



## City Research Online

### City, University of London Institutional Repository

---

**Citation:** Liu, W. (2012). Essays on economic value of intraday covariation estimators for risk prediction. (Unpublished Doctoral thesis, City University London)

This is the accepted version of the paper.

This version of the publication may differ from the final published version.

---

**Permanent repository link:** <https://openaccess.city.ac.uk/id/eprint/11792/>

**Link to published version:**

**Copyright:** City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

**Reuse:** Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Liu, Wei (2012). Essays on economic value of intraday covariation estimators for risk prediction.  
(Unpublished Doctoral thesis, City University London)



**CITY UNIVERSITY  
LONDON**

[City Research Online](#)

**Original citation:** Liu, Wei (2012). Essays on economic value of intraday covariation estimators for risk prediction. (Unpublished Doctoral thesis, City University London)

**Permanent City Research Online URL:** <http://openaccess.city.ac.uk/2736/>

#### **Copyright & reuse**

City University London has developed City Research Online so that its users may access the research outputs of City University London's staff. Copyright © and Moral Rights for this paper are retained by the individual author(s) and/ or other copyright holders. All material in City Research Online is checked for eligibility for copyright before being made available in the live archive. URLs from City Research Online may be freely distributed and linked to from other web pages.

#### **Versions of research**

The version in City Research Online may differ from the final published version. Users are advised to check the Permanent City Research Online URL above for the status of the paper.

#### **Enquiries**

If you have any enquiries about any aspect of City Research Online, or if you wish to make contact with the author(s) of this paper, please email the team at [publications@city.ac.uk](mailto:publications@city.ac.uk).

# Essays on the Economic Value of Intraday Covariation Estimators for Risk Prediction

Wei Liu

A thesis submitted for the degree of Doctor  
of Philosophy (Ph.D.) in Finance

CITY UNIVERSITY  
CASS BUSINESS SCHOOL  
FACULTY OF FINANCE

LONDON 2012

# CONTENTS

	<b>List of Tables</b>	iv
	<b>List of Figures</b>	v
	<b>Abbreviations</b>	vi
	<b>Acknowledgements</b>	vii
	<b>Declaration</b>	viii
	<b>Abstract</b>	ix
<b>Chapter 1</b>	<b>Introduction</b>	
	1.1 Overview of the Research	1
	1.2 Literature on Volatility Forecasting	2
	1.2.1 The Concepts of Volatility Forecasting	2
	1.2.2 Intraday Volatility Measures	3
	1.2.3 Economic Value of Volatility Forecasting	4
	1.3 Motivation and Objective of the Thesis	5
	1.4 Layout of the Thesis	8
<b>Chapter 2</b>	<b>Common Methodology</b>	
	2.1 Realized Volatility Estimators	11
	2.2 Realized Covariance Estimators	16
	2.2.1 Realized Covariance (RC)	16
	2.2.2 Realized Kernel Covariance (RKC)	16
	2.3 Univariate Volatility Models	18
	2.4 Multivariate Volatility Models	21
	2.4.1 Nonparametric EWMA Forecasts	21
	2.4.2 Parametric MGARCH Forecasts	22
<b>Chapter 3</b>	<b>Decision-Based Evaluation of Realized Covariance Forecasts</b>	
	3.1 Introduction	27
	3.2 Literature Review	30
	3.2.1 Market Timing Strategy	30
	3.2.2 Economic Value of Volatility Timing	31

3.3	Methodology	33
3.3.1	Volatility Timing Strategies	34
3.3.2	Economic Evaluation Framework	35
3.3.3	Transaction Costs and Rebalancing Frequency	36
3.3.4	Covariance Forecasting	37
3.3.5	Statistical Evaluation Methods	38
3.4	The Dataset	39
3.5	Empirical Results	42
3.5.1	Statistical Evaluation of Covariance Forecasts	42
3.5.2	Estimation Results of Covariance Forecasts	43
3.5.3	Economic Evaluation of Covariance Forecasts	43
3.5.4	The Impact of Rebalancing Frequency	46
3.5.5	Robustness Checks	48
3.6	Conclusion	49
<b>Chapter 4</b>	<b>Single Index and Portfolio Models for High-Frequency VaR Prediction</b>	
4.1	Introduction	78
4.2	Literature Review	81
4.2.1	Single Index versus Portfolio Models	81
4.2.2	VaR Backtesting Approaches	83
4.3	Methodology	84
4.3.1	VaR Modelling Approach	85
4.3.2	Volatility Forecasting	86
4.3.3	VaR Backtesting Framework	87
4.4	The Dataset	92
4.5	Empirical Results	94
4.5.1	Estimation Results of the Models	95
4.5.2	Out-of-Sample VaR Backtesting	96
4.5.3	Robustness Checks	100
4.6	Conclusion	102
<b>Chapter 5</b>	<b>The Economic Value of Combining Intraday Volatility Estimators and Implied Volatility</b>	
5.1	Introduction	113
5.2	Literature Review	116
5.2.1	Model Forecasted versus Implied Volatility	116
5.2.2	Forecast Combination	117
5.2.3	Economic Value of Forecast Combination	119
5.3	Methodology	120
5.3.1	Forecast Combination Techniques	120

5.3.2	Volatility Forecasting Framework	124
5.3.3	Forecast Evaluation Framework	126
5.4	The Dataset	130
5.5	Empirical Results	132
5.5.1	Statistical Evaluation of Combination Volatility Estimators	133
5.5.2	Economic Evaluation of Combination Volatility Estimators	134
5.5.3	Estimation Results of the Models	138
5.5.4	Robustness Tests	138
5.6	Conclusion	141
<b>Chapter 6</b>	<b>Conclusion</b>	
6.1	Concluding Remarks	150
6.2	Further Research	153
	<b>References</b>	156
	<b>Appendices</b>	
	Appendix 3.1: Index Explanation	171
	Appendix 3.2: Additional Covariance Ranking Figures	172

## List of Tables

Table 3.1:	Summary Statistics of Daily Index Returns	51
Table 3.2:	Statistical Evaluation of Covariance Estimators	52
Table 3.3:	Estimation Results of Daily Based Covariance Estimators	54
Table 3.4:	Estimation Results of Intraday Based Covariance Estimators	55
Table 3.5:	Summary Statistics of Portfolio Weights	56
Table 3.6:	Covariance Ranking (12% Target Return, Daily Rebalancing)	57
Table 3.7:	Covariance Ranking (12% Target Return, Weekly Rebalancing)	58
Table 3.8:	Covariance Ranking (12% Target Return, Monthly Rebalancing)	59
Table 3.9:	Covariance Ranking (12% Target Return, RD-5%)	60
Table 3.10:	Covariance Ranking (12% Target Return, RD-8%)	61
Table 3.11:	Covariance Ranking (12% Target Return, RD-10%)	62
Table 3.12:	Covariance Ranking (12% Target Return, RD-20%)	63
Table 3.13:	Covariance Ranking (12% Target Return, WD-25th)	64
Table 3.14:	Covariance Ranking (12% Target Return, WD-50th)	65
Table 3.15:	Covariance Ranking (16% Target Return, Daily Rebalancing)	66
Table 3.16:	Covariance Ranking (8% Target Return, Daily Rebalancing)	67
Table 3.17:	Covariance Ranking (12% Target Return, No Short-Selling)	68
Table 3.18:	Evaluation of Alternative Rebalancing Strategies	69
Table 4.1:	Descriptive Statistics for Returns and Volatility	105
Table 4.2:	Estimation Results of Single Index Models	105
Table 4.3:	Estimation Results of MGARCH Models	106
Table 4.4:	Estimation Results of Intraday Based MGARCH Models	107
Table 4.5:	$LR_{CC}$ Test for (R)GARCH, (R)DCC (Out-of-Sample)	108
Table 4.6:	DQ&DB Test for (R)GARCH, (R)DCC (Out-of-Sample)	108
Table 4.7:	$LR_{CC}$ , DQ and DB Test for Alternative Realized Models	108
Table 4.8:	$LR_{CC}$ , DQ&DB Test for (R)DCC, (R)GARCH, DCC	109
Table 4.9:	Linear Regression Test Results	109
Table 5.1:	Summary Statistics of 41 Realised Measures	143
Table 5.2:	Summary Statistics for the Combinations of Realized Measures	144
Table 5.3:	Performance of the Regression-Based Combination	144
Table 5.4:	Estimation Result and Robustness of Selected Models	144

## List of Figures

Figure 3.1:	Time Series Properties of Stock Indices	54
Figure 3.2.1:	Daily models' Optimal Portfolio Weights	55
Figure 3.2.2:	Realized models' Optimal Portfolio Weights	56
Figure 3.3:	Covariance Ranking (Daily Rebalancing)	57
Figure 3.4:	Covariance Ranking (Weekly Rebalancing)	58
Figure 3.5:	Covariance Ranking (Monthly Rebalancing)	59
Figure 3.6:	Covariance Ranking, 12% Target Return Return-Driven Rebalancing with a 10% Threshold	60
Figure 3.7:	Covariance Ranking, 12% Target Return Volume-Driven Rebalancing with 25th Percentile	61
Figure 4.1:	Rebased Index Prices	89
Figure 4.2:	Return of NASDAQ/Russell/S&P 500 Indices	90
Figure 4.3:	Volatility of NASDAQ/Russell/S&P 500 Indices	91
Figure 5.1:	Price and Volatility Patterns of the S&P 100 Index	124
Figure 5.2:	Statistical Evaluation of the Individual Measures and Their Combinations	125
Figure 5.3:	Empirical Performance of Option Strategies (0.25%)	126
Figure 5.4:	Empirical Performance of Option Strategies (0.20%)	127
Figure 5.5:	Empirical Performance of Option Strategies (0.30%)	128

## Abbreviations

- ADIs – Authorized Deposit-Taking Institutions
- AIC – Akaike Information Criterion
- AMAPE – Adjusted Mean Absolute Percentage Error
- APARCH – Asymmetric Power ARCH
- ARCH – Autoregressive Conditional Heteroskedasticity
- ARFIMA – Autoregressive Fractionally Integrated Moving Average
- ARIMA – Autoregressive Integrated Moving Average
- (R)BEKK – (Realized) Baba, Engle, Kraft and Kroner Model
- BQ – Binary Quantile
- BWSJ – Brownian Semimartingale Plus Jump Process
- CC – Conditional Coverage
- (R)CCC – (Realized) Constant Conditional Correlation
- DB – Dynamic Binary
- (R)DCC – (Realized) Dynamic Conditional Correlation
- DQ – Dynamic Quantile
- EGARCH – Exponential GARCH
- (R)GARCH – (Realized) Generalized Autoregressive Conditional Heteroskedasticity
- (R)GJRGARCH – (Realized) Glosten Jagannathan Runkle GARCH
- IV – Integrated Variance
- (R)MGARCH – (Realized) Multivariate GARCH
- MAE – Mean Absolute Error
- MIM – Multiple Indicators Model
- MSE – Mean Squared Error
- LB – Ljung-Box Test
- LR – Likelihood Ratio
- OLS – Ordinary Least Squares
- QV – Quadratic Variation
- RBP – Realized Bi-Power Variation
- RC – Realized Covariance
- RK – Realized Kernel
- RKC – Realized Kernel Covariance
- RK\_2ND – 2nd Order Realized Kernel
- RK\_EPA – Epanechnikov Realized Kernel
- RK\_TKH – Tukey-Hanning Realized Kernel
- (R)ROLL – (Realized) Rolling Window Conditional Covariance Estimator
- (R)PGARCH – (Realized) Power GARCH
- RPV – Realized Power Variation
- RR – Realized Range
- RV – Realized Variance
- SPDR – Standard and Poor’s Depository Receipts
- UC – Unconditional Coverage
- VaR – Value at Risk

# Acknowledgements

First and foremost I offer my sincerest gratitude to my supervisors, Prof. Ana-Maria Fuertes and Dr. Elena Kalotychou, who have supported me throughout my PhD with their patience and knowledge. Their kind help and warm-hearted encouragement were well beyond the call of duty.

I would like to express my very great appreciation to my viva examiners Prof. Stuart Hyde and Dr. Amir Alizadeh for their valuable and constructive comments.

I am grateful to Cass Business School for providing me a great research environment and teaching opportunities. I would like to thank participants at the 2009 Conference on Computational and Financial Econometrics (CFE) in Cyprus and the 2010 CFE in London, as well as seminar participants at Cass Business School for useful feedback and comments. I would also like to thank PhD officers Abdul Momin and Malla Pratt for their patience and help.

I am also grateful for all the help, time and support my parents have given to me in writing this thesis. Without their financial and spiritual support this thesis would not have come into existence.

Wei Liu  
London 2012

# Declaration

The author of this thesis grants power of discretion to the University Library to allow this thesis to be copied in whole or in part without further reference to him. This permission covers only single copies made for study or research purposes, subject to normal conditions of acknowledgement.

# Abstract

This thesis investigates the economic value of incorporating intraday volatility estimators into the volatility forecasting process. The increased reliance on volatility forecasting in the financial industry has intensified the need for more rigorous analysis from an economic perspective instead of merely statistical point of view. A better understanding of the available methods has implications for portfolio optimization, volatility trading and risk management. More recently, volatility of asset returns was once again under spotlight during the 2008-2009 financial crisis. One of the most visible indicators of the crisis that captured the attention of the financial industry was the extremely high level of asset return volatility. This uncertainty prompted much interest for a more accurate, yet practically applicable approach for volatility forecasting.

The study contributes to the extant volatility forecasting literature in three areas. First, it addresses the question of how to practically and effectively exploit intraday price information for variance and covariance modelling and forecasting. Second, it addresses the development of an ‘optimal’ intraday volatility model that accommodates market practitioners’ preferences. Third, it evaluates the economic value of combining realized (intraday) volatility estimators for utilizing unique information embedded in each estimator. The thesis is organised as follows.

Chapter 2 introduces the various realized volatility estimators, volatility forecasting procedures and their corresponding realized extensions used in our subsequent empirical investigations.

Chapter 3 evaluates the economic value of various intraday covariance estimation approaches for mean-variance portfolio optimization. Economic loss functions overwhelmingly favour intraday covariance matrix models instead of their daily counterparts. The constant conditional correlation (CCC) augmented with realized volatility produces the highest economic value when applied with a time-varying volatility timing strategy.

Chapter 4 compares the practical value of intraday based single index (univariate) and portfolio (multivariate) models through the lens of Value-at-Risk (VaR) forecasting. VaR predictions are generated from standard daily univariate or multivariate GARCH models, as well as GARCH models extended with ARFIMA forecasted realized measures. Conditional coverage test results indicate that intraday models, both univariate and multivariate ones, outperform their daily counterparts by providing more accurate VaR forecasts.

Chapter 5 investigates the economic value of combining intraday volatility estimators for volatility trading. The simulated option trading results indicate that a naive combination of an intraday estimator and implied volatility cannot be outperformed by the best individual estimator. In addition, trading performance can be further boosted by applying more complex combination models such as a regression based combination of 42 single volatility estimators.

# CHAPTER 1: INTRODUCTION

## 1.1 Overview of the Research

Volatility and correlation are key inputs to asset pricing, portfolio selection and risk management. Asset pricing models such as the Black-Scholes option pricing formula and its various extensions use volatility estimate as an important factor. Portfolio managers seek to determine the weights of assets that provide the optimal return/risk trade-off given a particular utility function. For hedging against risk and managing investment risk, accurate volatility forecasts are also crucial. Furthermore, the establishment of Basel II Accord in 2004, effectively, rendered volatility and correlation forecasting a mandatory risk-management exercise for many financial institutions.

More recently, volatility of asset returns was once again under spotlight during the 2008-2009 financial crisis, a major disruption to the financial sector and global economy. One of the most visible signs of the crisis that captured the attention of both the academia and financial industry was the extremely volatile stock markets. This uncertainty prompted much interest for more accurate, yet practically applicable frameworks for volatility modelling and forecasting.

Motivated by this interest, the thesis investigates how intraday information can be exploited to improve volatility forecasting models from an economic viewpoint. A number of features make the current study distinct. First, a systematic analysis of the economic value of realized (intraday) covariance matrix in portfolio optimization is carried out. Second, the practical value of intraday volatility measures is rigorously evaluated by Value-at-Risk (VaR) forecasting performance. Third, the economic value of realized forecast combination is analysed in the related context of volatility trading. The dataset for the empirical studies include real market tick-by-tick data for several important stock and commodity indices in the United States spanning the period from 06 January 1997 to 30 September 2011. The findings from this study will be useful for practitioners in financial

institutions, hedge funds and proprietary trading houses interested in the development of intraday volatility models for economic applications.

The remainder of this chapter is structured as follows. Section 1.2 discusses the concepts of volatility forecasting and its application. Section 1.3 describes the motivation and objective of this thesis. Finally, section 1.4 outlines the layout of the study.

## 1.2 Literature on Volatility Forecasting

### 1.2.1 The Concepts of Volatility Forecasting

Volatility is a crucial concept in finance but a challenging one because, unlike returns, volatility is latent or unobserved. Therefore, it needs to be proxied. An asset with relatively high volatility, hence greater uncertainty, is often regarded as a riskier investment, since larger price variation brings greater uncertainty to the investment returns.

Volatility of financial asset return series is central to financial economics. Indeed, as noted by Campbell, Lo and MacKinlay (1997):

*“... what distinguishes financial economics is the central role that uncertainty plays in both financial theory and its empirical implementation ... Indeed in the absence of uncertainty, the problems of financial economics reduce to exercises in basic microeconomics”.*

Given volatility’s central role in financial economics and practical applications, a vast literature on modelling the conditional variance of asset returns has emerged. Among the most popular are the *autoregressive conditional heteroskedasticity* (ARCH) models, *stochastic volatility* (SV) models, and *regime-switching* models; see comprehensive surveys by Bollerslev et al. (1994) for GARCH-type models, Ghysels et al. (1996) for stochastic volatility models, and Franses and van Dijk (2000) for regime-switching models. Due to the space limitation, this thesis focuses on GARCH models and their realized extensions.

The popularity of GARCH models, introduced by Engle (1982), is attributed to both their intuition and their ability to capture several stylized facts about asset return volatil-

ity, such as volatility clustering<sup>1</sup> and mean reversion. Furthermore, GARCH models can be flexibly augmented to measure additional stylized facts such as volatility asymmetry<sup>2</sup> and jumps, or to incorporate additional information such as trading volume and intraday volatility estimators.

GARCH models are not confined to a univariate context. They can be naturally extended to a multivariate setup which enables modelling and forecasting the conditional covariance matrix for a portfolio of assets. It is now widely accepted that financial volatilities move together over time across assets and markets. Recognizing this characteristic through a multivariate GARCH framework leads to more relevant empirical models than working with separate univariate GARCH models.

### 1.2.2 Intraday Volatility Measures

One limitation of standard univariate and multivariate GARCH models is that they employ only daily or lower frequency data of asset returns, which may limit the accuracy of volatility forecasting. First, although the squared daily return generates an unbiased estimate for the realized integrated volatility, it is a very noisy estimator, and foreseeable variation in the real latent volatility process is often dwarfed by measurement error (see, e.g., ABDE, 2001). Second, daily volatility models neglect possibly dramatic intraday price movements, such as those shown during the peak of the 2008-2009 financial crisis, omitting valuable information about the market sentiment, risk signal and the magnitude of intraday return variation.

The search for a better framework for the estimation and prediction of the conditional variance of asset returns has therefore led to the study of intraday volatility measures.

---

<sup>1</sup>Large innovations tend to be followed by large innovations in either direction, and small innovations tend to follow small innovations. It is a case of heteroskedasticity, where variances of the error terms are not equal. This phenomenon is widely found in volatility of financial asset returns.

<sup>2</sup>Negative shocks have typically a stronger effect on future volatility than positive shocks. This asymmetry is sometimes ascribed to a leverage effect. The theory explains that when price of a stock falls, its debt-to-equity ratio increases, adding the risk of returns to equity holders.

The most commonly used intraday estimator, realized volatility (RV), is the sum of finely-sampled squared return realizations over a fixed time interval. Many other nonparametric realized (intraday) volatility estimators have recently been proposed (see Andersen and Bollerslev, 1998; Andersen et al., 2001; Barndoff-Nielsen and Shephard, 2004).

One direction to improve the accuracy of daily GARCH models is to augment them with intraday volatility estimators, which is the main theme of this thesis. The effectiveness of this augmentation can be judged by two distinct types of evaluation methods. The first one consists of statistical loss functions, which is extensively applied in the literature. The second method uses economic criteria, which is relatively new and has not been explored in depth. This study focuses on the second framework, which is briefly introduced in the next subsection.

### **1.2.3 Economic Value of Volatility Forecasting**

The economic importance of volatility modelling and forecasting in many market applications implies that the success or failure of a volatility model will depend on its ability to generate economic value when applied to real market decision making. Therefore, it is arguably more relevant to evaluate volatility forecasts directly in an economic framework using real market data.

The economic value generated by a volatility model can be defined as the economic benefit of switching to it from another volatility specification, a base model. This economic benefit can be relevant to many important areas of the financial industry, such as portfolio optimization, proprietary trading and risk management. For example, the economic value can be quantified as the increase in investor's utility in mean-variance asset allocation, the additional clarity obtained about tomorrow's Value-at-Risk (VaR) of a portfolio, or the Sharpe ratio increase in volatility trading strategies.

## 1.3 Motivation and Objective of the Thesis

This thesis empirically investigates the forecasting performance of univariate and multivariate intraday volatility models and provides insights on their relative ability to generate economic value when applied to portfolio optimization, risk management and volatility trading. For this purpose, it utilizes a variety of intraday volatility measures, forecasting techniques as well as combination methods. The importance of volatility forecasting for portfolio managers and market practitioners, together with the need to fill existing gaps in the empirical high-frequency literature, constitute the primary motivation of the thesis.

One shortcoming of conventional GARCH models is that they use the squared daily return, an extremely noisy estimator of ex post volatility, as proxy for the day's variance in forecast evaluation. Andersen and Bollerslev (1998) show that when GARCH forecasts are compared with the sum of intraday squared returns, as the true volatility proxy, they appear far more accurate than when compared with the squared daily close-to-close returns.

Empirically, several studies have explored the issue of utilizing intraday information in forecasting daily volatility by incorporating intraday volatility estimator as an additional regressor of the daily GARCH models. For example, Fuertes, Izzeldin and Kalotychou (2009) incorporated realized volatility (RV), realised range (RR), realized power variation (RPV) and realized bipower variation (RBP) into the daily GARCH model. The authors find that intraday information increases statistical accuracy of GARCH forecasts and, in particular, RPV enhances the GARCH forecasting ability the most.

However, the extant literature on the merit of incorporating intraday data in volatility forecasting mostly rests on a univariate framework. There is very limited empirical research on using intraday information for modelling and forecasting the entire conditional covariance matrix. Covariance between asset returns is a crucial input in mean-variance portfolio selection. In this setting univariate intraday volatility models are simply not capable of achieving the goal.

In addition, although in some cases portfolio management decisions can be made based on univariate intraday models, it is worth to question whether the investment performance would further improve by utilizing multivariate intraday models. For example, some financial institutions are required to make Value-at-Risk (VaR) predictions for their investment portfolio according to the Basel II Accord. This can be done either using univariate models which regard the whole portfolio as a single asset, or by multivariate analysis which forms volatility forecasts using information of each portfolio constituent. So far no consensus has been reached about the ‘best’ approach using daily data, univariate or multivariate ones for VaR prediction (Berkowitz and O’Brien, 2002; McAleer and da Veiga, 2008; Dumitrescu, 2012). Furthermore, in an intraday context, the univariate and multivariate debate is yet to be informed empirically.

A number of other important issues in the empirical literature on intraday volatility modelling warrant further analysis. Above all, intraday volatility specifications have not been fully evaluated from an economic perspective, while most of the empirical studies that use intraday information to forecast volatility are evaluated statistically. Statistical evaluation of intraday models show overwhelming improvement in forecasting performance compared with daily specifications (Andersen et al., 2001; Barndoff-Nielsen and Shephard, 2004; Fuertes, Izzeldin and Kalotychou, 2009). But whether the intraday models would benefit market practitioners is still open to debate.

A priori, intraday models exploit more information and therefore are expected to outperform their daily counterparts. But in practice intraday data are harder to collect and proceed and intraday models are more demanding computationally. Given the complexity of intraday specifications, it is arguable that they are useful for fund managers and market practitioners only if they can generate significant economic value. Against this backdrop, it is important to empirically analyse the performance of intraday models for different financial applications such as portfolio optimization, risk management and volatility trading.

Another issue worth considering is forecast combination. The empirical literature on forecast combination mostly rests on daily or lower frequency data. Research shows that combination forecasts often outperform the forecasts generated by the best individual model (see Clemen, 1989; Makridakis and Hibon, 2000; Stock and Watson, 2004; and Becker and Clements, 2008). The benefit of forecast combination is three-fold. First, it integrates different models that may differ in the information they exploit and/or how they exploit it. Second, it mitigates the effect of structural breaks. Third, it reduces the impact of parameter estimation uncertainty and model uncertainty.

Given that forecast combination has been shown to be worthwhile in a variety of applications with low frequency data, it is natural to further the research on this topic in an intraday context. Over the last decade a variety of intraday volatility estimators has been proposed. Each of these estimators is formulated in a different way and can be sampled at different intraday frequencies. Combining the intraday estimators could incorporate unique pieces of information as well as reduce the effect of structural breaks suffered by individual forecasting model, therefore producing better forecasting performance. However, the literature of forecast combination has seldom reached intraday level, let alone the economic value of combining intraday information.

The objective of the thesis is to bridge the gaps that exist in the literature by addressing the aforementioned issues using three distinct empirical studies. First, in a portfolio optimization context, a variety of multivariate intraday volatility models are evaluated by the maximum return that an investor with quadratic utility function would be willing to sacrifice annually in order to capture the performance gains associated with the intraday covariance estimators. Second, the performances of univariate and multivariate intraday-based VaR models are compared. Third, in an options trading framework, the economic value of combining a battery of intraday volatility estimators is evaluated via the incremental Sharpe ratio brought by volatility trading based on realized combinations.

## 1.4 Layout of the Thesis

The remainder of the thesis is divided into 3 parts. Firstly, Chapter 2 presents a common methodology to be applied in the subsequent empirical studies. Secondly, Chapters 3 to 5 contain three distinct empirical studies on the economic value of intraday (co)variation estimators for different financial applications. Finally, Chapter 6 concludes the thesis and provides recommendations for further research.

Chapter 2 provides a methodological review of the background technical materials to be applied in our subsequent empirical research. First, a variety of realized variance and realized covariance matrix estimators are introduced. Second, univariate and multivariate volatility models and their corresponding realized extensions are explained. These estimators and models form the foundation of our three empirical essays.

Chapter 3 evaluates the economic value of intraday information in the context of portfolio optimization, in which the economic loss function chosen to compare their performance is the quadratic utility function implicit in mean-variance asset allocation. In this empirical study we propose a novel forecasting framework to employ intraday information for predicting the conditional covariance matrix. The purpose is to evaluate the proposed intraday estimators as well as to explore the optimal volatility timing strategy that best exploits the covariance forecasts for portfolio weighting decisions.

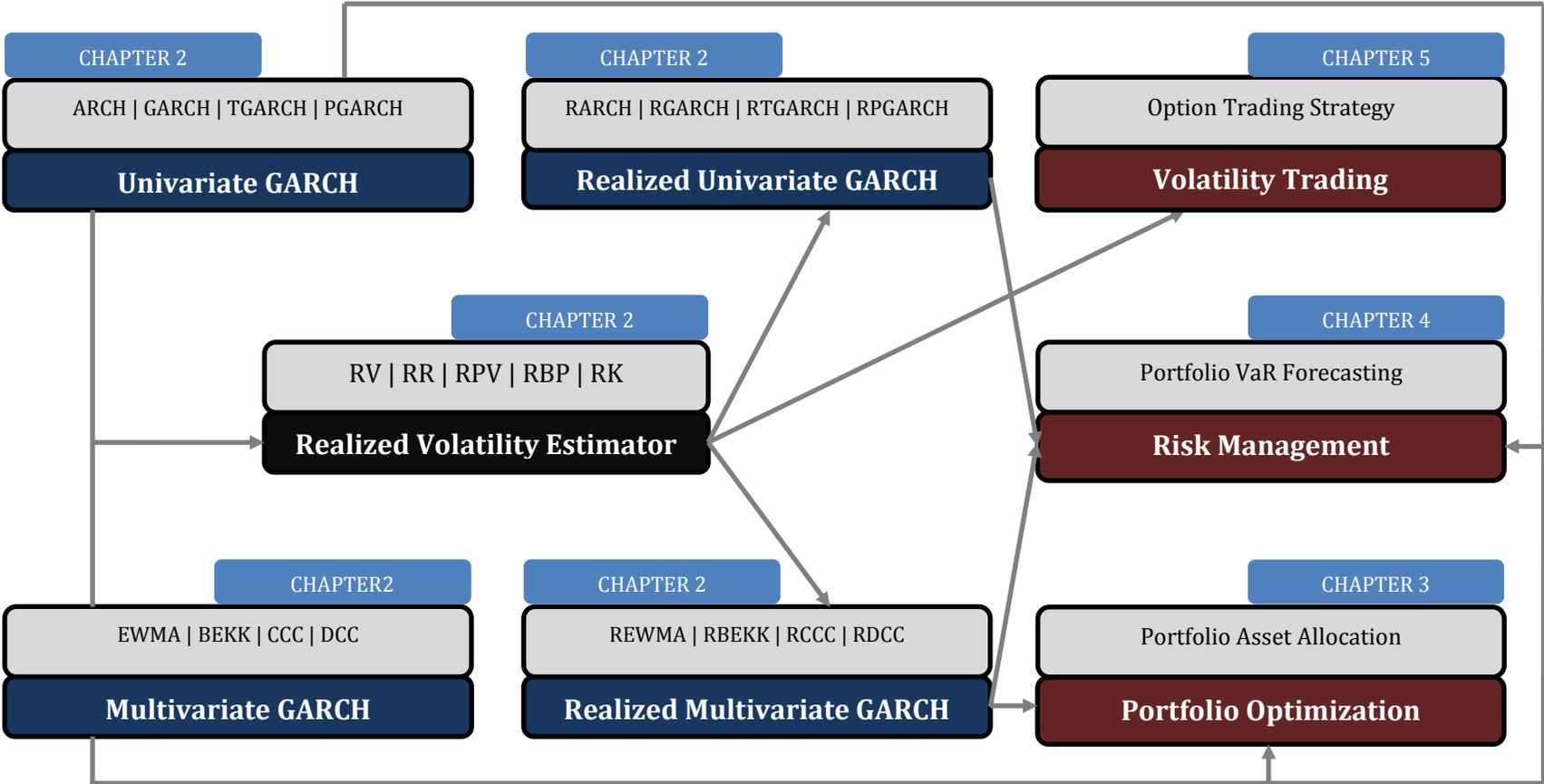
Chapter 4 compares univariate (single index) and multivariate (portfolio) intraday volatility models in a risk management setting. We assess the volatility estimators' performance in predicting portfolio Value-at-Risk (VaR) with single index models vis-à-vis portfolio models. The volatility models are also augmented with intraday information to form the realized (intraday-based) single index and portfolio specifications. Out-of-sample VaR predictions of the daily or intraday volatility models are evaluated by a number of conditional coverage tests.

Chapter 5 moves a step further by investigating the economic value of combining realized volatility estimators. The economic value of a combination estimator is assessed

directly through the return accrued and standard deviation in return occurred in the context of volatility trading. We address two issues in this chapter. First, we are interested in whether simple combination models, which combine a realized volatility estimator and the implied volatility, can outperform the best individual volatility measure. Second, can further combination models which combine a variety of intraday volatility estimators and the implied volatility using different combination methods provide additional economic gains compared with simple combination models.

Chapter 6 concludes the thesis by providing an overview of our study and a summary of the findings. Finally, the chapter recommends potential directions for future research.

# DESCRIPTIVE DIAGRAM OF THE THESIS



# CHAPTER 2: Common Methodology

## 2.1 Realized Volatility Estimators

This chapter lays out the background literature and methodology to be used throughout the thesis. We first introduce realized (intraday) univariate and multivariate volatility measures, which are used as major inputs to all the three empirical studies. Inspired by the work of Schwert (1989b) and Hsieh (1991), Andersen and Bollerslev (1998) define the realized volatility for the day as the sum of squared intraday returns. The idea is that if the sample path of volatility is continuous, then increasing the sampling frequency yields arbitrarily precise estimates of volatility at any given point in time (Merton, 1980). Therefore, the unobserved ex post volatility eventually becomes observable. A variety of realized volatility measures have been proposed and applied in academia and financial industry after Andersen and Bollerslev's (1998) initial work. A review of literature on commonly used realized volatility estimators are given by Andersen et al. (2008), and Barndorff-Nielsen and Shephard (2006).

Since the early 1990s intraday data have become increasingly available to academic research. This development has made possible a wide range of empirical studies investigating intraday information. The literature makes use of realized volatility measures and evaluate their properties for various asset classes. For instance, ABDE (2001) examine the Dow Jones Industrial Average, ABDL (2001) evaluate currencies, and Areal and Taylor (2002) study stock index futures. These empirical studies suggest that accurate measures of volatility can be generated using high frequency data, and realized volatility is a better proxy for spot volatility than lagged squared returns.

Each intraday based volatility measure may contain unique information that is not captured by competing estimators. This thesis therefore utilizes a wide range of realized volatility measures to exploit intraday information. In detail, five broad categories are considered: realised volatility (RV), realised range (RR), realised power variation (RPV),

realized bipower variation (RBP) and realized kernel (RK) measures. Different sampling frequencies are applied to these estimators. For RV, RR, RPV and RBP the sampling frequencies considered are 1, 5, 15, 30 and 60 minutes. For RK models the frequencies are set to 1 and 5 minutes. In total 41 realized volatility measures are applied in this study. The five classes of realized volatility estimators are set out below.

The realized variance (RV) is the most widely used estimator. It is defined as

$$RV_t = \sum_{j=1}^M r_{t,j}^2 \quad (2.1)$$

where  $M$  is the number of intra-daily returns used, and  $r_{t,j}$  is the  $j$ th return on day  $t$ . Andersen and Bollerslev (1998) note that RV provides a relatively accurate measure of volatility compared with daily squared returns. Theoretical properties of RV can be found in Andersen and Bollerslev (1998) and Barndorff-Nielsen and Shephard (2002).

Based on the range estimator of Parkinson (1980), Christensen and Podolskij (2005) propose to improve upon the RV estimator by replacing each intraday squared return with the high-low range. The proposed realized range (RR) estimator is defined as

$$RR_t = \frac{1}{4 \log 2} \left[ \sum_{j=1}^M 100 \times (\log(p_{t,j}^h) - \log(p_{t,j}^l))^2 \right] \quad (2.2)$$

where  $\log(p_{t,j}^h)$  and  $\log(p_{t,j}^l)$  are the high and low prices of the  $j$ th interval. Christensen and Podolskij (2005) find that RR is more efficient than other variance estimators based on squared returns in an ideal world without market frictions (no bid-ask bounce, discontinuous trading or jumps). The asymptotic variance of the RR measure is  $0.4 \int_{t-1}^t \sigma(u)^4 du$ , where the integral is defined as integrated quarticity, which is 5 times smaller than the variance of RV at  $2 \int_{t-1}^t \sigma(u)^4 du$ . Christensen and Podolskij (2005) and Martens and van Dijk (2006) suggest that, as  $M \rightarrow \infty$  the RR estimator converges in probability to the quadratic variation ( $RR_t \xrightarrow{p} QV_t$ ). However, the result does not hold in a jump-diffusion context as with jumps RR is not a consistent estimator of QV. Martens and van Dijk (2006) show that for plausible market frictions the optimal RR has a lower MSE than the optimal RV. But both RR and RV are upward biased, and infrequent trading may lead

to downward bias in RR but not RV.

The Realized Power Variation (RPV) was introduced by Barndorff-Nielsen and Shephard (2004) to accommodate return series with large jumps. A RPV of order  $z$  can be expressed as

$$RPV_t(z) = \mu_z^{-1} \delta^{1-z/2} \sum_{j=1}^M |r_{t,j}|^z \quad (2.3)$$

where  $0 < z < 2$ ,  $\mu_z = E|\mu|^z = 2^{z/2} \frac{\Gamma(\frac{1}{2}(z+1))}{\Gamma(\frac{1}{2})}$ ,  $\mu \sim N(0, 1)$ . In a special case of  $z = 1$  RPV becomes the realised absolute variation. Liu and Maheu (2005) examine the 1-day-ahead forecasting properties of RPV for different orders  $z = \{0.25, 0.5, \dots, 1.75\}$  and find that 0.5, 1, and 1.5 yield the lowest RMSE. Fuertes, Izzeldin and Kalotychou (2009) suggest that absolute return measures are more persistent than those squared estimators so RPV could outperform RV in forecasting financial risk. Furthermore, RPV may yield better forecasts than RV when the sample period contains large jumps. We therefore estimate RPV for different orders  $z = \{0.5, 1, 1.5\}$ .

The realised bipower variation (RBP), an intraday volatility measure proposed by Barndorff-Nielsen and Shephard (2006), can be expressed as

$$RBP_t = \mu_1^{-2} \sum_{j=1}^M |r_{t,j}| |r_{t,j-1}| \quad (2.4)$$

where  $\mu_1 = E(|\mu|) = \sqrt{2}/\sqrt{\pi}$  and  $\mu \sim N(0, 1)$ . The authors show that RBP converges in probability to the integrated variance ( $RBP_t \xrightarrow{p} IV_t$ ) so it is also immune to jumps.

A disadvantage of realized variance estimators is that they can be very sensitive to market frictions when applied to high frequency noisy data. Barndorff-Nielsen et al. (2008, 2009) propose a variety of realized variance measures of quadratic variation, the realized kernel (RK) estimators, to combat market frictions in noisy high-frequency data. In this paper we apply six popular RK estimators: Barlett, Epanechnikov, 2nd Order, Cubic, Parzen and Tukey-Hanning.

We consider the case where  $Y$  is a Brownian semimartingale plus jump process (BMSJ)

as

$$Y_t = \int_0^t a_u du + \int_0^t \sigma_u dW_u + J_t \quad (2.5)$$

where  $J_t = \sum_{i=1}^{N_t} C_i$  is a finite activity jump process.  $N_t$  is the total number of jumps occurred in time interval  $[0, t]$  and  $N_t < \infty$  for any  $t$ .  $a$  is a locally bounded drift,  $\sigma$  is a càdlàg<sup>1</sup> volatility process and  $W$  is a Brownian motion.

The quadratic variation of  $Y$  can be expressed as

$$[Y] = \int_0^T \sigma_u^2 du + \sum_{i=1}^{N_T} C_i^2 \quad (2.6)$$

where  $\int_0^T \sigma_u^2 du$  is the integrated variance that can be estimated from the observations  $X_{\tau_0}, \dots, X_{\tau_n}$ ,  $0 = \tau_0 < \tau_1 < \dots < \tau_n = T$ , where  $X_{\tau_j}$  is a noisy observation of  $Y_{\tau_j}$ ,

$$X_{\tau_j} = Y_{\tau_j} + U_{\tau_j} \quad (2.7)$$

Assume  $U$  is a noise term with  $E(U_{\tau_j}) = 0$  and  $Var(U_{\tau_j}) = \omega^2$ .  $U$  can be caused by market frictions such as liquidity effects, bid/ask spread and recording mistakes. We apply the realized kernels to estimate the process of  $Y_{\tau_j}$ . The estimators take on the following form:

$$K(X) = \sum_{h=-H}^H k\left(\frac{h}{H+1}\right) \gamma_h, \quad \gamma_h = \sum_{j=|h|+1}^n x_j x_{j-|h|} \quad (2.8)$$

where  $k(x)$  is a kernel weight function,  $x_j$  is the  $j$ th high frequency return over the interval  $\tau_{j-1} - \tau_j$  and  $h = -H, \dots, -1, 0, 1, 2, \dots, H$ .  $H$  is the optimized bandwidth. Barndorff-Nielsen et al. (2008) note that as  $n \rightarrow \infty$  if  $K(U) \xrightarrow{p} 0$  and  $K(Y) \xrightarrow{p} [Y]$  then

$$K(X) \xrightarrow{p} [Y] = \int_0^T \sigma_u^2 du + \sum_{i=1}^{N_T} C_i^2 \quad (2.9)$$

The weight functions of the 6 RK estimators to be applied are listed in the table below:

---

<sup>1</sup>A function defined on the real numbers (or a subset of them) that is everywhere right-continuous and has left limits everywhere.

	$k(x)$
Bartlett	$1 - x$
Epanechnikov	$1 - x^2$
2nd Order	$1 - 2x + x^2$
Cubic Kernel	$1 - 3x^2 - 2x^3$
Parzen	$1 - 6x^2 + 6x^3, \quad 0 \leq x \leq 1/2$
	$2(1 - x)^3, \quad 1/2 \leq x \leq 1$
Tukey-Hanning	$(1 - \cos(\pi x))/2$

Barndorff-Nielsen et al. (2008) note that  $H \propto n^{3/5}$  is the best trade-off between asymptotic bias and variance<sup>2</sup>. The optimal choice of bandwidth is

$$H^* = c^* \xi^{4/5} n^{3/5} \quad (2.10)$$

with  $c^* = \left\{ \frac{k''(0)^2}{k^{0,0}} \right\}^{1/5}$  and  $\xi^2 = \frac{\omega^2}{\sqrt{T \int_0^T \sigma_u^4 du}}$ , where  $k^{0,0} = \int_0^1 k(x)^2 dx$ .<sup>3</sup>  $c^*$  can be directly calculated for the RK measures, for example,  $c^* = \left( \frac{(12)^2}{0.269} \right)^{1/5} = 3.5134$  for the Parzen, 2.28 for Bartlett and 3.42 for 2nd order kernel. The bandwidth  $H^*$  then depends on the unknown quantities  $\omega^2$  and  $\int_0^T \sigma_u^4 du$ , where the latter is the integrated quarticity. We estimate  $\xi$  very simply by

$$\xi = \frac{\hat{\omega}^2}{\widehat{IV}} \quad (2.11)$$

where  $\hat{\omega}^2$  is an estimator of  $\omega^2$ , which is estimated by the realized variance using every trade or quote,  $\omega^2 = \frac{RV_{tick}}{2n}$ , where  $n$  is the number of non-zero returns that were used to compute  $RV_{tick}$ .  $\widehat{IV}$  is estimated by  $RV_{15 \text{ min}}$ , the realized variance based on 15 minute returns. The use of  $\widehat{IV}$  is motivated by the assumption that  $\sigma_u^2$  does not vary too much over the interval  $[0, t]$ .

<sup>2</sup>This means that  $K(X) \xrightarrow{P} [Y]$  at a rate  $n^{1/5}$ , which has the advantage of being non-negative with probability 1.

<sup>3</sup>The non-stochastic weight function  $k(x)$  is characterized by:

- (i)  $k(0) = 1, k'(0) = 0$ ;
- (ii)  $k$  is twice differentiable with continuous derivatives;
- (iii)  $k^{0,0}, k^{1,1}, k^{2,2} < \infty$ , where  $k^{0,0} = \int_0^1 k(x)^2 dx, k^{1,1} = \int_0^1 k'(x)^2 dx, k^{2,2} = \int_0^1 k''(x)^2 dx$ ;
- (iv)  $\int_{-\infty}^{\infty} k(x) \exp(ix\lambda) dx \geq 0$  for all  $\lambda \in \mathbb{R}$ .

