum cost is the fee that equalizes the total utilities of the two agents. More precisely, the problem is:

Find \( c \) such that: \( V^{LF} = V^{HF} \Rightarrow \sum_{t=1}^{T} \tilde{V}_t^{HF} = \sum_{t=1}^{T} \tilde{V}_t^{LF} \),

with \( \tilde{V}_t^{HF} = U((\bar{W} - c)R_{p,t}^{HF}) + \beta U((\bar{W} - c)(1 + \mu_p)) \)
and \( \tilde{V}_t^{LF} = U(\bar{W}R_{p,t}^{LF}) + \beta U(\bar{W}(1 + \mu_p)) \).  

(4.2.8)

In the problem 4.2.8, the subscripts \( HF \) and \( LF \) which stand for High-Frequency and Low-Frequency help differentiate the total value of the two types of investor. Notice that the term \(-T \beta \delta\) is omitted because it is assumed that the agents have the same attitude toward risk. The problem 4.2.8 is solved numerically because it is non-linear in \( c \).

4.3 Empirical Results

The data set, used in chapter 2, consists of the daily intraday prices for ten stocks with a period running from January 3, 2006 to April 30, 2012\(^1\). The in-sample for the initial estimation runs from January 4, 2006 to April 30, 2010 and has a size of 1079. With a size of 501 (approximately two years), the out-of-sample runs from May 3, 2010 to April 30, 2012. Starting on April 30, 2010, the agents construct their portfolios only with the initial in-sample data. Subsequently, the portfolios are rebalanced based on the new information after each day by forecasting the

\(^1\)The ticker symbols of the assets are AXP, BAC, CVX, KO, DD, GE, IBM, JPM, MSFT and XOM.
4.3. EMPIRICAL RESULTS

conditional covariance matrix with a rolling window of 1,079 data points.

The risk-free rate is set to the average of the interest rate on 3-month treasury bills over the forecasting period (annualized rate of 0.0875%). The weights in the portfolio are calculated by minimizing the portfolio variance subject to a target annualized return of 10%.

The summary statistics for the two subsamples are presented in Table 4.1. The averages are close to zero except for BAC, which shows negative average in both subsamples. The evolution of the Bayes-Stein estimator of the expected returns over the forecasting period, depicted in Fig. 4.1 to 4.5, shows positive expected returns for most of the assets except for BAC.

The volatility timing strategy implemented with the MGARCHX model introduces more variation of the portfolio weights than the DCC’s strategy does (Fig. 4.6 to 4.10). The reason may be found in the fact that intraday data allows the forecasts of conditional covariances to react faster to changes in market volatility.

4.3.1 Sharpe Ratio and Portfolio Volatility

The evolution of the forecasted annualized portfolio standard deviation (in percent) are plotted in Fig. 4.11. The graph shows similar dynamics for the two strategies. However, the expected volatility of the DCC portfolio appears slightly higher than the expected volatility implied by MGARCHX model portfolio. The densities of these annualized volatilities (Fig. 4.12) show also that the portfolio weights calculated with only daily data imply higher expected variances. This is confirmed by the averages, which are 7.82% for the portfolio using intraday data and 8.49% for the other strategy. The result is that the Sharpe ratio calculated with the covariance matrices generated by the DCC model (0.14) is lower than the one implied by the MGARCHX model (0.17). Thus, the risk-adjusted excess portfolio return is improved with the addition of the intraday information. Fig. 4.17 presents the results of a simulation exercise where the target return varies between 0% and 18%. For each target return in the range, the median of the portfolio variances for the forecasting period is reported as a measure of the volatility. Thus, the graph represents the relation between expected mean and ex-ante variance of the portfolio. The portfolio which includes intraday data in its design delivers a lower expected risk for the same level of expected return and the difference increases with the target return.

To shed some light on the difference of the two strategies we can look at the contribution of each asset to the total portfolio risk. We know that the portfolio variance is 

\[ \sigma_{p,t+1|t}^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{i,t} w_{j,t} \sigma_{ij,t+1} \sigma_{ij,t+1|t}, \]  

(4.3.1)

where \( w_{i,t} \) and \( w_{j,t} \) are the weights at time \( t \) of the \( i^{th} \) and \( j^{th} \) asset. The covariance between
the two assets is $\sigma_{ij,t+1|t}$. We can take the partial derivative of 4.3.1 with respect to $w_{i,t}$ to find the marginal contribution of the $i^{th}$ asset to the portfolio volatility. Doing so we get:

$$\frac{\partial \sigma_{p,t+1|t}^2}{\partial w_{i,t}} = 2w_{i,t}\sigma_{i,t+1|t} + 2\sum_{j=1}^{n} w_{j,t}\sigma_{ij,t+1|t}. \quad (4.3.2)$$

The first term on the right hand side of 4.3.2 goes to zero as the number of assets increases, and thus can be diversified. However, the second term is the contribution of the $i^{th}$ asset to the overall portfolio variance and is a linear function of the sum of the covariances with the other assets. The median values of the second term for the second period are presented in Table 4.2. When positive, all the marginal contributions to the overall portfolio risk are smaller for the MGARCHX strategy. Furthermore for BAC, the decrease of the portfolio variance is slightly larger for the intraday data inclusion strategy as well. Overall then, the intraday information allows to reduce further the portfolio expected volatility.