## 2.2 Realized Covariance Estimators

Two realized covariance matrix estimators, *realized covariance* (RC) and *realized kernel covariance* (RKC) are considered for the empirical studies.

### 2.2.1 Realized Covariance (RC)

Let us assume that log prices follow a multivariate continuous-time stochastic volatility process and let  $N$  denote the number of assets in a portfolio. Define  $\mathbf{S}_t$  as the value of the  $N \times N$  positive-definite diffusion matrix at time  $t$ . The integrated diffusion matrix  $\int_0^1 \mathbf{S}_{t+\tau} d\tau$  represents the latent covariance matrix of the vector of continuously compounded returns over daily time interval  $t$  to  $t + 1$ . Let us divide this interval into  $M$  subintervals of length  $\delta$ . The theory of quadratic variation implies that  $\sum_{j=1}^M \mathbf{r}_{t,j} \mathbf{r}'_{t,j} - \int_0^1 \mathbf{S}_{t+\tau} d\tau \rightarrow 0$  as  $\delta \rightarrow 0$ , where  $r_{t,j}$  represents the  $N \times 1$  returns vector over the  $j$ th (intraday) subinterval of day  $t$ . This motivates the model-free *realized covariance* estimator for day  $t$  defined as

$$\mathbf{RC}_t \equiv \sum_{j=1}^M \mathbf{r}_{t,j} \mathbf{r}'_{t,j} \quad (2.12)$$

which is asymptotically consistent for the latent covariance matrix.

### 2.2.2 Realized Kernel Covariance (RKC)

We follow the procedure of Barndorff-Nielsen et al. (2008) to calculate the *realized kernel covariance* (RKC). Suppose the  $k$ -dimensional efficient log-price  $Y(t)$  follows a continuous time diffusion process as

$$Y(t) = \int_0^t a(u) du + \int_0^t \Phi(u) dW(u) \quad (2.13)$$

where  $a(t)$  is a vector of drift components,  $\Phi(t)$  is the instantaneous volatility matrix, and  $W(t)$  is a vector of independent Brownian motions.<sup>4</sup> For reviews of the theory of this

---

<sup>4</sup>Jumps are not discussed in this work. Handling individual asset jumps and common jumps among several assets is an open question which we leave for future research.

type of process see Ghysels, Harvey and Renault (1996). Stochastic process theory (e.g. Protter, 2004) implies that the integrated covariance of  $Y(t)$ ,  $\int_0^1 \Phi(u)\Phi'(u)$ , is equal to its quadratic variation over the same interval,

$$[Y](1) = \text{plim}_{n \rightarrow \infty} \sum_{j=1}^n \{Y(t_j) - Y(t_{j-1})\} \{Y(t_j) - Y(t_{j-1})\}' \quad (2.14)$$

for any sequence of partitions  $0 = t_0 < t_1 < \dots < t_n = 1$  with  $\sup_j \{t_{j+1} - t_j\} \rightarrow 0$  for  $n \rightarrow \infty$ .

Barndorff-Nielsen et al. (2008) observe the log price process  $X = (X^{(1)}, X^{(2)}, \dots, X^{(k)})'$ , which is generated by  $Y$ , but is contaminated with market microstructure noise. Prices are quoted at different times and at different frequencies for different assets over the trading day,  $t \in [0, 1]$ . We apply the idea of refresh time used in Barndorff-Nielsen et al. (2008) to synchronize the data.

Suppose the observation times for the  $i$ -th stock are written as  $t_1^{(i)}, t_2^{(i)}, \dots, i = 1, 2, \dots, k$ . Let  $N_t^{(i)}$  count the number of distinct price observations for the  $i$ -th asset up to time  $t$ . The quoted prices for the day is  $X^{(i)}(t_j^{(i)})$ , for  $j = 1, 2, \dots, N_1^{(i)}$ . For example, the  $j$ -th price update for asset  $i$ ,  $X^{(i)}(t_j^{(i)})$  arrives at  $t_j^{(i)}$ .

Let  $\tau_1 = \max(t_1^{(1)}, \dots, t_1^{(k)})$  represents the first refresh time of the trading day and  $\tau_{j+1} = \max(t_{N_{\tau_j}^{(1)}+1}^{(1)}, \dots, t_{N_{\tau_j}^{(k)}+1}^{(k)})$  the subsequent refresh time.  $\tau_1$  is the time when the prices of all the assets have been updated at least once.  $\tau_2$  is then the first time when all the prices are again updated, etc. These refresh times  $\tau_j$  form a new time clock for each trading day to synchronize the data. The number of observations of the synchronized price vector is  $n + 1$ , which is no larger than the number of quotes of the asset with the fewest price observations. We then define the synchronized high frequency return vector as  $x_j = X(\tau_j) - X(\tau_{j-1})$ ,  $j = 1, 2, \dots, n$ , where  $n$  is the number of refresh return observations for the day.

The daily positive semi-definite realized kernel  $RKC_t$  is then calculated as

$$RKC_t = \sum_{h=-n}^n k \left( \frac{h}{H+1} \right) \Gamma_h \quad (2.15)$$

where the non-stochastic  $k(x)$  is one of the 6 RK weight functions introduced in section 2.1. The  $h$ -th realized autocovariance is

$$\Gamma_h = \begin{cases} \sum_{j=|h|+1}^n x_j x'_{j-h} & h \geq 0 \\ \sum_{j=|h|+1}^n x_{j-h} x'_j & h < 0 \end{cases} \quad (2.16)$$

For full details along with the calculation of bandwidth  $H$  see Barndorff-Nielsen et al. (2008). We calculate the multivariate realized kernel estimation for our intraday data, generating a series of daily  $RKC_t$  matrices, which will be used to augment the standard daily multivariate GARCH models.

## 2.3 Univariate Volatility Model

Starting with Engle's (1982) autoregressive conditional heteroskedasticity (ARCH) specification, a wide range of volatility models have been proposed, such as the generalized ARCH (GARCH) (Bollerslev, 1986), exponential ARCH (Nelson, 1991), stochastic volatility and implied volatility models, for modelling and forecasting the volatility of financial asset returns. Among the models, GARCH specifications are widely applied in both academia and industry. The popularity of GARCH models can be attributed largely to their ability to accommodate several stylised facts of financial data, such as time-varying volatility, volatility clustering and asymmetric responses to positive and negative surprises of equal magnitude.

While GARCH models are mainly estimated using daily return series in practice, over the past decade a number of studies have focused on improving the forecasts from GARCH models by exploiting intraday information (see Martens, 2001; Engle, 2002; Koopman et al., 2005). Intraday volatility estimators may contain unique information about asset return variations that is not captured by the daily squared return, a noisy estimator of ex post volatility.

The empirical literature indeed suggests that high-frequency information could be employed to improve daily volatility forecast accuracy. A popular method is use intraday

information as an additional regressor to extend daily GARCH models. Instances of augmentation variable are the daily high-low price range (Parkinson, 1980; Taylor, 1987), the number of intraday price changes (Laux and Ng, 1993), daily trading volume (Bessembinder and Seguin, 1993) and the standard deviation of intraday returns (Taylor and Xu, 1997). More recently, Fuertes, Izzeldin and Kalotychou (2009) investigate individual NYSE/Nasdaq stocks, and provide statistical evidence in favour of intraday based volatility models. They incorporate nonparametric estimators of daily price variability into a GARCH model. Four estimators are compared: RV, RR, RPV and RBP. Test results show that RR fits relatively well in the in-sample fit analysis and RPV provides the best out-of-sample performance.

This thesis applies four popular GARCH models, ARCH, GARCH, GJRGARCH and PGARCH, for modelling and forecasting the conditional variance of the portfolio returns. Alongside the original version of these four standard GARCH models based on daily returns we implement a realized version that incorporates intraday information into the volatility forecasts by adding a realized volatility estimator as an additional regressor into the daily specifications.

Let the conditional mean be captured by the following ARMA(p, q) equation

$$r_t = \theta_0 + \sum_{i=1}^p \theta_i r_{t-i} + \sum_{j=1}^q \lambda_j u_{t-j} + u_t, \quad u_t | \Omega_{t-1} \sim iid(0, h_t) \quad (2.17)$$

where  $r_t$  are daily returns,  $u_t$  are whitened returns and  $h_t$  is the conditional variance of returns. The lag orders of the conditional mean equation can be appropriately selected so as to remove all the return autocorrelation and volatility clustering. We use the Ljung-Box and ARCH LM test for these purposes. The conditional variance  $h_t$  is captured by the following GARCH models.

The autoregressive conditional heteroskedasticity (ARCH) model is proposed by Engle (1982), which assumes that the variance of the current innovation to be a function of the actual size of past innovations. An ARCH model is formulated as

$$h_t = w + \sum_{i=1}^r \alpha_i u_{t-i}^2 \quad (2.18)$$

and the realized ARCH is defined as

$$h_t = w + \sum_{i=1}^r \alpha_i u_{t-i}^2 + \gamma \hat{v}_t \quad (2.19)$$

where  $\hat{v}_t$  can be an ARFIMA forecasted realized estimator based on past intraday prices, or an ARFIMA forecasted combination of realized estimators.

The GARCH model, proposed by Bollerslev (1986), is a generalization of Engle's (1982) ARCH specification. A GARCH model is usually expressed as GARCH( $p, q$ ), where  $p$  gives the number of autoregressive lags and  $q$  shows the quantity of moving average component lags. The day  $t$  conditional variance is a weighted average of a constant mean, information about volatility and squared errors on the past. A GARCH specification is formulated as

$$h_t = w + \sum_{i=1}^r \alpha_i u_{t-i}^2 + \sum_{j=1}^s \beta_j h_{t-j} \quad (2.20)$$

and the corresponding realized GARCH model is

$$h_t = w + \sum_{i=1}^r \alpha_i u_{t-i}^2 + \sum_{j=1}^s \beta_j h_{t-j} + \gamma \hat{v}_t \quad (2.21)$$

Glosten, Jagannathan and Runkle (1992) augment the GARCH model to accommodate possible asymmetries between the effects of positive and negative shocks of the same magnitude on the conditional variance. The proposed GJRGARCH model adds a dichotomous dummy variable into the standard GARCH specification. The first-order threshold specification can be expressed as

$$h_t = w + \sum_{i=1}^r \alpha_i u_{t-i}^2 + \sum_{j=1}^s \beta_j h_{t-j} + \delta u_{t-1}^2 I_{t-1} \quad (2.22)$$

The dummy variable  $I(\cdot)$  stands for the indicator function, where  $I_{t-1} = 1$  if  $u_{t-1} < 0$  and  $= 0$  otherwise. The realized GJR-GARCH model is specified as

$$h_t = w + \sum_{i=1}^r \alpha_i u_{t-i}^2 + \sum_{j=1}^s \beta_j h_{t-j} + \delta u_{t-1}^2 I_{t-1} + \gamma \hat{v}_t \quad (2.23)$$

PGARCH is an asymmetric Power GARCH model proposed by Ding, Granger and Engle (1993) based on the standard deviation GARCH models of Taylor (1986) and

Schwert (1989a). The PGARCH model is given by

$$h_t^\eta = w + \sum_{i=1}^r \alpha_i (|u_{t-i}| - \tau_i u_{t-i})^\eta + \sum_{j=1}^s \beta_j h_{t-j}^\eta \quad (2.24)$$

where  $\eta$  is the power parameter that can be estimated rather than imposed, and  $\tau$  is included to capture the effects of asymmetric shocks. The realized PGARCH model is specified as

$$h_t^\eta = w + \sum_{i=1}^r \alpha_i (|u_{t-i}| - \tau_i u_{t-i})^\eta + \sum_{j=1}^s \beta_j h_{t-j}^\eta + \gamma \hat{v}_t \quad (2.25)$$

## 2.4 Multivariate Volatility Models

This section outlines the different covariance matrix forecasts considered in the thesis. On the one hand, following Foster and Nelson (1996) we deploy a parsimonious rolling window (ROLL) approach, which can be regarded as a special case of a multivariate GARCH specification, as the benchmark model. On the other hand, we apply three covariance modeling approaches within the MGARCH family: the scalar CCC model (Bollerslev, 1990), scalar DCC model (Engle, 2002) and the diagonal BEKK model (Engle and Kroner, 1995). Alongside the original version of these four covariance forecasting approaches based on daily returns we implement a realized version that embeds intraday return information in the covariance forecasts through the realized covariance matrix.

### 2.4.1 Rolling conditional covariance estimator

The rolling conditional covariance estimation approach advocated by Foster and Nelson (1996) constructs daily estimates of the covariance matrix using the following backward-looking rolling approach

$$\hat{\mathbf{C}}_t = \sum_{i=1}^{\infty} \mathbf{\Omega}_{t-i} \odot \mathbf{r}_{t-i} \mathbf{r}'_{t-i}, \quad (2.26)$$

where  $\mathbf{\Omega}_{t-i}$  is a symmetric ( $N \times N$ ) matrix of weights,  $\odot$  denotes element-by-element multiplication, and  $\mathbf{r}_{t-i}$  is an  $N \times 1$  dimensional vector of demeaned daily returns. Although the above equation admits a wide range of weighting schemes, following Fleming

et al. (2001; 2003) and Martens et al. (2008) we let the weights decline in an exponential fashion as the returns move further away into the past (i.e. as the magnitude of  $i$  increases) according to  $\mathbf{\Omega}_{t-i} = \alpha \exp(-\alpha i) \mathbf{1}\mathbf{1}'$  where  $\mathbf{1}$  denotes an  $N \times 1$  vector of ones. This approach is consistent with the results in Foster and Nelson (1996) showing that exponentially decaying weights produce the smallest MSE asymptotically. Thus the elements of  $\hat{\mathbf{C}}_t$  can be equivalently obtained as<sup>5</sup>

$$\hat{\mathbf{C}}_t = \exp(-\alpha) \hat{\mathbf{C}}_{t-1} + \alpha \exp(-\alpha) \mathbf{r}_{t-1} \mathbf{r}'_{t-1}. \quad (2.27)$$

As in Fleming et al. (2003) the realized rolling conditional covariance estimator, called RROLL hereafter, exploits the intraday returns by replacing the  $N \times N$  matrix  $\mathbf{r}_{t-1} \mathbf{r}'_{t-1}$  with the realized covariance matrix  $RC_{t-1}$ . Thus equation (2.27) becomes

$$\tilde{\mathbf{C}}_t = \exp(-\alpha) \tilde{\mathbf{C}}_{t-1} + \alpha \exp(-\alpha) \mathbf{RC}_{t-1}. \quad (2.28)$$

where  $\mathbf{RC}_{t-1}$  can be the *realized covariance*, equation (2.12), or the *realized kernel covariance*, equation (2.15). The decay factor ( $0 < \alpha < 1$ ) is optimally selected through a fine grid search in step of 0.001 so as to minimize the MSE over the in-sample period.<sup>6</sup> For this purpose, we initialize the rolling approach by estimating  $\hat{\mathbf{C}}_t$  over the first 100 days of the in-sample period and the window is rolled forward up to the last available in-sample day.

## 2.4.2 Multivariate GARCH forecasts

In contrast to univariate GARCH specifications, multivariate GARCH (MGARCH) models specify the risk of one asset as depending dynamically on its own past volatility as

---

<sup>5</sup>It can be shown that this nonparametric rolling window estimator is nested under the Engle and Kroner (1995) multivariate GARCH model, that is, the EWMA covariance estimator can be seen as a restricted M-GARCH model.

<sup>6</sup>On each in-sample day, from day 101 onwards, an overall squared error measure is obtained by aggregating the squared differences between each forecast (variance or covariance) and the “observed” realized values. The average of the overall squared error measures over the in-sample window provides the MSE.

well as on the historical volatility of other assets. Multivariate extensions from univariate volatility models enable modelling of the relationship between components of a portfolio and allow for scenario and sensitivity analysis.

This section lists out three alternative models, diagonal BEKK, scalar CCC and DCC, that can be applied to estimate the conditional variance of each asset and the conditional correlation of each pair of assets. In the MGARCH class of models, the return process is specified as

$$\begin{aligned} r_t &= \Pi_t(\Theta) + \epsilon_t \\ \epsilon_t &= \mathbf{C}_t^{1/2}(\theta)\mathbf{z}_t \end{aligned} \tag{2.29}$$

where  $r_t$  is an  $N \times 1$  vector denoting the day  $t$  asset returns,  $\Pi_t(\theta)$  are the conditional mean returns which are a function of a parameter vector  $\Theta$ , the innovation vector is  $\epsilon_t$ , and  $\mathbf{C}_t^{1/2}$  is an  $N \times N$  positive definite matrix such that  $\mathbf{C}_t$  is an estimator of the conditional variance matrix of  $r_t$ . The standardized innovations are assumed to be *iid* and uncorrelated across assets, that is, the mean of  $\mathbf{z}_t$  is the  $N \times 1$  zero vector, and its covariance is given by the identity matrix of order  $N$ .

The most basic multivariate GARCH model is the VEC specification, a direct generalization of the univariate GARCH model, proposed by Bollerslev et al. (1988). In the general VEC model, each component of the conditional covariance matrix  $\mathbf{C}_t$  is a linear function of the lagged squared errors, cross-products of errors and lagged values of the components of  $\mathbf{C}_t$ . There are two problems associated with the VEC. First, it does not guarantee positive definiteness of the conditional covariance matrix. Second, it needs to estimate a large number of parameters, which becomes computationally difficult when the number of assets in the assessing portfolio exceeds three or four.

A number of MGARCH models have been proposed to address the problems. We first introduce the diagonal Baba, Engle, Kraft and Kroner (BEKK) model proposed by Engle and Kroner (1995). The model parameters are estimated by the maximum likelihood procedure under the assumption of conditional multivariate normality. It reduces the parameters to be estimated and guarantees the positive definiteness of  $\mathbf{C}_t$ , which is

necessary for generating realistic covariance matrix forecasts. A BEKK(1,1) model parameterizes  $\mathbf{C}_t$  as a function of its own lagged values and lagged errors of the mean equation as follows

$$\mathbf{C}_t = \mathbf{S}'\mathbf{S} + \mathbf{A}'\boldsymbol{\epsilon}_{t-1}\boldsymbol{\epsilon}'_{t-1}\mathbf{A} + \mathbf{B}'\mathbf{C}_{t-1}\mathbf{B} \quad (2.30)$$

where  $\mathbf{S}$  is an upper  $N \times N$  triangular parameter matrix with a positivity restriction on its diagonal elements, and  $\mathbf{A}$  and  $\mathbf{B}$  are  $N \times N$  positive diagonal parameter matrices. The covariance matrices thus obtained are positive definite and stationary by default. The realized BEKK model (RBEKK hereafter) is defined as

$$\mathbf{C}_t = \mathbf{S}'\mathbf{S} + \mathbf{A}'\boldsymbol{\epsilon}_{t-1}\boldsymbol{\epsilon}'_{t-1}\mathbf{A} + \mathbf{B}'\mathbf{C}_{t-1}\mathbf{B} + \Gamma'\mathbf{R}\hat{\mathbf{C}}_t\Gamma \quad (2.31)$$

By imposing a diagonal restriction on parameter matrices the number of parameters in the diagonal BEKK model is  $3N + N(N + 1)/2$ . However, the number of unknown parameters is still high and as a consequence the model is rarely used when the number of series is larger than 3 or 4. This problem can be alleviated by using estimators within the Conditional Correlation (CC) family, described below, which can be easily computed in a two-step approach by decomposing the conditional covariance into conditional variance and conditional correlation matrices.

In the context of modelling conditional correlations rather than conditional covariances, Bollerslev (1990) introduces a Constant Conditional Correlation (CCC) model to reduce the number of unknown parameters in the BEKK formulation by assuming that the conditional correlations,  $Q = (\rho_{ij})$ , to be time constant. The CCC model estimates the conditional variances and correlations using a simple two-step routine. In this routine, the conditional variance of the  $i$ th series,  $h_{ii,t}$   $i = 1, \dots, N$ , is calculated in the first step by fitting univariate GARCH(1,1) models to each series. The second step involves calculating the covariances from the product of the corresponding conditional GARCH standard deviations and conditional correlations as

$$\mathbf{C}_t \equiv \mathbf{D}_t\mathbf{R}\mathbf{D}_t = (\rho_{ij}\sqrt{h_{ii,t}h_{jj,t}}) \quad (2.32)$$

where  $\mathbf{D}_t = \text{diag}(h_{11,t}^{\frac{1}{2}} \dots h_{NN,t}^{\frac{1}{2}})$  is a diagonal matrix constructed by the conditional variances of the  $N$  series, and  $Q = (\rho_{ij})$  is a symmetric positive definite matrix with off-diagonal elements filled by the long-run return correlations. Similar to the BEKK model, the CCC specification also guarantees the positive definiteness of the conditional covariance matrix. The realized CCC covariance matrix (RCCC hereafter) is calculated through an intraday-augmented GARCH(1,1) approach

$$h_{ii,t} = \omega_i + \alpha_i \varepsilon_{i,t-1}^2 + \beta_i h_{ii,t-1} + \gamma_i \hat{v}_{i,t}, \quad i = 1, \dots, N \quad (2.33)$$

where  $\hat{v}_{i,t}$  is the  $i$ th diagonal entry of the realized covariance matrix  $\mathbf{R}\hat{\mathbf{C}}_t$ .

As an extension of the CCC model, Engle (2002) introduced the Dynamic Conditional Correlation (DCC) model, which permits the conditional correlation between portfolio components to vary parsimoniously over time as  $Q_t = (\rho_{ij,t})$ . DCC is estimated from a two-step approach in a way similar to the CCC model, the first step involves fitting univariate GARCH(1,1) models to each series to obtain the standardized errors,  $z_{it} = \varepsilon_{it} / \sqrt{h_{ii,t}}$ ,  $i = 1, \dots, N$ . The second step calculates the time-varying correlation matrix  $Q_t = \text{diag}(p_{11,t}^{-\frac{1}{2}} \dots p_{NN,t}^{-\frac{1}{2}}) P_t \text{diag}(p_{11,t}^{-\frac{1}{2}} \dots p_{NN,t}^{-\frac{1}{2}})$ , where  $P_t = (p_{ij,t})$  is a  $N \times N$  symmetric positive definite matrix defined as

$$\mathbf{P}_t = (1 - \alpha - \beta) \bar{\mathbf{P}} + \alpha \mathbf{z}_{t-1} \mathbf{z}'_{t-1} + \beta \mathbf{P}_{t-1}. \quad (2.34)$$

$\alpha$  and  $\beta$  are non-negative scalar parameters satisfying  $(\alpha + \beta) < 1$ , and  $\bar{\mathbf{P}}$  is a  $N \times N$  unconditional covariance matrix of the standardized errors  $z_t$ . The DCC covariance matrix is estimated as follows

$$\mathbf{C}_t \equiv \mathbf{D}_t \mathbf{Q}_t \mathbf{D}_t \quad (2.35)$$

with  $D_t = \text{diag}(h_{1t}^{\frac{1}{2}} \dots h_{Nt}^{\frac{1}{2}})$ . The realized DCC covariance matrix (RDCC hereafter) is formulated by adding a lagged realized covariance matrix as an additional explanatory variable to the  $\mathbf{P}_t$  regression

$$\mathbf{P}_t = (1 - \alpha - \beta) \bar{\mathbf{P}} + \alpha \mathbf{z}_{t-1} \mathbf{z}'_{t-1} + \beta \mathbf{P}_{t-1} + \gamma \mathbf{R}\hat{\mathbf{C}}_t \quad (2.36)$$

where  $\mathbf{R}\hat{\mathbf{C}}_t$  is day  $t$  realized covariance matrix and  $\gamma$  is a non-negative scalar. A drawback of the DCC model (versus the BEKK approaches) is that all the conditional correlations are assumed to follow the same dynamics.

# CHAPTER 3: Decision-Based Evaluation of Realized Covariance Forecasts

## 3.1 Introduction

Over the last decade, the use of high frequency data in volatility forecasting has been advocated and many nonparametric realized volatility estimators have been proposed (see Andersen and Bollerslev, 1998; Andersen et al., 2001; Barndorff-Nielsen and Shephard, 2004). Consistent with Clark's (1973) propositions, a number of empirical findings (see Ebens, 1999; Areal and Taylor, 2002) provide support for intraday volatility estimators. For example, as noted in Fleming et al. (2003), realized volatility tends to be lognormally distributed and daily returns standardized by realized volatility follow a normal distribution. Furthermore, realized volatility exhibits long-memory dynamics, volatility clustering, and obeys precise scaling laws under temporal aggregation.

Although the literature supports the intraday volatility approach, it is mainly statistical in nature, which brings a separate question of whether the gains in accuracy are sufficient to have an economic value on decisions that depend on conditional volatility forecasts. Indeed, applications such as risk management should benefit since performance in this context depends largely on the statistical accuracy of the forecasts. However, whether using intraday information leads to better investment management decisions is a question yet to be fully addressed. It is possible that daily volatility models provide sufficient information of volatility characteristics for these purposes so that switching to intraday volatility yields only negligible benefits. In contrast, perhaps intraday volatility models could produce significant economic value by assisting investment decisions with an extra layer of information.

Despite growing popularity of intraday volatility models, few studies have explored the value of intraday data for modelling and forecasting the entire conditional covariance matrix and on the relative economic value of different multivariate covariance estimation

approaches (see Bauwens, Laurent and Rombouts, 2006, for a review on multivariate ARCH-type models). Since the initial study of Fleming et al. (2001) on the economic value of volatility timing, several papers have furthered the research by delving into issues such as the use of intraday data (Fleming et al., 2003), the choice of optimal sampling frequency for realized covariances and the best portfolio rebalancing frequency (Martens, van Dijk, and de Pooter, 2008). However, the literature that makes use of intraday price information is based only on nonparametric rolling window covariance estimators. On the other hand, a number of studies have provided evidence in favour of Multivariate GARCH (MGARCH, hereafter) estimators in the context of dynamic asset allocation (see Engle and Colacito, 2006; Kalotychou, Staikouras and Zhao, 2009; Della Corte, Sarno and Tsiakas, 2009).

This empirical study contributes to the literature in three directions. First, we assess the statistical vis-à-vis economic value of augmenting daily historical covariance matrices with additional information. By ‘economic value’ we refer to the maximum return that an investor with quadratic utility function would be willing to sacrifice each day in order to capture the performance gains associated with the augmented covariance estimators. To this end, we deploy three distinct MGARCH models and augment them with realized covariance measures (the resulting specifications are called RMGARCH hereafter). In this respect, our study is a direct extension of Fuertes, Izzeldin and Kalotychou (2009) by considering the entire covariance matrix, as opposed to variances only, and adopting a portfolio mean-variance framework instead of a volatility-trading one. The incremental role of intraday information is assessed by comparing the one-day-ahead covariance forecasts using, on the one hand, a battery of statistical forecast accuracy measures and, on the other, a dynamic asset allocation framework that accounts for transaction costs and short-selling restrictions. The main question of interest is whether the statistical advantage (if any) translates into improved performance in a volatility-timing framework.

Second, we address the research question that whether more sophisticated intraday

models, the realized MGARCH (RMGARCH hereafter) specifications, can outperform the naive intraday based rolling conditional covariance (RROLL hereafter) model. On the one hand, RMGARCH models may produce better results given that the model parameters are optimally estimated. On the other hand, if the dimension of the portfolio increases, RMGARCH models become harder to implement due to usually large number of parameters to be estimated. Given the numerical difficulties, the practical value of implementing RMGARCH models would be limited if there is no significant economic merit in switching from RROLL to RMGARCH. Third, we focus on the issue of rebalancing frequency by comparing several rebalancing strategies where the frequency of rebalancing is either time-constant (i.e. daily, weekly and monthly) or time-varying. In the latter case rebalancing is triggered by changes in asset returns or market conditions beyond a threshold level.

To preview our key results, a general phenomenon across all forecasting approaches, ROLL and MGARCH models, is that the covariance forecasts that exploit intraday information provide consistent superior performance to their daily counterparts which is robust to transactions costs and the presence of short-selling constraints. Statistical accuracy of covariance forecasts is not tantamount to forecasting ability for market timing purposes. Under time-constant rebalancing frequency strategies the RROLL produces the highest basis points hence the largest economic value. While under time-varying rebalancing strategies the RCCC generates the best performance from the same economic point of view. Regarding to the rebalancing frequency, we find that more economic benefit can be extracted from following an optimal time-varying rebalancing strategy rather than rebalancing the portfolio on a daily basis.

The remainder of the chapter is organized as follows. Section 3.2 summarizes the relevant literature. Section 3.3 presents the covariance forecasting methodology and the evaluation framework. Section 3.4 introduces the dataset. Section 3.5 discusses empirical results and Section 3.6 concludes.

## 3.2 Literature Review

### 3.2.1 Market Timing Strategy

Market timing strategy often refers to the investment decision that investors strategically make after observing and forecasting market conditions. For instance, investors can shift their portfolio completely between risky and risk-free assets based on the one-step forecasts of the excess return of the stocks (Pesaran and Timmermann, 1994), or reallocate asset weights in their portfolio according to one-day ahead volatility predictions (Fleming et al. 2001, 2003). The market timing literature (see Grant, 1977; Breen, 1989; Chen and Knez, 1996) suggests that simple market timing strategies often perform significantly better than buy-and-hold policies thanks to utilization of forecasting power in asset returns or volatilities.

The value of implementing market timing strategies is determined by the forecasting capability. Extensive empirical studies show that although asset returns may not follow random walk (see Fama and French, 1988; Poterba and Summers, 1988), they are extremely hard to model and predict. In contrast, the literature suggests that the volatility of asset returns can be predicted in a degree of accuracy using standard volatility models (see Andersen and Bollerslev, 1998). Inspired by the relative predictability of asset return variations, vast literature studied volatility timing, a market timing strategy based on volatility forecasts, over the past decade.

Volatility timing strategies for a portfolio of assets often employ a mean-variance optimization framework. The mean-variance approach proposed by Markowitz (1952) is a cornerstone of modern portfolio theory. The approach is still the most widely used risk-return analysis framework in both academic research and financial practice because it provides an easy to implement tool for investors to optimize asset allocation by considering the trade-off between risk and return.

A static mean-variance analysis will optimize investors' portfolio allocation for a sin-

gle upcoming period. This static framework has been extended to dynamic settings for discrete-time and continuous-time model, respectively (see Korn and Trautmann, 1995; Li and Ng, 2000; Leippold et al., 2004). For example, in daily dynamic mean-variance optimization, an investor will be concerned with strategies in which the portfolio is rebalanced to a specified allocation at the end of each trading day. The goal here is to minimize the portfolio volatility for a target return, or maximize the portfolio return for a target level of volatility.

### 3.2.2 Economic Value of Volatility Timing

A vast literature exists on statistical accuracy of volatility forecasts in a portfolio context. For a survey of such literature see (Poon and Granger, 2003, 2005), or for a more detailed comparison of forecasts, see Hansen and Lunde (2005) or Becker and Clements(2008). In contrast, few researches examine the economic value of volatility forecasting within the context of portfolio optimization. West, Edison, and Cho (1993) employ a utility based evaluation of the economic value of a variety of volatility forecasts. Graham and Harvey (1996) and Copeland and Copeland (1999) assess trading rules based on volatility changes.

Fleming et al. (2001) are among the first to explore the economic value of volatility timing in a mean-variance context. Conditional mean-variance analysis is used to assess the short-horizon gains of volatility timing. They optimize a portfolio that consists of four asset classes (stocks, bonds, gold and cash) with the assumption that the investor's objective is to maximize expected return (or minimize volatility) while matching a target volatility (or expected return). The finding is that the performance of a dynamic trading strategy, which estimates the optimal asset weights over time, is significantly better than the unconditionally efficient static portfolios with the same target expected return and volatility.<sup>1</sup>

---

<sup>1</sup>Using actual returns, Fleming et al. (2001) generate an artificial (bootstrap) sample of 4,000 observations from which the mean returns, volatilities, and covariances for each asset are estimated. For the

Chou and Liu (2008) use a range-based volatility (volatility calculated from the high-low range of daily price) model to study the economic value of volatility timing in the framework of Fleming et al (2001) and find that the proposed volatility model outperforms the return-based one. Della Corte, Sarno and Tsiakas (2009) investigate how dynamic correlations in exchange rate returns affect the optimal portfolio choice of risk-averse investors. The finding is that economic value in correlation timing is robust to reasonable transaction costs as well as model uncertainty and asymmetric correlations. Mancino and Sanfelici (2008) find significant economic value in Fourier covariance estimation approaches in the presence of strong microstructure effects.

Fleming et al. (2003) build upon Fleming et al. (2001) and emphasize the economic value of realized volatility relative to daily volatility estimates. The authors use intraday returns on three actively traded futures contracts (S&P 500 index, Treasury bonds, and gold) to demonstrate that a mean-variance efficient investor would be willing to pay 50 to 200 basis points per annum to capture the observed gains from the optimized portfolio based on intraday estimated covariances instead of choosing unconditionally efficient static portfolios. In addition, the realized volatility based estimator at the 5-minute sampling frequency also outperforms the daily returns based estimator. The result is robust to transaction costs, estimation risk regarding expected returns and the performance measurement horizon. In contrast, Liu (2004) studies the performance of the minimum variance portfolio and the minimum tracking error portfolio (tracking the S&P 500 index) using 5-minute intraday returns for the 30 Dow Jones index components and finds that an investor will not switch from daily to intraday returns to estimate the conditional covariance matrix if she rebalances her portfolio monthly and has more than

---

volatility-timing strategies, the inputs are the estimated mean returns and the daily covariance matrix estimates to determine the daily optimal portfolios. For the static strategies, the inputs are the estimated mean returns, volatilities, and covariances from the bootstrap sample to determine the unconditional optimal portfolio. For the volatility-timing strategies, daily rebalancing is carried out to track the time-varying portfolio weights whereas for the static portfolios it is need to maintain a constant weight in each asset.

6 months of past data available.

More recently, De Pooter et al. (2008) focus on the issue of determining the optimal sampling frequency as judged by the performance of S&P 500 equity portfolios where performance is defined in terms of switching fees as in Fleming et al. (2001). The finding is that the optimal sampling frequency ranges between 30 and 65 minutes, much lower than the often used 5-minutes. The optimal sampling frequency is lower as a result of the trade-off between more accurate covariance estimates and the bias caused by market microstructure effects. Their findings are robust to transaction costs and rebalancing frequency. They show that using realized volatility computed at 30- to 65-minute frequency and adopting a monthly rebalancing strategy yields the best portfolio performance. Clements and Silvennoinen (2012) find support for the use of intraday data for volatility timing and portfolio selection. They also find that the choice of loss function is important to form optimal portfolios.

### 3.3 Methodology

We compare different covariance models in a mean-variance efficient framework with a view to evaluate their relative economic merit. More specifically, we confront the performance of the daily- and intraday-based rolling conditional covariance estimators (Fleming et al., 2001, 2003) to that of MGARCH models with lagged realized volatility (RV) and realized covariance (RCV) as additional regressors. These MGARCH specifications, diagonal BEKK (Engle and Kroner, 1995), scalar CCC (Bollerslev, 1990) and DCC (Engle, 2002), extend the univariate GARCH study of Fuertes et al. (2009) to a multivariate context and an economic (as opposed to purely statistical) loss function. In order to incorporate intraday information in the MGARCH models we augment them by 5-minute sampled realized variances and covariances.

We evaluate the incremental gains of intraday information in covariance forecasting by assessing its impact on the performance of daily dynamic asset allocation strategies.

The evaluation builds on the framework developed in Fleming et al. (2001). Accordingly, we consider an investor who adopts a mean-variance optimization method to invest in a portfolio of stocks. The investor's objective is to minimize volatility while matching a target expected return. The investor treats expected return as constant as it is reasonable to assume that the expected returns do not change on a daily basis. The investor therefore adopts a volatility-timing strategy with daily rebalancing according to the changes in the covariance matrix estimates.

### 3.3.1 Volatility Timing Strategies

Consider an investor with a quadratic utility function who wants to minimize portfolio volatility while matching a specific expected return over a certain time horizon. Treating the expected return as constant, we can construct the optimal weights with one-step-ahead estimates of the conditional covariance matrix.

Let  $r_t \equiv (r_{1t}, r_{2t}, \dots, r_{Nt})'$  denote the day  $t$  return vector on  $N$  risky assets defined as the logarithmic open-to-close price. Define  $\xi_{t-1}$  to be the information set available at the end of day  $t - 1$ , then  $\boldsymbol{\mu}_t \equiv E[r_t \mid \xi_{t-1}]$  and  $\boldsymbol{\Sigma}_t \equiv E[(r_t - \boldsymbol{\mu}_t)(r_t - \boldsymbol{\mu}_t)' \mid \xi_{t-1}]$  are, respectively, the  $N \times 1$  conditional expected return vector and the  $N \times N$  conditional covariance matrix of  $r_t$ . At the open of day  $t$ , the investor formulates the *target return* quadratic optimization problem

$$\begin{aligned} \min_{w_t} \quad & \mathbf{w}_t' \boldsymbol{\Sigma}_t \mathbf{w}_t \\ \text{s.t.} \quad & \mathbf{w}_t' \boldsymbol{\mu}_t + (1 - \mathbf{w}_t' \mathbf{1}) r_f = \mu_p \end{aligned} \tag{3.1}$$

with closed-form solution

$$\mathbf{w}_t = \frac{(\mu_p - r_f) \boldsymbol{\Sigma}_t^{-1} (\boldsymbol{\mu}_t - r_f \mathbf{1})}{(\boldsymbol{\mu}_t - r_f \mathbf{1})' \boldsymbol{\Sigma}_t^{-1} (\boldsymbol{\mu}_t - r_f \mathbf{1})} \tag{3.2}$$

where  $\mathbf{1}$  is a  $N \times 1$  unit vector,  $\mathbf{w}_t \equiv (w_{1t}, w_{2t}, \dots, w_{Nt})'$  is an  $N \times 1$  vector of optimal weights on the risky assets,  $\mu_p$  is the target portfolio return and  $r_f$  denotes the risk-free rate of return. We set the target return between 8% and 16% per annum, similar to the

range assumed in Fleming (2001), and there are no limitations on short-selling. The case where short-selling is constrained is considered as a robustness check.

Expression (3.2) shows that the optimal weights change through time as the expected returns,  $\boldsymbol{\mu}_t$ , and covariance matrix,  $\boldsymbol{\Sigma}_t$ , vary. We estimate  $\boldsymbol{\Sigma}_t$  using one-day-ahead forecasts (denoted  $C_t$ ) in a rolling-window approach which yields a time-series of daily covariance estimates — each one day-ahead-forecast leads to the determination of a daily portfolio weight  $\mathbf{w}_t$ . The covariance estimators applied in the study are those described in Chapter 2.2. In line with previous studies we treat the  $N \times 1$  vector  $\boldsymbol{\mu}_t$  as constant with each element equal to the average return of the corresponding index over the out-of-sample period. This implies that the portfolio weights only depend on the daily covariance forecasts and in this sense we focus on pure volatility-timing strategies.

By evaluating the performance of volatility timing strategies, we can directly measure the economic gain of incorporating additional information (intraday returns) when predicting the daily covariance matrix. Fleming et al. (2003) investigate the merits of high-frequency intraday data when forming mean-variance efficient stock portfolios with daily rebalancing. For this purpose, they compare rolling covariance forecasts based on daily and on realized (intraday) data. We extend the analysis of the merits of additional (intraday return) information to three parametric (MGARCH) covariance forecasting approaches, BEKK, CCC and DCC outlined in Chapter 2. Alongside the dynamic portfolios, we also consider a passive Buy-and-Hold (BH) strategy which assumes a constant covariance matrix estimated unconditionally in-sample.

### 3.3.2 Economic Evaluation Framework

Following Fleming et al. (2003), we assess the performance of a given estimator of the conditional covariance matrix,  $\mathbf{C}_t$ , using a utility-based evaluation. Consider the quadratic utility function of a representative investor who strategically allocates an initial wealth

$W_0$  into the risk-free asset and  $N$  risky assets. The daily utility of the investment is

$$U(r_{p,t}) = W_0 \left[ (1 + r_{p,t}) - \frac{\gamma}{2(1 + \gamma)} (1 + r_{p,t})^2 \right] \quad (3.3)$$

where  $r_{p,t} \equiv (1 - \mathbf{w}'_t \mathbf{1})r_f + \mathbf{w}'_t r_t$  is the ex post portfolio return, and  $\gamma$  is the investor's relative risk aversion.

Thus a utility-based comparison of two conditional variance estimators,  $\mathbf{C}_{a,t}$  and  $\mathbf{C}_{b,t}$ , can be expressed as

$$\sum_{t=1}^T U(r_{p_{a,t}}) = \sum_{t=1}^T U(r_{p_{b,t}} - \Delta) \quad (3.4)$$

where  $r_{p_{a,t}}$  and  $r_{p_{b,t}}$  denote the portfolio returns associated with the two estimators over the investment horizon  $T$ , and  $\Delta$  is the maximum annualized fee in basis points (bp) that the representative investor would be willing to pay each day in order to switch from the daily  $\mathbf{C}_{a,t}$  to the realized  $\mathbf{C}_{b,t}$ . In line with the assumptions made in Fleming et al. (2003), we consider two levels of risk aversion  $\gamma = 1$  and  $\gamma = 10$ .

### 3.3.3 Transaction Costs and Rebalancing Frequency

With active rebalancing, it is important to bring transaction costs into the picture. The relative performance of the various trading strategies will depend considerably on their trade intensity and associated transaction costs. We account for transaction costs as follows. After rebalancing at the start of day  $t$  using the optimal weights from (3.2) based on the covariance forecasts for day  $t$ , the  $i$ th asset has a new weight  $w_{i,t}$  in the portfolio,  $i = 1, \dots, N$ . The portfolio return on day  $t$  is  $r_{p,t} = \sum_i^N w_{i,t} r_{i,t}$ . On the instant just before rebalancing at the start of day  $t$  (denoted  $t^-$ ) the actual weight of the  $i$ th stock in the portfolio is  $w_{i,t^-} = w_{i,t-1} \frac{1+r_{i,t-1}}{1+r_{p,t-1}}$ . The change in weight at the start of day  $t$  (after rebalancing) is given by  $w_{i,t} - w_{i,t^-}$ . Following Martens et al. (2008), we assume that the transaction cost is a fixed percentage  $c$  on each traded dollar. With an initial wealth  $W_0$ , the total \$ value of the transaction costs on day  $t$  is

$$C_t = W_0 \times c \sum_{i=1}^N |w_{i,t} - w_{i,t^-}| \quad (3.5)$$

and the net portfolio return is  $r_{t,p} - c_t$  where  $c_t \equiv C_t/W_0$  denotes the total transaction costs in percentage.<sup>2</sup> We adopt  $c = 20\%$  annualized percentage points, which amounts to a daily transaction cost of 8bp.

Daily rebalancing incurs large turnover and the impact of transaction costs on the performance of the market timing strategies is likely to be substantial so we also consider weekly and monthly rebalancing approaches where the portfolio is rebalanced on the first day of each week and a monthly rebalancing approach where the portfolio is rebalanced on the first day of each month. In addition to fixed rebalancing strategies we introduce a return-driven rebalancing strategy which serves two purposes. First, it introduces time-variation in the frequency of rebalancing throughout the trading period. Second, it reduces the daily turnover and thus transaction costs. Intuitively, it is not necessary to rebalance the portfolio on day  $t$  if the optimal  $N \times 1$  weight vector  $\mathbf{w}_t$  does not differ much from  $\mathbf{w}_{t-1}$ . If the actual weight change from day  $t - 1$  end to day  $t$  end in the portfolio is equal or higher than a cut-off  $\kappa$ , i.e.  $\frac{\sum_{i=1}^N |w_{i,t} - w_{i,t-1}|}{N} \geq \kappa$ , we rebalance the portfolio on day  $t$  using the optimal  $\mathbf{w}_t$  based on the day  $t + 1$  covariance forecasts. We adopt several values for the weight change cut-off,  $\kappa = \{10\%, 15\%, 20\%\}$ . In addition, we also introduce a volume-driven rebalancing strategy that rebalances on day  $t$  if transaction volume of the market (approximated by volume of the S&P500 stock index) exceeds a threshold of the in-sample period of estimation on day  $t - 1$ . We apply two values for the rebalancing threshold,  $S = \{25th, 50th\}$  percentiles, which means we rebalance on day  $t$  if the market volume exceeds  $S$  of the in-sample volume on day  $t - 1$ .

### 3.3.4 Covariance Forecasting

This empirical work applies four different covariance matrix forecasts, a rolling (ROLL) approach, equation (2.27), and three models within the MGARCH family: the scalar CCC

---

<sup>2</sup>The quantity  $\sum_{i=1}^N |w_{i,t} - w_{i,t-1}|$  is the % of the portfolio rebalanced at the end of day  $t$  and hence,  $W_0 \times \sum_{i=1}^N |w_{i,t} - w_{i,t-1}|$  is the \$ value of the portfolio rebalanced. By multiplying the latter by  $c$  (transaction costs per \$ traded) we get the total \$ transaction costs for day  $t$ .

model, equation (2.32), scalar DCC model, equation (2.35), and the diagonal BEKK model, equation (2.30). Alongside the original version of these four covariance matrix models based on daily returns we implement a realized version that incorporates intra-day information in the covariance forecasts through the realized covariance matrix,  $\mathbf{RC}$ , equation (2.12). See Chapter 2 for a detailed discussion of the realized covariance matrix models.

The sample is divided into an estimation period ( $T_0 = T - T_1$ ) of 1759 days, and a holdout window ( $T_1$ ) of 500 days. Each of the forecasting approaches described below is deployed over an initial window, denoted  $[1, T_0]$  and a 1-day-ahead covariance matrix is forecasted. The window is rolled forward,  $[2, T_0 + 1]$ , to obtain a second covariance matrix forecast and so forth over 500 iterations. This parameter updating approach offers a ‘shield’ against structural (breaks) in the covariance process during the out-of-sample period.

### 3.3.5 Statistical Evaluation Methods

Several statistical loss functions have been deployed in the literature to gauge the out-of-sample forecast performance of volatility models. For instance, in a recent volatility forecast competition using high-frequency realized measures, Fuertes et al. (2009) adopt the following measures:

$$\begin{aligned}
 \text{Mean absolute error} \quad MAE &= \frac{1}{T_1} \sum_{t=1}^{T_1} |\tilde{\sigma}_t^2 - h_{t,m}| \\
 \text{Mean squared error} \quad MSE &= \frac{1}{T_1} \sum_{t=1}^{T_1} (\tilde{\sigma}_t^2 - h_{t,m})^2
 \end{aligned} \tag{3.6}$$

where  $\tilde{\sigma}_t^2$  denotes the day  $t$  realized variance and  $h_{t,m}$  denotes the variance forecast from the  $m$ th model. MAE and MSE belong to the family of symmetric loss functions, in the sense that they equally penalize over- and under-predictions. The most widely adopted, MSE, proposed by Bollerslev et al. (1994) is based on a quadratic loss function and so it is particularly good where large forecast errors are disproportionately more worrisome than smaller errors. MAE is less sensitive to severe mispredictions than MSE.

In the present study we extend these volatility forecast accuracy measures in order to compute “overall” versions for the entire covariance matrix  $C_t$ . Accordingly, we aggregate the corresponding losses associated with the different variance and covariance forecasts on each day of the forecasting period. For instance, the MSE of the covariance matrix forecasts from model  $m$  is calculated as  $MSE = \frac{1}{T_1} \sum_{t=1}^{T_1} \left( \sum_{i,j=1}^N (i \leq j) \frac{(\hat{\sigma}_{t,ij}^2 - h_{t,ij}^m)^2}{N(N+1)/2} \right)$  where  $T_1 = 500$  days and  $N$  is the number of portfolio constituents.

### 3.4 The Dataset

A portfolio of three popular financial indices (NASDAQ 100, Russell 2000 equity indexes and the CRB commodity index) is studied in this chapter. The three indices are chosen for the purpose of diversification. The NASDAQ 100 is a large-cap index which incorporates large (non-financial) companies both inside and outside the US. Empirical literature suggests that large cap stocks exhibit lower volatility in returns while small cap companies yield higher profit. The Russell 2000 is therefore selected to diversify our portfolio by including small cap securities. The CRB index diversifies the portfolio further as commodity prices tend to have low correlation with equity prices. The main characteristics of the three indices are outlined in Appendix 3.1.

The data consists of tick-by-tick price quotes spanning the period from 4 January 1999 to 31 December 2007, a total of  $T = 2259$  trading days.<sup>3</sup> Although NASDAQ and NYSE have a pre-market session from 7:00am to 9:30am and a post-market session (in these session the securities are traded between banks and dealers, without stock exchange control) from 4:00pm to 8:00pm, we focus on the normal trading session from 9:30am to 4:00pm (390 minutes) as usually major transactions are done during the normal trading times. We convert the daily prices to returns  $r_t$  by the conventional log-differencing method:

$$r_t = 100 \times \log(p_t/p_{t-1}), \quad t = 1, \dots, T \text{ days}, \quad (3.7)$$

---

<sup>3</sup>The data are obtained from the Disk Trading database. See <http://www.is99.com/disktrading/>.

where  $P_{t-1}$  and  $P_t$  denote, respectively, the opening and closing price on trading day  $t$ .<sup>4</sup> Our proxy for the risk free rate  $r_f$  is 4.5% per annum which is equal to the average 3-month Treasury Bill rate over the in-sample period.

In order to construct intraday returns, the trading day [9:30am-4:00pm] is divided into  $M$  intervals of 5-minute length. The 5-minute sampling interval has been shown to be small enough to accurately capture price dynamics and large enough to dampen down the adverse effects of market microstructure frictions.<sup>5</sup> The price at the start of the  $j$ th intraday interval is computed as the average of the closing and opening prices of intervals  $j - 1$  and  $j$ , respectively. The  $j$ th intraday return on day  $t$  for the  $i$ th asset is therefore computed as

$$r_{t,j} = \left( \frac{\log(p_{t,j}^c) + \log(p_{t,j+1}^o)}{2} - \frac{\log(p_{t,j-1}^c) + \log(p_{t,j}^o)}{2} \right), j = 2, \dots, M - 1 \quad (3.8)$$

where  $p_{t,j}^c$  ( $p_{t,j}^o$ ) is the closing (opening) price of the  $j$ th intraday interval. The extreme-interval returns are  $r_{t,1} = \left( \frac{\log(p_{t,1}^c) + \log(p_{t,2}^o)}{2} - \log(p_{t,1}^o) \right)$  and  $r_{t,M} = \left( \log(p_{t,M}^c) - \frac{\log(p_{t,M-1}^c) + \log(p_{t,M}^o)}{2} \right)$ . The price  $p_{t,M}^c$  is the closing price on day  $t$ , simply denoted  $p_t$  in (3.7), defined as the last price observed before 4:00pm. Likewise,  $p_{t,1}^o$  is the observed opening price on day  $t$ , simply denoted  $p_{t-1}$  in (3.7), defined as the first price recorded after 9:00am. The intraday closing price  $p_{t,j}^c$  is similarly defined as the last seen tick before the  $j$ th 5-min mark; likewise for  $p_{t,j}^o$  with reference to the 5-min mark  $j-1$ . The aggregation of all  $M$  intraday returns gives the daily open-to-close return  $r_t = \sum_{j=1}^M r_{t,j} = \log\left(\frac{p_{t,M}^c}{p_{t,1}^o}\right) = \log\left(\frac{p_t^c}{p_t^o}\right)$ . Typically, we have  $M = 78$  intraday returns with the exception of days with delayed openings and/or early

---

<sup>4</sup>In this study the daily return is defined as open-to-close (instead of close-to-close) because we seek to avoid the inclusion of the overnight return. A practical problem with the inclusion of the overnight return is that they weight it should deserve in realised volatility (covariance) measures is somewhat arbitrary as Hansen and Lunde (2006) and Engle et al. (2006) emphasize. It is well known that the overnight return is more volatile than the intraday 5-min returns, which could introduce extra “noise” into the realised covariance estimators. See also Ahoniemi et al. (2012).

<sup>5</sup>Andersen et al. (2001), Barndoff-Nielsen and Shephard (2002), and Taylor and Xu (1997), inter alios, advocate this grid also because daily returns standardized by 5-min realised volatility are approximately normal. Forecasting studies that use 1-, 5-, 15- and 30-min data report mixed results but overall they also tend to favour the 5-min sampling (Martens and van Dijk, 2006; Pong et al., 2004; Ghysels et al., 2006).

closures of the exchanges. The inter-daily (close-to-close) return can be decomposed as the sum of the overnight return and the daily return, i.e.  $\log(\frac{p_t^c}{p_{t-1}^c}) = \log(\frac{p_t^o}{p_{t-1}^o}) + \log(\frac{p_t^c}{p_t^o})$ . The present study focuses on comparing the effectiveness of different models to forecast the covariance of daily (as opposed to intraday) returns. Choosing the daily return, which excludes the overnight return, as the modelling object of interest, allows us to sidestep the practical problem of having to determine the weight that the overnight return should deserve in realised covariance measures, a somewhat arbitrary choice (Hansen and Lunde, 2006).

Figure 3.1, Panels A to D, plots the daily returns, realized volatilities, dynamic rolling 30-day window correlations and realized covariances. The figures provide prima facie evidence that, although the returns of the indices are essentially iid, the variances, correlations and covariances exhibit significant clustering. From 2000 to 2002, there was a period of high volatility driven by the September 11th terrorist attacks and the burst of the dot-com bubble. In particular, NASDAQ 100 show extreme price movements during the period as it consists of many IT and internet companies. During the period between 2003 and 2007 all three indices exhibit low volatilities. Correlations among portfolio components show clear time-variation. The daily correlation between NASDAQ 100 and Russell 2000 is relatively high averaging 0.79 over the sample period whereas CRB and NASDAQ 100 (or Russell 2000) are not highly correlated.

Table 3.1 reports summary statistics of the return distribution alongside the Jarque-Bera normality test and the Ljung-Box autocorrelation test for daily returns ( $r_t$ ), squared daily returns ( $r_t^2$ ) and the cross-products of daily returns (e.g.  $r_t^{\text{Nasdaq}} \times r_t^{\text{CRB}}$ ). The table corroborates the main stylized facts of the daily returns distribution, namely, large deviations from normality particularly in the form of fat tails, and strong memory in the second moment of the returns distribution.

Over the sample period CRB index yields the highest return (6.8% annually) and lowest variance, backed by the long-lasting commodity boom. In contrast, NASDAQ 100

return stays relatively flat and has a higher volatility. Normality of all the returns are rejected, which is expected since non-normality is a common characteristic of financial asset returns. Panel B of Table 3.1 presents the historical correlation matrix. NASDAQ 100 and Russell 2000 have a correlation of 0.79 percent, not a surprise as both are US stock indices. The inclusion of CRB index in the portfolio is justified by the low correlation between itself and the NASDAQ 100 (0.014) or Russell 2000 (0.072) indices.

## 3.5 Empirical Results

### 3.5.1 Statistical Evaluation of Covariance Forecasts

In this section we evaluate the statistical accuracy of the competing covariance forecasts. In particular, we rank them by the mean loss over the 500-day out-of-sample period using the statistical loss functions, namely the mean square error (MSE) and the mean absolute error (MAE), outlined in Section 3.3. Table 3.2 sets out the results. The covariance forecasts are assessed by two common loss functions: the Mean Square Error (MSE) and the Mean Absolute Error (MAE). 500 daily forecasts of the three indices are generated and the average MSE and MAE of the variance and covariance terms are taken. The in-sample modelling period is 4th January 1999 to 4th January 2006 and the 500-day out-of-sample forecasting period is 5th January 2006 to 31st December 2007.

The results of statistical evaluation in Table 3.2 suggest that three realized models, namely RROLL, RCCC and RDCC, produce similarly superior performances compared with other competing volatility models. Pairwise comparison between each daily estimator and its intraday counterpart reveals similar results: all the realized models yield smaller forecast errors than their daily counterparts, showing that realized models can provide statistically more accurate volatility forecasts by utilizing intraday information. However, in practice, the ranking of covariance forecasting models is useful to investors only if it leads to tangible economic gains. The economic value of the volatility forecasts is evaluated in the section below.

### 3.5.2 Estimation Results of the Conditional Covariance Models

The left columns in Table 3.3 and 3.4 report the resulting daily and intraday based covariance models' parameter estimates, together with robust standard errors in parentheses and p-values in square brackets. In general, the coefficients of the covariance models are significant at 5% significance levels. However, a few insignificant parameters are detected. For instance, the correlation coefficients between CRB and Russell 2000 ( $\rho_{12}$ ), and between CRB and NASDAQ 100 ( $\rho_{13}$ ), are not always significant in the CCC, DCC, RCCC and RDCC estimations. This is because the actual correlations between the commodity index and stock indices were not particularly strong, making the correlation coefficient statistically indifferent from zero. For the RMGARCH models, the estimation results confirm the significance of all the coefficients of the realized matrix at 5% significance level. The results suggest that the realized variance and covariances have explanatory power in the corresponding RMGARCH equations.

### 3.5.3 Economic Evaluation of Covariance Forecasts

In this section the optimal-weight selection problem is solved iteratively using as inputs the daily covariance matrix forecasts  $\hat{C}_t$  over the 500-day out-of-sample period. Thus we compare the competing forecasting approaches by assessing the performance of volatility-timing strategies and the incremental value of switching from one forecasting method to another. Throughout the analysis, we assume a known constant conditional mean vector,  $\boldsymbol{\mu}$ , equals to the long-run or unconditional mean of each index over the in-sample period. We start by assuming daily rebalancing with zero transaction costs ( $c = 0$ ) which are subsequently increased to an annualized percentage points of 20% ( $c = 20\%$ ). For each strategy we report the portfolio turnover (TO), the average short selling and the percentage of rebalancing days over the 500-day trading window. TO is defined as the proportion of the portfolio value that is rebalanced in total over the out-of-sample period which, following the notation of Section 3.3, in annualized terms is given by  $TO =$

$$\frac{\sum_{t=1}^{500} (\sum_{i=1}^N |w_{i,t} - w_{i,t-1}|)}{500} \times 252.$$

The optimal weights over the 500-day out-of-sample period obtained through the daily-returns-based ROLL, BEKK, DCC and CCC forecasts and realized-covariance-based forecasts are plotted in Figure 3.2, Panel A and B, respectively. The graphs suggest smaller weight variations for the daily models and, on this basis, one would expect that the asset allocation strategies based on the realized covariances are more trade-intensive and, in turn, more heavily penalized by transactions costs. This issue will be examined below.

Table 3.5 shows the weights that minimize conditional volatility while setting the expected return equal to 12%. The sample period is 5th January 2006 through 31st December 2007. The portfolio weights are constructed based on the one-step-ahead forecasts of the conditional covariance matrix. As expected, the sign and magnitude of each of the weights depends on the forecasted expected returns and the conditional volatility and correlation estimates. For instance, the weight in NASDAQ 100 is generally negative because the average return on the stock index was negative in the forecast period. The swings in the weights are significant in all the models as a high level of risk exposure is needed to match the target return of 12%. However, the swings are more pronounced in realized covariance models. By incorporating intraday information the realized models appear to be more sensitive compared with their daily based counterparts. The table also presents the implicit weights in cash. A negative cash weight implies that the corresponding position in the underlying asset is levered.

The performance of a target return portfolio with a 12% annual target return is presented in Table 3.6 and Figure 3.3.A<sup>6</sup> (for a detailed numeric display of the model performance see Appendix 3.2). The results show that realized forecasting approaches which exploit intraday information perform consistently better than their daily-return-based counterparts. Figure 3.3.B reports the annualized basis points fee ( $\Delta$ ) that an investor

---

<sup>6</sup>The static portfolio falls far behind all the active rebalancing strategies in terms of the return and volatility generated. The performance of the static portfolio is not displayed in the figures hereafter for the purpose of showing the comparison among the competing covariance forecasts in a larger scale.

with quadratic utility and a relative risk aversion level ( $\gamma$ ) would pay to switch from the covariance models based exclusively on daily data to the realized-augmented models. In a zero transaction costs setting, an investor with low relative risk aversion ( $\gamma = 1$ ) would be willing to pay at least 44.7 basis points per year to switch from the daily ROLL covariance matrix estimate to the RROLL one. An investor with high relative risk aversion ( $\gamma = 10$ ) would be willing to pay even more (142.7bp) to switch from the daily to the realized framework.

What do we learn from these results? First, the active rebalancing strategies are better than the passive buy & hold (BH) one in terms of Sharpe ratio and economic value. An BH strategy produces 12% annualized return and yields a high standard deviation of 33.2%, delivering inferior performance compared with all the competing active strategies. Second, adding realized covariance does improve the performance (in terms of economic gains) of the covariance forecasts. The economic value produced by switching from daily models to corresponding realized specifications are all positive.

Third, there is no economic value in switching from the realized rolling covariance model to more sophisticated MGARCH models when risk aversion is relatively high. For example, an investor with high relative risk aversion ( $\gamma = 10$ ) would be willing to pay 60.34bp to switch from the RBEKK to RROLL. Forth, among the parametric forecasting approaches, the RCCC model gives the best overall result with the highest return and lowest volatility. These findings are robust to the presence of transaction costs. In the context of an annualized 20% daily transaction cost as shown in the right panel of Table 3.6 and Figure 3.3.C, an investor with high relative risk aversion ( $\gamma = 10$ ) would be willing to pay at least 141.51 basis points per year to switch from the daily ROLL covariance matrix to the RROLL counterpart.

By comparing the DCC and RDCC model we find disagreement between the Sharpe ratio and switching fee. The Sharpe ratio of DCC is higher than RDCC while there is still economic value in switching from DCC to RDCC. This is because the quadratic

utility function penalizes risk more than Sharpe ratio when standard deviation of return is relatively high. In this case although the return of DCC model is higher but a risk averse investor will still choose RDCC for risk reduction.

### **3.5.4 The Impact of Rebalancing Frequency**

The analysis thus far is based on daily rebalancing. The performance of the models with a weekly rebalancing strategy is shown in Table 3.7 and Figure 3.4. According to this strategy we rebalance the portfolio every Monday according to the calculated optimal portfolio weights and then keep the weights unchanged until next Monday. The proportion of rebalancing days has reduced significantly from 100% to 20% for the dynamic models and turnover has lowered to around 16% from 22%. Net of transaction costs, the RROLL now occupies the top forecasting approach generating a Sharpe ratio at 0.513. Interestingly the results of economic value are very similar to the daily rebalancing case, where the realized measures dominate the table according to the positive switching fees. The RROLL are now the best under both the low and high risk aversion assumptions.

Table 3.8 and Figure 3.5 present the model performances under target return strategy with monthly rebalancing. We rebalance the portfolio at the beginning of each month according to the optimal portfolio weights and then keep the weights unchanged until next month. The pattern between economic value of the realized versus daily models remains the same as the realized models once again outperform their daily counterparts. The proportion of rebalancing days has reduced further from 20% to 5% for the dynamic models and turnover has lowered to around 13%. Now the difference lies in the inter-comparison between realized models. From daily rebalancing to monthly rebalancing we see a gradual improvement of RROLL model. The model is outperformed by the RDCC and RCCC under a low risk aversion assumption under daily rebalancing strategy. With weekly rebalancing the RROLL provides the best economic value among all the realized models under both risk aversion assumptions. Now with monthly rebalancing the RROLL

not only dominates the economic value but also provides the highest Sharpe ratio.

Table 3.11 and Figure 3.6 show the model performances under a return-driven strategy with a rebalance threshold of 10%<sup>7</sup>, which means that we rebalance the portfolio on day  $t$  only if the overall percentage change in returns exceeds 10% on day  $t - 1$ . The results pattern is similar to the daily rebalancing case. All the realized measures perform better than their counterparts according to the switching fee. Generally the RCCC model has the highest basis points and the RROLL ranks the second. Again the results are similar with or without transaction costs. As the rebalancing threshold increases from 5% to 20%<sup>8</sup> we see a gradual reduction in the proportion of rebalancing days. Cases with alternative rebalancing thresholds at 5% (Table 3.9), 8% (Table 3.10) and 20% (Table 3.12) show similar results.

The results of a volume driven strategy with 25th percentile threshold<sup>9</sup> are presented in Table 3.13 and Figure 3.7. We rebalance on day  $t$  if transaction volume on day  $t - 1$  exceeds the 25th percentile of the in-sample period. S&P 500 volume is used as a representative of the market volume. The model performances are consistent with those adopting other strategies. Realized models outperform their daily counterparts again and the RCCC and RROLL are the top performers under both the low and high risk aversion scenarios according to the switching fee criteria. Table 3.14 presents the results of a volume driven strategy with 50th percentile threshold and shows similar results.

From the results shown in Table 3.6 to 3.14 and Figure 3.3 to 3.7 we can conclude that under time-constant (i.e. daily, weekly and monthly) rebalancing frequency strategy the RROLL model generally produce the highest basis points hence the largest economic

---

<sup>7</sup>The 10% case is presented as a representative for the return-driven strategy. Model performances under target return strategies with the rebalance threshold of 5%, 8%, 10% and 15% are evaluated and shown in Appendix 3.2.

<sup>8</sup>Covariance ranking under a return-driven strategy with a 20% rebalance threshold is presented in Appendix 3.3.1

<sup>9</sup>The 25th percentile threshold case is presented as a representative for the volume-driven strategy. Model performances under target return strategy with a rebalance threshold of 50th percentile is also evaluated and shown in Appendix 3.2. In addition, covariance ranking of the 50th percentile case is presented in Appendix 3.3.2

value; while under time-varying ones (i.e. return-driven and volume-driven) the RCCC shows the best performance from the same economic point of view.

Table 3.18 provides a comparison of the economic value of the different rebalancing strategies after transaction costs deduction. The annualized basis points fee an investor would pay to switch from the daily rebalancing strategy to less frequent ones are presented. As the realized models outperform their daily counterparts throughout the empirical study, we compare the realized models only. For all the models, generally a less frequent rebalancing strategy performs better as in 60 out of 72 cases lower frequency rebalancing strategies outperform their daily counterparts. For example, the results for RROLL and RDCC show clearly that almost all the listed frequencies outperform their daily counterparts. In addition, the results for RBEKK and RCCC suggest that the time-varying rebalancing strategies are economically better than the daily rebalancing one.

The return-driven strategy with a 20% threshold works well for the RROLL and RBEKK while a 5% threshold gives relatively good results for the RCCC and RDCC model. If there is an overall best rebalancing strategy it is the return-driven with 10% to 20% threshold. Under this strategy we have positive basis points in all cases across the four realized models under two different risk aversion levels. This result indicates that more economic benefit can be extracted from following an appropriate time-varying rebalancing strategy rather than rebalancing the portfolio on a daily basis.

### **3.5.5 Robustness Checks**

Back to the daily rebalancing we do robustness checks by setting a range of target returns from 8% (Table 3.16) to 16% (Table 3.15). In line with Fleming et al. (2001) we find that the pattern of model performances remains largely the same across the different target returns. With higher target return the standard deviation is increased as more short selling transactions are needed to meet the target return. As a result of high risk and

greater difference in the standard deviations produced by competing modes, the switching fees in a high target return case are generally larger than those in a low target return one. As an example, the switching fees an investor is willing to switch from ROLL to RROLL are 76.4 and 198.5 under low and high risk aversion assumptions respectively for a target return of 16%, while the fees are significantly lower, 20.2 and 87.5 correspondingly, for a target return of 8%. The high switching fees generated in the 16% target return case show clearly that with a high target return the realized models are better suited than their daily counterparts economically.

We carry the realized and daily performance comparison further by imposing a no short-selling constraint. Table 3.17 shows the performance of the models. Although the target return of 12% does not change, we can see that the returns are reduced significantly as a result of the constraint. However, the pattern that realized models perform better than their daily counterparts remains unchanged. All the realized models yield positive switching fees, signalling relatively high economic value.

Strong performances of the realized models in the cases of different target return rates and no short-selling constraint further confirm the economic value of intraday covariance estimators. In short, realized models perform better than their daily counterparts and a risk averse investor would be willing to pay to switch from the daily models to corresponding realized ones.

## 3.6 Conclusion

This chapter presents an empirical study to compare different daily and intraday covariance models in a mean-variance efficient framework with a view to evaluate their relative economic value. By "economic value" we refer to the maximum return that an investor with quadratic utility function would be willing to sacrifice each day in order to capture the performance gains associated with the augmented covariance estimators. The economic loss function chosen to compare their performance is the quadratic utility function

implicit in mean-variance asset allocation. A dynamic optimal-weight portfolio strategy is deployed, which is based on one-day-ahead covariance matrix forecasts from a model-free ROLL approach and three competing MGARCH formulations (BEKK, CCC, DCC). In order to exploit intraday information, we extend each of these estimators with realized covariances into RMGARCH models (RBEKK, RCCC, RDCC).

The study contributes to the literature in three directions. First, we evaluate the statistical vis-à-vis economic value of augmenting daily historical covariance models with additional information. The main question of interest is whether the statistical advantage (if any) translates into improved performance in a volatility-timing framework. Second, we address the research question that whether more sophisticated intraday models, the RMGARCH models, can outperform the naive intraday based rolling approach. Although a number of studies have provided evidence in favour of MGARCH specifications, the literature that makes use of intraday information for volatility timing is based only on naive rolling window covariance estimators. This chapter seeks to complement the literature in this respect. Third, we focus on the issue of rebalancing frequency by comparing several rebalancing strategies where the frequency of rebalancing is either time-fixed or time-varying. In the latter case rebalancing is triggered by changes in asset returns beyond a threshold level and by market conditions.

A general result across all forecasting approaches, rolling estimators and MGARCH models, is that the covariance forecasts that exploit intraday information provide consistent superior performance to their daily counterparts which is robust to transactions costs and the presence of short-selling constraints. In addition, all the active rebalancing strategies perform better than the passive buy & hold one in terms of Sharpe ratio and economic value, a result that is in line with Fleming et al. (2001, 2003). We also find that statistical accuracy of covariance matrix forecasts is not tantamount to forecasting ability for market timing purposes.

Under time-constant rebalancing frequency strategies (i.e. daily, weekly and monthly

ones) we find strong economic evidence in support of the RROLL estimator based on the realized covariance matrix. While under time-varying rebalancing strategies (i.e. return-driven and volume-driven) the RCCC stands out in the same economic point of view. The highest Sharpe ratio is generated by the RCCC forecasts with a volume-driven rebalancing strategy with a 50th percentile threshold.

Regarding to the rebalancing frequency, we find that more economic value can be generated by following an optimal flexible rebalancing strategy rather than rebalancing the portfolio on a daily basis. The return-driven approach with 10-20% threshold is generally the best rebalancing strategy as it consistently yields positive basis points (the annualized fee for switching from daily rebalancing strategy to the return-driven one) for all cases across the four realized models under both high and low risk aversion levels.

TABLE 3.1  
Summary Statistics of Daily Index Returns

A. Individual Statistics			
	NASDAQ 100	Russell 2000	CRB
Mean	0.004	0.026	0.027
Median	0.094	0.095	0.026
Maximum	16.71	5.79	3.74
Minimum	-11.64	-7.54	-3.46
StDev	2.18	1.30	0.76
Skewness	0.22	-0.14	-0.11
Kurtosis	7.22	4.12	3.85
Jarque-Bera	1696***	126***	72***
LB Statistic			
$r_t$ (daily return)	20.66**	12.99**	2.45
$r_t^2$ (Squared return)	549.54***	483.60***	63.916***
$r_t^{NASDAQ,CRB}$		14.03**	
$r_t^{Russell,NASDAQ}$		467.96**	
$r_t^{Russell,CRB}$		5.27	
B. Correlation Matrix			
	NASDAQ 100	Russell 2000	CRB
NASDAQ	1	0.794	0.013
Russell 2000	0.794	1	0.072
CRB	0.014	0.072	1

Notes: The table provides summary statistics for the returns of the NASDAQ 100, Russell 2000 and CRB indices. The sample period is 4th January 1999 to 31st December 2000. LB denotes the Ljung-Box test statistic for the null hypothesis of no autocorrelation in daily returns up to 5 days. \*, \*\* and \*\*\* denote significance at the 10%, 5% and 1% level respectively. LB(squared returns) gives an indication of volatility persistence.  $r_t^{NASDAQ,CRB}$  is the LB test applied to the cross product of daily returns and gives an indication of autocorrelation in covariances.

TABLE 3.2  
Statistical Evaluation of Covariance Estimators

A. Mean Square Error							
	C11	C22	C33	C12	C13	C23	Overall
Static	1.907	0.127	22.726	0.003	6.584	0.000	5.225
ROLL	1.194	0.352	0.790	0.059	1.293	0.031	0.620
RROLL	1.076	0.118	0.775	0.000	1.310	0.000	0.547
BEKK	1.151	0.271	0.809	0.015	1.714	0.002	0.660
RBEKK	1.168	0.293	0.783	0.016	1.381	0.006	0.608
DCC	1.159	0.291	0.802	0.015	1.286	0.004	0.593
RDCC	1.087	0.266	0.779	0.001	1.131	0.000	0.544
CCC	1.159	0.291	0.802	0.004	1.278	0.000	0.589
RCCC	1.091	0.249	0.772	0.003	1.153	0.000	0.544

B. Mean Absolute Error							
	C11	C22	C33	C12	C13	C23	Overall
Static	1.128	0.234	4.677	0.053	2.566	0.021	1.446
ROLL	0.686	0.507	0.544	0.174	1.005	0.143	0.510
RROLL	0.674	0.451	0.549	0.011	1.134	0.001	0.470
BEKK	0.742	0.456	0.605	0.099	1.168	0.035	0.518
RBEKK	0.718	0.472	0.566	0.094	1.082	0.063	0.499
DCC	0.733	0.475	0.554	0.098	1.047	0.052	0.493
RDCC	0.716	0.454	0.551	0.079	1.031	0.014	0.474
CCC	0.733	0.475	0.554	0.059	1.044	0.006	0.479
RCCC	0.719	0.452	0.546	0.064	1.053	0.006	0.473

Notes: The table provides statistic evaluation of the covariance forecasting models. C11, C22 and C33 represent the forecasted variances of Russell 2000, CRB and NASDAQ 100 respectively. C12, C13 and C23 represent the forecasted covariances of the three indices. The top panel of the table reports the estimated loss associated to each covariance forecasting model using Mean Square Error. The bottom panel shows the estimated loss associated to each model using Mean Absolute Error. The models with a prefix of "R-" represent realized models that incorporated intra-day information; other models are based on daily data. The in-sample modeling period is 4th January 1999 to 4th January 2006 and the 500-day out-of-sample forecasting period is 5th January 2006 to 31st December 2007.

TABLE 3.3  
Estimation Results of Daily Based Covariance Estimators

	BEKK			CCC			DCC		
$c_1$	0.156	(0.021)	[0.000]	0.035	(0.011)	[0.002]	0.035	(0.011)	[0.002]
$\alpha_1$	0.246	(0.019)	[0.000]	0.071	(0.011)	[0.000]	0.071	(0.011)	[0.000]
$\beta_1$	0.961	(0.006)	[0.000]	0.905	(0.015)	[0.000]	0.905	(0.015)	[0.000]
$c_2$	0.032	(0.017)	[0.052]	0.004	(0.003)	[0.159]	0.004	(0.003)	[0.159]
$\alpha_2$	0.202	(0.017)	[0.000]	0.040	(0.011)	[0.000]	0.040	(0.011)	[0.000]
$\beta_2$	0.977	(0.004)	[0.000]	0.959	(0.011)	[0.000]	0.959	(0.011)	[0.000]
$c_3$	0.021	(0.015)	[0.163]	0.002	(0.002)	[0.382]	0.002	(0.002)	[0.382]
$\alpha_3$	0.093	(0.012)	[0.000]	0.023	(0.014)	[0.099]	0.023	(0.014)	[0.099]
$\beta_3$	0.995	(0.001)	[0.000]	0.973	(0.018)	[0.000]	0.973	(0.018)	[0.000]
$\rho_{12}$		-	-	0.057	(0.022)	[0.011]	0.055	(0.029)	[0.058]
$\rho_{13}$		-	-	-0.012	(0.021)	[0.567]	-0.014	(0.027)	[0.611]
$\rho_{23}$		-	-	0.774	(0.008)	[0.000]	0.779	(0.009)	[0.000]
$a$		-	-		-	-	0.016	(0.005)	[0.002]
$b$		-	-		-	-	0.931	(0.024)	[0.000]
Log $L$	-8962.222			-8892.272			-8882.360		
Obs	1758			1758			1758		

Notes: The table provides estimation results of the conventional daily data based covariance forecasting models.  $c_1$ ,  $c_2$  and  $c_3$  represent the diagonal elements of the intercept matrix in the covariance equation for CRB, Russell 2000 and NASDAQ 100 respectively.  $\alpha_1$ ,  $\alpha_2$  and  $\alpha_3$  represent the coefficients of the previous squared errors and cross-product of errors.  $\beta_1$ ,  $\beta_2$  and  $\beta_3$  represent the coefficients of the previous conditional variances and covariances.  $\rho_{12}$ ,  $\rho_{13}$  and  $\rho_{23}$  represent the estimated correlation coefficients. The left column reports the resulting parameter estimates for BEKK, CCC and DCC, together with robust standard errors in parentheses and p-values in square brackets. The modeling period is 4th January 1999 to 4th January 2006.

TABLE 3.4  
Estimation Results of Intraday based Covariance Estimators

	RBEKK			RCCC			RDCC		
$c_1$	0.129	(0.032)	[0.000]	0.013	(0.011)	[0.238]	0.013	(0.011)	[0.238]
$\alpha_1$	0.003	(0.001)	[0.015]	0.001	(0.031)	[0.965]	0.001	(0.031)	[0.965]
$\gamma_1$	0.423	(0.076)	[0.000]	0.157	(0.117)	[0.180]	0.157	(0.117)	[0.180]
$\beta_1$	-0.928	(0.028)	[0.000]	0.879	(0.066)	[0.000]	0.879	(0.066)	[0.000]
$c_2$	-0.001	(0.000)	[0.001]	0.037	(0.016)	[0.021]	0.037	(0.016)	[0.021]
$\alpha_2$	0.238	(0.017)	[0.000]	0.056	(0.010)	[0.000]	0.056	(0.010)	[0.000]
$\gamma_2$	-0.228	(0.042)	[0.000]	0.116	(0.056)	[0.038]	0.116	(0.056)	[0.038]
$\beta_2$	0.946	(0.011)	[0.000]	0.863	(0.038)	[0.000]	0.863	(0.038)	[0.000]
$c_3$	0.000	(0.000)	[0.003]	0.000	(0.000)	[0.057]	0.000	(0.000)	[0.057]
$\alpha_3$	0.211	(0.016)	[0.000]	0.000	(0.000)	[0.000]	0.000	(0.000)	[0.000]
$\gamma_3$	-0.183	(0.037)	[0.000]	0.161	(0.055)	[0.003]	0.161	(0.055)	[0.003]
$\beta_3$	0.956	(0.010)	[0.000]	0.872	(0.042)	[0.000]	0.872	(0.042)	[0.000]
$\rho_{12}$		-	-	0.058	(0.029)	[0.045]	0.052	(0.071)	[0.465]
$\rho_{13}$		-	-	-0.014	(0.047)	[0.772]	-0.023	(0.029)	[0.424]
$\rho_{23}$		-	-	0.770	(0.023)	[0.000]	0.774	(0.027)	[0.000]
$a$		-	-		-	-	0.012	(0.005)	[0.027]
$b$		-	-		-	-	0.915	(0.032)	[0.000]
Log $L$	-8852.815			-8844.324			-8856.448		
Obs	1758			1758			1758		

Notes: The table provides estimation results of the intraday data based covariance forecasting models.  $c_1$ ,  $c_2$  and  $c_3$  represent the diagonal elements of the intercept matrix in the covariance equation for Russell 2000, CRB and NASDAQ 100 respectively.  $\alpha_1$ ,  $\alpha_2$  and  $\alpha_3$  represent the coefficients of the previous squared errors and cross-product of errors.  $\beta_1$ ,  $\beta_2$  and  $\beta_3$  represent the coefficients of the previous conditional variances and covariances.  $\gamma_1$ ,  $\gamma_2$  and  $\gamma_3$  represent the coefficients of the realized covariance matrix. The left column reports the resulting parameter estimates for RBEKK, RCCC and RDCC, together with robust standard errors in parentheses and p-values in square brackets. The modeling period is 4th January 1999 to 4th January 2006.



TABLE 3.6  
Covariance Ranking · 12% Target Return, Daily Rebalancing

	<i>No transaction cost</i>					<i>With transaction cost</i>						
	Return	StDev	Sharpe	Switching fee		Return	StDev	Sharpe	Switching fee		TO	Rebalan. Days
				$\gamma=1$	$\gamma=10$				$\gamma=1$	$\gamma=10$		
Static	12	33.2333	0.3611			12	33.2333	0.3611			0%	0%
ROLL	15.1189	29.066	0.5202			13.2392	29.066	0.4555			23%	100%
RROLL	14.4917	27.4159	0.5286	44.6540	142.7182	12.9919	27.4159	0.4739	44.7950	141.5113	19%	100%
BEKK	13.8927	29.1632	0.4764			12.043	29.1632	0.413			22%	100%
RBEKK	15.0299	28.8983	0.5201	9.2465	53.4762	13.3158	28.8983	0.4608	9.3450	52.1004	20%	100%
				<i>-38.0405</i>	<i>-60.3435</i>				<i>-38.1545</i>	<i>-60.3898</i>		
DCC	15.0146	28.5317	0.5262			13.4498	28.5317	0.4714			19%	100%
RDCC	13.9065	27.2926	0.5095	31.3707	120.1193	12.8065	27.2926	0.4692	30.2753	118.1861	33%	100%
				<i>0.4344</i>	<i>2.0924</i>				<i>-0.5144</i>	<i>1.1994</i>		
CCC	14.1591	28.4474	0.4977			12.7195	28.4474	0.4471			17%	100%
RCCC	15.0388	27.1163	0.5546	37.7453	126.6908	13.02	27.1163	0.4802	36.4750	124.6168	26%	100%
				<i>7.5642</i>	<i>10.5038</i>				<i>6.9455</i>	<i>9.8460</i>		

Notes: The table summarizes the model performances under a target return strategy. Each day a new set of portfolio weights is obtained by solving a portfolio optimization problem in which the expected return for each index equals its in-sample mean return and the conditional covariance matrix is estimated out of sample using different covariance forecasting models. Each day, we calculate the optimal portfolio weights that minimize conditional volatility subject to a target return of 12%. Except the Static model, which keeps its first day optimal portfolio weights with no rebalancing throughout the forecasting period, all other 8 models adopts a active volatility timing strategy that rebalances portfolio weights daily. The left panel reports the model performance in terms of return, standard deviation, Sharpe ratio and the switching fee. The switching fee is an annualized basis points fee ( $\Delta\gamma$ ) an investor with quadratic utility and relative risk aversion of  $\gamma$  would pay to switch from a daily returns covariance matrix forecast to its corresponding realized one. We use  $\gamma=1$  and  $\gamma=10$  to represent investors with low and high risk aversion respectively. The middle panel report the results with transaction costs which is 20% annualized percentage points. The right panel provides additional information on annualized turnover and the percentage of days of rebalancing. Switching fees in italics represent basis points that an investor would be willing to pay to switch from RROLL to another realized model.

TABLE 3.7  
Covariance Ranking · 12% Target Return, Weekly Rebalancing

	No transaction cost					With transaction cost						
	Return	StDev	Sharpe	Switching fee		Return	StDev	Sharpe	Switching fee		TO	Rebalan. Days
				$\gamma=1$	$\gamma=10$				$\gamma=1$	$\gamma=10$		
Static	12.0000	33.2333	0.3611			12.0000	33.2333	0.3611			0%	0%
ROLL	15.2625	29.1419	0.5237			14.4588	29.1419	0.4961			17%	20%
RROLL	14.0602	27.3908	0.5133	46.5650	146.9651	13.3484	27.3908	0.4873	46.5043	146.2733	17%	20%
BEKK	13.7178	29.2384	0.4692			13.1204	29.2384	0.4487			14%	20%
RBEKK	15.0258	29.0417	0.5174	7.3374	44.9778	14.4417	29.0417	0.4973	7.3500	44.5829	14%	20%
				-42.3650	-68.3104				-42.1279	-68.0294		
DCC	15.1390	28.8131	0.5254			14.5801	28.8131	0.5060			14%	20%
RDCC	13.1396	27.9567	0.4700	22.5315	98.9529	12.5504	27.9567	0.4489	22.4786	98.4857	16%	20%
				-14.7208	-21.2475				-14.4872	-20.9663		
CCC	13.9835	28.7647	0.4861			13.4695	28.7647	0.4683			13%	20%
RCCC	13.6263	27.9610	0.4873	22.6659	96.2026	13.0602	27.9610	0.4671	22.5549	95.6583	16%	20%
				-14.3378	-20.9182				-14.0839	-20.6179		

Notes: The table presents the model performances under target return strategy with weekly rebalancing. We rebalance the portfolio at the beginning of the week according to the calculated optimal portfolio weights and then keep the weights unchanged until the following week. See note to table 3.6.

TABLE 3.8  
Covariance Ranking · 12% Target Return, Monthly Rebalancing

	No transaction cost					With transaction cost						
	Return	StDev	Sharpe	Switching fee		Return	StDev	Sharpe	Switching fee		TO	Rebalan. Days
				$\gamma=1$	$\gamma=10$				$\gamma=1$	$\gamma=10$		
Static	12.0000	33.2333	0.3611			12.0000	33.2333	0.3611			0%	0%
ROLL	16.5640	29.4036	0.5633			16.1388	29.4036	0.5489			14%	5%
RROLL	14.9355	27.3532	0.5460	53.9863	161.0680	14.5352	27.3532	0.5314	53.8421	160.5401	14%	5%
BEKK	14.5776	29.4333	0.4953			14.2710	29.4333	0.4849			12%	5%
RBEKK	15.7424	29.2631	0.5380	6.4437	41.9942	15.4482	29.2631	0.5279	6.5075	42.0373	12%	5%
				-50.0696	-82.4761				-49.8471	-82.1900		
DCC	14.0448	28.9411	0.4853			13.7653	28.9411	0.4756			12%	5%
RDCC	11.0019	27.9656	0.3934	24.6554	104.4340	10.7316	27.9656	0.3837	24.6606	104.2465	13%	5%
				-18.9475	-26.0240				-18.6886	-25.7069		
CCC	13.0638	28.7185	0.4549			12.8076	28.7185	0.4460			12%	5%
RCCC	11.1110	27.9107	0.3981	21.0177	93.8545	10.8440	27.9107	0.3885	21.0064	93.6700	13%	5%
				-17.4538	-23.8175				-17.2012	-23.5117		

Notes: The table presents the model performances under target return strategy with weekly rebalancing. We rebalance the portfolio at the beginning of the week according to the calculated optimal portfolio weights and then keep the weights unchanged until the following month. See note to table 3.6.