Finally, we look at the behavior of the Global Minimum Variance Portfolio (GMVP) of the two strategies. The GMVP solves the following problem:

$$\min_{w_t} \quad w_t'\Sigma_{t+1|t}w_t,$$

s.t. $w_t'r_{p,t+1|t} = 1. \quad (4.3.3)$

The advantage of the GMVP is that the weights calculated depend only on the forecasted covariance matrix. With the series of GMVP and given that the true covariance matrix is unobserved, two ex-post measures of portfolio volatility are computed. The two measures use a different proxy for the true covariance matrix. In the first one, the outer product of the vector of the actual returns is used and the ex-post portfolio variance is:

$$\hat{\sigma}_{p,t+1} = w_t'r_{t+1}w_t.$$ 

The second utilizes the realized kernel at $t + 1$ as a proxy and the ex-post variance is:

$$\hat{\sigma}_{p,t+1} = w_t'X_{t+1}w_t.$$ 

The time series of the annualized standard deviations calculated from these ex-post variances are presented in Fig. 4.13 to 4.16. These figures hint at the fact that the addition of the high-frequency information yields smaller ex-post portfolio volatility. The average of the annualized standard deviation for the GMVP constructed with the MGARCHX model is 10.28% and 11.48% when the proxy is respectively the squared returns and the realized kernel. The GMVP from the DCC model however has an average of 11.54% and 11.90%. Patton and Sheppard [2009]
have shown that if the weights are constructed from the true covariance, then the variance of a portfolio computed from any other covariance matrix must be larger. Therefore, in our case the GMVP with the smaller ex-post volatility is the closest to the true GMVP. We can rank the two strategies by using the Diebold and Mariano [2002] and West [1996] test (DMW) on the difference of ex-post volatilities. The DMW yields a t-statistic of 2.68 and 4.08 when the proxy are the square returns and the realized kernel². Thus, including the intraday information allows to reduce the ex-post variance of the portfolio.

4.3.2 The Value of High-Frequency Information

For the initial analysis, the coefficient of RRA, $\gamma$ is set to 2 and we find that the maximum annual fee that the agent with access to intraday data is willing to pay is 174 annual basis points. This means that the investor would need to forgo a return on investment which is approximately equal to an annualized rate of 1.74% to achieve the same level of satisfaction as the agent without access to the high-frequency information.

To assess the robustness of the result, we repeat the exercise above for a range of target returns and coefficients of RRA. We allow the return to vary between 8 and 12% and the coefficient of RRA to vary between 1 and 5. Fig. 4.18 presents the 3-D plot of the results.

As expected the more risk averse the investor is, the higher the price she/he is willing to pay for the high-frequency information. Also, the cost is positively correlated with the target return. That is, the higher the target return, the higher the price the investor is willing to pay. The cost difference between the less risk averse agent ($\gamma = 1$) and the more risk averse one ($\gamma = 5$) is about 67 annual basis points when the target return is 8%. The difference almost goes up to 78 basis points when the target return is set to 12%.

4.4 Conclusion

This chapter has presented empirical evidence of an economic gain for investor in a situation of portfolio allocation.

In an empirical exercise, we have presented the economic value of access to intraday information to agents facing a simple asset allocation problem. The agents rebalance daily their portfolio of ten risky assets from the S&P 100 and a risk-free asset. Using the method of Fleming [2003], the maximal cost that an investor is willing to pay for high-frequency information was evaluated at 174 annual basis points when the target return is 10% and the relative risk aversion is 2 in a power utility function. It is also found that the cost is increasing in both the relative risk aversion and the target return.

²The difference is the volatility from DCC model minus the volatility for the MGARCHX model
Finally, the risk adjusted excess return is higher for the strategy that uses intraday data to forecast conditional covariance. The empirical analysis of the contribution of each asset to the overall portfolio risk shows that diversification with the use of the high-frequency information results in smaller portfolio expected variance.
### 4.5 Tables and Figures

Table 4.1: Summary Statistics for the Daily Log Returns: In-Sample and Out-of-Sample Period

<table>
<thead>
<tr>
<th>Stocks</th>
<th>In-Sample</th>
<th></th>
<th>Out-of-Sample</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>St. dev</td>
<td>max</td>
<td>min</td>
</tr>
<tr>
<td>AXP</td>
<td>-0.00</td>
<td>3.34</td>
<td>18.78</td>
<td>-19.38</td>
</tr>
<tr>
<td>BAC</td>
<td>-0.07</td>
<td>4.88</td>
<td>30.32</td>
<td>-34.08</td>
</tr>
<tr>
<td>CVX</td>
<td>0.04</td>
<td>2.14</td>
<td>18.93</td>
<td>-13.36</td>
</tr>
<tr>
<td>KO</td>
<td>0.04</td>
<td>1.40</td>
<td>12.96</td>
<td>-9.10</td>
</tr>
<tr>
<td>DD</td>
<td>0.01</td>
<td>2.18</td>
<td>10.84</td>
<td>-11.68</td>
</tr>
<tr>
<td>GE</td>
<td>-0.04</td>
<td>2.51</td>
<td>17.96</td>
<td>-13.66</td>
</tr>
<tr>
<td>IBM</td>
<td>0.05</td>
<td>1.60</td>
<td>10.90</td>
<td>-6.10</td>
</tr>
<tr>
<td>JPM</td>
<td>0.02</td>
<td>3.66</td>
<td>22.38</td>
<td>-23.22</td>
</tr>
<tr>
<td>MSFT</td>
<td>0.02</td>
<td>2.11</td>
<td>17.07</td>
<td>-12.46</td>
</tr>
<tr>
<td>XOM</td>
<td>0.02</td>
<td>2.00</td>
<td>15.87</td>
<td>-15.02</td>
</tr>
</tbody>
</table>

The log returns are in percentage point. The In-sample period is January 3, 2006 - March 31, 2010. The Out of sample period is May 3, 2010 - April 30, 2012.
Table 4.2: Median Contribution to Total Portfolio Risk of a Single Risky Asset

<table>
<thead>
<tr>
<th></th>
<th>MGARCHX Portfolio</th>
<th>DCC Portfolio</th>
</tr>
</thead>
<tbody>
<tr>
<td>AXP</td>
<td>0.0036</td>
<td>0.0042</td>
</tr>
<tr>
<td>BAC</td>
<td>-0.0012</td>
<td>-0.0011</td>
</tr>
<tr>
<td>CVX</td>
<td>0.0058</td>
<td>0.0066</td>
</tr>
<tr>
<td>KO</td>
<td>0.0057</td>
<td>0.0065</td>
</tr>
<tr>
<td>DD</td>
<td>0.0050</td>
<td>0.0058</td>
</tr>
<tr>
<td>GE</td>
<td>0.0019</td>
<td>0.0020</td>
</tr>
<tr>
<td>IBM</td>
<td>0.0065</td>
<td>0.0076</td>
</tr>
<tr>
<td>JPM</td>
<td>0.0037</td>
<td>0.0042</td>
</tr>
<tr>
<td>MSFT</td>
<td>0.0041</td>
<td>0.0047</td>
</tr>
<tr>
<td>XOM</td>
<td>0.0045</td>
<td>0.0053</td>
</tr>
</tbody>
</table>

The contribution is the median of 2 times the sum of product of the weights of asset j and the covariance with asset j. The value are multiplied by 100.

Figure 4.1: Time Evolution of the Bayes-Stein Estimator of the Expected Return in %: AXP and BAC
4.5. TABLES AND FIGURES

<table>
<thead>
<tr>
<th>Period</th>
<th>Apr10</th>
<th>Jul10</th>
<th>Oct10</th>
<th>Jan11</th>
<th>Apr11</th>
<th>Jul11</th>
<th>Oct11</th>
<th>Jan12</th>
<th>Apr12</th>
<th>Jul12</th>
</tr>
</thead>
<tbody>
<tr>
<td>%</td>
<td>0.03</td>
<td>0.035</td>
<td>0.04</td>
<td>0.045</td>
<td>0.05</td>
<td>0.055</td>
<td>0.06</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **CVX**
- **KO**

Figure 4.2: Time Evolution of the Bayes-Stein Estimator of the Expected Return in %: CVX and KO

<table>
<thead>
<tr>
<th>Period</th>
<th>Apr10</th>
<th>Jul10</th>
<th>Oct10</th>
<th>Jan11</th>
<th>Apr11</th>
<th>Jul11</th>
<th>Oct11</th>
<th>Jan12</th>
<th>Apr12</th>
<th>Jul12</th>
</tr>
</thead>
<tbody>
<tr>
<td>%</td>
<td>-0.01</td>
<td>0</td>
<td>0.01</td>
<td>0.02</td>
<td>0.03</td>
<td>0.04</td>
<td>0.05</td>
<td>0.06</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **DD**
- **GE**

Figure 4.3: Time Evolution of the Bayes-Stein Estimator of the Expected Return in %: DD and GE
4.5. TABLES AND FIGURES

Figure 4.4: Time Evolution of the Bayes-Stein Estimator of the Expected Return in %: IBM and JPM

Figure 4.5: Time Evolution of the Bayes-Stein Estimator of the Expected Return in %: MSFT and XOM
Figure 4.6: Optimal Portfolio Weights in % Constructed with the MGARCHX Model: AXP, BAC and the Risk-Free Rate Asset

Figure 4.7: Optimal Portfolio Weights in % Constructed with the MGARCHX Model: CVX, KO and the Risk-Free Rate Asset
Figure 4.8: Optimal Portfolio Weights in % Constructed with the MGARCHX Model: DD, GE and the Risk-Free Rate Asset

Figure 4.9: Optimal Portfolio Weights in % Constructed with the MGARCHX Model: IBM, JPM and the Risk-Free Rate Asset
4.5. TABLES AND FIGURES