TABLE 3.9  
Covariance Ranking · 12% Target Return, Return-Driven Rebalancing with a 5% Threshold

	No transaction cost					With transaction cost						
	Return	StDev	Sharpe	Switching fee		Return	StDev	Sharpe	Switching fee		TO	Rebalan. Days
				$\gamma=1$	$\gamma=10$				$\gamma=1$	$\gamma=10$		
Static	12.0000	33.2333	0.3611			12.0000	33.2333	0.3611			0%	0%
ROLL	15.3192	29.0817	0.5268			13.9553	29.0817	0.4799			21%	47%
RROLL	14.5130	27.4205	0.5293	44.6921	143.2771	13.4810	27.4205	0.4916	45.0209	142.7352	18%	44%
BEKK	13.9058	29.1630	0.4768			12.6936	29.1630	0.4353			20%	39%
RBEKK	15.2270	28.9128	0.5267	8.9326	51.7525	14.1480	28.9128	0.4893	9.0892	51.1493	18%	36%
				-38.1689	-60.7485				-38.3005	-60.9525		
DCC	14.9808	28.5569	0.5246			13.9990	28.5569	0.4902			17%	42%
RDCC	13.5902	27.2713	0.4983	31.9309	121.3004	12.2000	27.2713	0.4474	30.7707	119.5057	31%	62%
				0.4618	2.0719				-0.5858	1.0561		
CCC	14.2664	28.4471	0.5015			13.5223	28.4471	0.4754			15%	30%
RCCC	15.1742	27.1368	0.5592	36.8064	125.9039	13.5814	27.1368	0.5005	35.7978	124.2471	24%	58%
				7.3267	10.1284				6.6289	9.3761		

Notes: The table presents the model performances under target return strategy with return-driven rebalancing. We rebalance the portfolio on day t if the overall percentage change in returns exceeds 5% on day t-1. See note to table 3.6.

TABLE 3.10  
Covariance Ranking · 12% Target Return, Return-Driven Rebalancing with a 8% Threshold

	No transaction cost						With transaction cost					
	Return	StDev	Sharpe	Switching fee		Return	StDev	Sharpe	Switching fee		TO	Rebalan. Days
				$\gamma=1$	$\gamma=10$				$\gamma=1$	$\gamma=10$		
Static	12.0000	33.2333	0.3611			12.0000	33.2333	0.3611			0%	0%
ROLL	15.3535	29.1752	0.5263			14.2925	29.1752	0.4899			19%	28%
RROLL	15.4561	27.4402	0.5633	47.5632	147.7752	14.6987	27.4402	0.5357	47.8656	147.4046	17%	22%
BEKK	15.5626	29.0538	0.5356			14.6035	29.0538	0.5026			18%	23%
RBEKK	15.2851	28.9389	0.5282	3.2218	32.0355	14.4281	28.9389	0.4986	3.2060	30.8111	17%	22%
				-39.3597	-62.3367				-39.4890	-62.4880		
DCC	15.2465	28.5037	0.5349			14.5918	28.5037	0.5119			15%	21%
RDCC	13.9097	27.2084	0.5112	29.3146	116.3517	12.9301	27.2084	0.4752	28.0352	114.4533	29%	32%
				-0.9289	0.4789				-2.1020	-0.7646		
CCC	14.3530	28.3899	0.5056			13.8606	28.3899	0.4882			13%	15%
RCCC	15.7393	27.1547	0.5796	35.3619	122.5257	14.5709	27.1547	0.5366	34.3405	120.8126	22%	29%
				6.9895	9.8226				6.2551	8.9563		

Notes: The table presents the model performances under target return strategy with return-driven rebalancing. We rebalance the portfolio on day t if the overall percentage change in returns exceeds 8% on day t-1. See note to table 3.6.

TABLE 3.11  
Covariance Ranking · 12% Target Return, Return-Driven Rebalancing with a 10% Threshold

	No transaction cost					With transaction cost						
	Return	StDev	Sharpe	Switching fee		Return	StDev	Sharpe	Switching fee		TO	Rebalan. Days
				$\gamma=1$	$\gamma=10$				$\gamma=1$	$\gamma=10$		
Static	12.0000	33.2333	0.3611			12.0000	33.2333	0.3611			0%	0%
ROLL	14.5817	29.2299	0.4989			13.6459	29.2299	0.4668			19%	21%
RROLL	15.2501	27.4928	0.5547	48.1607	147.9089	14.6106	27.4928	0.5314	48.4175	147.5448	16%	16%
BEKK	15.2672	29.1051	0.5246			14.4368	29.1051	0.4960			17%	17%
RBEKK	15.3939	28.8589	0.5334	7.5779	50.8744	14.6441	28.8589	0.5074	7.6463	50.3172	16%	16%
				-35.5131	-56.1571				-35.5766	-56.1909		
DCC	15.8186	28.5689	0.5537			15.3009	28.5689	0.5356			14%	13%
RDCC	14.0384	27.2981	0.5143	29.1973	116.7622	13.0192	27.2981	0.4769	28.1168	115.2520	27%	21%
				-0.6750	0.6374				-1.6453	-0.3710		
CCC	14.4845	28.3793	0.5104			14.1041	28.3793	0.4970			12%	10%
RCCC	15.6794	27.1386	0.5778	35.2997	122.6868	14.7028	27.1386	0.5418	34.4420	121.3014	20%	20%
				8.7572	12.2941				8.1061	11.5171		

Notes: The table presents the model performances under target return strategy with return-driven rebalancing. We rebalance the portfolio on day t if the overall percentage change in returns exceeds 10% on day t-1. See note to table 3.6.

TABLE 3.12  
Covariance Ranking · 12% Target Return, Return-Driven Rebalancing with a 20% Threshold

	No transaction cost						With transaction cost					
	Return	StDev	Sharpe	Switching fee		Return	StDev	Sharpe	Switching fee		TO	Rebalan. Days
				$\gamma=1$	$\gamma=10$				$\gamma=1$	$\gamma=10$		
Static	12.0000	33.2333	0.3611			12.0000	33.2333	0.3611			0%	0%
ROLL	17.8490	29.8267	0.5984			17.3344	29.8267	0.5812			16%	6%
RROLL	14.1863	27.6536	0.5130	55.8844	166.4286	13.9163	27.6536	0.5032	56.2350	166.5287	13%	3%
BEKK	14.0804	29.0210	0.4852			13.6045	29.0210	0.4688			14%	6%
RBEKK	16.0170	28.6744	0.5586	12.3354	62.7870	15.5733	28.6744	0.5431	12.4832	62.8859	14%	6%
				<i>-24.7350</i>	<i>-39.8172</i>				<i>-24.8771</i>	<i>-39.9271</i>		
DCC	14.4575	28.2368	0.5120			14.2546	28.2368	0.5048			11%	2%
RDCC	13.2213	27.1791	0.4865	22.8442	104.6326	12.6011	27.1791	0.4636	22.1241	103.8163	22%	5%
				<i>5.0873</i>	<i>9.9452</i>				<i>4.4821</i>	<i>9.3706</i>		
CCC	13.4398	28.3476	0.4741			13.2587	28.3476	0.4677			11%	2%
RCCC	13.2227	27.1168	0.4876	33.4566	119.6990	12.7329	27.1168	0.4696	32.9216	118.8033	16%	5%
				<i>11.6951</i>	<i>17.1886</i>				<i>11.1879</i>	<i>16.5626</i>		

Notes: The table presents the model performances under target return strategy with return-driven rebalancing. We rebalance the portfolio on day t if the overall percentage change in returns exceeds 20% on day t-1. See note to table 3.6.

TABLE 3.13  
Covariance Ranking · 12% Target Return, Volume-Driven Rebalancing with 25th Percentile

	No transaction cost					With transaction cost						
	Return	StDev	Sharpe	Switching fee		Return	StDev	Sharpe	Switching fee		TO	Rebalan. Days
				$\gamma=1$	$\gamma=10$				$\gamma=1$	$\gamma=10$		
Static	12.0000	33.2333	0.3611			12.0000	33.2333	0.3611			0%	0%
ROLL	15.0360	29.0563	0.5175			13.3267	29.0563	0.4587			22%	83%
RROLL	14.0392	27.4378	0.5117	53.0881	191.0566	12.6452	27.4378	0.4609	51.7310	189.7150	19%	83%
BEKK	13.8028	29.1537	0.4735			12.1262	29.1537	0.4159			21%	83%
RBEKK	14.6918	28.9000	0.5084	15.6504	131.5145	13.1874	28.9000	0.4563	14.2798	130.2482	19%	83%
				<i>-37.4377</i>	<i>-59.5422</i>				<i>-37.4512</i>	<i>-59.4668</i>		
DCC	15.1735	28.5398	0.5317			13.7412	28.5398	0.4815			18%	83%
RDCC	13.5810	27.2801	0.4978	53.9404	193.8074	12.0039	27.2801	0.4400	51.6422	191.5727	32%	83%
				<i>0.8522</i>	<i>2.7507</i>				<i>-0.0888</i>	<i>1.8577</i>		
CCC	14.3191	28.4536	0.5032			12.9982	28.4536	0.4568			17%	83%
RCCC	14.3965	27.1199	0.5308	60.8938	201.9942	12.5085	27.1199	0.4612	58.9364	200.0089	25%	83%
				<i>7.8057</i>	<i>10.9376</i>				<i>7.2055</i>	<i>10.2939</i>		

Notes: The table presents the model performances under target return strategy with volume-driven rebalancing. We rebalance on the day t if transaction volume on day t-1 exceeds 25<sup>th</sup> percentile of the in-sample period. S&P500 volume is used as it is more representative of the market the three indices we have. See note to table 3.6.

TABLE 3.14  
A.3.2.9 Covariance Ranking · 12% Target Return, Volume-Driven Rebalancing with 50th Percentile

	No transaction cost					With transaction cost						
	Return	StDev	Sharpe	Switching fee		Return	StDev	Sharpe	Switching fee		TO	Rebalan. Days
				$\gamma = 1$	$\gamma = 10$				$\gamma = 1$	$\gamma = 10$		
Static	12.0000	33.2333	0.3611			12.0000	33.2333	0.3611			0%	0%
ROLL	15.7627	29.0942	0.5418			14.5335	29.0942	0.4995			20%	50%
RROLL	15.5948	27.4475	0.5682	45.0951	143.7685	14.5952	27.4475	0.5318	45.1651	142.9180	17%	50%
BEKK	13.9221	29.2002	0.4768			12.8602	29.2002	0.4404			17%	50%
RBEKK	15.1528	28.9465	0.5235	8.9492	52.0978	14.1509	28.9465	0.4889	9.0550	51.6095	16%	50%
				<i>-39.7064</i>	<i>-62.8045</i>				<i>-39.6796</i>	<i>-62.8193</i>		
DCC	15.0907	28.6268	0.5272			14.1163	28.6268	0.4931			16%	50%
RDCC	14.1175	27.5968	0.5116	24.1622	103.9231	13.2919	27.5968	0.4816	23.3476	102.5630	29%	50%
				<i>-8.7694</i>	<i>-11.1850</i>				<i>-9.4845</i>	<i>-11.8272</i>		
CCC	13.8901	28.5527	0.4865			13.0111	28.5527	0.4557			15%	50%
RCCC	18.5757	27.2433	0.6818	40.7682	129.8060	17.2538	27.2433	0.6333	40.0852	128.3085	22%	50%
				<i>7.8939</i>	<i>9.9955</i>				<i>7.4766</i>	<i>9.5432</i>		

Notes: The table presents the model performances under target return strategy with volume-driven rebalancing. We rebalance on the day t if transaction volume on day t-1 exceeds the average of the in-sample period. S&P500 volume is used as it is more representative of the market the three indices we have. See note to table 3.6.

TABLE 3.15  
Covariance Ranking · 16% Target Return, Daily Rebalancing

	No transaction cost						With transaction cost					
	Return	StDev	Sharpe	Switching fee		Return	StDev	Sharpe	Switching fee		TO	Rebalan. Days
				$\gamma=1$	$\gamma=10$				$\gamma=1$	$\gamma=10$		
Static	16.0000	44.3529	0.3607			16.0000	44.3529	0.3607			0%	0%
ROLL	20.1625	38.7912	0.5198			17.4441	38.7912	0.4497			33%	100%
RROLL	19.3254	36.5890	0.5282	76.3752	198.5478	17.1264	36.5890	0.4681	76.3444	196.6214	28%	100%
BEKK	18.5260	38.9210	0.4760			15.8218	38.9210	0.4065			32%	100%
RBEKK	20.0436	38.5674	0.5197	15.9136	78.3554	17.5353	38.5674	0.4547	16.0755	76.4688	30%	100%
				<i>-67.7171</i>	<i>-181.0207</i>				<i>0.4088</i>	<i>0.4088</i>		
DCC	20.0232	38.0781	0.5258			17.7240	38.0781	0.4655			28%	100%
RDCC	18.8931	36.4012	0.5190	55.6749	168.5080	16.0911	36.4012	0.4420	54.0616	165.2739	46%	100%
				<i>1.9290</i>	<i>6.2500</i>				<i>-6.0360</i>	<i>-6.0360</i>		
CCC	18.8816	37.9656	0.4973			16.7467	37.9656	0.4411			26%	100%
RCCC	20.0556	36.1892	0.5542	63.9377	177.1802	17.1796	36.1892	0.4747	62.5894	174.1339	37%	100%
				<i>12.4945</i>	<i>20.0021</i>				<i>0.0532</i>	<i>0.0532</i>		

Notes: The table summarizes the model performances under a target return strategy with a target return of 16%. See note to table 3.6.

TABLE 3.16  
Covariance Ranking · 8% Target Return, Daily Rebalancing

	No transaction cost						With transaction cost					
	Return	StDev	Sharpe	Switching fee		Return	StDev	Sharpe	Switching fee		TO	Rebalan. Days
				$\gamma=1$	$\gamma=10$				$\gamma=1$	$\gamma=10$		
Static	8.0000	22.1138	0.3618			8.0000	22.1138	0.3618			0%	0%
ROLL	10.0753	19.3408	0.5209			8.9115	19.3408	0.4608			14%	100%
RROLL	9.6580	18.2428	0.5294	20.1640	87.4659	8.7410	18.2428	0.4791	20.3461	86.7929	12%	100%
BEKK	9.2595	19.4055	0.4772			8.1294	19.4055	0.4189			13%	100%
RBEKK	10.0161	19.2292	0.5209	4.2793	29.2767	8.9678	19.2292	0.4664	4.3784	28.8298	12%	100%
				-6.8226	-18.9338				-10.5200	-12.0668		
DCC	10.0059	18.9852	0.5270			9.0558	18.9852	0.4770			11%	100%
RDCC	8.9506	18.1251	0.4938	13.2724	72.6508	7.8055	18.1251	0.4306	12.5416	71.2259	21%	100%
				-1.5225	-0.2911				-1.8263	-1.6597		
CCC	9.4367	18.9291	0.4985			8.5720	18.9291	0.4528			10%	100%
RCCC	10.0221	18.0434	0.5554	17.0403	77.0286	8.7573	18.0434	0.4853	16.5070	75.6957	16%	100%
				1.7701	3.9944				2.0796	2.3663		

Notes: The table summarizes the model performances under a target return strategy with a target return of 8%. See note to table 3.6.

TABLE 3.17  
Covariance Ranking · 12% Target Return, No Short-Selling, Daily Rebalancing

	No transaction cost					With transaction cost						
	Return	StDev	Sharpe	Switching fee		Return	StDev	Sharpe	Switching fee		TO	Rebalan. Days
				$\gamma=1$	$\gamma=10$				$\gamma=1$	$\gamma=10$		
Static	3.1643	13.2763	0.2383			3.1643	13.2763	0.2383			0%	0%
ROLL	4.0217	16.5611	0.2428			3.1670	16.5611	0.1912			7%	60%
RROLL	5.1866	12.3005	0.4217	57.3084	157.3148	4.5881	12.3005	0.3730	57.4036	163.8000	5%	100%
BEKK	3.8975	16.6563	0.2340			2.8721	16.6563	0.1724			9%	76%
RBEKK	3.9783	16.4475	0.2419	3.5589	25.5960	3.0732	16.4475	0.1868	3.6641	25.1581	8%	79%
				-23.4921	-610.9757				-33.4460	-37.1827		
DCC	3.9372	16.5079	0.2385			3.1769	16.5079	0.1924			6%	87%
RDCC	2.8990	15.3465	0.1889	17.0900	76.5661	1.5079	15.3465	0.0983	16.0272	74.3115	14%	100%
				-17.9516	-36.2506				-25.6131	-28.1863		
CCC	3.9934	16.1704	0.2470			3.3358	16.1704	0.2063			4%	100%
RCCC	5.4436	15.0536	0.3616	18.6005	76.4251	4.2556	15.0536	0.2827	18.0696	75.4985	11%	98%
				-13.7212	-29.9051				-20.2181	-22.4738		

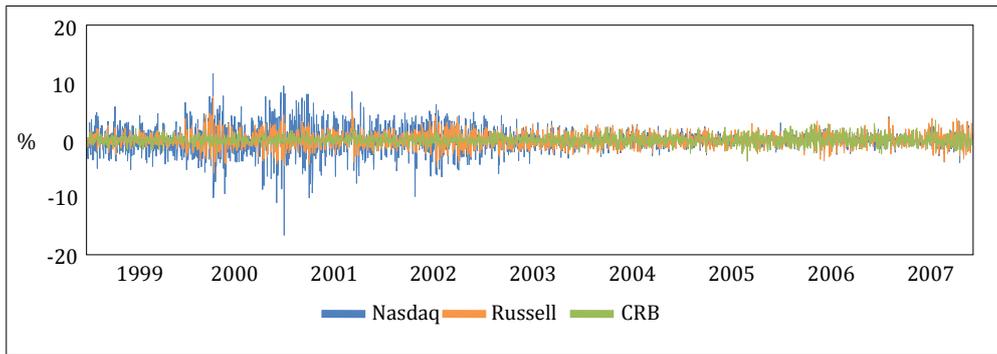
Notes: The table summarizes the model performances under a target return strategy with a no short-selling constraint. See note to table

TABLE 3.18  
Evaluation of Alternative Rebalancing Strategies Relative to Daily Rebalancing

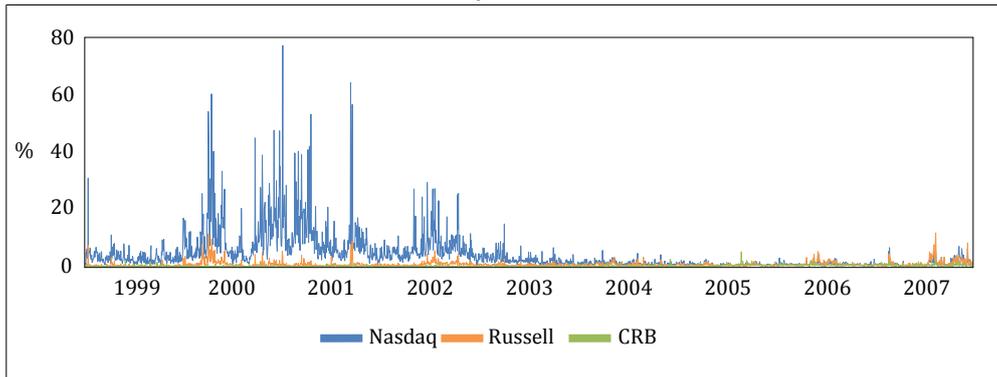
	RROLL		RBEKK		RDCC		RCCC	
	$\gamma=1$	$\gamma=10$	$\gamma=1$	$\gamma=10$	$\gamma=1$	$\gamma=10$	$\gamma=1$	$\gamma=10$
Weekly	9.724	12.761	0.875	-1.037	1.667	2.520	-4.215	-4.631
Monthly	17.062	27.028	0.033	-3.582	3.849	8.281	-5.764	-6.619
RD 5%	8.241	9.223	2.614	5.530	9.959	23.540	9.028	23.958
RD 8%	11.085	13.893	-3.269	-14.809	7.224	18.488	7.571	20.524
RD 10%	11.637	14.033	1.171	4.698	7.305	19.287	7.672	21.013
RD 15%	11.636	15.284	1.652	2.074	5.162	17.042	10.294	24.218
RD 20%	19.455	33.017	6.008	17.266	1.313	7.851	6.152	18.515
VOL 25th	6.979	12.336	0.534	3.785	2.281	3.242	4.779	11.978
VOL 50th	8.385	9.406	2.580	5.990	2.536	6.598	13.315	28.020
VOL 75th	15.620	19.700	-1.931	-7.811	7.084	-8.840	3.859	-3.937

Notes: The table presents the annualized basis points fee ( $\Delta\gamma$ ) an investor with quadratic utility and relative risk aversion of  $\gamma$  would pay to switch from the daily rebalancing strategy to weekly, monthly, or time-varying rebalancing approaches. We use  $\gamma=1$  (blue bar) and  $\gamma=10$  (orange bar) to represent investors with low and high risk aversion respectively. The transaction cost is 20% annualized percentage points. RD stands for a return-driven volatility timing strategy. The percentage following each strategy name indicates the rebalancing threshold assigned to the corresponding strategy. For instance, a RD 5% approach rebalances the portfolio weights on day  $t$  if the absolute return of the portfolio on day  $t - 1$  exceeds 5%. VOL stands for a volume-driven strategy. The 25<sup>th</sup>, 50<sup>th</sup> and 75<sup>th</sup> indicate rebalancing conditions based on historical trading volumes explained in section 3.3.

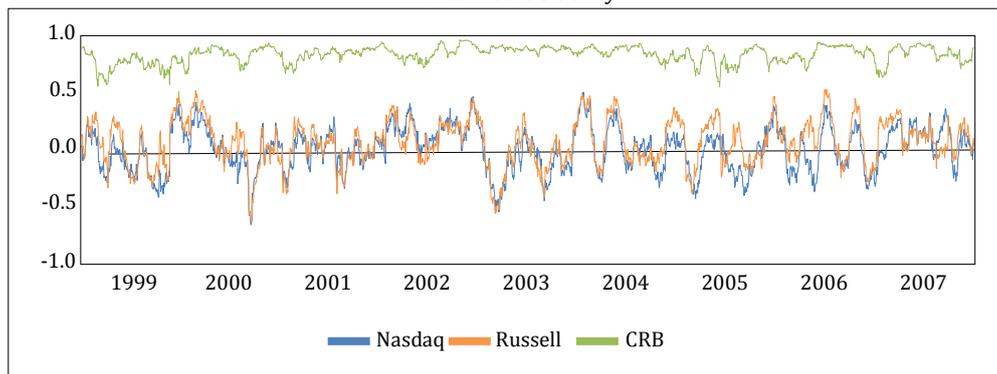
FIGURE 3.1  
Time Series Properties of Stock Indices



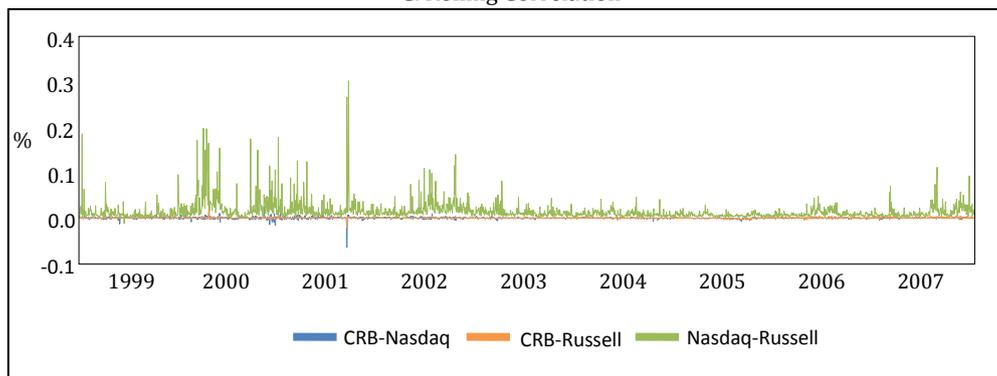
A. Daily Return Plot



B. Realized Volatility



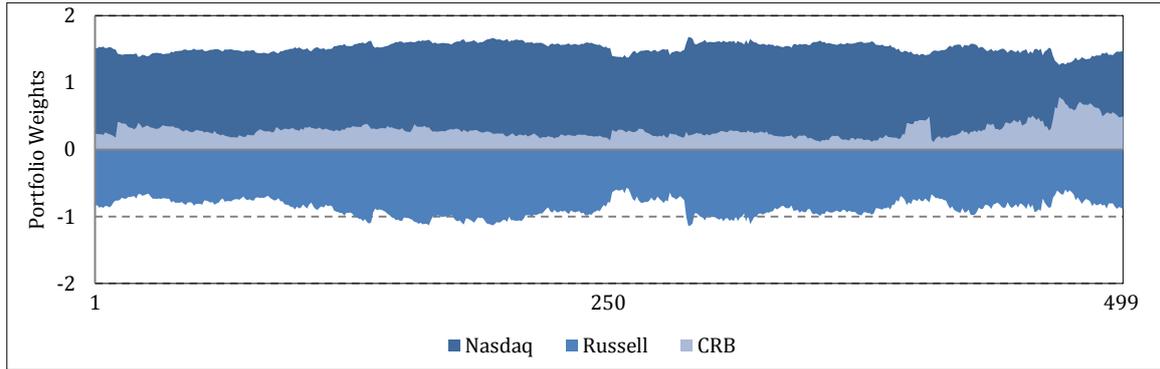
C. Rolling Correlation



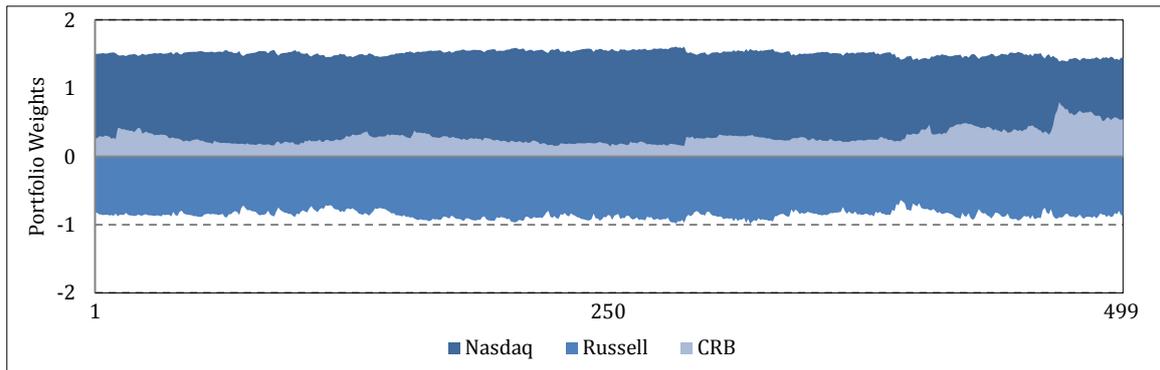
D. Realized Covariances

Notes: Panel A shows the daily return of NASDAQ 100, Russell 2000 and CRB indices over the whole sample period. Panel B and C represent the realized variances cross-market dynamic correlations. The dynamic correlation each day is calculated by a rolling window correlation of the past month between the indices. Panel D shows the realized covariance.

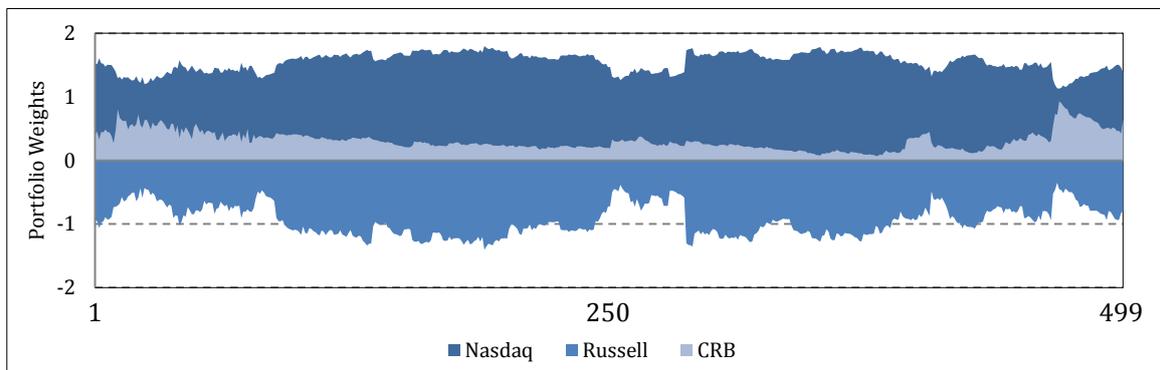
FIGURE 3.2.1  
Daily models Portfolio Weights from Daily Target Return Portfolio Optimization



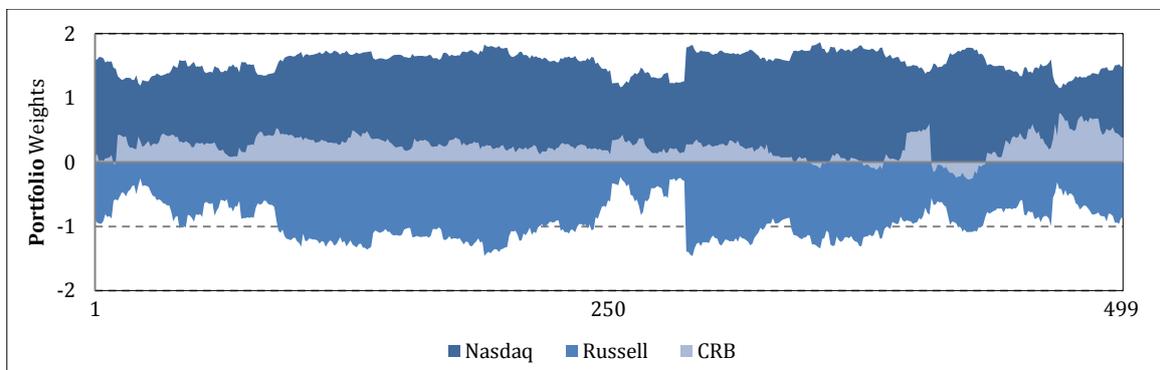
A) Portfolio weights of the DCC Model



B) Portfolio weights of the CCC Model



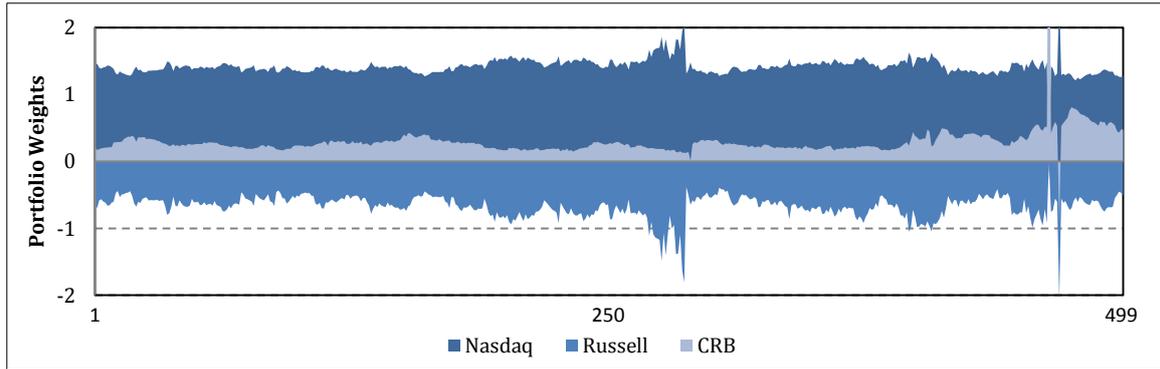
C) Portfolio weights of the BEKK Model



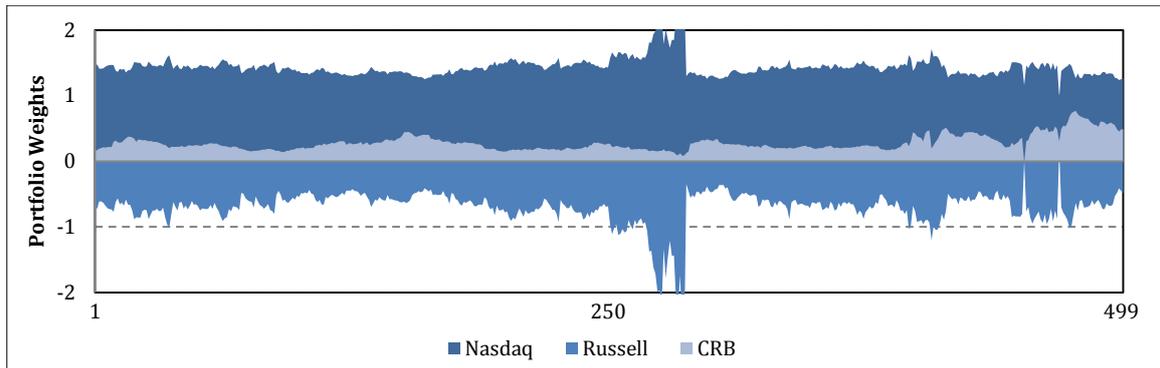
D) Portfolio weights of the ROLL Model

Notes: The figures show optimal weights generated by the daily models DCC, CCC, BEKK and ROLL. Negative weights are allowed as we adopt an unconstrained optimization approach.

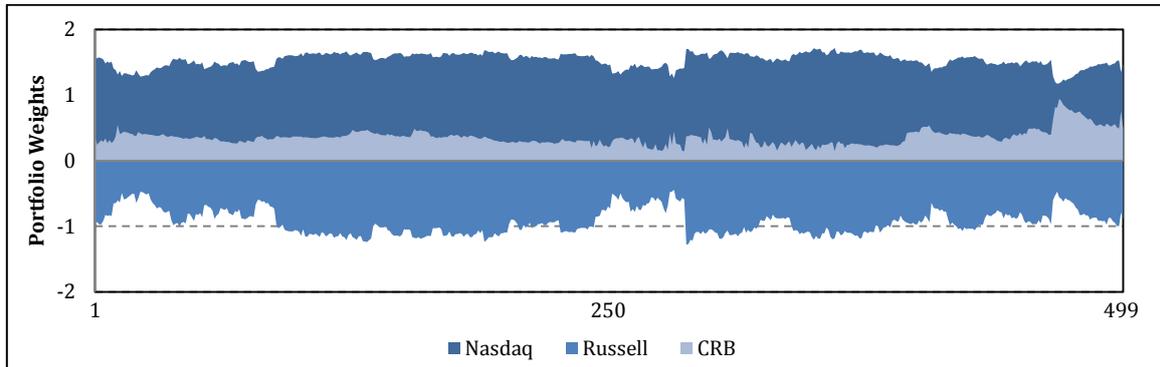
FIGURE 3.2.2  
 Realized Models' Portfolio Weights from Daily Target Return Portfolio Optimization



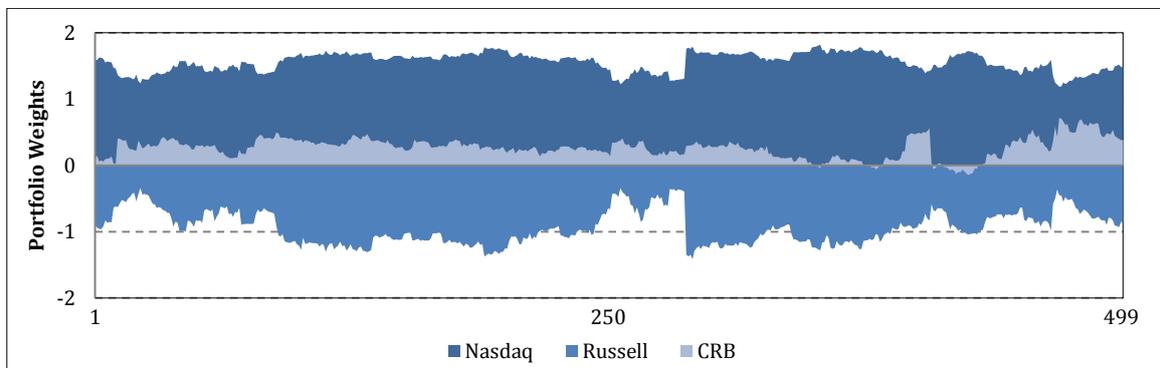
A) Portfolio weights of the realized DCC Model



B) Portfolio weights of the realized CCC Model



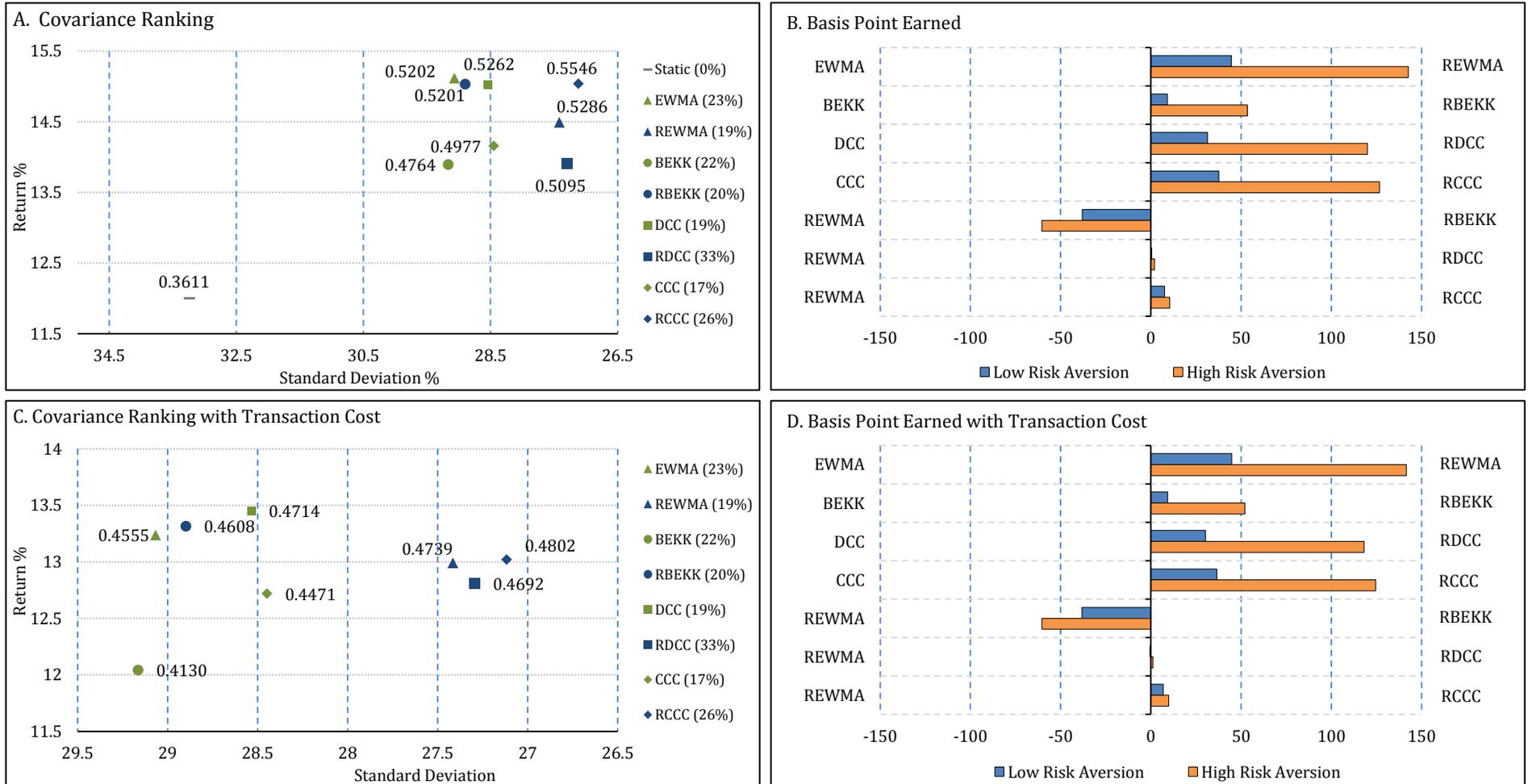
C) Portfolio weights of the realized BEKK Model



D) Portfolio weights of the realized RROLL Model

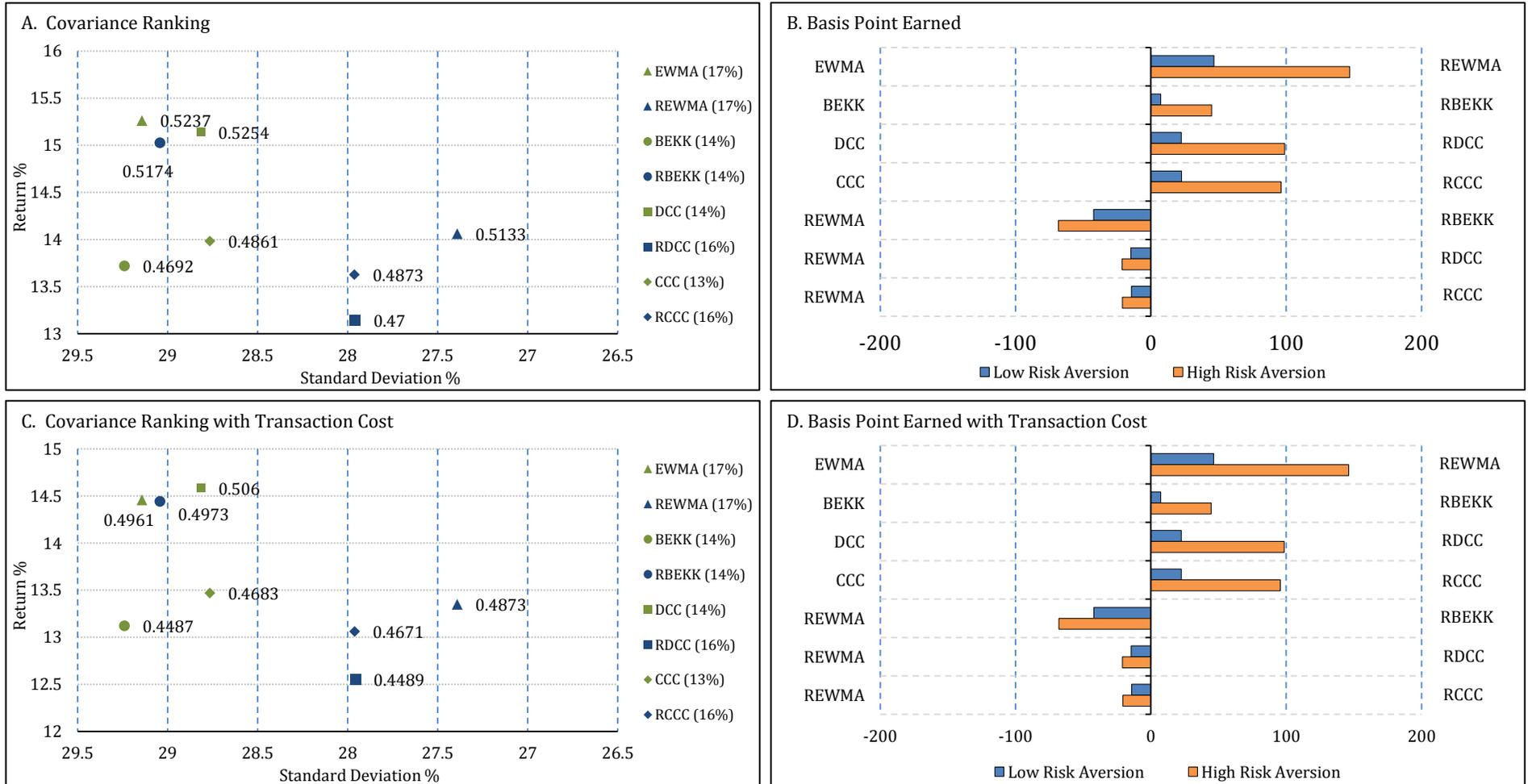
Notes: The figures show optimal weights generated by the realized models RDCC, RCCC, RBEKK and RROLL.