Figure 4.10: Optimal Portfolio Weights in % Constructed with the MGARCHX Model: MSFT, XOM and the Risk-Free Rate Asset

Figure 4.11: Forecasted Portfolio Annualized Standard Deviation in %
4.5. TABLES AND FIGURES

The ex-post portfolio variance is calculated with the outer product of the vector of actual returns. That is, 
\[ \sigma_{p,t+1} = w_t'r_{t+1}r_{t+1}'w_t. \]
4.5. TABLES AND FIGURES

Figure 4.14: Density of the Ex-post Portfolio Annualized Standard Deviation Calculated with Outer Product of Vector of Actual Returns

The ex-post portfolio variance is calculated with the outer product of the vector of actual returns. That is, $\sigma_{p,t+1} = w_t' r_{t+1} r_{t+1}' w_t$.

Figure 4.15: Ex-post Portfolio Annualized Standard Deviation in % Calculated with the Realized Kernel

The ex-post portfolio variance is calculated with the realized kernel. That is, $\sigma_{p,t+1} = w_t' X_{t+1} w_t$.
4.5. TABLES AND FIGURES

Figure 4.16: Density of the Ex-post Portfolio Annualized Standard Deviation Calculated with the Realized Kernel

The ex-post portfolio variance is calculated with the realized kernel. That is, $\sigma_{p,t+1} = w_t' X_{t+1} w_t$.

Figure 4.17: Portfolio Expected Mean and Variances

Each annualized volatility is the average of the portfolio variances for 501 periods given an expected mean.
Figure 4.18 : Sensitivity of Cost to Target Return and Relative Risk Aversion
Chapter 5

Conclusion

In recent years, the ease of access to high-frequency information has opened the field of multivariate modeling of volatility to new horizons. A new strand of the literature which attempts to incorporate the new source of information to existing models of volatility has emerged. Studies such as those of Shephard and Sheppard [2010], Hansen et al. [2012], Gourieroux et al. [2009], Jin and Maheu [2012], Noureldin and Sheppard [2012] and more recently Hansen et al. [2014] are just a few examples. This dissertation contributes to that literature in two ways. The first contribution is methodological as we propose two frameworks that are relatively easy to implement. The second addition is the economic evaluation of the use of intraday data in portfolio allocation problem.

The first model proposed, called the Realized Regime Switching for Dynamic Correlations (R-RSDC) model builds on the Regime Switching for Dynamic Correlations (RSDC) of Pelletier [2006]. It is easy to implement as we circumvent the curse of dimensionality by using the Expectation-Maximization (EM) algorithm of Dempster et al. [1977]. The methodology proposed in chapter 2 effectively uses the information contained in the daily log returns and the realized covariance calculated with the multivariate kernel of Barndorff-Nielsen et al. [2011]. We describe the variance of each asset in the analysis with a modified Generalized Autoregressive Conditional Heteroskedasticity (GARCH) of Bollerslev [1986] with the modification consisting in replacing the lagged square return by the lagged value of the realized variance. Then, on a second level the parameters of the regime-switching correlation model are estimated with the EM algorithm. We have presented empirical evidence of the statistical gain in incorporating high-frequency information to forecast to conditional covariance.

The Multivariate GARCHX (MGARCHX) model is the second specification that we propose. This is a straightforward extension of the univariate modified GARCH. A simple multivariate measurement equation is also introduced to completely identify the parameters of the model and to allow multistep ahead forecast. The estimation method is inspired by the Flexible
GARCH model of Ledoit et al. [2003] where the computational burden of large-scale covariance matrices estimation is simplified by a decentralized estimation procedure. A Comparison to a recent model of joint modeling of returns and realized covariance, the scalar HEAVY model of Noureldin and Sheppard [2012], has shown that statistically the MGARCHX is superior. However, limited comparison in a Value-at-Risk context predicts similar performance.

The last contribution of this dissertation is the computation of the economic value of incorporating high-frequency data in asset allocation problem. We have not only found that the cost that an investor is willing to pay for the intraday information is sizable, but it is also increasing in both the relative risk aversion and the target return. Furthermore, chapter 4 has shown evidence that the risk adjusted excess return is higher for the strategy that uses intraday data to forecast conditional covariance. Finally, there is empirical evidence that diversification with the use of the high-frequency information results in smaller portfolio expected variance.

Numerous interesting directions for future research remain. First, the R-RSDC as presented does not account for asymmetry in the univariate specification and assumes multivariate normal distribution. For future research, the model can be extended to include asymmetry and fat tailed distribution.

Second, the assumption of Wishart distribution for the realized covariance matrix implies that the degree of freedom must be greater than the number of assets minus one. The likelihood in our first step estimation also implies that the estimation of the degree of freedom has no effect on the univariate parameters. However, the expressions for the updating equations for the EM algorithm suggest that the degree of freedom will affect considerably the contribution of the lower frequency information. We postpone for future work the analysis of the exact effect of the degree of freedom on the R-RSDC model results and estimation.

Third, the two models depend heavily on the fact that the realized covariance is positive definite and not ill-conditioned. The multivariate kernel used in both specifications breaks down when the number of assets is very large. The reason can be found in the synchronization procedure which looses too many transactions in large dimensional case. A variation of the multivariate kernel, called the Blocking and Regularisation and introduced by Hautsch et al. [2012], could be used instead for larger number of assets.
BIBLIOGRAPHY


83


F. C. Klebaner, Introduction to stochastic calculus with applications. Imperial College Press, 2005. [Online]. Available: http://books.google.com/books?hl=en&id=JYzW0uqQxB0C&oi=fnd&pg=PR5&dq=Introduction+to+Stochastic+Calculus+with+Applications\&ots=SrqK8iWVY7\&sig=TOA54GMsrprQIVuQ6kS0mJDi0Js


Appendix A

Appendix

A.1 Proof of Eq. 2.17 and 2.18

The log-likelihood functions in the conditional variance case is:

\[ L_c = \frac{1}{2} \sum_{t=1}^{T} \left\{ -n \ln(2\pi) - \ln|H_t| - Y_t' H_t^{-1} Y_t \right\} \]

While the realized covariances implies the following:

\[ L_r = \frac{1}{2} \sum_{t=1}^{T} \left\{ (\nu - n - 1) \ln|X_t| - (n\nu) \ln(2) - \nu \ln|\Sigma_t/\nu| \\
- 2 \ln \Gamma_n(\nu/2) - \text{tr}\left( X_t \Sigma_t^{-1} \right) \right\} \]

Now, replacing \( H_t \) by its decomposition in \( L_c \) we have:

\[ \frac{1}{2} \sum_{t=1}^{T} \left\{ -n \ln(2\pi) - 2 \ln|D_t| - \hat{U}_t' \hat{U}_t \right\} + \frac{1}{2} \sum_{t=1}^{T} \left\{ \hat{U}_t' \hat{U}_t - \ln|\Gamma_t| - \hat{U}_t' \Gamma_t^{-1} \hat{U}_t \right\} \]

where \( \hat{U}_t = D_t^{-1} Y_t \) is the diagonal matrix of standardized residual returns. \( \hat{U}_t \) consists of the elements of \( Y_t \) in Eq. 2.1 divided by the corresponding square root of the conditional variance.
A.1. PROOF OF EQ. 2.17 AND 2.18

Noticing that \(2 \ln |D_t| = \sum_{i=1}^{n} \ln h_{i,t}\) and \(\hat{U}_t^t \hat{U}_t = \sum_{i=1}^{n} \frac{y_{i,t}^2}{h_{i,t}}\), the final expressions is:

\[
\mathcal{L}_c = \mathcal{L}_c(\text{univariate}) + \mathcal{L}_c(\text{multivariate})
\]

\[
\mathcal{L}_c(\text{univariate}) = \frac{1}{2} \sum_{t=1}^{T} \left\{ \sum_{i=1}^{n} \left( \frac{n}{2} \ln \left( \frac{\nu}{2} \right) - \nu \ln \Gamma_t - \frac{\nu}{2} \ln |D_t| \right) \right\}
\]

\[
\mathcal{L}_c(\text{multivariate}) = \frac{1}{2} \sum_{t=1}^{T} \left\{ \hat{U}_t^t \hat{U}_t - \ln |\Gamma_t| - \hat{U}_t^t \Gamma_t^{-1} \hat{U}_t \right\}
\]

Similarly to \(H_t\), the realized covariance matrix can be decomposed into realized standard deviations and correlation matrices:

\[
X_t = V_t S_t V_t
\]  

(A.1.1)

where \(V_t = \text{diag}(\sqrt{x_{1,t}}, \sqrt{x_{2,t}}, ..., \sqrt{x_{n,t}})\) and \(S_t\) is the realized correlation matrix. Therefore, \(\{x_{i,t}\}_{t=1}^{T}\) are the daily realized variances for the \(i^{th}\) asset.

If \(X_t\) and \(\Sigma_t\) are replaced by their decompositions, \(\mathcal{L}_r\) becomes:

\[
\mathcal{L}_r = \sum_{t=1}^{T} \left\{ \frac{\nu - n - 1}{2} 2 \ln |V_t| + \frac{\nu - n - 1}{2} 2 \ln |S_t| + \frac{\nu}{2} \ln (\nu/2) - \frac{\nu}{2} 2 \ln |\tilde{D}_t| \right\}
\]

\[
- \frac{\nu}{2} \ln |\Gamma_t| - \ln \Gamma_n(\nu/2) - \frac{\nu}{2} \text{tr}(\tilde{S}_t \Gamma_t^{-1})\right\}
\]