FIGURE 3.3  
Covariance Ranking · 12% Target Return, Daily Rebalancing



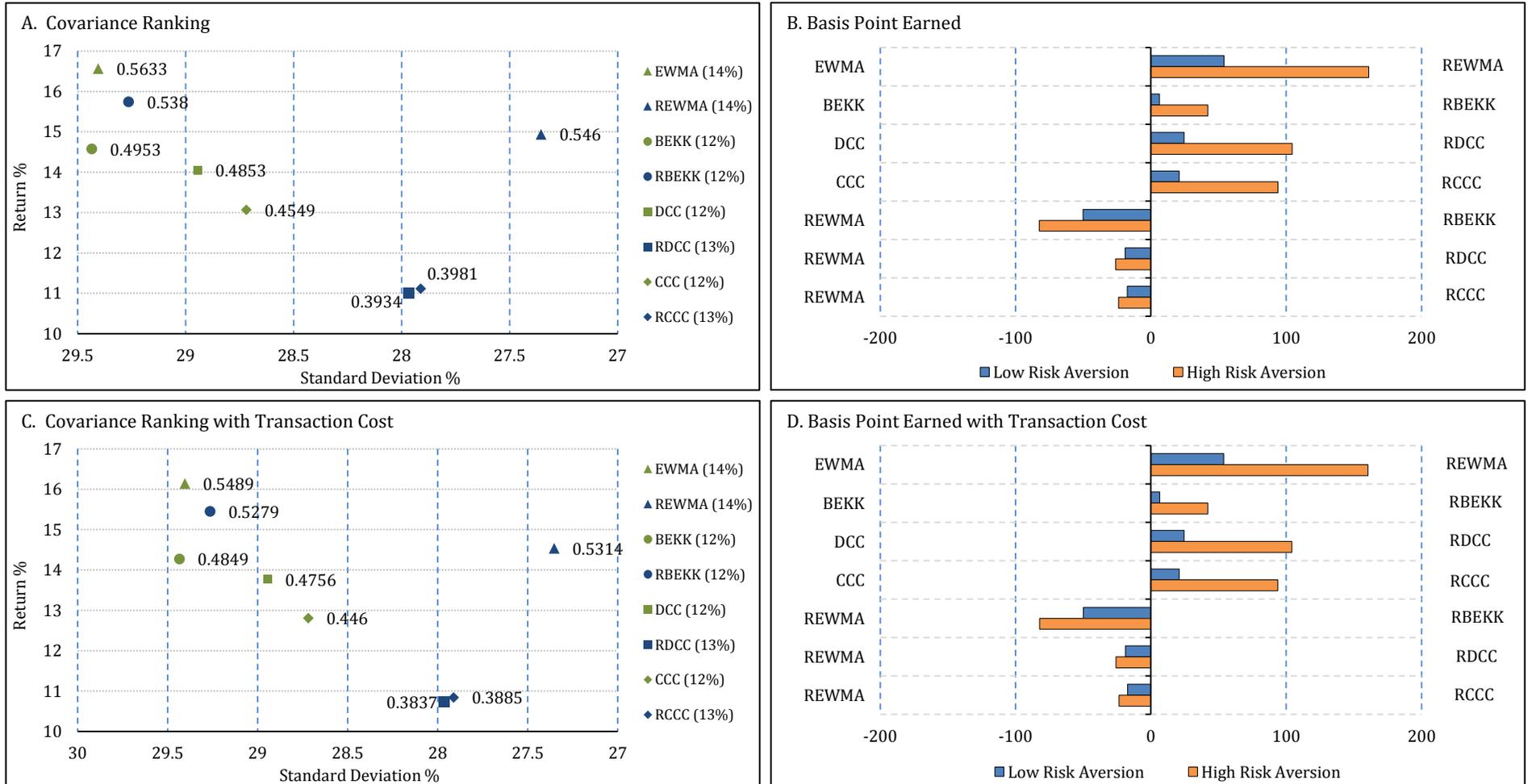
Notes: Each day the optimal portfolio weights are allocated to minimize conditional volatility subject to a target return of 12%. The left panels report the model performance in terms of return, standard deviation and Sharpe ratio. The right panels show switching fees, an annualized basis points fee an investor with quadratic utility and constant relative risk aversion of  $\gamma$  would pay to switch from a covariance matrix forecast on the left hand side to one on the right. We use  $\gamma=1$  (blue bar) and  $\gamma=10$  (orange bar) to represent investors with low and high risk aversion respectively. The transaction cost is 20% annualized percentage points. The brackets after each model name provides additional information on annualized turnover.

FIGURE 3.4  
Covariance Ranking · 12% Target Return, Weekly Rebalancing



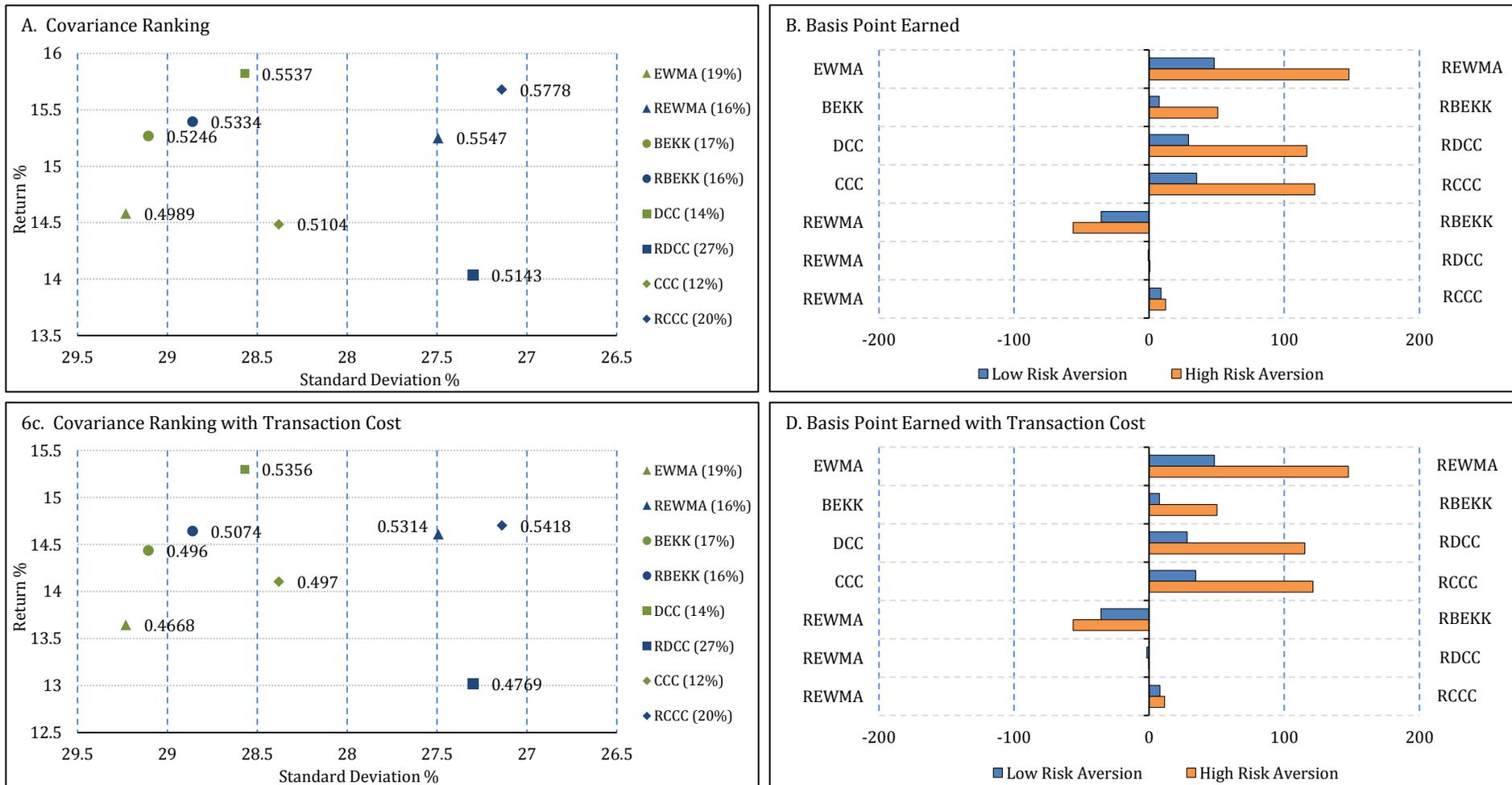
Notes: The portfolio is rebalanced at the beginning of the week according to the calculated optimal portfolio weights and then keep the weights unchanged until the following week. See note to Figure 3.3.

FIGURE 3.5  
Covariance Ranking · 12% Target Return, Monthly Rebalancing



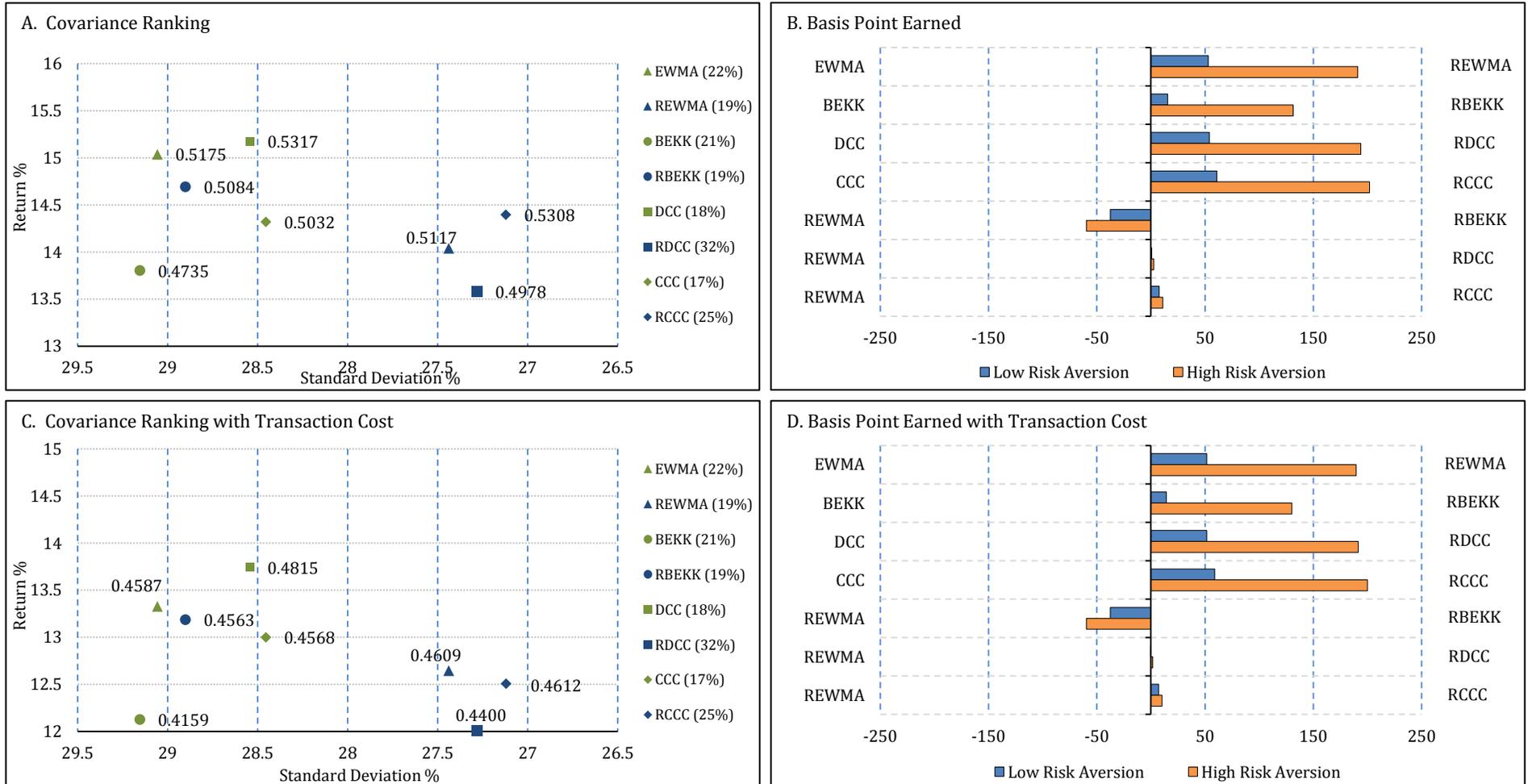
Notes: The portfolio is rebalanced at the beginning of the month according to the calculated optimal portfolio weights and then keep the weights unchanged until the following month. See note to Figure 3.3.

FIGURE 3.6  
Covariance Ranking • 12% Target Return, Return-Driven Rebalancing with a 10% Threshold



Notes: We rebalance the portfolio on day  $t$  if the overall percentage change in returns exceeds 10% on day  $t - 1$ . See note to Figure 3.3.

FIGURE 3.7  
Covariance Ranking • 12% Target Return, Volume-Driven Rebalancing with a 25<sup>th</sup> Percentile



Notes: We rebalance on the day  $t$  if transaction volume on day  $t - 1$  exceeds the 25<sup>th</sup> percentile of the in-sample period. S&P500 volume is used as a volume proxy since it is more representative of the market the three indices we have. See note to Figure 3.3.

# CHAPTER 4: Single Index and Portfolio Models for High-Frequency VaR Prediction

## 4.1 Introduction

Value-at-risk (VaR) is a widely used tool, which indicates the possible loss of an asset portfolio at a certain risk level over a certain period, for estimating and reporting financial risk (see Jorion, 2000, for a detailed interpretation of VaR). The technique has become increasingly important since the 1995 amendment to the Basel Accord that permitted banks and other authorized deposit-taking institutions (ADIs) to use internal models to calculate VaR. After that, numerous VaR estimation and evaluation approaches have been proposed and widely applied in risk management. Hence, it was supposed that VaR estimation has been well developed and can provide reliable portfolio risk assessment.

However, the 2008-2009 global financial crisis called the entire risk management system into question. Reputable financial institutions with perceived sounding risk management faced survival difficulties. For instance, Lehman Brothers filed for bankruptcy protection, Merrill Lynch was sold, and Fannie Mae and Freddie Mac were taken over by the US federal government. Given that VaR is the industrial standard risk measure, the dramatic financial turmoil made it necessary to at least investigate the adequacy of VaR's estimation procedure and the accuracy and robustness of its predictions.

From a statistical point of view, calculation of VaR thresholds of a portfolio involves estimating the variance of portfolio returns, which can be modelled either by (1) treating the portfolio returns as a univariate process or (2) forecasting the conditional variances and covariances of the portfolio components. The former is termed as single index model and the latter portfolio model.<sup>1</sup>

The single index (univariate) volatility model is by and large the conventional method

---

<sup>1</sup>A comprehensive review of alternative single index and portfolio models can be found in McAleer and da Veiga (2008) and Kuester et al (2006).

for predicting daily VaR. Among numerous single index volatility models, the ARCH specifications are widely-used by academicians and market practitioners. Starting with Engle's (1982) original ARCH model, a variety of ARCH volatility models have been developed, such as GARCH (Bollerslev, 1986), EGARCH (Nelson, 1991) and GJRARCH models, seeking to improve forecast accuracy. These ARCH specifications alongside Gaussian or Student- $t$  quantiles have been widely used by financial managers for daily VaR estimation; e.g. see the recent RiskMetrics methodology (Zumbach, 2007).

From both statistical and practical aspects, it is also necessary to analyse and evaluate portfolio models. Single index models may produce limited accuracy by ignoring the complicated correlation among the component returns. As a result, univariate specifications give out less information on the source of risk. Furthermore, portfolio models could be used as a complement to single index models and could help in assessing the accuracy of VaR predictions made by single index models. Nevertheless, financial institutions are only starting to consider portfolio models due to the computational difficulties in the covariance and joint density estimation.

Since both the single index and portfolio models can be used to estimate VaR, the question that emerges is: which model makes more accurate VaR predictions? The empirical literature has studied the univariate and multivariate contest extensively (see Berkowitz and O'Brien, 2002; Brooks and Persaud, 2003; Bauwens et al., 2006; McAleer and da Veiga, 2008; Christoffersen, 2009; McAleer, 2009; Dumitrescu, 2012; Santos et al., 2012), but a clear conclusion is yet to be drawn. As the accuracy of VaR predictions is vital for risk management, the contest between single index and portfolio models warrants further study. In addition, the contest has mainly been studied using daily asset returns, while rarely has the question been investigated in a high-frequency context. Given that intraday volatility estimators are suggested by numerous studies (see Andersen et al., 2008; Barndorff-Nielsen and Shephard, 2006) to be statistically superior to their daily counterparts, it is worth to reassess the contest based on intraday information.

This chapter contributes to the literature by comparing single index and portfolio models in a high frequency context. With intraday information incorporated, single index models may improve forecast accuracy by adding an extra dimension of information while no much noise is added (much less parameters are added compared with portfolio models). In contrast, the effect of adding intraday information into portfolio models could be mixed. On one hand, high-frequency forecasts provide more relevant information about the volatility and correlation of the portfolio components; on the other hand, more parameter estimation adds estimation noise and potential microstructure problems to the model. Therefore the additional trade-offs between noise and signal is not clear at a first glance. In addition, intraday based covariance is needed for estimating high-frequency portfolio models, which is a potential source of additional microstructure problems.

We propose to extend both the single index and portfolio GARCH models with a realized (co)variation estimator. For single index models, we incorporate a realized kernel measure by adding an extra variable into the standard GARCH, GJRGARCH and PGARCH specifications. For portfolio models, we add a realized kernel covariance (RKC) estimator into multivariate GARCH specifications such as BEKK, CCC, and DCC. The performance of these augmented GARCH class models will be compared with the benchmark ones, single index or portfolio models using daily data. To make the comparison between single index and portfolio models clear, we first analyse two representative models, the standard GARCH and DCC, and their realized versions, RGARCH and RDCC. Next, we provide backtesting results for other single index and portfolio models to verify further the conclusion made from assessing the representative models.

To preview the key results, we show that in a high-frequency context, both the single index and portfolio models provide adequate VaR forecasts. However, no group of models is statistically superior based on the results of the backtesting tests. Generally the portfolio models yield more accurate coverage ratios while the single index models provide smaller average and maximum absolute deviation of violations. The intraday models in

both groups are statistically superior to their daily counterparts, which is in line with the findings of Fuertes and Olmo (2012). Nevertheless, the result adds credit to the application of realized single index models as they are more parsimonious and computational friendly.

The rest of the paper is organized as follows. Section 4.2 explains the VaR and backtesting methods, volatility models and combination techniques. Section 4.3 discusses the data to be used in the empirical test. Section 4.4 presents the empirical results, and Section 4.5 concludes.

## **4.2 Literature Review**

### **4.2.1 Single Index versus Portfolio Models**

The main concern about the choice between single index (univariate) and portfolio (multivariate) model for VaR estimation is the signal-to-noise ratio (see McAleer and da Veiga, 2008). The Single index model needs only a univariate specification to estimate the variance of the single index, obviating the need for covariances or correlations. There is little signal, meaning that univariate models may yield low predictive accuracy, but there is also little noise. Portfolio models have more signals because they exploit covariances and correlations, but there is more estimation noise, possibly compromising their forecasting ability.

In addition, each of these two alternatives has pros and cons computationally. First, single index models have to be estimated again whenever the vector of weights changes, as the change produces a different univariate time series of portfolio returns. This requirement is not necessary when a portfolio model is employed. But, if the dimension of the portfolio increases, portfolio models become harder to implement due to usually large number of parameters to be estimated. Therefore the trade-offs between noise and signal, and between forecast accuracy and numerical difficulty, are not clear at a first glance.

McAleer and da Veiga (2008) find mixed results using daily data of four international

stock market indices. They show that the single index model leads to excessive and often serially dependent violations (hence monetary penalty), while the portfolio model leads to too few violations (hence more capital charges). Dumitrescu et al. (2012) compare the VaR forecasting performance between the univariate GARCH model and multivariate DCC model. Judged by dynamic binary (DB) backtesting criteria, they find that the two volatility models are too similar for the backtests to discriminate between them. In other words, the correlation amongst assets does not have any impact on the calculation of the VaR and implicitly on its validity.

Few studies argue in favour of either the single index or portfolio approaches. Berkowitz and O'Brien (2002) evaluate the performance of banks' trading risk models by examining the statistical accuracy of the multivariate VaR forecasts made by 6 large U.S. banks. Bank's multivariate risk models usually conduct structural modelling that measure the joint distribution of all material risks conditional on current information, or employs approximations to reduce computational burdens if portfolios are large and complex. However despite the detailed information employed in the banks models, VaR forecasts generated by banks do not outperform the forecast made by a standard univariate GARCH model, which provides lower VaRs and is better at predicting changes in volatility. They conclude that the results may reflect substantial computational difficulties in constructing large-scale structural models of trading risks for large and complex portfolios. Christoffersen (2009) also agrees that single index models are more appropriate if the purpose is calculating VaR, whereas portfolio models are more suitable for portfolio selection. Bauwens et al. (2006) conjecture that, under the present state of the art, it is better to adopt univariate framework when facing large and complex portfolios.

In contrast, Santos et al. (2012) compare multivariate and univariate GARCH models to forecast portfolio VaR in the context of large and diversified portfolios. The results of predictive performance for one-step-ahead VaR obtained with both Monte Carlo simulations and with real market data show that multivariate GARCH models outperform

competing univariate models on an out-of-sample basis.

### 4.2.2 VaR Backtesting Approaches

Various statistical tests of predictive accuracy to compare alternative VaR forecasts have been developed in the literature. Dumitrescu et al. (2012) suggest that a proper VaR backtesting method needs to address three main issues. First, the power of the backtesting technique, which is the probability of rejecting a model that is not valid, should be reasonably high. Hurlin and Tokpavi (2008) show that some backtesting procedures are too optimistic in the sense that they do not reject the validity of a model as often as they should. Second, the evaluation methodology needs to be model-free. Evaluation must be implementable whatever the volatility model applied to estimate the series of VaR. Third, estimation risk must be evaluated since risk of estimation error present in the estimates of corresponding model parameters pollutes VaR forecasts.

A number of backtesting tests have been proposed to take into account the three criteria over the past 20 years, and among them three popular approaches can be applied to examine the validity of VaR forecasts generated by single index and portfolio models. The first one is the *LR* conditional coverage (*LR<sub>CC</sub>*) test introduced by Christoffersen (1998). The method examines the quality of VaR forecasts through parameter restrictions on the transition probability matrix associated with a two-states Markov chain model (violation/no violation). Two elements of the test, the unconditional coverage and the independence hypotheses, are derived from the martingale difference hypothesis.

The Second one is the Dynamic Quantile (*DQ*) test proposed by Engle and Manganelli (2004). The method is based on a linear regression model which projects VaR violations onto a set of independent variables and subsequently examines different restrictions on the coefficients of the regression model.

The third technique is the non-linear Dynamic Binary (*DB*) model developed by Dumitrescu et al. (2012). They argue that in view of the dichotomic character of the series

of violations, a non-linear model seems more appropriate. The DB model is a nonlinear extension of the  $DQ$  test. The extended model is based on a dynamic binary regression model which links the sequence of violations to a set of explanatory variables including the lagged VaR and the lagged violations in particular. Monte-Carlo experiments show that the DB test exhibits good small sample properties in realistic sample settings (5% coverage rate with estimation risk).

### 4.3 Methodology

In this section, we present a brief definition of VaR, a description of the volatility forecasting procedure, and the evaluation methods used for assessing the forecast accuracy of a number of alternative volatility models. Three commonly applied single index models, GARCH, equation (2.20), GJRGARCH, equation (2.22), and PGARCH, equation (2.24), and their corresponding realized extensions, RGARCH, equation (2.21), RGJRGARCH, equation (2.23), and RPGARCH, equation (2.24), are studied in the empirical study. Alternatively, three standard portfolio models, diagonal BEKK, equation (2.30), scalar CCC, equation (2.32), and DCC, equation (2.35), and their corresponding realized extensions, RBEKK, equation (2.31), RCCC, equation (2.33), and RDCC, equation (2.36), are also applied.

The intraday based regressor selected for each realized single index model is the ARFIMA forecasted realized kernel (RK) estimator with a Parzen kernel weight function (equation 2.8). Correspondingly, the intraday based covariance estimator deployed in each realized portfolio model is the ARFIMA forecasted realized kernel covariance (RKC) with a Parzen kernel weight function (equation 2.15). The choice of RK and RKC estimators are motivated by their statistical consistency in estimating noisy high-frequency data (see Barndorff-Nielsen et al., 2008, 2009). The single index and portfolio models will be applied on a portfolio of three financial assets. We assume an equal weighted portfolio, which has been commonly used in the empirical literature; see, for instance, Zaffaroni

(2007), DeMiguel et al. (2009) and Santos et al. (2012).

### 4.3.1 VaR Modelling Approach

Our VaR modelling approach builds upon the contributions of Clements et al. (2008) and Fuertes and Olmo (2012). Let us assume that log return at day  $t$  on a single asset,  $r_t$ , follows a pure multiplicative process  $r_t = \sqrt{\sigma_t^2} \varepsilon_t$ , where  $\sigma_t^2$  is either a GARCH-type conditional variance of the daily return, a realized volatility conditional expectation, or a combination of both.  $\varepsilon_t$  is an *iid* unit variance random variable with probability distribution  $F_\varepsilon$ ,  $\varepsilon_t \sim F_\varepsilon(\cdot)$ . The VaR of  $r_t$  is defined as the  $\alpha$ -percentage quantile of the conditional distribution of asset returns given the investment manager's information set  $\Omega_{t-1}$ . The forecasted 1-day-ahead VaR, a measure of the worst case 1-day-ahead loss, is calculated as

$$VaR_{t+1,\alpha} \equiv \sqrt{\hat{\sigma}_{t+1}^2} \left( \hat{\theta}_t \right) \hat{F}_\varepsilon^{-1}(\alpha) \quad (4.1)$$

where  $\hat{\theta}_t$  is a consistent estimator of the model parameters needed to calculate  $\hat{\sigma}_{t+1}^2$ , and  $\hat{F}_\varepsilon^{-1}(\alpha)$  is the  $\alpha$ -quantile estimate associated with  $F_\varepsilon$ .

Equation (4.1) implies that the adequacy of VaR forecasts impinges, in principle, on two factors: the volatility model selected to produce the daily volatility forecasts,  $\hat{\sigma}_{t+1}^2$ , and the approach used for the  $\alpha$ -quantile calculation. The candidates that we consider for belongs to one of the two classes of estimation method for forecasting portfolio volatility, the single index and portfolio models. The single index models considered are daily GARCH models and GARCH models augmented with intraday volatility estimators. The portfolio models applied are multivariate GARCH specifications and their augmented versions with intraday volatility estimators.

In this study we choose  $\alpha$ , the level of significance chosen for the VaR, at 1%. Two probability distributions are adopted for the  $\alpha$ -quantile calculation. First,  $F_\varepsilon(\cdot)$  is specified as the standard Gaussian density so that jumps are not properly accounted for in the quantile estimate  $\hat{F}_\varepsilon^{-1}(\alpha)$ . Second, standardized (unit variance) student- $t$  densities

with degrees of freedom parameters estimated by maximum likelihood from the in-sample returns are used to account for fat tails.

Nonparametric distributions are not considered as they are beyond the scope of this empirical study. For example, Engle and Gonzalez-Rivera (1991) calculate quantiles from the empirical distribution function (edf) of the innovation series of a GARCH model; McNeil and Frey (2001) employ a semiparametric approach that combines the edf and an extreme value density which includes the parametric Student- $t$  distribution.

### 4.3.2 Volatility Forecasting

The out-of-sample daily volatility forecasts for the RK and RKC estimators are made by rolling window autoregressive fractionally integrated moving average (ARFIMA) estimations. A stochastic process  $s_t$  is called an ARFIMA( $p, d, q$ ) process if the fractionally differenced process is an autoregressive moving-average (ARMA) process as

$$(1 - \phi L - \dots - \phi_p L^p)(\Delta^d X_t) = (1 - \theta_1 L - \dots - \theta_q L^q)(\varepsilon_t) \quad (4.2)$$

where  $d$  is not restricted to integral values. ARFIMA is associated with long-memory process, a stationary series whose autocorrelation function decays slowly. An ARFIMA specification yields a parsimonious parameterization of long-memory processes that nests the ARMA model, which is commonly used for short-memory processes. The ARFIMA model can also be considered as an extension of the ARIMA model as the former allows for fractional degrees of integration.

As noted in Fuertes and Olmo (2012), the ARFIMA modelling framework has been successfully used in the literature to capture the stylized slow, less than exponentially, decay in autocorrelations of daily realized volatilities. For example, Bhardwaj and Swanson (2006) investigated the usefulness of ARFIMA models in practical prediction-based applications, and find evidence that such specifications often outperform a wide class of the benchmark non-ARFIMA models, including AR, ARMA, ARIMA, random walk, and related models. This paper focus on the homoscedastic ARFIMA (1,  $d$ , 0) model, which

has been employed in Andersen et al. (2003), Koopman et al. (2005), and Fuertes and Olmo (2012). The model is proved to be an efficient competitor to other time series methods of forecasting realized volatility. The model specification is

$$(1 - \phi L)(1 - L)^d(s_t - w) = \varepsilon_t \quad (4.3)$$

where  $s_t$  is one of the the daily realized estimators,  $w$  is the unconditional mean of  $s_t$ , and  $L$  is the lag operator ( $Ls_t = s_{t-1}$ ). Estimation of the coefficients in the equation including  $d$  is carried out by exact maximum likelihood under the normal assumption and forecasts are calculated from the autoregressive and moving average representation of the process (see Fuertes and Olmo, 2012). In the ARFIMA models fitted to the log measures, the forecasted volatility is obtained through the common exponential transformation (see Forsberg and Ghysels, 2007; Clements et al., 2008).

### 4.3.3 VaR Backtesting Framework

Three main criteria are used to compare the forecasting performance of the conditional volatility models, namely: (1) the LR conditional coverage test ( $LR_{CC}$ ); (2) the dynamic quantile ( $DQ$ ) test; (3) the dynamic binary response ( $DB$ ) model. We also complement the backtesting by providing the result of the linear regression approach introduced by Pagan and Schwert (1990), and statistics of the percentage of violation, average and maximum absolute deviation of violations.

#### 4.3.3.1 Likelihood Ratio Test

Christoffersen (1998) proposes likelihood ratio ( $LR$ ) tests of unconditional coverage, serial independence and conditional coverage, which are subsequently used by Lopez (1998) to assess VaR forecasts. An appropriate VaR model should display the property that the unconditional coverage, estimated as the number of observed violations divided by the sample size, should equal  $\alpha$ , the level of significance selected for the VaR. The probability

of having  $X$  violations in  $T$  trading days is formulated as

$$\Pr(X) = C_X^{250} (0.01)^x (0.99)^{T-x} \quad (4.4)$$

The  $LR$  statistic for examining whether the unconditional coverage is equal to alpha is:

$$LR_{UC} = 2[\log(\alpha^x (1 - \alpha)^{N-x}) - \log((0.01^x)(0.99^{N-x}))] \quad (4.5)$$

where  $\alpha = x/N$ ,  $x$  is the number of violations and  $N$  is the number for forecasts. The  $LR$  statistic is asymptotically distributed as  $\chi(1)$ .

But a model with correct unconditional coverage may still be sub-optimal if the violations are serially dependent, as consecutive large losses may cause bankruptcy. Independence can be tested by the  $LR$  statistic with a null hypothesis of serial independence against the alternative of first-order Markov dependence as follows.

Assume the violation is dependent over time and can be described as a first order Markov sequence with transition probability matrix

$$\Pi_1 = \begin{bmatrix} \pi_{00} & \pi_{01} \\ \pi_{10} & \pi_{11} \end{bmatrix} \quad (4.6)$$

where  $\pi_{ij}$  is the transition probability

$$\pi_{ij} = P[I_t = i \text{ and } I_{t+1} = j] \quad (4.7)$$

where  $I_t$  is an indicator variable which equals to 1 if there is a violation at time  $t$  and 0 otherwise. For example,  $\pi_{01}$  is the probability of a violation after a non-violation, and  $\pi_{11}$  is the probability of two consecutive violations. With a total sample of  $T$  observations, the likelihood function of the first-order Markov process is

$$L(\Pi_1) = \pi_{00}^{T_{00}} \pi_{01}^{T_{01}} \pi_{10}^{T_{10}} \pi_{11}^{T_{11}} \quad (4.8)$$

where  $T_{ij}$  is the number of observations with a  $j$  following an  $i$ .

Taking first derivatives with respect to  $\pi_{01}$  and  $\pi_{11}$  and equating to zero, the maximum likelihood estimates are

$$\begin{aligned} \hat{\pi}_{01} &= \frac{T_{01}}{T_{00} + T_{01}} \\ \hat{\pi}_{11} &= \frac{T_{11}}{T_{10} + T_{11}} \end{aligned} \quad (4.9)$$

Since  $T_{i0} + T_{i1} = T_i$ , the estimated matrix can be written as

$$\hat{\Pi}_1 = \begin{bmatrix} T_{00}/T_0 & T_{01}/T_0 \\ T_{10}/T_1 & T_{11}/T_1 \end{bmatrix} \quad (4.10)$$

We can then test the independence hypothesis using a likelihood ratio test as

$$LR_{IND} = -2 \log[L(\hat{\Pi})/L(\hat{\Pi}_1)] \sim \chi_1^2 \quad (4.11)$$

where  $\hat{\Pi}$  is the transition matrix under independence (with  $\pi_{01} = \pi_{11}$ ),

$$\hat{\Pi} = \begin{bmatrix} 1 - T_1/T & T_1/T \\ 1 - T_1/T & T_1/T \end{bmatrix} \quad (4.12)$$

which can be calculated in a similar fashion as the calculation of  $\hat{\Pi}_1$ .

The conditional coverage ( $LR_{CC}$ ) test can finally be applied to jointly assessing unconditional coverage and independence. The  $LR_{CC}$  statistic is formulated as the sum of the  $LR_{UC}$  statistic and the  $LR_{IND}$  statistic, which is asymptotically distributed as  $\chi(2)$ . The  $LR_{CC}$  test can be denoted as

$$LR_{CC} = LR_{UC} + LR_{IND} \quad (4.13)$$

#### 4.3.3.2 Dynamic Quantile ( $DQ$ ) Test

The  $DQ$  test is a linear regression based model introduced by Engle and Manganelli (2004) to assess the quality of VaR forecasts. Let  $I_t(\alpha)$  be the binary variable associated with the ex-post observation of an  $\alpha\%$  VaR violation at time  $t$ , where

$$I_t(\alpha) = \begin{cases} 1 & \text{if } r_t < -VaR_{t|t-1}(\alpha) \\ 0 & \text{otherwise} \end{cases} \quad (4.14)$$

Let  $Hit_t(\alpha) = I_t(\alpha) - \alpha$  denotes the demeaned process of violation. From the definition of the VaR, the conditional expectation of  $Hit_t(\alpha)$  given the information known on day  $t - 1$  must be zero. Engle and Manganelli (2004) suggest that the conditional coverage assumption can be tested in the following linear regression as

$$Hit_t(\alpha) = \delta + \sum_{k=1}^K \beta_k Hit_{t-k}(\alpha) + \sum_{k=1}^K \gamma_k g[Hit_{t-k}(\alpha), Hit_{t-k-1}(\alpha), \dots, z_{t-k}, z_{t-k-1}, \dots] + \varepsilon_t \quad (4.15)$$

where  $g[\cdot]$  is a function of past violations and of the instrument variables  $z_{t-k}$  belonging to the entire informational set available  $\Omega_{t-1}$ , and  $\varepsilon_t$  is a *i.i.d.* process. In this study the instruments used in the *DQ* specification are a constant, the VaR forecast and the first four lagged hits. Testing for the null hypothesis of conditional coverage is then implemented by testing the joint nullity of the parameters  $\beta_k, \gamma_k, \forall k = 1, \dots, K$ , and that of the intercept  $\delta$  as

$$H_0 : \delta = \beta_1 = \dots = \beta_k = \gamma_1 = \dots = \gamma_k = 0, \forall k = 1, \dots, K \quad (4.16)$$

If we denote  $\Psi = (\delta, \beta_1, \dots, \beta_k, \gamma_1, \dots, \gamma_k)'$  as the vector of the  $2K + 1$  coefficients of the model and  $Z$  as the matrix of independent variables of the regression, the test statistic of *DQ* test satisfies the following relation:

$$DQ = \frac{\hat{\Psi}' Z' Z \hat{\Psi}}{\alpha(1 - \alpha)} \xrightarrow{T \rightarrow \infty} \chi^2(2K + 1) \quad (4.17)$$

#### 4.3.3.3 Dynamic Binary (*DB*) Model

Dumitrescu et al. (2012) extend the *DQ* method by considering a non-linear approach, the dynamic binary (*DB*) model, to accommodate the dichotomic character of the series of violations. let us consider a dynamic binary response model as

$$\Pr[I_t(\alpha) = 1 | \Omega_{t-1}] = E[I_t(\alpha) | \Omega_{t-1}] = F(\pi_t) \quad (4.18)$$

where  $F(\cdot)$  denotes a cumulative distribution function and  $\pi_t$  is a index function. Assume  $\pi_t$  satisfies the following autoregressive representation:

$$\pi_t = c + \sum_{j=1}^{q_1} \beta_j \pi_{t-j} + \sum_{j=1}^{q_2} \delta_j I_{t-j}(\alpha) + \sum_{j=1}^{q_3} \psi_j l(x_{t-j}, \varphi) + \sum_{j=1}^{q_4} \gamma_j l(x_{t-j}, \varphi) I_{t-j} \quad (4.19)$$

where  $l(\cdot)$  is a function of a finite number of lagged values of observables, and  $x_t$  is a vector of explicative variables. We apply four *DB* specifications proposed by Dumitrescu

et al. (2012), denoted by  $DB_1$  to  $DB_4$ :

$$\begin{aligned}
DB_1 : \pi_t &= c + \beta_1 \pi_{t-1}, \\
DB_2 : \pi_t &= c + \beta_1 \pi_{t-1} + \delta_1 I_{t-1}(\alpha), \\
DB_3 : \pi_t &= c + \beta_1 \pi_{t-1} + \delta_1 I_{t-1}(\alpha) + \psi_1 VaR_{t-1}, \\
DB_4 : \pi_t &= c + \beta_1 \pi_{t-1} + \delta_1 I_{t-1}(\alpha) + \psi_1 VaR_{t-1} + \gamma_1 VaR_{t-1} I_{t-1}.
\end{aligned} \tag{4.20}$$

The first two equations correspond to a  $DB$  model including the lagged index as an explanatory variable and some additional information through the past values of the violation process. The third model is derived from the autoregressive quantile specifications proposed by Engle and Mangenelli (2004) and the fourth introduces an asymmetry in the response of the index to past VaR.

The general form of the model is  $\Pr_{t-1}(I_t(\alpha) = 1) = F(\pi_t)$ , where  $\pi_t$  is one of the four  $DB$  models. The log-likelihood function can be expressed as:

$$\ln L(\theta; I(\alpha) \ln F(\pi_t(\theta, Z_t)) + (1 - I_t(\alpha)) \ln(1 - F(\pi_t(\theta, Z_t)))) \tag{4.21}$$

where  $\theta = [\beta', \delta', \psi', \gamma']$  is the vector of coefficients of the model except for the intercept and  $Z_t$  is the vector of explanatory variables at time  $t$  corresponding to a certain equation from the four models applied in the study. The models are estimated through constrained maximum likelihood estimation<sup>2</sup>.

Similar to the  $DQ$  test, a  $DB$  model tests the nullity of the coefficients of the regression to assess the null hypothesis of conditional coverage ( $\Pr[I_t = 1 | \Omega_{t-1}] = F(F^{-1}(\alpha)) = \alpha$ ):

$$H_0 : \beta = 0, \delta = 0, \psi = 0, \gamma = 0 \text{ and } c = F^{-1}(\alpha) \tag{4.22}$$

A dynamic binary LR ( $DB_{LRCC}$ ) test can be used in such a context.  $DB_{LRCC}$  tests the efficiency assumption of the VaR model as:

$$DB_{LRCC} = -2 \left\{ \ln L(0, F^{-1}(\alpha); I(\alpha), Z_t) - \ln F(\hat{\theta}, \hat{c}; I_t(\alpha), Z_t) \right\} \xrightarrow[T \rightarrow \infty]{d} \chi^2(\dim(Z_t)) \tag{4.23}$$

where  $\hat{\theta}$  is the the vector of estimated coefficients of the binary-choice specification (under the alternative hypothesis, by maximum-likelihood) and  $\hat{c}$  is the estimated intercept of the specification.

---

<sup>2</sup>See Dumitrescu et al. (2012) for a detailed discussion on constrained maximum likelihood estimation.

#### 4.3.3.4 Linear Regression Method

Pagan and Schwert (1990) introduce a forecast evaluation approach whereby the volatility forecasts are regressed on the actual volatility. In this study, Realized Variance (RV) sampled at 5-minute frequency is used as a proxy for the actual volatility. RV is one of the most commonly used realized volatility measure and the 5-minute sampling frequency has been shown to be small enough to accurately capture price dynamics while large enough to dampen down the adverse effects of microstructure frictions. The auxiliary regression formula is then given by:

$$y_t = \alpha + \beta \hat{\sigma}_t^2 + \varepsilon_t \quad (4.24)$$

Where  $y_t$  is the actual volatility of day  $t$  and  $\hat{\sigma}_t^2$  is the forecasted volatility. In this auxiliary regression, a perfect forecast will produce an  $\alpha$  of zero and a  $\beta$  of 1. Forecast accuracy is compared on the basis of the  $R^2$  criterion.

#### 4.3.3.5 Magnitude of Violation

For each model we present the percentage of violations, the maximum and average absolute deviations of violations from the VaR forecasts. A set of adequate VaR predictions would violate about  $\alpha N$  times spanning the forecasting period, produce relatively low average absolute deviation and small maximum absolute deviation. In addition to from comparing each volatility models based on these three measures, we will also compare the average of the three measures produced by each group of estimators, i.e. the single index and portfolio volatility specifications.

## 4.4 The Dataset

Daily prices for three major stock market indices, namely S&P 500, Russell 2000 and NASDAQ Composite, are used in the empirical study. The data were obtained from *DiskTrading* database for the period 8 January 1998 to 30 September 2011, which yields

3449 observations. The three indices are chosen for diversification purposes. S&P 500 is one of the most commonly followed equity indices, which mainly concentrates on US based large publicly held companies. NASDAQ composite is another large-cap index that incorporates large companies both in and outside the US. Majority of the stocks included in the index are technology and growth ones. Russell 2000 is a small-cap index of the bottom 2000 stocks listed in the US. The portfolio of these three equity indices largely represents the broad stock market of US listed companies.

Tick-by-tick price quotes for the same period are collected from the same source and subjected to scrubbing methodologies to detect and repair bad ticks. At the time the data were collected, this period was the longest for which high-frequency data on all three indices were available. For each trading day, we use data in the normal trading session from 9:30am to 04:00pm (390 minutes) and exclude the pre-market session from 7:00am to 9:30am and post-market session from 4:00pm to 8:00pm.