Upon noticing also that \(2 \ln |\tilde{D}_t| = \sum_{i=1}^{n} \ln \tilde{h}_{i,t}\) and that \(\text{tr}(\nu \tilde{S}_t) = \nu \sum_{i=1}^{n} \frac{x_{i,t} x_{i,t}}{\tilde{h}_{i,t}}\), the final expression for the log-likelihood function is obtained:

\[
\mathcal{L}_r = \mathcal{L}_r(\text{univariate}) + \mathcal{L}_r(\text{multivariate})
\]

\[
\mathcal{L}_r(\text{univariate}) = \sum_{t=1}^{T} \left\{ \sum_{i=1}^{n} \frac{\nu}{2} \ln (\nu/2) - \ln \Gamma_t(\nu/2) \right\}
\]

\[
- \frac{\nu}{2} \ln (\tilde{h}_{i,t}) + \frac{\nu - 2}{2} \ln x_{i,t} - \frac{\nu}{2} \ln x_{i,t}^2 \tilde{h}_{i,t} \right\}
\]

\[
\mathcal{L}_r(\text{multivariate}) = \sum_{t=1}^{T} \left\{ \frac{1 - n}{2} 2 \ln |V_t| - \frac{\nu}{2} \ln |\Gamma_t| \right\}
\]

\[
+ n \ln \Gamma(\nu/2) - \ln \Gamma_n(\nu/2) + \frac{\nu}{2} \text{tr}(\tilde{S}_t [I_n - \Gamma_t^{-1}])\right\}
\]
Therefore combining \( L_c(\text{univariate}) \) and \( L_r(\text{univariate}) \) yields expression (2.17). And combining \( L_c(\text{multivariate}) \) and \( L_r(\text{multivariate}) \) yields expression (2.18).

### A.2 Proof of Eq. 2.19 and 2.20

#### A.2.1 Notations

Let \( X = \{X_t\}_{t=1}^T \), \( Y = \{Y_t\}_{t=1}^T \) be the sets of observed realized covariance matrices and daily log returns. Let \( \mathcal{S} = (s_T, s_{T-1}, ..., s_1) \) be the set of unobserved states. We assume that \( s_t \) can take a particular value in \( \{1, 2, ..., M\} \). Additionally, we assume that the transition between states follows a first order Markov chain such that

\[
\pi_{ij} = P(s_t = j | s_{t-1} = i).
\]

The objective is to find the values of the parameters \( \theta_2 = \{\Gamma_1, \Gamma_2, ..., \Gamma_M; \pi_{11}, \pi_{12}, ..., \pi_{MM}\} \) such that the likelihood, \( f(Y, X; \theta_2) \), is maximized. \( S \) is not observed and we use a hidden markov process framework to estimate the likelihood by the multiple integral

\[
\int_{s \in S} f(Y, X, S; \theta_2) dS.
\]

To implement the EM - algorithm, first define \( Q(\theta_2^{k+1} | \theta_2^k) \), the intermediate quantity of EM, such that:

\[
Q(\theta_2^{k+1} | \theta_2^k) = \int_{s \in S} \ln f(X, Y, \mathcal{S}; \theta_2^{k+1}) f(\mathcal{S}|X, Y; \theta_2^k) dS \tag{A.2.1}
\]

For the following derivation it is important to notice that:

\[
f(Y, X, \mathcal{S}; \theta_2) = f(Y_T, X_T|Z_t; \theta_2) p(s_T|s_{T-1}; \theta_2) f(Y_{T-1}, X_{T-1}|Z_{T-1}; \theta_2) p(s_{T-1}|s_{T-2}; \theta_2) \]

\[
... p(Y_2, X_2|Z_1; \theta_2) p(s_2|s_1; \theta_2) f(Y_1, X_1|Z_0; \theta_2) \rho_0 \tag{A.2.2}
\]

The function \( f(Y_t, X_t|Z_{t-1}; \theta_2) \) is the conditional distribution of \( X_t \) at time \( t \) given the information \( Z_{t-1} = (s_{t-1}, s_{t-2}, ..., s_1) \) and \( \rho_0 \) is the initial distribution of states.

#### A.2.2 Proof of Eq. 2.19

To implement the Maximization step of the EM algorithm we maximize \( Q(\theta_2^{k+1} | \theta_2^k) \) with respect to \( \pi_{ij} \) and subject to the condition that \( \sum_{j=1}^M \pi_{ij} = 1 \) by introducing a Lagrange multiplier \( \lambda_i \).

We have then the maximization problem and the first order conditions:
A.2. PROOF OF EQ. 2.19 AND 2.20

\[ \mathcal{L} = Q(\theta_2^{k+1} \mid \theta_2^k) - \lambda_i \left( \sum_{j=1}^{M} \pi_{ij}^{k+1} - 1 \right) \]

\[ \frac{\partial Q(\theta_2^{k+1} \mid \theta_2^k)}{\partial \pi_{ij}^{k+1}} = \lambda_i \] (A.2.3)