Figure 4.1 plots the price evolution of the three indices selected. For comparison purposes we rescale each series to a base of 100 on 8 January 1998. NASDAQ Composite decoupled from the other two indices during the first five years of our sample. NASDAQ increased 300% from 1998 to 2000 before its subsequent crash, which is famously known as the collapse of the dot-com bubble, around the millennium. S&P 500 and Russell 2000 were not affected too much as the emerging IT and internet companies were mostly listed in NASDAQ. Moving through the timeline we can see synchronized tumbles experienced by all the three indices at the end of 2008 caused by the credit crunch and global financial crisis. Since 2003, Russell 2000, the small cap index, mildly outperformed the two alternatives, showing strong momentum in start-up companies. However, S&P 500 bottomed the performance rank during the same period, which has barely increased value since 2008, signalling a broad stagnation in the main economy.

The plots of returns are shown in Figures 4.2A-C. Each of the index returns exhibits clustering effect, which will be accounted by appropriate time series models. The de-

scriptive statistics for the return series of the three indices are given in Table 4.1.A. All returns have similar means and medians, which are close to zero, minima which differ between -9.482 and -12.578, and maxima vary between 8.786 and 13.232. The three standard deviations differ slightly, ranging from 1.350 for S&P 500 to 1.830 for NASDAQ Composite. Skewness varies among all three indices while kurtosis is similar for all series. Jarque-Bera test strongly rejects the null hypothesis of normality for all the series, which may be caused by fat-tail positioned, or extreme, observations.

Figure 4.3 plots daily volatilities over the sample period using RV sampled at 5 minute frequency. All the series exhibits volatility clustering, which will be captured by GARCH class models. The volatility of all indices appears to be high during the 08-09 financial crisis, while only NASDAQ show a high level during the millennium. The descriptive statistics for the volatility of the three series are given in Table 4.1.B. NASDAQ Composite gives the largest mean (median) volatility at 3.348 (0.790), while S&P 500 displays the lowest mean (median) volatility at 1.822 (0.409). The maxima of the three volatility series vary significantly, with Nasdaq Composite having the highest maxima and S&P 500 showing the lowest. All volatility series display high degree of kurtosis and substantial skewness, indicating the existence of extreme observations.

## 4.5 Empirical Results

This section compares the forecasting performance of the single index and portfolio models and their intraday based extensions. The benchmark models are univariate GARCH and multivariate GARCH specifications based on daily data. The realized models are the benchmark ones augmented with a realized kernel (RK) estimator. We first assess the volatility models' out-of-sample forecasting performance.

Although the previous chapter suggests that an actively managed portfolio produces higher economic value, in this empirical study it is assumed that the portfolio weights are equal and constant over time. There are two reasons behind the simplification. First,

with the simplified assumption, we can compare our empirical results directly with the empirical literature. The case of a long position in an equally-weighted portfolio has been extensively applied in the literature; see, for instance, Veiga and McAleer (2005), DeMiguel et al. (2009) and Santos et al. (2009).

Second, in the case of an actively managed portfolio, a model's volatility forecasts not only affect the VaR predictions but also the portfolio weights, hence producing a unique portfolio in terms of asset weights for every volatility specification. Since it is more straightforward to compare the VaR adequacy of different volatility models based on a portfolio with identical assets and weights, we assume equal weights to avoid additional complications. Nevertheless, these assumptions can be relaxed to jointly assess a volatility model's performance on economic value and VaR adequacy, although that is beyond the scope of this empirical study.

To make the comparison of univariate and multivariate models clear, we first compare the performance of two representative models from each group of specifications, namely GARCH, RGARCH, DCC, and RDCC. Backtesting results of the alternative single index (GJRGARCH, PGARCH, RGJRGARCH, and RPGARCH) and portfolio models (CCC, DCC, RCCC, and RDCC) will be provided as supplement information for the comparative study.

#### **4.5.1 Estimation Results of the Single Index and Portfolio Models**

The left columns in Table 4.2 report the resulting univariate GARCH models' parameter estimates and regression statistics. The numbers in square brackets are the p-values of the corresponding parameters. In estimation we do not impose artificial constraints on the signs of individual coefficient. As a result, some of the parameters of the lagged squared error in the conditional variance equations are shown to be negative. However, none of the negative parameters are proved to be statistically significant. All other estimated parameters are strictly positive, including those of the asymmetric term and realized

volatility. Leverage effects are detected in both the TGARCH and RTGARCH, implying that a negative surprise will have greater impact on volatility than a positive shock of the same magnitude. Highly significant coefficients of the realized volatility in all the realized models statistically justify the use of intraday information.

Table 4.3 and 4.4 show the resulting MGARCH and RMGARCH models' parameter estimates, together with robust standard errors in parentheses and p-values in square brackets. The subscripts 1, 2 and 3 denote S&P 500, Russell 2000 and NASDAQ 100, respectively. The coefficients of all the models are largely significant at 5% significance level, showing satisfactory estimation results. Different from the results in Chapter 3, correlation estimates for both the MGARCH and RMGARCH models in this study are highly significant. This is because the data now include three highly liquid stock indices in the U.S. market, while not including the previously used CRB commodity index. The correlations between these stock market indices have been pretty strong and the multivariate covariance models are capable at estimating them. In line with the estimation results of Chapter 3, the estimation output confirms the significance of all the coefficients of the realized covariance matrix at 5% significance level, suggesting that the intraday information have significant explanatory power in the corresponding RMGARCH equations. Nevertheless, for some models the parameters of intercept and previous squared errors are not statistically significant, which is similar to the single-index case.

### **4.5.2 Out-of-Sample VaR Backtesting**

The forecast period includes the last 947 days of our sample, covering a period from 2 January 2008 to 30 September 2011. The reason to produce forecasts for the prolonged period is to include the days with dramatic market volatility during the worst financial crisis (2008-2009) since the great depression. The spectacular price movements in the wake of the financial crisis challenges the adequacy of VaR forecasts predicted by the volatility models studied in this chapter. A consistent VaR forecasting performance

for the volatile forecasting period would enhance our confidence in recommending the corresponding volatility model.

Table 4.5 presents the conditional coverage test results for the four representative volatility models (GARCH, DCC, RGARCH and RDCC) considering a coverage rate of 1%. Both the GARCH and DCC specifications fail the conditional coverage test by providing high test statistics far exceeding the 1% significance level following a  $\chi(2)$  distribution. They lead to significantly greater number of violations than expected, ranging from DCC's 2.64% to GARCH's 2.85%. The models also fail the conditional coverage test for the same reason.

The performance of the daily models is coincident with what happened during the 2008-2009 financial crisis. The risk taken by the largest banks and investment firms in much of the Western world were too excessive which threatened to bring down the entire financial system. Inadequately calculated VaR is possibly one of the causes of the financial turmoil. VaR models have been criticized for ignoring the possibility of a systemic financial meltdown and the magnitude of losses when the 1% or 5% (commonly applied levels of significance chosen for the VaR) events occur. Although the validity of VaR as a risk measure is beyond the scope of this study, we can at least find that standard daily VaR models generate biased VaR estimates for the portfolio returns and therefore are indeed inadequate in forecasting daily VaR over the volatile forecasting period.

In contrast, the two realized models, RGARCH and RDCC, both produce adequate VaR predictions according to the conditional coverage test results, signalling practical value in incorporating realized estimator into standard GARCH-type models. Both specifications produce insignificant  $LR_{CC}$  test statistics at 1% significance level. Therefore we do not reject the null hypothesis that the number of observed violations divided by the number of trading days is equal to  $\alpha$ , the level of significance chosen for the VaR. In terms of the percentage of violations, the conditional coverage tests suggest that the daily models leads to excessive violations (2.85% for GARCH and 2.64% for DCC) while

the realized models leads to too few violations (0.64% for GARCH and 0.74% for DCC).

The test results indicate that VaR forecasts generated by both the single index and portfolio models could be improved significantly if intraday information is utilized. And the risk measurement for the forecast period is now, statistically speaking, adequate. Given its relative superior forecast accuracy, VaR estimated by intraday based models is arguably a useful tool for risk management. It could at least be used to assist risk managers understand what they should expect to happen on a daily basis in an environment that is roughly the same. In addition, when VaR started to "miss" on a regular basis, it might imply that the market is experiencing a structural change, hence signalling need for risk-off adjustments (to reduce the proportion of risky assets in an investment portfolio).

Let us now focus on the contest between single index and portfolio models in a high-frequency context. The finding that RDCC produces slightly more accurate percentage of hits out-of-sample (0.74%), better than RGARCH's 0.64%, is not statistically significant according to the results of the  $LR_{CC}$  test. This conclusion is in line with the performance of daily GARCH and DCC studied by Dumitrescu et al. (2012), who find that the two volatility models statistically identical. In other words, the correlation amongst assets does not have any impact on the calculation of the VaR and implicitly on its validity.

Furthermore, the average absolute deviation of violations for RGARCH (0.24%) is lower than that of the RDCC (0.52%), and we observe the same pattern in the maximum and median absolute deviation of violations. Therefore, the result of the comparison between RGARCH and RDCC is not a clear cut. This result is in line with the findings of Berkowitz and O'Brien (2002). Their comparison of single index and portfolio models in a daily context show that despite the detailed information used in the multivariate specifications, VaR forecasts generated by portfolio models do not outperform the forecast made by a standard univariate GARCH model, which provides lower VaRs and is better at predicting changes in volatility.

Table 4.6 gives the corresponding  $p$ -values of the  $DQ$  and  $DB$  test results for the four

representative models considering a coverage rate of 1%. The results vary severely from one specification to another, but the conclusion is the same: The realized models provide superior VaR forecasts compared with their daily counterparts. All the  $DQ$  and  $DB$  test results show highly significant  $p$ -values for GARCH and DCC models, rejecting the null hypothesis of conditional coverage at 1% significance level. So the daily models are again proved to be incapable at forecasting VaR in the period of extraordinary volatility. In contrast, the validity of RGARCH and RDCC model is reconfirmed by the tests, with  $p$ -values surpassing 26% for all the cases, showing that the two realized models did a good job describing the evolution of the left tail for the portfolio under study.

For the comparison of RGARCH and RDCC, the results of the  $DQ$  and  $DB$  tests are in line with what we found in the  $LR_{CC}$  test. RDCC fares better in general according to the  $p$ -values, but not statistically significant, which means that we cannot make a definite conclusion on the comparison between the two realized volatility models.

Next, we extend our study by testing more realized specifications to see whether the conclusion drawn on the representative intraday based models is universal. Table 4.7 represents the  $p$ -values of the  $LR_{CC}$ ,  $DQ$  and  $DB$  tests for four additional realized models: two single index ones (RGJRGARCH and RPGARCH) and two portfolio models (RBEKK and RCCC) All the additional realized specifications pass the 6 conditional coverage tests with  $p$ -values higher than 30%, signalling validity of the models in forecasting VaR. Similar to the result of GARCH and DCC, the validity of the daily versions of the additional models, namely GJRGARCH, PGARCH, BEKK and CCC, is rejected by the conditional tests. All the daily models tend to provide excessive and often serially dependent violations according to the test results of out-of-sample forecasting. The results of daily models are therefore not provided here given that our main focus is on the comparison of realized single index and portfolio models.

In terms of the superiority of VaR forecasts, the test results show that the four realized models are too similar for the  $p$ -values to discriminate among them. Again, the correlation

amongst assets does not play an important role on the calculation of the VaR forecasts. In conclusion, the difference between the VaR forecasts generated by realized single index specifications and portfolio models is not statistically significant.

### 4.5.3 Robustness Checks

Robustness tests are composed of in-sample VaR prediction for the whole sample period of 3449 trading days, and out-of-sample VaR forecasting for the last 947-day sample period using forecasts relying on Student- $t$  rather than Normal densities for the quantile computation. The in-sample fitting consists of 10 years of data which is aimed at providing a broader assessment of the forecasting power of different volatility models. Student- $t$  distribution is often used in VaR prediction since the distribution contains fatter tails that are suitable for modelling volatility of financial asset returns. At last, the out-of-sample forecasting performance of the volatility models are assessed by a linear regression method.

#### 4.5.3.1 In Sample Fitting

Table 4.8.A provides the  $p$ -values of the  $LR_{CC}$ ,  $DQ$  and four  $DB$  tests for the four representative models, GARCH, DCC, RGARCH and RDCC based on in-sample fitting. The in-sample period covers 10 years from 1998 to 2007, which is relatively calmer than the volatile out-of-sample period since 2008. The in-sample fitting performance judged by the conditional coverage tests show that both the daily and realized models are able to describe the evolution of the left tail for the portfolio under study. The daily models performed better than in the out-of-sample case by providing  $p$ -values surpassing the 5% critical value for all the 6 conditional coverage tests. This result shows that the daily models are capable at modelling VaR for a relatively less volatile period.

The realized models once again provide superior performance compared with the daily models in the in-sample fitting. However, the contrast between the two groups becomes smaller. This is partly due to an improvement in the performance of daily models, and

partly attributed to too few violations in the VaR forecasts made by intraday-based portfolio models over the 10 years period. Nonetheless, the conditional coverage tests for RGARCH and RDCC produce  $p$ -values higher than 30%, implying validity of the realized models. In a high-frequency context, RGARCH and RDCC produce very similar quality of in-sample forecasts, which once again cannot be separated statistically according to the backtesting evaluation. This result is in line with the conclusion drawn on the case of out-of-sample forecasting.

#### 4.5.3.2 Student- $t$ Equity Returns

It is well-known that the Normal density tends to underestimate the probability of extreme variations (the collapse of the hedge fund, Long Term Capital Management, was a particular reminder of this; see Jorion (2000)). Consequently we evaluate the VaR forecasts relying on Student- $t$  densities, the most commonly used fat-tailed distribution as a model for asset returns, for the quantile  $\hat{F}_\varepsilon^{-1}(\alpha)$  computation. In detail, we use standardized (unit variance) Student- $t$  distribution with degrees of freedom parameter (4.04) estimated by ML from the in-sample standardized returns. The model estimation is now based on student- $t$  densities. Table 4.8.B presents the model performance judged by  $LR_{CC}$ ,  $DQ$  and four  $DB$  conditional coverage tests for the out-of-sample forecasting period. Daily models once again generate too many violations for the period and therefore are incapable at prediction accurate VaR. Both the GARCH- $t$  and DCC- $t$  model fail all the tests.

Table 4.8.B shows that there is significant improvement in VaR predictions from using daily models to intraday models. RGARCH- $t$  and RDCC- $t$  produce adequate VaR forecasts for the forecasting period and pass all the backtesting criteria. But, no one specification is statistically superior between the two realized models, which leads to the same conclusion drawn under normal densities for the quantile  $\hat{F}_\varepsilon^{-1}(\alpha)$  computation.

### 4.5.3.3 Linear Regression Test

Table 4.9 presents the out-of-sample forecasting performance of the 4 representative models assessed by the linear regression test. The test is applied to complement the backtesting procedures by directly assessing the forecasting power of the volatility models. The P-values in Table 4.9 shows the probability values of the joint coefficient test of  $\alpha = 0$  and  $\beta = 1$ . Both the GARCH and DCC test results show insignificant p-values. In contrast, both intraday based models, RDCC and RGARCH present highly significant p-values and as a result the null hypothesis of unbiasedness is rejected. However, Realized specifications outperform their daily counterparts according to the  $R^2$  generated. RDCC performs the best by generating a  $R^2 = 0.489$ . RGARCH ranks the second by producing a  $R^2$  of 0.432. In contrast, the 2 daily models, GARCH and DCC, generate substantially lower  $R^2$ s (0.226 and 0.218 for DCC and GARCH respectively) hence cannot compete with their intraday counterparts. This inconsistency in test results shows that the intraday based volatility models are not perfect in theory. However, judged by practice value, the realized models are the better ones as they produce more accurate volatility forecasts. Indeed, the strong forecast power is translated into more adequate VaR predictions. The result of the linear regression test further confirms the practical value of the two intraday based volatility specifications from a different viewpoint.

## 4.6 Conclusion

This study compares the quality of VaR forecasts between single index models and portfolio models in a high-frequency context. A number of daily univariate and multivariate GARCH class models, namely GARCH, GJRGARCH, PGARCH, BEKK, CCC and DCC, are used as the benchmark models in the chapter. The first three models are single index ones and the remaining are portfolio models. These daily models are augmented with a realized kernel (RK) estimator to construct the realized models, namely RGARCH, RGJRGARCH, RPGARCH, RBEKK, RCCC and RDCC. Three main backtesting crite-

ria are used to compare the forecasting performance of the conditional volatility models: the  $LR$  conditional coverage ( $LR_{CC}$ ) test, the dynamic quantile ( $DQ$ ) test, and the dynamic binary response ( $DB$ ) model. In order to make the comparison clear, we focused our analysis on two representative daily models, GARCH and DCC, and two intraday based models, RGARCH and RDCC.

We first study whether intraday volatility models can significantly improve the predictive power of classical daily GARCH models. An intraday volatility estimator may contain unique information about past price variations that are not captured by daily squared returns. A volatility model incorporated with a realized (intraday) estimator therefore could provide superior forecasting performance. Our finding confirms this hypothesis. The realized models outperform their daily counterparts by providing adequate VaR forecasts for a prolonged and volatile out-of-sample forecasting period. The four intraday based volatility models all produce reasonable VaR predictions according to the conditional coverage test result, signalling practical value of incorporating realized estimator into standard GARCH-type models. Daily models, in contrast, fail all the conditional coverage tests, implying that they are incapable at forecasting VaR for the volatile out-of-sample forecasting horizon.

The comparison between realized single index and portfolio models, however, shows mixed results. For example, the RDCC model produces slightly more accurate percentage of hits out-of-samples (0.74%) compared with RGARCH's 0.64%. However, the advantage is not statistically significant according to the results of the conditional coverage tests. This conclusion is similar to the result of daily GARCH and DCC comparison studied by Dumitrescu (2012). In addition, the average absolute deviation of violations for RGARCH (0.24%) is lower than that of the RDCC (0.52%), and we observe the same pattern in the maximum absolute deviation of violations. This result is in line with the findings of Berkowitz and O'Brien (2002). In a word, the comparison between RGARCH and RDCC is not a clear cut.

To sum up, augmented single index models generally perform as good performance compared to their portfolio counterparts. It seems that the correlation amongst assets does not have significant impact on the calculation of the VaR and implicitly on its validity. This is possibly because while assessing more information and estimating more parameters, portfolio models are also subject to greater noise in individual asset variances and covariances. In addition, further noise in data is added by the incorporating of realized variances and covariances into the portfolio models. In conclusion, both standard univariate and multivariate GARCH class models augmented with a well-structured realized estimator can provide adequate out-of-sample VaR forecasts. Nevertheless, given the parsimonious nature of the realized single index models, they are probably more suited for forecasting VaR in daily practices.

TABLE 4.1  
Descriptive Statistics for Returns and Volatility

A. Returns			
	S&P 500	Russell 2000	NASDAQ Composite
Mean	0.004	0.011	0.012
Median	0.061	0.074	0.115
Maximum	10.771	8.786	13.232
Minimum	-9.482	-12.578	-10.207
Std. Dev.	1.350	1.622	1.830
Skewness	-0.178	-0.321	-0.010
Kurtosis	9.722	7.347	7.110
Jarque-Bera	6514	2776	2429
B. Volatility of Returns			
	S&P 500	Russell 2000	NASDAQ Composite
Mean	1.822	2.629	3.348
Median	0.409	0.742	0.790
Maximum	116.020	158.200	175.091
Minimum	0.000	0.000	0.000
Std. Dev.	5.379	6.619	8.276
Skewness	10.371	9.038	7.908
Kurtosis	161.143	138.057	103.209
Jarque-Bera	3657955	2669798	1479899

Notes: Panel A provides summary statistics for the returns of the S&P 500, Russell 2000 and NASDAQ Composite indices. Panel B provides summary statistics for the volatility of returns of the S&P 500, Russell 2000 and NASDAQ Composite indices. The sample period is 8 January 1998 to 30 September 2011 (3449) Observations.

TABLE 4.2  
Estimation Results of Single Index Models

	GARCH		RGARCH		TGARCH		RTGARCH		PGARCH		RPGARCH	
$c$	0.015	[0.000]	0.038	[0.002]	0.015	[0.000]	0.044	[0.000]	0.016	[0.000]	0.105	[0.002]
$\alpha$	0.069	[0.000]	-0.010	[0.410]	-0.006	[0.413]	-0.109	[0.000]	0.071	[0.000]	-0.021	[0.130]
	0.921	[0.000]	0.728	[0.000]	0.939	[0.000]	0.750	[0.000]	0.922	[0.000]	0.723	[0.000]
$\gamma$	-	-	-	-	0.109	[0.000]	0.176	[0.000]	-	-	-	-
$\rho$	-	-	-	-	-	-	-	-	1.851	[0.000]	1.482	[0.000]
$rv$	-	-	0.465	[0.000]	-	-	0.434	[0.000]	-	-	0.343	[0.000]
<i>S.E.</i>	1.212		1.212		1.212		1.212		1.212		1.212	
<i>Log L</i>	-3794		-3754		-3757		-3726		-3794		-3751	
<i>DW</i>	2.005		2.006		2.006		2.006		2.005		2.006	
<i>AIC</i>	3.011		2.981		2.982		2.960		3.012		2.979	
<i>Obs</i>	2523		2523		2523		2523		2522		2522	

Notes: The table provides estimation results of the univariate volatility forecasting models.  $c$  represents the intercept of the conditional variance equation.  $\alpha$  represents the coefficients of the previous squared errors.  $\beta$  represents the coefficients of the previous conditional variances.  $\rho$  represents the power coefficient of the PGARCH specification. The upper panel reports the resulting parameter estimates for GARCH, TGARCH and PGARCH and their realized counterparts, together with robust standard errors in parentheses and p-values in square brackets. The lower panel reports the corresponding regression outputs. The modeling period is 8 January 1998 to 30 September 2011.

TABLE 4.3  
Estimation Results of MGARCH Models

	BEKK			CCC			DCC		
$c_1$	0.093	(0.012)	[0.000]	0.014	(0.005)	[0.004]	0.035	(0.011)	[0.002]
$\alpha_1$	0.267	(0.013)	[0.000]	0.086	(0.011)	[0.000]	0.071	(0.011)	[0.000]
$\beta_1$	0.962	(0.004)	[0.000]	0.907	(0.012)	[0.000]	0.906	(0.015)	[0.000]
$c_2$	0.073	(0.007)	[0.000]	0.036	(0.008)	[0.000]	0.004	(0.003)	[0.159]
$\alpha_2$	0.254	(0.011)	[0.000]	0.099	(0.011)	[0.000]	0.040	(0.011)	[0.000]
$\beta_2$	0.963	(0.003)	[0.000]	0.884	(0.012)	[0.000]	0.959	(0.011)	[0.000]
$c_3$	0.033	(0.009)	[0.000]	0.011	(0.004)	[0.004]	0.002	(0.002)	[0.382]
$\alpha_3$	0.222	(0.012)	[0.000]	0.083	(0.013)	[0.000]	0.023	(0.014)	[0.099]
$\beta_3$	0.973	(0.002)	[0.000]	0.915	(0.012)	[0.000]	0.973	(0.018)	[0.000]
$\rho_{12}$		-	-	0.856	(0.004)	[0.000]	0.862	(0.029)	[0.000]
$\rho_{13}$		-	-	0.827	(0.006)	[0.000]	0.807	(0.038)	[0.000]
$\rho_{23}$		-	-	0.847	(0.005)	[0.000]	0.842	(0.033)	[0.000]
$a$		-	-		-	-	0.015	(0.005)	[0.006]
$b$		-	-		-	-	0.983	(0.008)	[0.000]
Log $L$	-11687.818			-11861.571			-11616.427		
$Obs$	2523			2523			2523		

Notes: The table provides estimation results of the conventional daily data based covariance forecasting models.  $c_1$ ,  $c_2$  and  $c_3$  represent the diagonal elements of the intercept matrix in the covariance equation for S&P 500, Russell 2000 and NASDAQ Composite indices respectively.  $\alpha_1$ ,  $\alpha_2$  and  $\alpha_3$  represent the coefficients of the previous squared errors and cross-product of errors.  $\beta_1$ ,  $\beta_2$  and  $\beta_3$  represent the coefficients of the previous conditional variances and covariances.  $\rho_{12}$ ,  $\rho_{13}$  and  $\rho_{23}$  represent the estimated correlation coefficients. The left column reports the resulting parameter estimates for BEKK, CCC and DCC, together with robust standard errors in parentheses and p-values in square brackets. The modeling period is 8 January 1998 to 30 September 2011.

TABLE 4.4  
Estimation Results of Intraday Based MGARCH Models

	RBEKK			RCCC			RDCC		
$c_1$	0.215	(0.037)	[0.000]	0.015	(0.008)	[0.068]	0.015	(0.008)	[0.068]
$\alpha_1$	0.252	(0.020)	[0.000]	0.000	(0.000)	[0.999]	0.000	(0.000)	[0.999]
$\gamma_1$	0.257	(0.044)	[0.000]	0.383	(0.060)	[0.000]	0.383	(0.060)	[0.000]
$\beta_1$	0.917	(0.017)	[0.000]	0.701	(0.044)	[0.000]	0.701	(0.044)	[0.000]
$c_2$	0.072	(0.047)	[0.129]	0.055	(0.015)	[0.000]	0.055	(0.015)	[0.000]
$\alpha_2$	0.266	(0.027)	[0.000]	0.073	(0.013)	[0.000]	0.073	(0.013)	[0.000]
$\gamma_2$	0.306	(0.083)	[0.000]	0.201	(0.059)	[0.001]	0.201	(0.059)	[0.001]
$\beta_2$	0.922	(0.028)	[0.000]	0.799	(0.034)	[0.000]	0.799	(0.034)	[0.000]
$c_3$	0.000	(0.000)	[0.280]	0.032	(0.017)	[0.056]	0.032	(0.017)	[0.056]
$\alpha_3$	0.240	(0.043)	[0.000]	0.009	(0.022)	[0.679]	0.009	(0.022)	[0.679]
$\gamma_3$	0.317	(0.151)	[0.036]	0.423	(0.145)	[0.003]	0.423	(0.145)	[0.003]
$\beta_3$	0.920	(0.055)	[0.000]	0.707	(0.080)	[0.000]	0.707	(0.080)	[0.000]
$\rho_{12}$		-	-	0.848	(0.021)	[0.000]	0.856	(0.026)	[0.000]
$\rho_{13}$		-	-	0.824	(0.022)	[0.000]	0.826	(0.025)	[0.000]
$\rho_{23}$		-	-	0.836	(0.021)	[0.000]	0.847	(0.023)	[0.000]
$a$		-	-		-	-	0.034	(0.006)	[0.000]
$b$		-	-		-	-	0.905	(0.028)	[0.000]
Log $L$	-11562.855			-11807.476			-11605.577		
Obs	2523			2523			2523		

Notes: The table provides estimation results of the intraday data based covariance forecasting models.  $c_1$ ,  $c_2$  and  $c_3$  represent the diagonal elements of the intercept matrix in the covariance equation for S&P 500, Russell 2000 and NASDAQ Composite indices respectively.  $\alpha_1$ ,  $\alpha_2$  and  $\alpha_3$  represent the coefficients of the previous squared errors and cross-product of errors.  $\beta_1$ ,  $\beta_2$  and  $\beta_3$  represent the coefficients of the previous conditional variances and covariances.  $\gamma_1$ ,  $\gamma_2$  and  $\gamma_3$  represent the coefficients of the realized covariance matrix. The left column reports the resulting parameter estimates for RBEKK, RCCC and RDCC, together with robust standard errors in parentheses and p-values in square brackets. The modeling period is 8 January 1998 to 30 September 2011.

TABLE 4.5  
LR<sub>CC</sub> Test for GARCH, RGARCH, DCC and RDCC (Out-of-Sample)

	LR <sub>CC</sub>	p-value	% of Violations	V.Mean	V.Maxima	V.Median
RDCC	0.710	0.701	0.007	0.521	1.736	0.352
RGARCH	1.265	0.531	0.006	0.241	0.720	0.128
DCC	17.770	0.000	0.026	0.540	2.432	0.400
GARCH	21.884	0.000	0.029	0.627	3.578	0.325

Notes: LR<sub>CC</sub> stands for the LR conditional coverage test. % of Violation is the percentage of violation occurred based on the VaR forecasts generated by the corresponding model. V.Mean stands for the average absolute deviation of violation, V.Maxima stands for the maximum absolute deviation of violation, and V.Median stands for the median absolute deviation of violation.

TABLE 4.6  
DQ and DB Test for GARCH, RGARCH, DCC and RDCC (Out-of-Sample)

	DQ	DB <sub>1</sub>	DB <sub>2</sub>	DB <sub>3</sub>	DB <sub>4</sub>
RDCC	0.690	0.517	0.696	0.806	0.527
RGARCH	0.451	0.467	0.565	0.696	0.480
DCC	0.000	0.001	0.001	0.001	0.000
GARCH	0.000	0.000	0.000	0.000	0.000

Notes: DQ stands for the dynamic quantile test and DB<sub>1</sub> to DB<sub>4</sub> are the different versions of the dynamic binary test. The table gives the corresponding *p*-values of the test results for the models considering a coverage rate of 1%.

TABLE 4.7  
LR<sub>CC</sub>, DQ and DB Test for RGJRGARCH, RPGARCH, RCCC and RBEKK (Out-of-Sample)

	LR <sub>CC</sub>	DQ	DB <sub>1</sub>	DB <sub>2</sub>	DB <sub>3</sub>	DB <sub>4</sub>
RCCC	0.637	0.590	0.516	0.672	0.632	0.611
RBEKK	0.583	0.516	0.489	0.562	0.617	0.575
RGJRGARCH	0.569	0.496	0.486	0.636	0.586	0.564
RPGARCH	0.411	0.377	0.399	0.406	0.412	0.308

Notes: The table provides the LR<sub>CC</sub>, DQ and DB conditional coverage test results for out-of-sample VaR forecasts generated by four realized models RGJRGARCH, RPGARCH, RCCC and RBEKK. See note to Table 4.5.

TABLE 4.8  
LRCC, DQ and DB Test for RDCC, RGARCH, DCC and GARCH

A. In Sample						
	LR <sub>CC</sub>	DQ	DB <sub>1</sub>	DB <sub>2</sub>	DB <sub>3</sub>	DB <sub>4</sub>
RDCC	0.482	0.316	0.514	0.550	0.562	0.511
RGARCH	0.368	0.256	0.353	0.344	0.526	0.410
DCC	0.153	0.121	0.162	0.171	0.140	0.162
GARCH	0.125	0.094	0.136	0.162	0.133	0.183
B. Out-of-Sample, Student- <i>t</i> Densities						
	LR <sub>CC</sub>	DQ	DB <sub>1</sub>	DB <sub>2</sub>	DB <sub>3</sub>	DB <sub>4</sub>
RDCC- <i>t</i>	0.609	0.355	0.596	0.537	0.501	0.577
RGARCH- <i>t</i>	0.438	0.312	0.408	0.511	0.450	0.393
DCC- <i>t</i>	0.000	0.000	0.000	0.000	0.000	0.001
GARCH- <i>t</i>	0.000	0.000	0.000	0.000	0.000	0.000

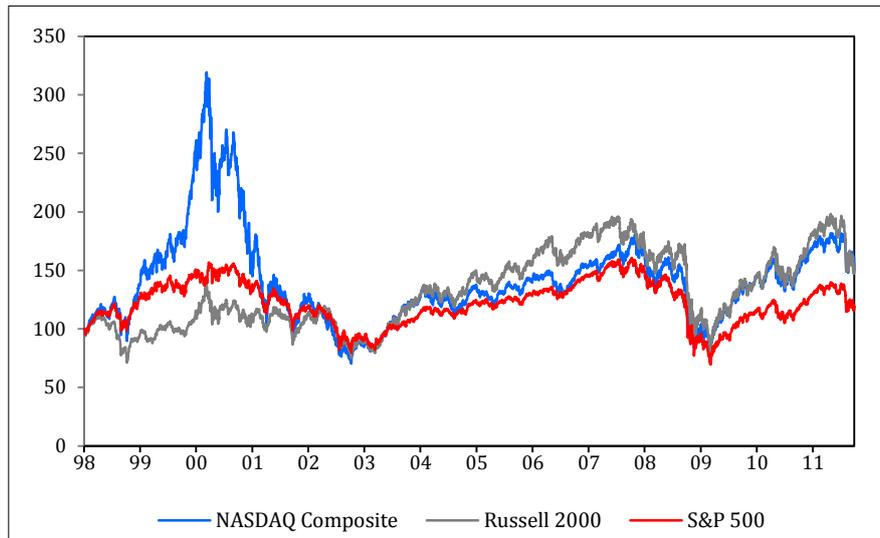
Notes: Panel A provides the LR<sub>CC</sub>, DQ and DB conditional coverage test results for in-sample VaR forecasts generated by four representative models RDCC, RGARCH, DCC and GARCH. Panel B provides the LR<sub>CC</sub>, DQ and DB conditional coverage test results for out-of-sample VaR forecasts generated by four representative models RDCC-*t*, RGARCH-*t*, DCC-*t* and GARCH-*t* based on the student-*t* densities. See note to Table 4.5.

TABLE 4.9  
Linear Regression Test Results for GARCH, RGARCH, DCC and RDCC (out-of-sample)

	alpha	<i>t</i> -ratio	beta	<i>t</i> -ratio	P-value	R <sup>2</sup>
RDCC	0.141	0.565	1.163	4.204	[0.000]	0.489
RGARCH	-0.038	-0.142	1.389	7.505	[0.002]	0.433
DCC	0.183	0.531	1.033	0.526	[0.553]	0.227
GARCH	0.275	0.797	0.979	-0.346	[0.714]	0.218

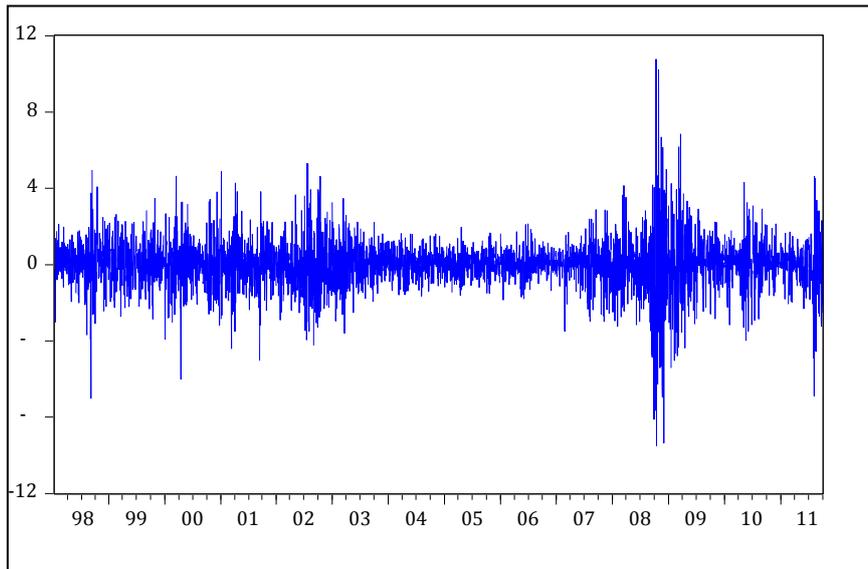
Notes: See equation (4.24) for the linear regression formula. The null hypothesis  $\alpha = 0$  is tested against the alternative  $\alpha \neq 0$ . The null hypothesis  $\beta = 1$  is tested against the alternative  $\beta \neq 1$ . The P-value represents the results of a joint coefficient test. The null hypothesis  $\alpha = 0$  and  $\beta = 1$  is tested against the alternative  $\alpha \neq 0$  and  $\beta \neq 1$ .

FIGURE 4.1  
Rebased Index Prices

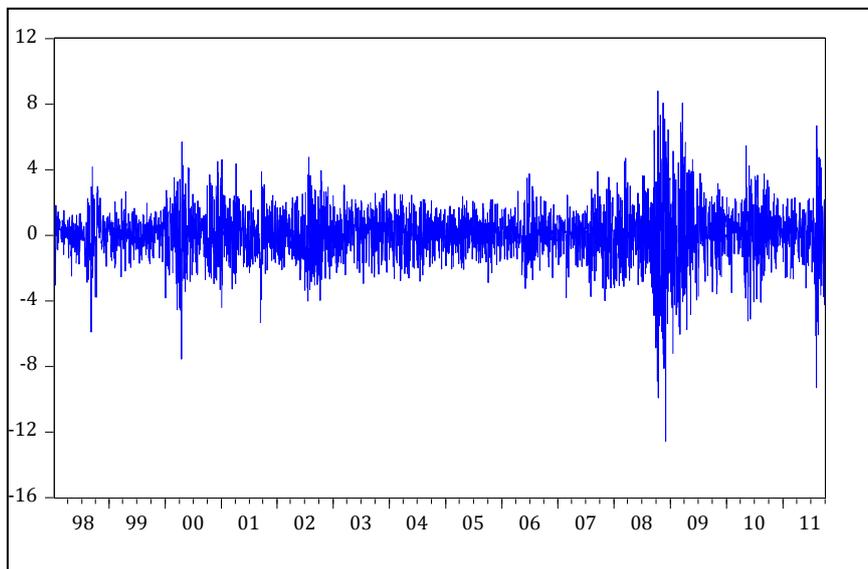


Notes: The figure illustrates the daily return of NASDAQ Composite, Russell 2000 and S&P 500 indices over the whole sample period. Indices prices are rebased at 100 at the beginning of 1998 for comparative purposes.

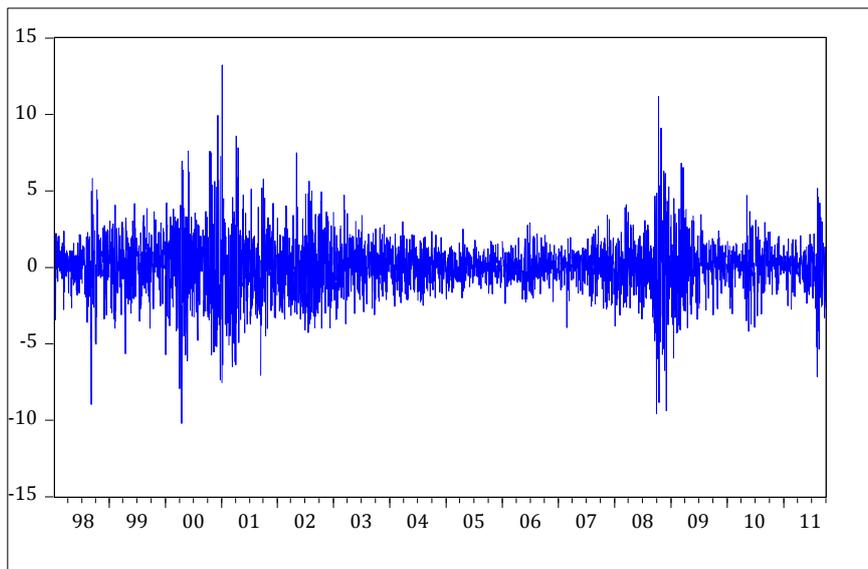
FIGURE 4.2  
Return Series of NASDAQ Composite, Russell 2000 and S&P 500 Indices



A) S&P 500 Return



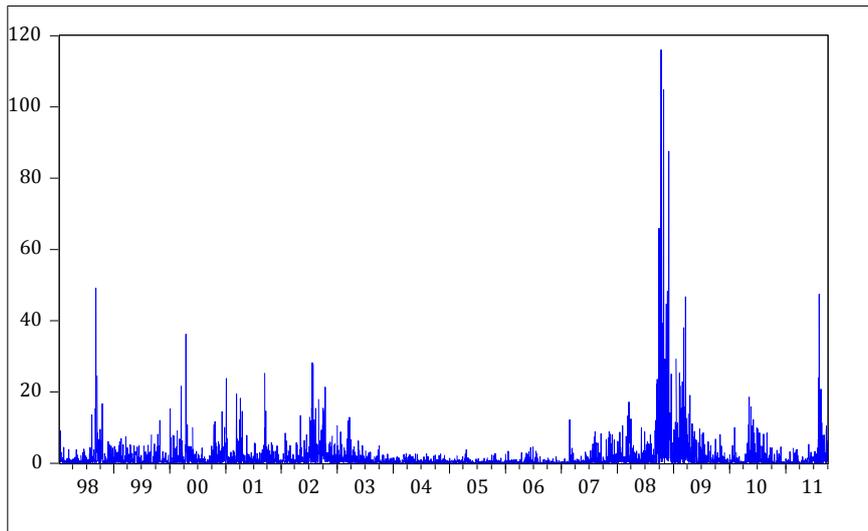
B) Russell 2000 Return



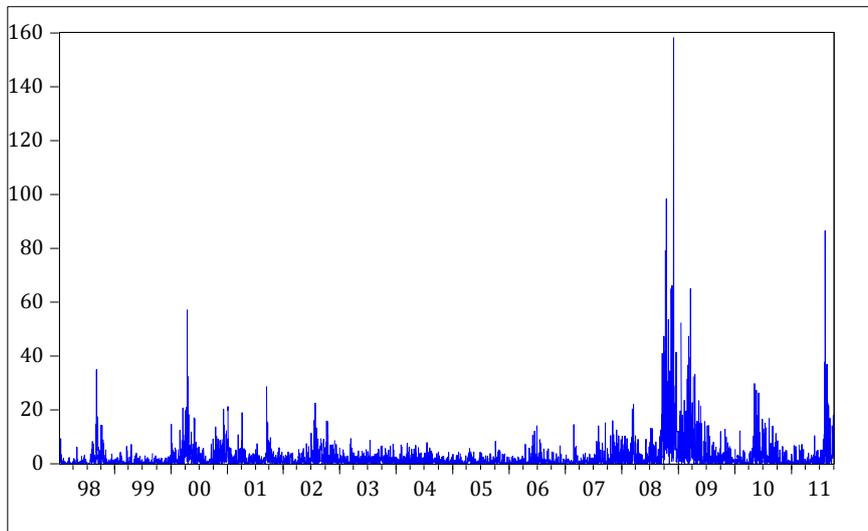
C) NASDAQ Composite Return

Notes: Daily return plot of the three indices from 8 January 1998 to 30 September 2011 (3449) Observations.

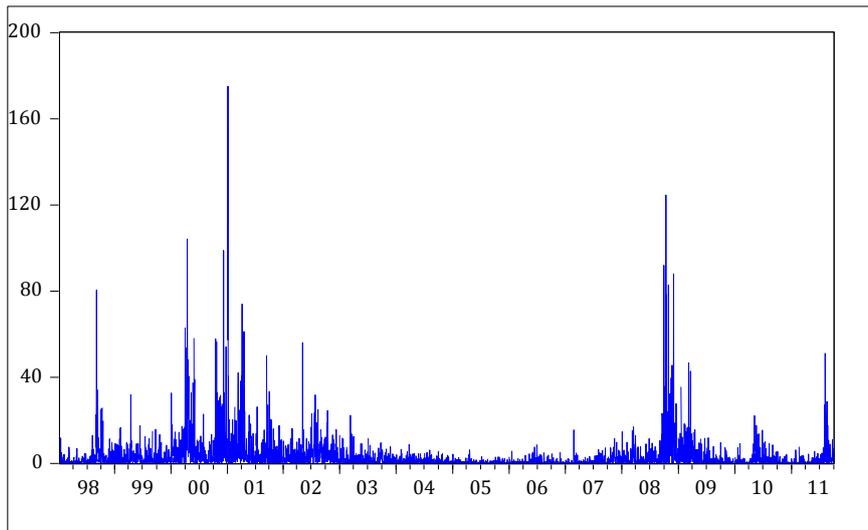
FIGURE 4.3  
Volatility of NASDAQ Composite, Russell 2000 and S&P 500 Indices



A) S&P 500 Volatility



B) Russell 2000 Volatility



B) NASDAQ Composite Volatility

Notes: Daily squared return plot of the three indices from 8 January 1998 to 30 September 2011 (3449) Observations.

# CHAPTER 5: The Economic Value of Combining Intraday Volatility Estimators and Implied Volatility

## 5.1 Introduction

The largest source of edge in volatility trading is in trading our predicted volatility against the market's volatility expectation. The predicted future volatility can be obtained either by volatility forecasts generated by econometric models, or implied volatility derived from option pricing models. Given the practical importance of volatility forecasting, the contest between the two forecast categories, model predicted and implied volatility, has attracted great attention in both academia and the financial industry.

Empirical studies have provided mixed results for the volatility forecasting contention. Some suggest that implied volatility estimates generally provide better volatility predictions, and model predicted volatility is unable to add to the information already contained in market forecasts (see Poon and Granger, 2003; Kinlay, 2005). In contrast, several papers indicate that implied volatility forecasts show evidence of a consistent and substantial bias, and they do not contain any information relevant to future volatility beyond that reflected in model-forecasted volatility (see Kinlay, 2005; Becker, Clements and White, 2007).

Given the mixed stories shown in the literature, it might be worth tackling the problem from a different perspective, which is forecast combination. If model-forecasted and implied volatility each contains unique information about future volatility, a combination estimator that combines the two individual volatility forecasts may provide superior statistical accuracy than any of the individual models. A couple of studies have examined this combination recently. For instance, Becker and Clements (2008) combine the implied volatility and econometric model forecasts and indeed find the combination forecasts, equal weighted and regression-based, to be statistically superior.

Forecast combination has become a popular research subject since the initial work

of Bates and Granger (1969), and has been shown to be able to produce more accurate volatility forecasts than individual models by many empirical studies. Just as a board of diverse people tends to make better decisions than each individual does, forecast combination tends to perform better than a single forecast. Hendry and Clements (2004) suggest that forecast combination offers a “shield” against misspecification and structural breaks.

The literature to date on combining forecasted volatility and implied volatility is mainly based on daily data. However, it is worthwhile to examine the forecasting performance in a realized context as recent studies show a significant improvement in forecast accuracy when intraday information is incorporated into the daily volatility models (see Fuertes, Kalotychou and Izzeldin, 2009).

Furthermore, the literature has, to our knowledge, evaluated only the statistical accuracy of combination estimators that combine model-forecasted volatility and implied volatility. A separate question is whether the advantages in statistical accuracy are sufficient to have a meaningful impact on decisions that based on conditional volatility estimates. Given the central role of volatility forecasting in volatility trading, a direct profit and risk evaluation of the trading strategies using different volatility forecasts would arguably be a better alternative to statistical loss functions.

This study bridges the gap between the literatures by assessing the economic value of a comprehensive set of intraday volatility measures, implied volatility and their combinations. Over 12 years of tick-by-tick data for the S&P 100 index price is used and 42 individual realized and implied volatility measures are applied in the empirical study. The economic value of the individual and combination volatility estimators are assessed out-of-sample according to their performances in simulated volatility trading exercises.

The forecast combinations used in the chapter can be divided into two broad categories: simple combinations and further combinations. A simple combination forecast combines ARFIMA forecasted realized volatility measure and predicted implied volatility estimator

using ordinary least squares (OLS) regression, whereas a further combination forecast combines 42 individual estimators including 41 distinct intraday volatility estimators and the implied volatility. Five different combination approaches are used to determine the weight of each individual volatility estimator in a further combination estimator: the equal weighted, geometrical mean, regression-based, Akaike information criterion (AIC) based, and multiple indicators model (MIM).

Two research questions are addressed in the empirical study. The first one is that whether forecast combination can outperform the best possible single volatility measure in terms of economic value, after transaction cost being deducted. On the one hand, an optimal combination of the realized volatility measures with implied volatility might improve trading performance as more relevant information about price variations is incorporated into the trading signal generation process. On the other hand, the economic value of implementing forecast combination would be limited if there is no significant profit gain after transaction cost.

The second research question is that whether more sophisticated further combination models can outperform simple combination models judged by economic loss functions. The fundamental question on which the literature of forecast combination focuses on is to determine the optimal weights attributed to combination forecasts. The optimal weights depend on each model's out of sample performance, and thus should vary over time according to the changes in forecast errors. However, a significant number of empirical studies (Clemen, 1989; Stock and Watson, 1999, 2001, 2004; Hendry and Clements, 2004; Smith and Wallis, 2009; Huang and Lee, 2010; Aiolfi et al., 2010) find that the equal weighted forecast combination turns out to be hard for most sophisticated competing procedures to beat. This finding is often referred as 'forecast combination puzzle'. In this empirical study we will investigate if the puzzle exists in a realized context.

The rest of the chapter is organized as follows. Section 5.2 provides a brief literature review on forecast combination. Section 5.3 explains the volatility estimators, combination

models and the economic evaluation methods. Section 5.4 discusses the data to be used in the paper. Section 5.5 presents the results, and Section 5.6 concludes.

## **5.2 Literature Review**

The broad volatility literature can be grouped in three strands: model-forecasted, implied, and stochastic volatility. Our review below focuses primarily on model-forecasted and implied volatilities, which are both applied in this empirical study. The statistical accuracy and economic value of forecast combination are also discussed in the section.

### **5.2.1 model-forecasted Volatility versus Implied Volatility**

Volatility forecasts can either be generated by econometric models given historical data, or derived from option prices using implied volatility. The former are based on statistical adequacy and the latter is drawn on market prediction. Surveys of standard econometric volatility models can be found in Campbell, Lo and MacKinlay (1997) and Gouriéroux and Jasiak (2001). A comprehensive discussion of implied volatility is given by Jorion (1995) and Poon and Granger (2003, 2005).

Forecasting performance of econometric models and implied volatility has received a great deal of research attention. Fleming (1998) corrects for serial dependence and goes on to find that the implied volatility from S&P 100 index options produces superior forecasts of future realized volatility compared with forecasts made using historical volatility. Blair, Poon and Taylor (2001) claim almost all the useful predictive information is in option prices when forecasting S&P 100 index volatility one or more days into the future. In a similar vein, Corrado and Miller (2005) show that implied volatility of indices dominate historical index volatility in providing forecasts of future volatility in the S&P 100 and NASDAQ 100 indices. Poon and Granger (2003, 2005) summarise the results of 93 papers that study the performance of volatility forecasts. They find that implied volatility estimates generally provide better volatility predictions.

In contrast, Becker, Clements and White (2006) study the volatility forecasts in a different context and find contradicting results. They test whether the VIX, implied volatility index derived from S&P 500 option prices, contains any information relevant to future volatility beyond that reflected in forecasts generated by econometric models. The conclusion is that the VIX does not contain any such information. In addition, Taylor and Xu (1997) compare the forecast accuracy of implied and historical volatility for the Deutschemark/Dollar exchange rate and find that when intraday 5-minute returns are used, historical volatility outperforms implied volatility in predicting realized volatility. Martens and Zein (2004) suggest that high-frequency forecasts do have incremental information over that contained in implied volatility for the S&P 500 index futures and options. Pong, Shakelton, Taylor and Xu (2004) indicate that model-forecasted volatility have incremental information not found in implied volatility, for forecast horizon of up to one week.

Becker and Clements (2007) examine the contention between model-forecasted volatility and implied volatility from a different perspective by combining the implied volatility and econometric model forecasts using S&P 500 data. Two combination forecasts, equal weighted and regression-based, are applied. The result indicates that combination volatility forecasts are statistically superior to individual approaches, indicating that the VIX<sup>1</sup> cannot simply be regarded as a combination of various forecasts produced by econometric models. Therefore, each class of volatility forecasts may contain different yet relevant information in predicting future volatility.

### **5.2.2 Forecast Combination**

There is now a growing body of literature on forecast combination, since the initial work of Bates and Granger (1969) suggests that combination of forecasts is an effective approach

---

<sup>1</sup>VIX is a trademarked ticker symbol for the Chicago Board Options Exchange Market Volatility Index, a widely used measure of the implied volatility of S&P 500 index options. VIX represents the market's expectation of stock market volatility over the next 30 day period.

of improving the accuracy of forecasts regarding a certain target variable. Combination procedures, such as shown in Clemen (1989), Makridakis & Hibon (2000), Stock and Watson (2004), Becker and Clements (2008) and Bjornland et al. (2012), generally outperform the forecast generated by the best individual model. For reviews of forecast combination from a classical perspective see Clemen (1989), Newbold and Harvey (2002), and Timmermann (2006).

Timmermann (2006) summarizes three main benefits of using forecast combination. First, a well-combined forecast integrates different pieces of information attached in each individual estimator. Second, the neutralizing nature of forecast combination reduces effect of structural breaks suffered by individual forecasting models. Third, forecast combination reduces the possibility of parameter estimation errors and model misspecification. Furthermore, time series data often exhibits time-varying conditions such as regime switching and parameter drifts, make choosing the best model among various individual approaches extremely hard. Forecast combination can practically reduce these unfavourable effects since it relies on a pool of different models.

Since the introduction of realized volatility by Andersen and Bollerslev (1998), a number of recent studies have explored the benefit of combining intraday volatility estimators. Given there is no consensus about a "best" measure of volatility, Engle and Gallo (2006) propose to jointly consider absolute daily returns, daily high-low range and daily realized volatility to develop a forecasting system based on their conditional dynamics. They run a regression with VIX as the dependent variable and generate the one-month-ahead volatility forecasts according to the system specifications as independent variables. The result shows that one-month-ahead forecasts match well the market-based volatility measure provided by the VIX index.

Fuertes, Izzeldin and Kalotychou (2009) examine individual NYSE/NASDAQ stocks, and provide statistical evidence in favour of high frequency data. They incorporate non-parametric estimators of daily price variability into a GARCH model. Four estimators are

compared: realised volatility (RV), realised range (RR), realised power variation (RPV) and realized bipower variation (RBP). Their results show that forecast combination of the four intraday volatility estimates is worth considering as the combination of all four intraday measures produces the smallest forecast error in about half of their sampled stocks.

Patton and Sheppard (2009) find that a simple equally-weighted average of 32 different realized estimators generally cannot be outperformed by any individual estimator. They also show that none of the individual estimators encompasses the information in all other measures, providing further support for realized combination. Fuertes and Olmo (2012) investigate the practical importance of several volatility forecasting issues addressed in a Value-at-Risk (VaR) context. The empirical results show that combined GARCH and ARFIMA forecasts produce quite competitive VaR measures.

### **5.2.3 The Economic Value of Forecast Combination**

While the statistical accuracy of volatility forecasts has been analysed extensively, the direct profitability of them is a relatively new area of interest. The idea is that optimal volatility measures assessed by statistical evaluation methods are not guaranteed to generate superior returns when used in volatility trading activities. Harvey and Whaley (1992) study the profitability of a trading rule based on forecasting implied volatility and find supportive evidence on option market efficiency after allowing for transaction costs. In contrast, Noh and Engle (1994) argue that a straddles trading strategy on the S&P 500 index based on GARCH model volatility forecasts can make significant profits after transaction costs. Martens and Zein (2002) note that intraday based estimators do have incremental information over that contained in implied volatilities for futures and options on the S&P 500 index.

Kinlay (2005) investigates the potential for generating abnormal profits using a simple straddle trading strategy, based on ARFIMA and GARCH generated volatility forecasts,

taking into account both transaction and (delta) hedging costs. The empirical results show that forecasted volatility is unable to add to the information already contained in implied volatility. However, model-forecasted volatility correctly foresees the direction of volatility change about 62% of the time whereas implied volatility have very poor direction prediction ability. Simulation in purchase or sale of at-the-money straddles based on the volatility forecasts yields a net annual compound return of 18.64% over 4 years from 2000 to 2003, during which the annual return on the S&P 500 index itself was -7.24%.

## 5.3 Methodology

In this section, we first present the forecast combination techniques to be applied. Next, we introduce the ARFIMA based volatility forecasting framework. Finally, we study the loss functions for statistical evaluation, and trading strategies for economic assessment of the combination estimators.

### 5.3.1 Forecast Combination Techniques

This chapter applies five combination techniques for the empirical study: equal weights, geometric mean, regression-based, Akaike information criterion based (AIC) and the multiple indicators model (MIM). In total 42 individual volatility estimators are available for forecast combination, including the implied volatility and 41 realized volatility measures (see Chapter 2 for a detailed explanation) constructed using different methods and sampled with a range of time frequencies from 1-minute to 60-minute.

#### 5.3.1.1 Equal Weights

The benchmark combination approach is the equal weights model, which assumes equal probability to each individual estimator, and considers a simple average of the predictions from each estimator. Thus, the weight assigned to each volatility estimator is

$$w_i = \frac{1}{N} \tag{5.1}$$

where  $N$  is the number of individual estimators to be combined.

### 5.3.1.2 Geometric Mean

Similar to the equal weights, the geometric mean is an averaging model, except that the numbers are multiplied and then the  $n$ th root of the calculated product is taken. The weight for individual estimator  $s_i$  is

$$w_i = \left( \prod_{i=1}^N s_i \right)^{\frac{1}{N}} \quad (5.2)$$

### 5.3.1.3 Regression-Based

Stock and Watson (2004) applied the method of discounted mean squared forecast error (MSFE) to combine a large number of individual forecasts. The regression-based combination is a special case of the discounted MSFE with a unit discount factor, which also corresponds to the optimal weighting scheme introduced in Bates and Granger (1969). In detail, the regression-based model is a standard parametric combination estimator, namely a linear combination:

$$y_t = \hat{w}_0 + \sum_{i=1}^N \hat{w}_i s_{i,t} + \varepsilon_t, \quad \varepsilon_t | \Omega_{t-1} \sim iid(0, \sigma^2) \quad (5.3)$$

where  $s_{i,t}$  is the  $i$ th realized estimator at time  $t$ . We consider an unconstrained combination framework as a constraint of nonnegative weights distorts the optimality of the parameter estimation.  $y_t$  is the realize volatility (RV) in the future 22 trading days sampled at 5-minute frequency, and is considered as the actual monthly volatility. For a six-and-half hour market, the monthly volatility is therefore the sum of the 1716 intraday squared 5-minute returns. There are two underlying concerns relate to the approximation of actual volatility. First, the 5-minute RV is generally considered an unbiased representative of real volatility by empirical literatures. Second, since we are measuring the real volatility over a long period, the difference in estimation among standard volatility estimators is often negligible.

#### 5.3.1.4 Akaike Information Criterion Based (AIC)

The AIC-based combination is introduced by Kapetanios et al. (2008) using Monte Carlo experiments on UK inflation data. The authors find that the model performs as well as or better than Bayesian averaging. A brief explanation of the AIC-based approach is provided in below.

In a classical view of probability, there is a true model, albeit one that may be varying through time. The hard part is how to estimate it. Given this true model, the uncertainty lies with the accuracy of data and parameter estimation. In contrast, instead of assuming there is a true model, Bayesian probabilities measure the degree of belief that an object has in an event. Parameters themselves are random variables with a probability distribution, rather than the estimated parameters being fitted around a given value. Bayes' theorem shows how new information can be used to update the conditional probability of a state occurring. In our study, the information comes from historical returns and the state is a future value of monthly volatility. Our forecast is conditional on the past data and our initial guess. Moreover, there is uncertainty over models. As none of our models is the "true" model, we characterise our views by means of probabilities related to each model. A high probability therefore signals a strong belief in the model. Given the probabilities, we can form the average forecast.

Bayesian model averaging has a key notion that the conditional probability of a model  $s_i$  being the true model, given the data,  $D_t$ , is  $pr(s_i|D_t)$ . But there is a frequentist analogue, and a weight scheme based on this has been studied by Akaike (1979). Akaike's suggestion derives from the Akaike information criterion (AIC). AIC is an asymptotically unbiased measure of minus twice the log likelihood of a given model. It is an unbiased estimator of the Kullback and Leibler (KL, 1951) distance of a given model where the KL distance is formulated as

$$I(f, g) = \int f(x) \log \left( \frac{f(x)}{g(x|\hat{\theta})} \right) dx \quad (5.4)$$

where  $f(x)$  is the unknown true model generating the data,  $g(x|\cdot)$  is the entertained model and  $\hat{\theta}$  is the estimate of the parameter vector for  $g(x|\cdot)$ . The KL distance is important in model selection. Within a given set of models, the difference of the AIC for two different models is an estimate of the difference between the KL distance for the two models. Further,  $\exp(-1/2\Psi_i)$  is the relative likelihood of model  $i$  where  $\Psi_i = AIC_i - \min_j AIC_j$  for  $j = 1, \dots, N$  and  $AIC_i$  is the AIC of the  $i$ th model in the model space  $N$ . Therefore,  $\exp(-1/2\Psi_i)$  can be thought of as the odds for the  $i$ th model to be the best KL distance model in  $N$ . In other words this quantity can be viewed as a crucial difference from a Bayesian Analysis, in which it is assumed that a model in  $N$  is the true model while in the Bayesian view the models must span the complete set.

We can obtain the optimal weights by normalising  $\exp(-1/2\Psi_i)$  as

$$w_i = \frac{\exp(-1/2\Psi_i)}{\sum_{i=1}^N \exp(-1/2\Psi_i)} \quad (5.5)$$

where  $\sum_i w_i = 1$ .

The  $w_i$  can be regarded as model probabilities under noninformative priors giving a parallel to Bayesian analysis. However, this analogy should not be taken literally as these model weights are firmly based on frequentist ideas and do not make explicit reference to prior probability distributions about either parameters or models.

### 5.3.1.5 Multiple Indicators Model (MIM)

The multiple indicators model (MIM) is a distinct approach to combine realised volatility measures proposed by Engle and Gallo (2006). It jointly considers different volatility measures to form a forecasting model based on their conditional dynamics. The MIM is a two-step approach. We first forecast directly the realized volatility measures themselves. For a non-negative valued process  $x_t$ , we describe it by a multiplicative error model (MEM, as studied in Engle, 2002). That is,  $x_t$  is the product of a time varying scale factor and a standard positive valued random variable. The formula is therefore

$$x_t = \mu_t \varepsilon_t, \quad \varepsilon_t | \Omega_{t-1} \sim iid.D(1, \varphi) \quad (5.6)$$

where  $\mu_t$  is a time varying scale factor which depends upon the recent history of the series and  $\varepsilon_t$  is a standard positive valued random variable.  $\mu_t$  is specified as

$$\mu_t = \tau + \sum_{i=1}^p \alpha_i x_{t-i} + \sum_{i=1}^p \beta_i \mu_{t-i} + c' z_t \quad (5.7)$$

where  $z_t$  summarizes further terms that can signal the dependence of the series on weakly exogenous variables included in the information set available at time  $t - 1$ . For example, the basic MEM specification for realized volatility estimator  $i$  can be expressed as

$$\mu_{i,t} = w_i + \alpha_i r_{t-1}^2 + \beta_i \mu_{i,t-1} \quad (5.8)$$

and here we extend the basic model to account for asymmetry by applying the APARCH model

$$\mu_{i,t} = w_i + \alpha_i r_{i,t-1}^2 + \beta_i \mu_{i,t-1} + \gamma_i r_{i,t-1}^2 d_{t-1} + \delta_i r_{t-1} \quad (5.9)$$

where  $d_{t-1}$  is a customary dummy variable for negative returns  $d_t = I(r_t < 0)$ .

After obtaining MEM forecasts for different realised volatility measures, we merge them together as a MIM system of equations. The last step is to regress each of the  $\mu_{i,t}$  on  $y_t$ , the actual volatility in the upcoming calendar month, to generate the optimal weight for each MEM model in the MIM combination.

### 5.3.2 Volatility Forecasting Framework

The out of sample monthly volatility forecasts for each realized volatility measure are made by rolling window autoregressive fractionally integrated moving average (ARFIMA) estimations. The monthly volatility is equivalent to the quadratic price variation in 22 trading days, therefore for each model we carry out a 22-period ahead variance forecast. The 22 daily forecasted variances are then summed up to form the forecasted monthly variance. The square root of the forecasted monthly variance is a comparable measure to the VXO, implied volatility of S&P 100 index for the coming calendar month. We also use the ARFIMA model to forecast VXO since long-range dependence is a typical

stylized factor of implied volatility (see Corsi, 2004; Koopman et al., 2005; Bandi and Perron, 2006). The ARFIMA approach is briefly discussed in below.

A stochastic process  $s_t$  is called an ARFIMA( $p, d, q$ ) process if the fractionally differenced process is an autoregressive moving-average (ARMA) series as

$$(1 - \phi L - \dots - \phi_p L^p)(\Delta^d X_t) = (1 - \theta_1 L - \dots - \theta_q L^q)(\varepsilon_t) \quad (5.10)$$

where  $d$  is not restricted to integral values. ARFIMA concerns long-memory process, a stationary series whose autocorrelation function decays slowly. An ARFIMA specification yields a parsimonious parameterization of long-memory processes that nests the ARMA model, which is commonly used for short-memory processes. The ARFIMA model can also be considered as an extension of the ARIMA model as the former allows for fractional degrees of integration.

As noted in Fuertes and Olmo (2012), the ARFIMA modelling framework has been successfully used in the literature to capture the stylized slow, less than exponentially decay in autocorrelations of daily realized volatilities. For example, Bhardwaj and Swanson (2006) investigate the usefulness of ARFIMA models in practical prediction-based applications. The authors find evidence that such specifications often outperform a wide class of the benchmark non-ARFIMA models, including AR, ARMA, ARIMA, random walk, and related models. This chapter focuses on the homoscedastic ARFIMA (1,  $d$ , 0) model, which has been employed in Andersen et al. (2003), Koopman et al. (2005), and Fuertes and Olmo (2012). The model is proved to be an efficient competitor to other time series methods of forecasting realized volatility. The model specification is

$$(1 - \phi L)(1 - L)^d(s_t - w) = \varepsilon_t \quad (5.11)$$

where  $s_t$  is one of the the daily realized estimators,  $w$  is the unconditional mean of  $s_t$ , and  $L$  is the lag operator ( $LS_t = s_{t-1}$ ). Estimation of the coefficients in the equation including  $d$  is carried out by exact maximum likelihood under the normal assumption and forecasts are calculated from the autoregressive and moving average representation of the

process (see Fuertes and Olmo, 2012). In the ARFIMA models fitted to the log measures, the forecasted volatility is obtained through the common exponential transformation (see Forsberg and Ghysels, 2007; Clements et al., 2008; Liu and Maheu, 2009).

After obtaining the ARFIMA forecasted monthly volatility based on individual estimators, 47 forecast combinations will be constructed. The forecast combinations can be divided into two broad categories: simple combinations and further combinations. A simple combination forecast combines an ARFIMA forecasted realized volatility estimator and ARFIMA predicted implied volatility using OLS regression. For each simple combination the formula is

$$h_{t,i} = \alpha_i + \beta_{1,i}V\hat{X}O_t + \beta_{2,i}\hat{s}_{i,t} \quad (5.12)$$

where  $h_t$  is the forecasted monthly volatility,  $V\hat{X}O_t$  is the ARFIMA forecasted VXO and  $\hat{s}_{i,t}$  is the ARFIMA forecasted volatility for realized estimator  $i$ .

A further combination combines 42 volatility estimators including the 41 intraday volatility estimators and implied volatility using 5 different combination approaches introduced in Section 5.3.1: the equal weighted, geometrical mean, regression-based, Akaike information criterion (AIC) based, and multiple indicators model (MIM).

### 5.3.3 Forecast Evaluation Framework

Though the main purpose of the empirical study is to examine the economic value of the individual and combination volatility estimators, statistical evaluation is also implemented for completeness. The statistical assessment consists of symmetric loss functions and the economic value is evaluated by the performance of option trading strategies.

#### 5.3.3.1 Statistical Evaluation

Individual volatility estimators and their combinations are evaluated by three popular symmetric loss functions: mean absolute error (MAE), mean squared error (MSE) and adjusted mean absolute percentage error (AMAPE). Proposed by Bollerslev et al. (1994),

MSE is the most widely applied forecast evaluation method, which is based on a quadratic loss function. It is best used when large forecast errors are disproportionately more concerned than smaller ones. MAE measures the actual error in absolute values, which is less sensitive to large mispredictions than MSE does. AMAPE, proposed by Makridakis (1993), is an absolute error divided by the average of the forecast and actual values. The specification of the loss functions are shown in below.

$$\begin{aligned}
\text{Mean Absolute Error} & \quad MAE = \frac{1}{T} \sum_{t=1}^T |\tilde{\sigma}_t^2 - y_{t,i}^2| \\
\text{Mean Squared Error} & \quad MSE = \frac{1}{T} \sum_{t=1}^T (\tilde{\sigma}_t^2 - y_{t,i}^2)^2 \\
\text{Adjusted Mean Absolute Percentage Error} & \quad AMAPE = \frac{1}{T} \sum_{t=1}^T \left| \frac{\tilde{\sigma}_t^2 - y_{t,i}^2}{\tilde{\sigma}_t^2 + y_{t,i}^2} \right|
\end{aligned} \tag{5.13}$$

where  $\tilde{\sigma}_t^2$  is the conditional variance and its proxy  $y_{t,i}^2$  for forecast evaluation is the 5-minute realized variance for the upcoming 22 trading days. The statistical accuracy of model  $i$  forecasts,  $y_{t,i}^2$  is evaluated by the loss functions.

### 5.3.3.2 Economic Evaluation

We analyse the potential profits using a straddle trading strategy based on different volatility forecasts. Hull and White (1987) note that, when volatility is constant, the Black-Scholes implied volatility of an at-the-money option approximately equals the expected future volatility over the life of the option. The implied volatility, VXO, is derived from the at-the-money call and put options based on the Black Scholes option pricing formula

$$\begin{aligned}
C &= SN(d_1) - Xe^{-rt}N(d_2) \\
P &= Ke^{-rt}N(-d_2) - SN(-d_1)
\end{aligned} \tag{5.14}$$

$$d_1 = \frac{\ln(S/X) + (r + \sigma^2/2)t}{\sigma\sqrt{t}}, \quad d_2 = d_1 - \sigma\sqrt{t} \tag{5.15}$$

where  $S$  is the S&P 100 index price, assumed to follow a log-normal distribution with a constant volatility  $\sigma$ ,  $K$  is the option strike price,  $t$  is the time to maturity and  $r$  is the risk-free interest rate and  $N(\cdot)$  represents the Gaussian density.

The economic value from volatility trading is tested as follows. We collect the price and option delta for near-to-the money options on the S&P 100 index for each trading day

at market close. We then simulate purchase or sale of at-the-money straddles, based on whether the forecasted volatility to expiration is greater or less than the implied volatility of the option contracts by a certain degree. The resulting long or short option position is then held to maturity.

A straddle is a non-directional option trading strategy involves buying (long straddle) or selling (short straddle) a call option and a put option approximately at-the-money at the same time. The two options are about at the same strike price and expire at the same time. In our case the options are at-the-money ones therefore the strike price is the underlying price. A short straddle is constructed if our forecasted volatility falls below the VXO by a certain percentage. This position is an unlimited risk, since the profit is capped to the premiums of the two options sold but large loss occurs when the S&P 100 index goes very high up or very low down.

A long straddle strategy makes a profit if the S&P 100 price moves far away from the strike price, meaning that we expect high volatility during the remaining life of the options bought. According to our trading strategy, a long straddle will be implemented if our forecasted volatility exceeds the VXO by a certain degree, or trading range. This position is attached with limited risk, as the most we may lose are the premiums of both options. In contrast, the profit potential is unlimited. The payoff from a long position straddle is listed in the table in below.

Payoff from a Long Straddle			
Range of Stock Price	Payoff from Call	Payoff from Put	Total Payoff
$S_T \leq K$	0	$K - S_T$	$K - S_T$
$S_T > K$	$S_T - K$	0	$S_T - K$

We construct a straddle on day  $t$  only if the difference between the VXO and our forecasted volatility on the trading day exceeds a preset trading range. For instance, if the trading range is 0.5%, it means that the forecasted volatility needs to be greater than

the VXO by at least 50 basis points (bp) to trigger a long straddle and to be smaller by at least 50bp for a short straddle. The larger the trading range, the smaller the chance a straddle will be constructed. This is because a large trading range requires a substantial difference between the forecasted volatility and the VXO, a triggering condition that is hard to meet. If the trading range is wide, it might be difficult to make large profit since only a limited number of trades will be implemented. However, if the trading range is too narrow, many transactions will be made and therefore a significant amount of transaction costs will be generated. In the empirical study we test a variety of trading ranges spanning from 0% to 1% to examine the robustness of the trading strategy.

Each implemented short or long position is held to maturity and delta-hedged at the end of each trading day using the net of the closing deltas of the two option contracts. In terms of the S&P 100 options, the delta is the measure of how the value of an option changes with respect to changes in the value of the S&P 100 index. Delta, represented in absolute value, can be regarded as an approximation of the probability that an option will finish in-the-money. Delta lies between 0 and 1 for a call option, and between 0 and -1 for a put option. Given the information on the closing deltas of the call and put option contracts, the option contracts bought or sold is delta hedged at the end of each trading day. The delta ( $\Delta$ ) can be calculated as  $\Delta = \frac{\partial C}{\partial S}$ . Assume that the delta hedging is achieved by trading SPDRs<sup>2</sup>, contracts at the American Stock Exchange at 1/10 the size of the S&P 100 index. Hedge contracts are signed at the reported closing offer or bid price.

In terms of transaction costs, we assume the trading costs at \$1 per round turn for each contract and at \$0.005 for each SPDR bought or sold. The transaction costs are higher than would be incurred by large-scale fund trading on electronic exchanges. We do not consider market impact, as the impact should be relatively minor in highly liquid instruments up to notional capital amounts of \$500 or so (see Kinlay, 2005).

---

<sup>2</sup>A S&P 100 exchange traded fund (ETF), before expenses, seeks to closely match the returns and characteristics of the S&P 100 index.

## 5.4 The Dataset

Tick by tick data for the S&P 100 cash index from November 06/01/1997 to 31/12/2010 were collected from multiple data feeds via *Disk Trading* database and subjected to scrubbing methodologies to detect and repair bad ticks. The estimation sample spans 10 years of trading data from 06/01/1997 to 29/06/2006, containing 2509 daily observations. The forecast period runs from 03/01/2007 to 31/12/2010, with 1008 daily samples spanning 4 years. Rolling window forecasts for 22 trading days are carried out based on parameters of the models estimated in the sampling period.

Figure 5.1.A shows a roller-coaster like price pattern of the S&P 100 index prices over the past 14 years. An upward price trend lasted 4 years from 1997 to 2000 and doubled the equity index from around 400 to 800. This buoyant momentum was driven by strong economic data and technology boom in the United States in the late 1990s. However, the collapse of the dot-com bubble and September 11th attacks lead to an overwhelmingly pessimistic market sentiment which dragged the S&P 100 index all the way down to 400 within 2 years from 2001 to 2002. Since 2003 the stock market recovered gradually and climbed up consistently for about 5 years supported by stable economic conditions and strong growth in emerging countries. But since late 2007 a rise in subprime mortgage delinquencies and foreclosures evolved into a fully-fledged subprime mortgage crisis, which destroyed financial liquidity and investors' confidence and brought the S&P 100 prices down to almost 300 from 700 in a year. After 2008 a significant amount of government bail-outs and stimulus shored up the global financial system and supported a rebound of the S&P 100 prices to above 500 at the end of 2010.

Figure 5.1.B presents the monthly volatility of the S&P 100 index using realized variance sampled at 5-minute frequency. The level of volatility differ dramatically throughout the sample period. From 1997 to 2002 the volatility remained at reasonable levels as the collapse of the dot-com bubble did not affect the S&P 100 index much, since it did not contain a large portion of internet stocks. From 2003 to 2007 the volatility was at rel-

atively low levels driven by stable market conditions. However, from 2008 onwards the volatility soared up substantially with a spike shown in late 2008, reflecting the peak of the 2008-2009 global financial crisis.

41 realized estimators introduced in Chapter 2 are calculated for the empirical study. Table 5.1 shows the summary statistics of the estimators. The largest average is estimated by RR\_30min (a realized range estimator sampled at 30-minute frequency). The measure gauges the average monthly volatility of the S&P 100 at 6.22%, equivalent to an annual rate of 21.6%. This is no surprise since RR employs the extreme values occurred in the sampling period. The smallest mean value is estimated by RK Cubic\_5min (a realized kernel measure with a cubic kernel weight function sampled at 5-minute frequency). The measure estimates the average monthly volatility at 3.25%, corresponding to an annual volatility of 11.27%. The variation in volatility estimated by the models is due to their distinct specifications, which interpret and sample volatility in different ways. This variation is also noted in previous empirical studies such as in Patton and Sheppard (2009). None of the estimators have a minimum value that is non-positive, including the realised kernel estimators, which do not ensure non-negativity of the estimates. The smallest standard deviation is obtained by the RPV\_15min and the largest is the RR\_5 minute. All the realized measures exhibit positive skewness and high kurtosis, therefore do not follow a normal distribution. This is commonly seen in financial time-series data, which have a high concentration around the mean and a fat tail extending both ends.

Daily VXO data from 06/01/1997 to 31/12/2010 are collected from the Chicago Board Options Exchange (CBOE). VXO is the ticker symbol for the CBOE Market Volatility Index, a popular measure of the implied volatility of S&P 100 index based on trading of S&P 100 (OEX) options. The VXO is generated by the Black-Scholes option pricing model and calculated from both call options and put options. It is a widely used measure of market risk and is often referred to as the "investor fear gauge". The implied volatility is quoted in percentage points; it is the square root of the par variance swap rate for a

30-day term initiated today. After the monthly volatility is derived, the CBOE convert the volatility to an annually rate by multiplying the monthly volatility with  $\sqrt{12}$ . For example, an annualized VXO of 15 implies that the market believes the S&P 100 prices will vary around 15% over the coming year. To make the VXO and our realized measures comparable, we convert both to monthly standard deviation.

In common with parallel research, we find that during the sample period the average VXO, which shows an annualized rate of 23.5%, is greater than the mean of any realized volatility measures. Kinlay (2005) suggests that, unlike model forecasts, implied volatility forecasts show evidence of a consistent and substantial bias. The relatively high average of VXO is caused by a risk premium factor. Option sellers bear significant downside risk so they need to charge a premium for volatility to profit and hedge. Therefore VXO contains a risk premium that is not incorporated in the realized measures.

Table 5.2 presents the summary statistics of the combinations of realized measures. Similar to the individual estimators, volatilities estimated by each combination are rescaled to monthly standard deviations. The largest average is estimated by the regression-based combination, which estimates the average monthly volatility of S&P 100 at 6.47%, corresponding to an annual volatility of 22.4%. The smallest mean is estimated by the geometric mean estimator, which gauges average monthly volatility at 4.66%, corresponding to an annual volatility of 16.14%. Similar to the individual estimators, none of the combination estimators has a minimum value that is non-positive and all of them exhibit positive skewness and high level of kurtosis.

## 5.5 Empirical Results

This section compares the performance of the individual and combination volatility estimators in the context of volatility trading. We first present the result of statistical evaluation, which is used as a reference for comparing the individual and combination volatility estimators. Then we evaluate the economic value of combining a large number

of individual volatility estimators. Finally, a robustness test is implemented by assessing the volatility estimators' performance in alternative settings of option-trading strategy.

### **5.5.1 Statistical Evaluation of Combination Volatility Estimators**

Monthly volatility of the S&P 100 index are forecasted by 42 individual and five combination estimators, which are ranked according to their MAE, MSE and AAPE statistics in Figure 5.2. The 42 individual estimators include the ARFIMA forecasted implied volatility (VXO hereafter) and 41 intraday-based volatility measures. The combination estimators combine the 41 individual realized volatility measures using five distinct combination methods, namely, equal weighted, geometric mean, regression-based, AIC based and MIM. For the combination estimators, the implied volatility is not selected as a combination component as we intend to compare the forecast power of the realized estimators and the implied volatility. The absolute forecast accuracy of an estimator is judged by comparing its forecasts with actual volatilities, which are approximated by the 5-minute realized variance (RV\_5min) of the upcoming month. The forecast period lasts four years from 2007 to 2010.

In detail, Figure 5.2 shows that the VXO, implied volatility of the S&P 100 index, generates unparalleled performance according to all the three statistical criteria, suggesting that implied volatility outperforms the best possible individual or combined realized estimators. The result is in line with the findings of Blair, Poon and Taylor (2001), and Miller (2003), which show that implied volatility estimates generally provide better volatility forecast statistically.

Under a rational expectations assumption, market participants use all the available information to form their expectations about future volatility, and therefore the market option price reflects the market's 'true' volatility forecasts. If the market is weakly efficient, the market's forecasts, the implied volatility, should be the best possible forecast given the currently available information. The US stock market, which contains the

S&P 100 components, is widely believed to be efficient in the weak form (see, Williamson, 1972). The result of statistical evaluation for VXO is therefore consistent with the theory.

Two realized combinations, Regression-Based and MIM, produce the second and third best results for all the three loss functions, outperforming all the individual realized estimators and suggesting statistical value of combining realized volatility estimators. This result is in line with the literature (see, for example, Patton and Sheppard, 2009), which shows that combination estimators often produce better forecast performance compared with individual forecasts. A well-combined volatility estimator takes the advantage of distinct information provided by its components, produces forecasts that cannot generally be out-performed by the best individual component.

### **5.5.2 Economic Evaluation of Combination Volatility Estimators**

The empirical studies presented in this section are designed to examine the economic value of combining the implied volatility and intraday-based volatility estimators. Two broad types of forecast combinations are considered. First, 41 simple combination estimators combine the implied volatility and an individual volatility measure using a regression-based approach. Second, five further combination measure combines the implied volatility and 41 distinct realized volatility estimators. The five further combination estimators differ from those in the statistical evaluation as they now have the implied volatility as a combination component. Trading performances of volatility trading strategies based on different combination forecasts are assessed to determine whether the simple combination forecasts would provide superior economic value compared with the best individual volatility estimator, ARFIMA forecasted VXO (VXO hereafter), and whether the trading performance can be further improved by carrying out further combinations, which combines a large number of individual volatility forecasts.

### 5.5.2.1 Simple Combination of Implied and Realized Volatility

Figure 5.3 shows the simulated trading results of trading strategies based on different volatility estimators and implemented on a trading range of 0.25%, which means that a straddle is bought or sold on day  $t$  only if the absolute difference between model-forecasted volatility and the implied volatility for day  $t$  exceeds 0.25%. The model presented in grey is the benchmark measure, VXO, the best individual estimator, the 41 models in blue are the simple combination estimators, and the five models in orange are further combinations. Panel C presents the largest win and loss, where the former is the biggest profit made by a single option transaction and the latter the largest loss made by a single trade.

We first assess the economic value in combining model-forecasted and implied volatility by comparing the simple combination estimators with the best single volatility estimator, VXO. The performance of the further combination estimators will be discussed in the next subsection.

Panel A of Figure 5.3 lists out the annualized return (in brighter colour) and volatility (in darker colour) generated by each estimator. The volatility models in Panel A are ranked by the Sharpe ratios generated, which are shown in brackets after the name of the corresponding estimator. The benchmark model, VXO, ranks the 28th among the 47 volatility estimators according to the criterion. It generates an annualized return of 12.56% while exhibits volatility of 6.69%, yielding a Sharpe ratio of 1.28.

In contrast, the best performing simple combination measure, the 5-minute realized range (RR\_5min), ranks the 4th with an annual return of 14.64%, volatility of 6.53%, and Sharpe ratio of 1.63. In fact, 21 simple combination measures, which combine the VXO and one of the 41 realized estimators, outperform the benchmark model which utilizes historical VXO information alone. The result indicates that a simple combination of a realized estimator and the implied volatility is more preferable than the best individual volatility estimator in terms of economic value.

This finding contradicts to Poon and Granger's (2003) suggestion that model predicted

volatility is unable to add the information already contained in market forecasts. One possible explanation is that intraday volatility estimators may contain unique information that is not captured by both the implied volatility and daily volatility measures. A simple regression based combination of the implied and intraday volatility estimators could effectively combine the distinct information contained in both, hence providing better economic performance.

Panel B evaluates the accuracy of directional predications by the percentage of winning and the percentage of option trades made a profit. Generally, the volatility models are more accurate than a random walk as even the worst performing estimator, 5-minute sampled realized power variation with  $Z=1$  (RPV\_1\_5min), has 52.12% of transactions ended up with a profit. The benchmark estimator VXO ranked the 10th with a percentage of winning of 54.05%, surpassed by 6 simple combination estimators. RPV\_1.5\_1min is the best simple combination estimator, which generates positive returns for 54.94% of its straddle trades.

It is interesting that the accuracy of directional prediction does not always translate into profit. For example, while ranks the 4th in terms of the accuracy of directional predictions, RPV\_1.5\_1min is placed the 22nd according to the Sharpe ratio produced. Thus an estimator with more accurate directional predictions is not guaranteed to outperform in terms of less over- or under-prediction. Since the performance of the straddle trading strategy is particularly based on the size of variation between the forecasted and actual volatility, accurate directional prediction alone does not guarantee profitability.

At last we take a look on the annual volatility generated by the trading strategies. The RPV\_1\_60min tops the simple combination category with an annual volatility of 6.49%, smaller than the 6.69% provided by the benchmark model VXO. In this setting a combination of implied and realized volatilities reduces the risk of the trading strategy. In summary, a simple combination of implied and realized volatility can generating economic value that cannot be outperformed by the single best estimator.

### 5.5.2.2 Further Combination of Individual Volatility Estimators

The second issue to be analysed is that whether the economic value can be further improved by adopting more complex combination approaches, which combine a large number of individual estimators. The five further combinations applied are the geometric mean, AIC-based, equal weights, regression-based and the MIM. Ranked by Sharpe ratio, all the five combination models provide superior performance than the simple combinations and the benchmark VXO model. The MIM and regression-based combinations top the ranking table. By generating an annual return of 17.16% and volatility of 5.88%, the MIM combination provides the strongest performance with a Sharpe ratio of 2.24. Regression-based combination also produces a Sharpe ratio exceeding two, significantly higher than the 1.63 obtained by the best single estimator, 60 minute sampled realized range (RR\_60min).

This result indicates that further combinations which utilize the information contents in a variety of volatility estimators from distinct classes can generate significant economic value in the context of volatility trading. The advantage of carrying out further combination lies in enriched information. Intraday estimators formed by different specifications may contain distinct yet relevant information for forecasting future volatility. A well-combined estimator of different realized measures and the implied volatility utilizes the information to enhance its forecasting power in predicting future volatility and producing better economic value.

Ranked by the percentage of winning trades (shown in Panel B of Figure 5.3), three further combination estimators, geometric mean, AIC-based and equal weights top the table with over 55% of trading activities being profitable, before transaction costs. The three combinations outperform the best estimator in the simple combination category, RPV\_1.5\_1min, which has a percentage of winning of 54.94%. Two further combination estimators, the regression-based and MIM, lagged behind with a percentage of winning at 53.37%. Interestingly, MIM and regression-based are the top performers according to the Sharpe ratio rankings. The result again indicates that accurate directional prediction

does not always translate into economic value.