From the equation of the expansion of \( f(Y,X,\mathcal{S};\theta_2) \) above,

\[ \frac{\partial \ln f(Y,X,\mathcal{S};\theta_2^{k+1})}{\partial \pi_{ij}^{k+1}} = \frac{1}{\pi_{ij}^{k+1}} \sum_{t=1}^{T} \mathbf{1}_{\{s_t = j, s_{t-1} = i\}} \]

\[ \frac{\partial Q(\theta_2^{k+1} \mid \theta_2^k)}{\partial \pi_{ij}^{k+1}} = \int_{s \in \mathcal{S}} \frac{1}{\pi_{ij}^{k+1}} \sum_{t=1}^{T} \mathbf{1}_{\{s_t = j, s_{t-1} = i\}} f(\mathcal{S} \mid X,Y;\theta_2^k) dS \]

\[ \frac{\partial Q(\theta_2^{k+1} \mid \theta_2^k)}{\partial \pi_{ij}^{k+1}} = \frac{1}{\pi_{ij}^{k+1}} \sum_{t=1}^{T} f(s_t = j, s_{t-1} = i \mid X,Y;\theta_2^k) \] (A.2.4)

To obtain (A.4), we use the Fubini theorem and the fact that

\[ \int_{s \in \mathcal{S}} \mathbf{1}_{\{s_t = j, s_{t-1} = i\}} f(\mathcal{S} \mid X,Y;\theta_2^k) dS = f(s_t = j, s_{t-1} = i \mid X,Y;\theta_2^k). \]

Using (A.4) in (A.3) yields:

\[ \sum_{t=1}^{T} f(s_t = j, s_{t-1} = i \mid X,Y;\theta_2^k) = \pi_{ij}^{k+1} \lambda_i \] (A.2.5)

Summing over all \( j \)'s:

\[ \sum_{j=1}^{M} \sum_{t=1}^{T} f(s_t = j, s_{t-1} = i \mid Y,X;\theta_2^k) = \sum_{j=1}^{M} \pi_{ij}^{k+1} \lambda_i \]

\[ \Rightarrow \sum_{t=1}^{T} f(s_{t-1} = i \mid X,Y;\theta_2^k) = \lambda_i \] (A.2.6)

Using (A.6) in (A.5) to substitute for \( \lambda_i \) and solving for \( \pi_{ij}^{k+1} \) gives Eq. 2.19.
A.2.3 Proof of Eq. 2.20

We maximize (A.1) with respect to the parameter $\theta \in \theta_2$. We then have the first order condition for the $i^{th}$ state:

$$\frac{\partial Q(\theta_{2}^{k+1}|\theta_{2}^{i})}{\partial \Gamma_{s_{t}=i}^{k+1}} = 0$$  \hspace{1cm} (A.2.7)

From (A.2) we have:

$$\frac{\partial \ln f(Y, X, S; \theta_{2}^{k+1})}{\partial \Gamma_{s_{t}=i}^{k+1}} = \sum_{t=2}^{T} \frac{\partial \ln f(Y_{t}, X_{t}|Z; \theta_{2}^{k+1})}{\partial \Gamma_{s_{t}=i}^{k+1}}$$  \hspace{1cm} (A.2.8)

Using (A.1),(A.7) and (A.8) and the Fubini theorem one more time we get:

$$\sum_{t=2}^{T} \sum_{i=1}^{M} \frac{\partial \ln f(Y_{t}, X_{t}|Z; \theta_{2}^{k+1})}{\partial \Gamma_{s_{t}=i}^{k+1}} f(s_{t} = i|X, Y; \theta_{2}^{k}) = 0$$  \hspace{1cm} (A.2.9)

$f(X, Y, S; \theta_{2}^{k+1})$ here is the joint probability distribution of the returns and the realized covariance defined in expression (2.14). We can express the log-likelihood to maximize (omitting terms which are fixed in the maximization problem) as $\mathcal{L}_{c}(\text{multivariate}) + \mathcal{L}_{r}(\text{multivariate})$ because $\Gamma_t$ does not appear in the univariate part. Differentiating w.r.t. $\Gamma_t$:

$$\frac{\partial \mathcal{L}_{c}(\text{multivariate})}{\partial \Gamma_t} = -\frac{1}{2} \Gamma_t^{-1} + \frac{1}{2} \Gamma_t^{-1} \hat{U}_t \hat{U}_t' \Gamma_t^{-1}$$  \hspace{1cm} (A.2.10)

$$\frac{\partial \mathcal{L}_{r}(\text{multivariate})}{\partial \Gamma_t} = -\frac{\nu}{2} \Gamma_t^{-1} + \nu \Gamma_t^{-1} \tilde{S}_t \Gamma_t^{-1}$$  \hspace{1cm} (A.2.11)

Omitting the superscript $k + 1$ for notation clarity, and combining (A.10) and (A.11) we get:

$$\frac{\partial \ln f(Y_{t}, X_{t}|Z; \theta_{2}^{k+1})}{\partial \Gamma_{s_{t}=i}^{k+1}} = -\frac{1}{2} \Gamma_{s_{t}=i}^{-1} + \frac{1}{2} \Gamma_{s_{t}=i}^{-1} \hat{U}_t \hat{U}_t' \Gamma_{s_{t}=i}^{-1} - \frac{\nu}{2} \Gamma_{s_{t}=i}^{-1} \hat{S}_t \Gamma_{s_{t}=i}^{-1}$$

$$= -\frac{\nu + 1}{2} \Gamma_{s_{t}=i}^{-1} + \frac{1}{2} \Gamma_{s_{t}=i}^{-1} \hat{U}_t \hat{U}_t' + \nu \tilde{S}_t \Gamma_{s_{t}=i}^{-1}$$  \hspace{1cm} (A.2.12)
A.3. DETAILS OF THE PROBABILITIES COMPUTATION

Applying (A.12) to (A.9) we have:

$$
\sum_{t=2}^{T} \left\{ -\frac{\nu + 1}{2} \Gamma_{s_{t-1}}^{-1} + \frac{1}{2} \Gamma_{s_{t-1}}^{-1} [\hat{U}_{t} \hat{U}_{t}'+ \nu \hat{S}_{t}] \Gamma_{s_{t-1}}^{-1} \right\} f(s_{t} = i|X,Y; \theta^k) = 0 \quad (A.2.13)
$$

Solving for $\Gamma_{s_{t-1}}$ gives Eq. 2.20 with observed $C_{t} = \frac{1}{\nu + 1} \{\hat{U}_{t} \hat{U}_{t}'+ \nu \hat{S}_{t}\}$.

A.3 Details of the Probabilities Computation

The Probabilities are computed by the algorithm of Hamilton [1989] and Kim [1994]. Inference on the unobserved state of the Markov Chain is given by the following equations:

$$
\hat{\xi}_{t|t} = \frac{\hat{\xi}_{t|t-1} \odot \eta_{t}}{1'(\hat{\xi}_{t|t-1} \odot \eta_{t})}, \quad (A.3.1)
$$

$$
\hat{\xi}_{t+1|t} = \Pi^{'} \hat{\xi}_{t|t}, \quad (A.3.2)
$$

$$
\eta_{t} = \begin{bmatrix}
        f(Y_{t}, X_{t}|\mathcal{F}_{t-1}, s_{t} = 1; \theta) \\
        \vdots \\
        f(Y_{t}, X_{t}|\mathcal{F}_{t-1}, s_{t} = M; \theta)
      \end{bmatrix}, \quad (A.3.3)
$$

where $\hat{\xi}_{t|t}$ is an $(M \times 1)$ vector which contains the probability of occurrence of each regime at time $t$ conditional on the observations up to time $t$. The $(M \times 1)$ vector $\hat{\xi}_{t+1|t}$ gives the forecasted probabilities of time $t + 1$ conditional on observations up to time $t$. The $m$-th element of the $(M \times 1)$ vector $\eta_{t}$ is the density of $(Y_{t}, X_{t})$ conditional on past observations and the regime $m$ at time $t$. The $(M \times 1)$ vector of 1s, and $\odot$ denotes elements-by-elements multiplication.

Given a starting value $\hat{\xi}_{1|0}$ and parameter values $\theta$, one can iterate over (A.3.1) and (A.3.2) for $t = 1, \ldots, T$. The log-likelihood is obtained as a by-product of this algorithm:

$$
\mathcal{L}(\theta) = \sum_{t=1}^{T} \log \left( 1'(\hat{\xi}_{t|t-1} \odot \eta_{t}) \right). \quad (A.3.4)
$$

Smooth inference on the state of the Markov chain can also be computed by the backward algorithm developed by Kim [1994]. The probability of being in each regime at time $t$ conditional on observations up to time $T > t$ is given by the following equation:

$$
\hat{\xi}_{t|T} = \hat{\xi}_{t|t} \odot \left\{ \Pi^{'} \left[ \hat{\xi}_{t+1|T} (\odot) \hat{\xi}_{t+1|t} \right] \right\} \quad (A.3.5)
$$
where \((\div)\) denotes element-by-element division. One would start iterating over (A.3.5) with \(t = T\), where \(\hat{\xi}_{T|T}\) is given by (A.3.1) until \(t = 1\) is reached. \(\hat{\xi}_{t|T}\) is therefore the vector of smoothed probability, that is the probability of a particular state at time \(t\) given all the available information, \(\mathcal{F}_T\).

### A.4 Proof of Eq. 2.28

Solving forward Eq. 2.4, we get for a \(L\)-ahead period:

\[
 h_{i,T+L} = \sum_{i=0}^{L-1} \omega_i \beta_i^j + \beta_i^L h_T + \sum_{k=0}^{L} \alpha_i \beta_i^{L-1} x_{T+L-1-k} \quad (A.4.1)
\]

Taking the expected value at time \(T\) of (A.4.1) we get

\[
 E_T h_{i,T+L|T} = h_{i,T+L|T} = \sum_{j=0}^{L-1} \omega_i \beta_i^j + \beta_i^L h_T + \sum_{k=0}^{L-1} \alpha_i \beta_i^k x_{T+L-1-k|T} \quad (A.4.2)
\]

The \(T+L-1-k\) step ahead expected value of the realized variance given the time \(T\) information is simply:

\[
 x_{T+L-1-k|T} = E_T x_{T+L-1-k} = \gamma_0 + \gamma_1 E_T h_{T+L-1-k} \quad (A.4.3)
\]