Ranked by the annual volatility generated in the trading simulation, three further combination estimators stand out. The MIM, regression-based and geometric mean produce the smallest volatility with standard deviation of 5.88%, 6.43% and 6.49% respectively. The best model in the simple combination class, RPV\_1\_60min has an annual standard deviation of 6.49%, which is identical to the value for geometric mean. In summary, further combination of realized estimators produces higher return, lower volatility and larger percentage of winning. These empirical results show that there is substantial economic value in carrying out further combinations.

### **5.5.3 Estimation Results of the Volatility Models**

The left columns in Table 5.4 report the parameter estimates and regression statistics of selected models. The Numbers in square brackets are the p-values of the corresponding parameters. The dependent variable of the models is the actual volatility based on 5-min realized variance. Further combination models, regression-based and MIM, use the corresponding combination estimators as independent variables together with the VXO. The simple combination model, RV\_5min, uses the ARFIMA forecasted RV\_5min as the independent variable alongside the VXO. All estimated parameters are significant and positive, including those of the realized volatility (RV) and the realized combination (regression-based and MIM). Highly significant coefficients of the realized combination in all the further combination models statistically justify the use of intraday information.

### **5.5.4 Robustness Tests**

It is interesting to assess whether consistent trading results can be generated using alternative trading strategies. Figure 5.4 presents the empirical performance of an option trading strategy based on a relatively narrow trading range of 0.2%, which is a filtering strategy where each agent trades only when the absolute difference between the model-

forecasted volatility for day  $t$  and the implied volatility in day  $t$  is greater than 0.2%. Two further combination models, the MIM and regression-based again top the ranking table with Sharpe ratios of 2.15 and 1.82, respectively. However, other models in the combination group, the AIC-based, equal weights and geometric mean are no longer superior to the best simple combination estimator in terms of Sharpe ratio produced. RPV\_1\_1min outperforms the three further combination models by generating a Sharpe ratio of 1.6. Similar to the previous case, the benchmark model VXO ranks the 38th with a Sharpe ratio of 1.02 only, outperformed by most of the simple and further combination models.

Ranked by annual volatility generated in the option trading simulation, MIM again produces the top result with a standard deviation of 5.82%. The best model in the simple combination class, RR\_1min, has an annual standard deviation of 6.38%, which outperforms all the four other further combination estimators.

Although the further combinations, MIM and regression-based, are again the top models in terms of return produced, standard deviation generated, and the number of winning trades made, they are no longer the best performers according to the percentage of winning (see Panel B of Figure 5.4). Simple combinations such as the 1-minute sampled Tukey-Hanning realized kernel estimator (RK\_TKH\_1min) and RR\_5min produce a percentage of winning over 55%. All the models in evaluation once again achieve a percentage of winning over 50%, showing that the estimators are better suited to produce directional forecasts than a random walk model.

Figure 5.5 shows the performance of trading strategies based on a relatively wide trading range of 0.3%. Two further combination models, MIM and regression-based once again outperform all other volatility estimators by producing Sharpe ratios of 1.81 and 1.32, respectively. However, these Sharpe ratios are smaller than previous values obtained from trading strategies with smaller trading ranges. This is because the higher the trading range, the larger the difference between model-forecasted volatility and VXO is required to generate a trading signal. Therefore fewer transactions will be implemented based

on a wide trading range, and consequently inactive trading may lead to lower profit. Three other further combinations, the AIC-based, equal weights and geometric mean are outperformed by most of the simple combination models. The benchmark model VXO performed slightly better than in the case of narrow trading range by ranking the 13th with a Sharpe ratio of 0.62.

Finally we choose one of the top ranking further combination estimators, the regression-based, to study the model performance under different trading ranges. Table 5.3 lists out information of the Sharpe ratio, number of winning trades, percentage of winning, annual return and volatility generated with different trading ranges. The table shows a bell shaped performance pattern in terms of Sharpe ratios, in which the ratios are smaller at both the low and high end of the trading range. In addition, the pattern of annual returns follows a similar shape to the Sharpe ratio one. The returns are lower at both ends of the trading range, which are 13.21% in the case of a 0.05% trading range and 9.6% in the case of a wide range at 1%. The highest annual return is obtained with a medium trading range at 0.5%, which is 17.84%.

These Sharpe ratio and return variations may be caused by a trade-off between the accuracy of prediction and frequency of trading. The narrower the trading range, the more options will be bought or sold, incurring more inaccurate predictions and higher transaction costs. In contrast, the wider the trading range, the lower the number of options will be bought or sold, leading to less profitable opportunities. In this study a trading range between 0.2% and 0.5% provides the right balance in regards to the trade-off by producing the highest Sharpe ratio.

The riskiness of the strategy (annual volatility of returns) decreases as trading range increases. This is because when we increase the trading range, we become implicitly more cautious in implementing an option transaction. A straddle is bought or sold only when the discrepancy between markets' perceived volatility and our forecasted volatility is significantly large in the case of a wide trading range, which in turn reduces the risk of

the trading strategy.

By observing the output of the empirical tests with different trading strategies, we provide three main findings. First, two further combination models, MIM and regression-based, consistently outperform all other estimators according to annualized return, standard deviation and the number of winning trades, showing significant economic value in the context of volatility trading. Second, there are always a few simple combination models capable of providing superior performances compared with the benchmark model VXO, the best single estimator. Hence, employing implied or realized volatilities alone cannot outperform their combinations judged by economic value. Finally, within trading ranges from 0.05% to 1%, all the combination models generate significant positive returns after hedging and transaction costs. This result further confirms the economic value of realized combination.

## 5.6 Conclusion

This chapter presents an empirical study on the economic value of combining intraday-based and implied volatility estimators. Both simple combinations of the implied volatility and an intraday volatility estimator, and further combinations of the implied volatility and 41 intraday estimators are considered. The sample spans 14 years of price data for S&P 100 index from 06/01/1997 to 31/12/2010. The economic value of the individual estimators and their combinations are assessed out-of-sample according to their performances based on straddle option trading strategies. A brief statistical evaluation on the individual and combination estimators has also been implemented.

The empirical study contributes to the literature in two directions. First, we address the research question that whether forecast combination can outperform the best possible single volatility measure in terms of economic value, after transaction cost being deducted. Second, we investigate if the ‘forecast combination puzzle’, which refers to the phenomenon that a naive equal weighted forecast combination turns out to be difficult

for most sophisticated competing combination models to beat, exists in a realized context by comparing the volatility trading performance of the simple and further realized combination models.

Our first finding is that, in general, a simple combination of a realized estimator and implied volatility cannot be outperformed by the best single estimator, the implied volatility, in terms of economic value. The VXO, implied volatility of S&P 100 index, is surpassed by a large number of simple and further combination estimators in terms of the Sharpe ratio generated. This is because an intraday volatility estimator may contain distinct information that is not captured by implied volatility, and a simple combination of the two could take advantage of the relevant information contained in both and leads to better performance economically.

In addition, two further combination models, MIM and regression-based which combines the implied volatility and 41 individual intraday volatility estimators, outperform all other estimators consistently according to the annualized return, standard deviation and Sharpe ratio criteria. This result indicates that more sophisticated forecast combinations, by benefiting from the information contained in different volatility estimators from distinct classes, are able to produce significant economic value judged by the performance of volatility trading.

Finally, the statistical evaluation results are largely in line with the economic ones. Statistically the VXO, MIM combination, and regression-based combination top the ranking table according to their out-of-sample performances. In terms of economic value the further combination models, MIM and regression-based, which both have the VXO as a component, also generate the best performance.

In conclusion, the empirical results shown in the chapter suggest that there is economic value in combining different volatility estimators introduced in the literature to date. The economic value is robust to hedging and transaction costs and further proves the economic benefit in combining realized and implied volatility measures.

TABLE 5.1  
Summary Statistics of 41 Realised Measures

	Mean	Median	Minimum	Std. Dev.	Skewness	Kurtosis
RV_1min	1.042	0.553	81.096	0.040	2.279	16.373
RV_5min	1.298	0.719	73.618	0.036	2.455	11.690
RV_15min	1.338	0.687	54.198	0.034	2.571	8.137
RV_30min	1.303	0.660	54.922	0.019	2.555	8.495
RV_60min	1.109	0.545	58.286	0.007	2.116	10.324
RR_1min	1.041	0.551	81.268	0.041	2.278	16.459
RR_5min	1.290	0.716	74.491	0.034	2.413	12.115
RR_15min	1.311	0.681	54.822	0.029	2.455	8.319
RR_30min	1.224	0.643	59.126	0.016	2.284	9.363
RR_60min	1.049	0.515	28.824	0.002	1.797	6.235
BPV_1min	0.909	0.473	81.786	0.034	2.247	17.659
BPV_5min	1.159	0.621	51.366	0.023	2.200	9.162
BPV_15min	1.122	0.583	55.998	0.026	2.238	9.514
BPV_30min	1.031	0.521	37.668	0.016	2.033	8.597
BPV_60min	0.739	0.347	26.260	0.002	1.439	8.268
RK_2nd_1min	1.286	0.717	96.045	0.037	2.660	16.494
RK_2nd_5min	1.289	0.678	64.302	0.039	2.487	9.568
RK_Bartlett_1min	1.291	0.723	98.165	0.036	2.690	16.881
RK_Bartlett_5min	1.296	0.683	68.494	0.038	2.521	10.107
RK_Cubic_1min	1.298	0.715	83.445	0.037	2.575	13.262
RK_Cubic_5min	1.268	0.660	57.019	0.028	2.432	8.877
RK_Epa_1min	1.302	0.727	92.632	0.035	2.651	15.448
RK_Epa_5min	1.292	0.681	55.917	0.031	2.434	8.462
RK_Parzen_1min	1.287	0.705	76.504	0.038	2.509	11.917
RK_Parzen_5min	1.261	0.667	57.888	0.034	2.398	9.034
RK_Th_1min	1.306	0.728	94.269	0.035	2.665	15.685
RK_Th_5min	1.277	0.671	56.915	0.035	2.438	8.734
RPV_0.5_1min	2.593	2.442	9.049	0.851	0.733	1.925
RPV_0.5_5min	0.890	0.856	2.700	0.268	0.247	1.376
RPV_0.5_15min	0.399	0.383	1.200	0.122	0.113	1.336
RPV_0.5_30min	0.239	0.229	0.664	0.080	0.070	1.279
RPV_0.5_60min	0.154	0.147	0.478	0.048	0.048	1.248
RPV_1_1min	1.592	1.332	17.462	0.282	1.036	4.139
RPV_1_5min	0.897	0.776	7.691	0.151	0.552	3.136
RPV_1_15min	0.535	0.457	4.296	0.088	0.338	2.989
RPV_1_30min	0.381	0.326	2.630	0.055	0.246	2.908
RPV_1_60min	0.294	0.245	2.404	0.028	0.203	3.142
RPV_1.5_1min	1.118	0.771	35.099	0.101	1.391	8.345
RPV_1.5_5min	1.017	0.734	23.353	0.090	1.138	6.047
RPV_1.5_15min	0.804	0.561	15.324	0.065	0.920	5.197
RPV_1.5_30min	0.677	0.471	11.571	0.032	0.787	5.081

Notes: The table presents basic summary statistics on the 41 different realized measures considered in the empirical study. RV stands for realized variance, RR stands for realized range, BPV stands for realized bi-power variation, RPV stands for realized power variation, and RK stands for realized kernel. 1-60min denotes the corresponding sampling frequency.

TABLE 5.2  
Summary Statistics for the Combinations of Realized Measures

	Equal Weights	Geometric Mean	Regression Based	AIC Based	MIM
Mean	4.955	4.660	6.468	3.935	6.405
Median	4.254	4.117	6.185	3.494	6.118
Maximum	50.696	33.482	20.909	27.611	21.349
Minimum	1.298	0.000	2.172	1.109	-3.928
Std. Dev.	3.166	2.689	2.210	2.075	2.209
Skewness	4.122	2.751	1.601	2.893	1.616
Kurtosis	33.447	16.779	8.207	18.139	8.438

Notes: The table presents basic summary statistics on the 5 different realised combinations considered in the empirical study.

TABLE 5.3  
Performance of the G&B Combination with Different Trading Ranges

<i>Granger &amp; Bates</i>	Trading Range				
	0.05%	0.25%	0.5%	0.75%	1%
Sharpe Ratio	1.34	2.12	2.13	1.34	0.92
Annual Return	13.23%	17.62%	17.84%	12.63%	9.60%
Annual Volatility	6.95%	6.43%	6.32%	6.16%	6.07%
No. of Winning Trades	167	159	132	71	40
% Winning	51.09%	53.37%	58.33%	56.80%	61.54%

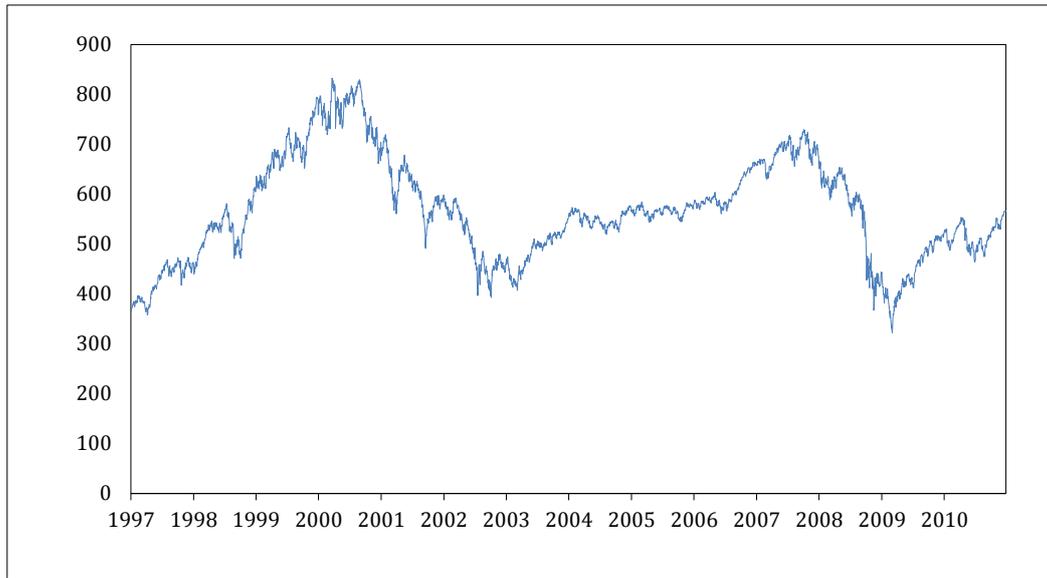
Notes: The table lists out information of the Sharpe ratio, number of winning trades, the percentage of winning, annual return and volatility with different trading ranges. The number of winning trades is the total number of profitable option deals made. The percentage of winning is the percentage of option trades made a profit. The annual return is the annualized profit rate between the 4 years trading period and the annual volatility is the annual standard deviation of the returns produced by the option strategies.

TABLE 5.4  
Estimation Result and Robustness of Selected Models

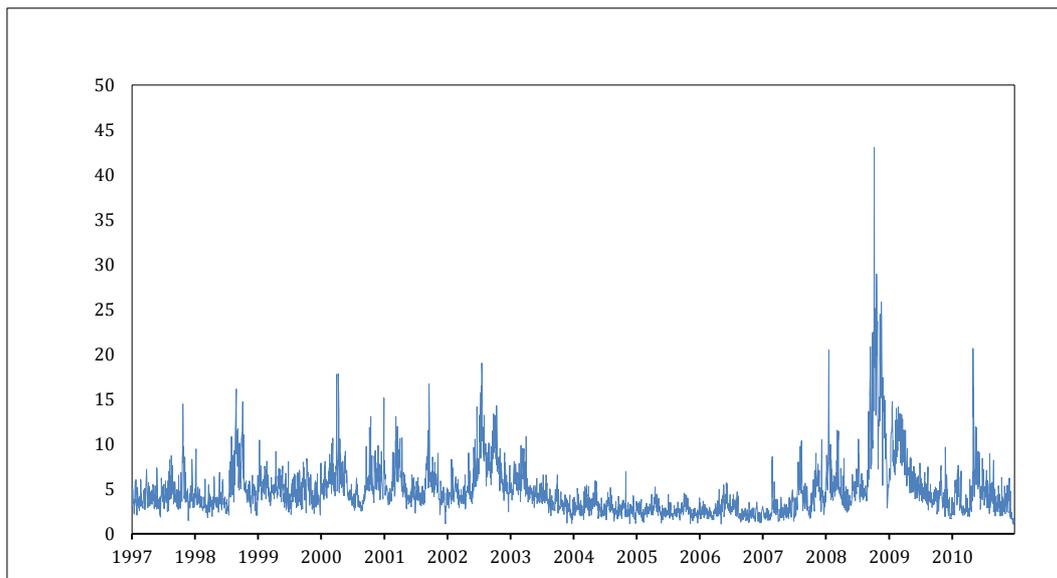
	Regression-Based		MIM		RV 5min	
<i>c</i>	0.494	[0.000]	0.731	[0.000]	0.854	[0.000]
<i>vxo</i>	0.754	[0.000]	0.857	[0.001]	0.853	[0.000]
<i>combi.</i>	0.177	[0.000]	0.071	[0.000]	-	-
<i>rv</i>	-	-	-	-	0.041	[0.001]
$\bar{R}^2$	0.837		0.833		0.833	
<i>S.E.</i>	1.099		1.112		1.112	
<i>Log L</i>	-5285		-5328		-5328	
<i>DW</i>	2.288		2.107		2.196	
<i>AIC</i>	3.027		3.051		3.052	
<i>Obs</i>	2993		2993		2993	

Notes: The table provides estimation results of selected forecasting models. *c* represents the intercept, *vxo* represents the coefficients of the ARFIMA forecasted VXO. *combi.* represents the coefficients of the combination estimator. *rv* represents the coefficient of the realized volatility. The upper panel reports the resulting parameter estimates, together with robust standard errors in parentheses and p-values in square brackets. The lower panel reports the corresponding regression outputs.

FIGURE 5.1  
Price and Volatility Patterns of the S&P 100 Index



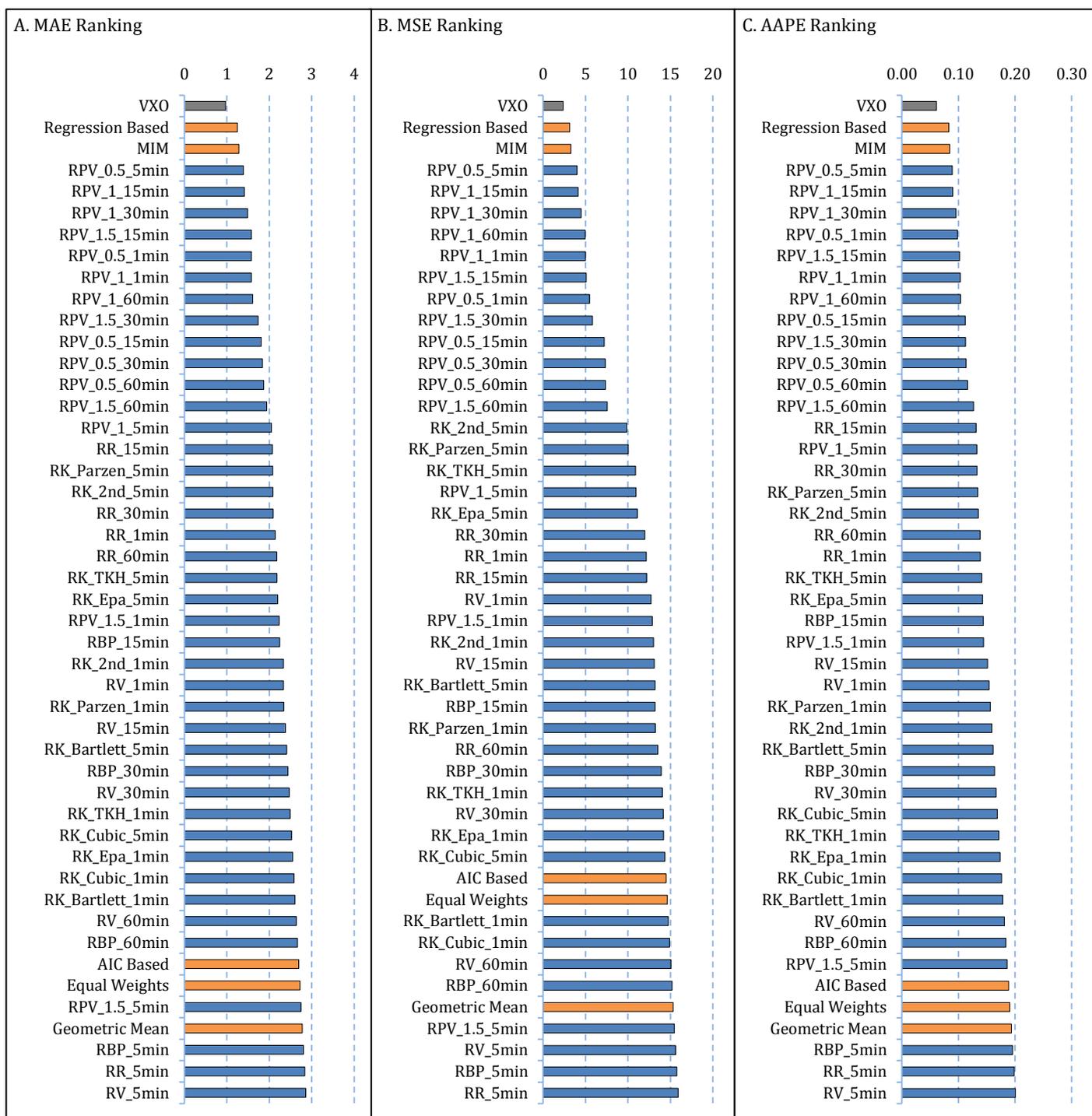
A. S&P 100 Price



B. S&P 100 Volatility

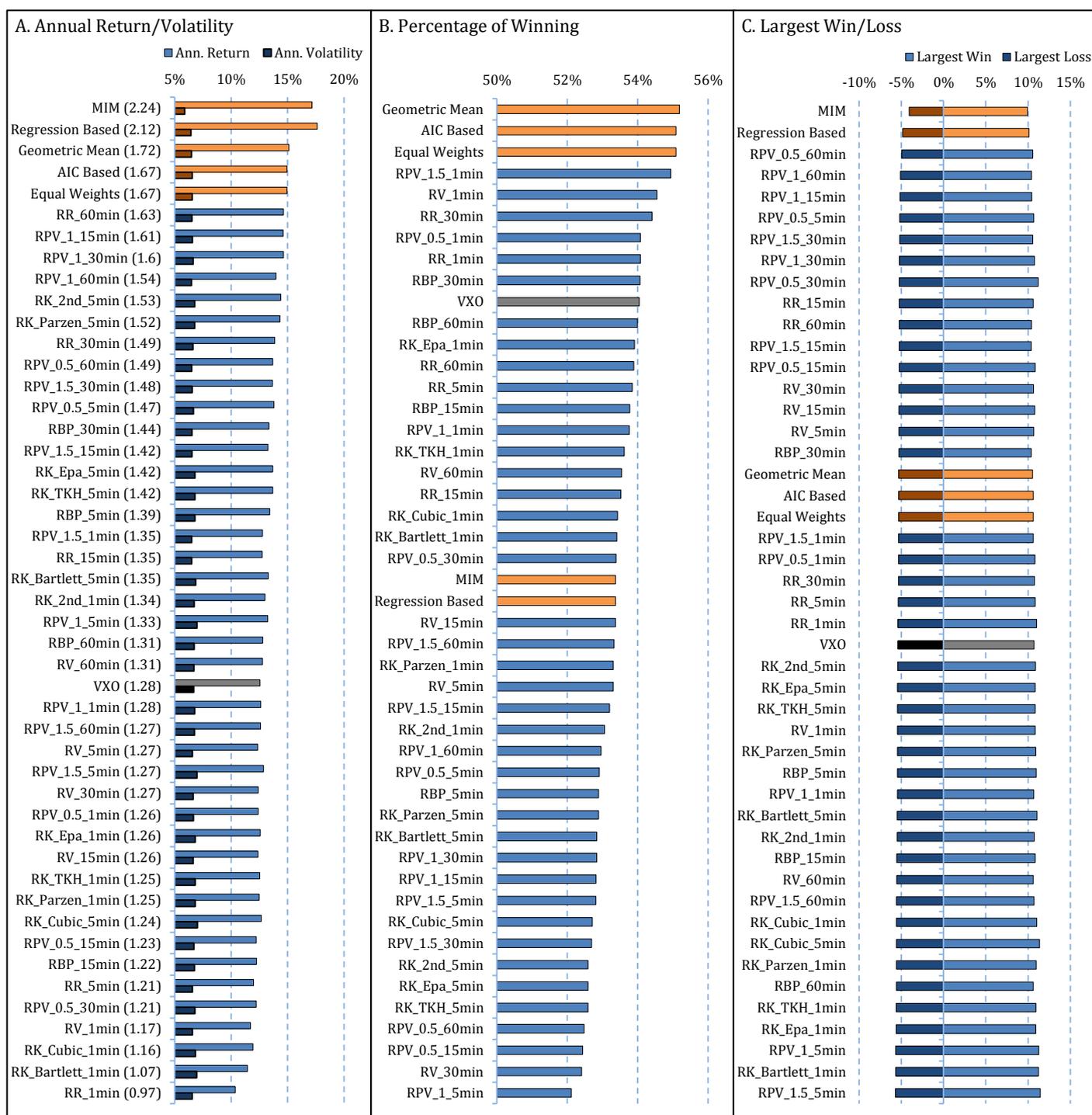
Notes: The figure plots the S&P100 index price and volatility over the period January 1997 to December 2010. The volatility is calculated using realized volatility based on 5-minute calendar-time trade prices, converted to monthly using the formula  $\sigma_t = \sqrt{22 \times RV_{5min}}$ .

FIGURE 5.2  
Statistical Evaluation of the Individual Measures and Their Combinations



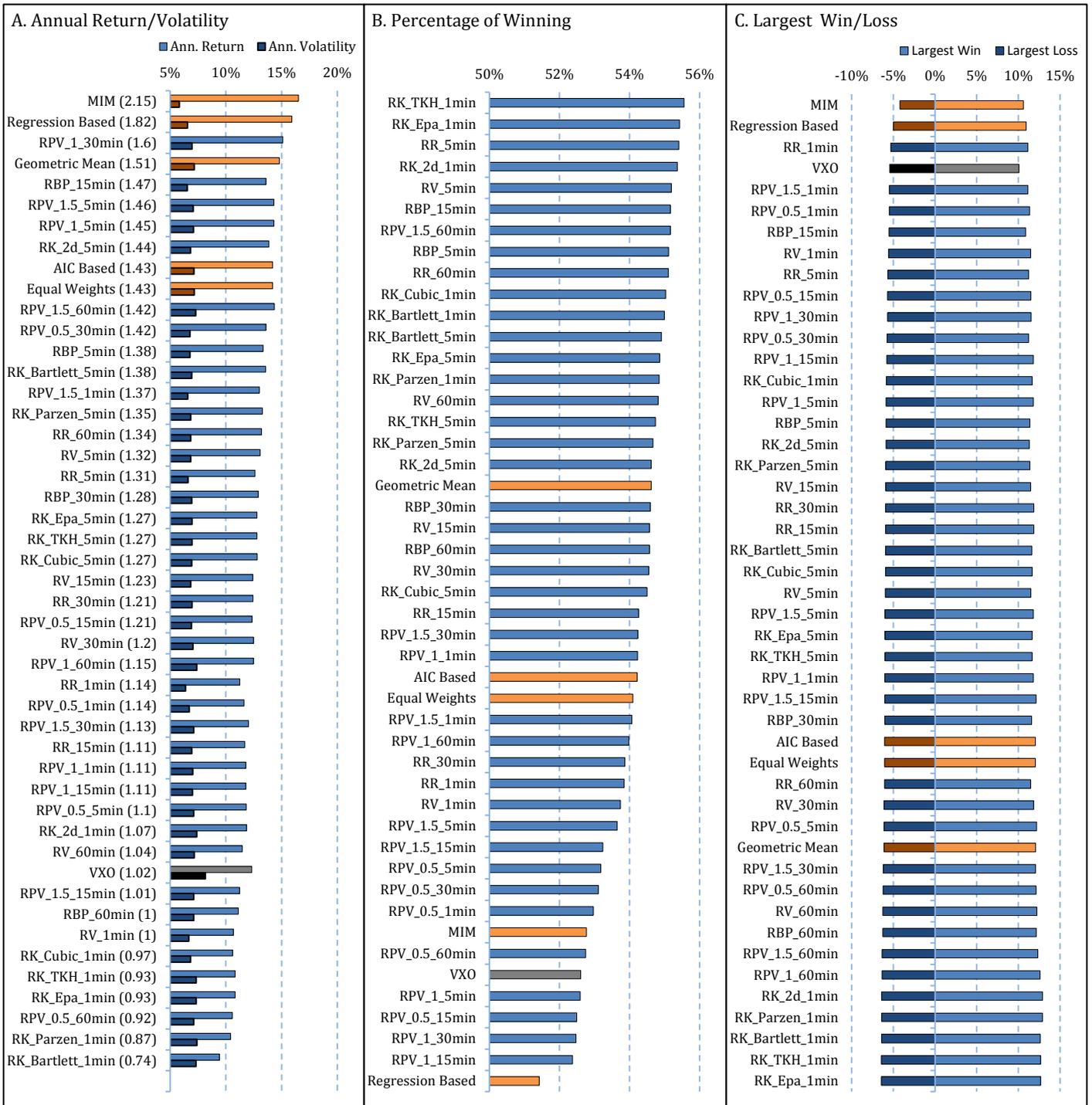
Notes: RV stands for Realized Volatility, RR for Realized Range, RBP for Realized Bi-power Variation, RPV for Realized Power Variation and RK for Realized kernels. The 6 RK specifications are Barlett (BAR), Epanechnikov (Epa), 2nd Order (2nd), Cubic (CUB), Parzen and Tukey-Hanning (TKH). Bars in ■ show corresponding test statistics for individual realized measures, and bars in ■ show that of the combination estimators. VXO is the S&P 100 implied volatility.

FIGURE 5.3  
Empirical Performance of Option Strategies (0.25%)



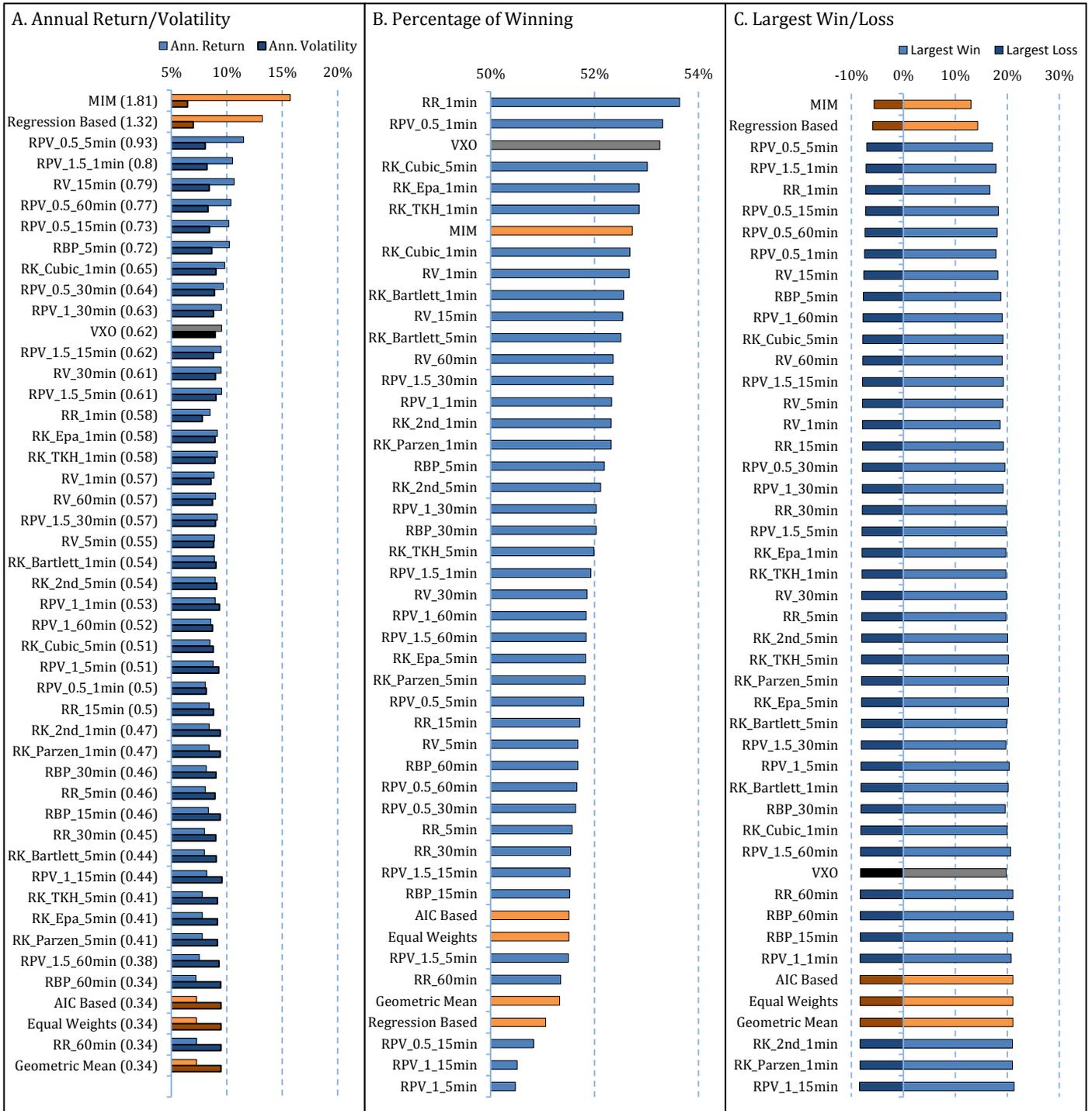
Notes: The simulated trading results are based on different volatility estimators and implemented on a trading range of 0.25%. Models shown in ■ and ■ are individual realized measures, and shown in ■ and ■ are combination estimators. VXO is the benchmark estimator, implied volatility of the S&P 100 index. 5.3.A lists out annualized return (in darker color) and volatility for each estimator. The annual return is the annualized profitability during the 4 years trading period and the annual volatility is the standard deviation of the returns produced by the option strategies. The figure is ranked by Sharpe ratio, which is shown in brackets after the name of the corresponding estimator. 5.3.B shows the percentage of winning, which is the percentage of option trades made a profit. 5.3.C presents the largest win and loss, where the largest win is the largest profit made by a single trade and the largest loss is the largest loss made by a single transaction.

FIGURE 5.4  
Empirical Performance of Option Strategies (0.20%)



Notes: The simulated trading results are based on different volatility estimators and implemented on a trading range of 0.2%. See note to Figure 5.3.

FIGURE 5.5  
Empirical Performance of Option Strategies (0.30%)



Notes: The simulated trading results are based on different volatility estimators and implemented on a trading range of 0.3%. See note to Figure 5.3.

# CHAPTER 6: CONCLUSION

## 6.1 Concluding Remarks

Volatility modelling and forecasting has been one of the most exciting and successful areas of research in financial econometrics in recent decades. This thesis provides a comprehensive investigation on the economic or practical value of volatility forecasting, with a distinct focus on the benefit of utilizing intraday (co)variation estimators. The thesis presents three empirical studies covering the three main practical applications of volatility forecasting: portfolio optimization, risk management and volatility trading. In general each empirical work follows a similar structure. First, univariate (variance) or multivariate (covariance matrix) intraday volatility estimators are constructed using real market data. Second, the intraday volatility estimators are used to augment univariate GARCH or multivariate GARCH class models in order to extract the economic value of intraday information. Finally, this economic value (if any) is compared and assessed by one of the three aforementioned economic criteria.

The empirical research is motivated by three main factors. First, given the central role of volatility forecasting in financial economics and practical applications, it is always tempting to search for a better volatility model. The second reason is that while the statistical accuracy of volatility forecasts has been analysed extensively, the economic value of the predictions is a relatively new area of interest. Since statistical accuracy does not always translate into profit in the real market, it is arguably more efficient to evaluate volatility forecasts directly in an economic framework. Furthermore, most empirical studies with economic value in mind focus on only daily or lower frequency data, leaving an enormous space for analysing the topic in a realized context.

We start by analysing the economic value of incorporating a realized covariance estimator into standard daily return based multivariate GARCH (MGARCH) models in the context of volatility timing. Although a number of studies have provided evidence

in favour of MGARCH specifications, the literature that makes use of intraday information for volatility timing is based only on nonparametric rolling window covariance estimators. Chapter 3 seeks to complement the literature in this respect. We deploy a dynamic optimal-weight portfolio strategy based on one-day-ahead covariance matrix forecasts from a model-free EWMA approach and three competing MGARCH models, BEKK, CCC and DCC. Each of these covariance models are augmented with realized covariance matrix to construct the realized covariance models, REWMA, RBEKK, RCCC and RDCC. The analysis is applied to a portfolio of three indices, NASDAQ 1000, Russell 2000 and the CRB commodity index.

An economic loss function based mean-variance portfolio optimization strongly suggests that intraday based covariance models outperform their daily return based counterparts. The ranking of realized covariance models depends on the decision-maker's portfolio rebalancing strategy. The nonparametric benchmark model, realized exponential weighted moving average (REWMA), is well suited for time-fixed rebalancing strategies, while more sophisticated parametric realized MGARCH (RMGARCH) models are better choices for time-varying rebalancing strategies. In general the realized CCC approach with a volume-driven rebalancing strategy produces the highest economic value, which is quantified as the maximum annualized fee in basis points that a representative investor would be willing to pay in order to switch from competing volatility estimators. Our results in Chapter 3 provide further evidence, which is new in the context of covariance forecasting, that realized covariance models, especially those sophisticated parametric ones, do have incremental economic value compared with volatility specifications based on data sampled at lower frequencies.

Another important practical application of volatility forecasting is risk management. The literature has evaluated both single index (univariate) and portfolio (multivariate) volatility models in terms of the forecast power of value-at-risk (VaR) forecasts. However, the univariate and multivariate model comparison has yet to be addressed in a high-

frequency context. Building on the RGARCH (realized GARCH) models proposed by Fuertes, Kalotychou and Izzeldin (2009) and the RMGARCH models studied in Chapter 3, Chapter 4 complements the literature by comparing the practical value of intraday based single index and portfolio models through the lens of VaR forecasting. VaR predictions are generated from standard daily univariate or multivariate GARCH class models, and GARCH class models extended with ARFIMA forecasted realized measures. Out-of-sample VaR predictions are assessed by a number of conditional coverage tests. Two research questions are addressed in the chapter. The first one is that whether realized volatility models can significantly improve the prediction power of classic daily GARCH class models. The second one is that, in a high-frequency context, which group of forecasts, single index or portfolio volatility models, delivers the best VaR prediction.

The empirical results provide further support for realized volatility estimators through the angle of VaR prediction, where the intraday based volatility models outperform their daily counterparts by providing more adequate VaR forecasts for a prolonged out-of-sample forecasting period. With regard to the univariate and multivariate comparison, we find that both the realized single index and realized portfolio models generate adequate VaR predictions. However, neither of them is statistically superior based on the backtesting results. The realized portfolio models provide more accurate coverage ratio while the realized single index specifications show smaller average and maximum absolute deviation of violations. Nevertheless, given the parsimonious nature of the realized single index models, they are probably more suited for forecasting VaR in daily practices.

Builds on the empirical study of realized univariate and multivariate volatility models in Chapter 3 and 4, Chapter 5 moves one step further by analysing forecast combination. The forecasting performance of a number of intraday-based realized volatility estimators, implied volatility and combinations of them are assessed in a volatility trading framework. Volatility forecast are generated by the ARFIMA model for each volatility estimator. Economic gain of a volatility estimator is assessed through the return accrued using volatility

trading strategies based on the corresponding volatility forecasts. The study is applied to 14 years of tick-by-tick data of S&P 100 index and its daily option prices. Evaluations are made on two fronts. First, we are interested in whether simple combination models, which combines a realized volatility estimator and implied volatility, can outperform the best individual volatility measure. Second, can further combination models, which combine 42 individual volatility estimators from 6 distinct classes of specifications, provide further economic gains compared with simple combination models.

The test results show that, in general, a simple combination of a realized estimator and implied volatility cannot be outperformed by the best single estimator in terms of economic value. A realized estimator may contain unique information that is not captured by implied volatility, a combination of the two can benefit from the relevant information provided by both measures. In addition, two further combination models, the multiple indicators model (MIM) and regression based, outperform all other estimators consistently according to annualized return, standard deviation and Sharpe ratio generated. Hence there is significant economic value in combining a variety of volatility estimators sampled at different time frequencies.

In conclusion, the thesis provides a comprehensive study of economic value of intraday based volatility estimators. The test results of all three empirical works show significant economic value when intraday information is incorporated into the volatility forecasting process. The findings are proved to be robust to transaction cost and dataset changes. Furthermore, the result is valid on both the univariate and multivariate volatility modelling frameworks. Furthermore, Chapter 5 shows that the economic value can be further improved via combining realized volatility estimators from distinct types of specifications.

## **6.2 Further Research**

It is important to recognize the complexity of future volatility of asset returns in the sense that return variations are determined by the interaction of economic fundamentals

and a range of unobserved factors – political changes, natural disasters, and physiological attributes could equally determine the magnitude of asset price swings. The search for better approaches to forecast volatility and, more broadly, to deploy volatility forecasts in practical applications such as portfolio optimization, risk management and volatility trading, will almost surely remain a challenging yet exciting area of research.

The results of our empirical studies point to some interesting directions for future study. One possibility is to evaluate newly proposed intraday (co)variation estimators to see whether forecasting accuracy can be further improved. For instance, a number of novel realized covariance matrix estimators have been introduced over the past two years. Estimators addressing both the non-synchronicity and the microstructure noise have been proposed by Zhang (2010), Barndorff-Nielsen, Hansen, Lunde and Shephard (2011) and Ait-Sahalia, Fan and Xiu (2010). Most recently, Park and Linton (2012) propose a new estimator of multivariate ex-post volatility that is robust to microstructure noise and asynchronous data timing. The method is based on Fourier domain techniques, which have been widely applied in discrete time series modelling. The proposed Fourier Realized Kernel (FRK) estimator is shown to outperform the realized covariance estimator especially when two assets are traded very asynchronously and with different liquidity and when estimating the high dimensional integrated covariance matrix. It would be interesting to see whether these newly proposed realized estimators would further improve the economic value of the realized volatility models.

In addition, instead of incorporating realized volatility into the standard volatility models, one can add other pieces of intraday information into the specifications. For instance, intraday trading volumes may possess distinct information in revealing future volatility. In this area, the multiplicative error model (MEM) introduced by Engle (2002) serves as a workhorse for the modelling of non-negative, serially dependent intraday data. MEM estimates its parameters over long estimation horizon in order to increase estimation efficiency. The model performs well in modelling financial duration data and intraday

trading volumes, see, e.g., Manganelli (2005), Brownlees et al. (2011) and Hautsch et al. (2011), among others. Most recently, Hardle et al. (2012) propose a local adaptive MEM accommodating time-varying parameters. A data-driven optimal length of local windows is selected, yielding adaptive forecasts at each point in time. By analysing one-minute cumulative trading volumes of five large NASDAQ stocks in 2008, they show that a local window of approximately 3 to 4 hours are reasonable to capture parameter variations while balancing modelling bias and estimator efficiency. These intraday volume estimators could be augmented into the realized models to provide an extra layer of information.

Last but not least, another potential avenue is to incorporate realized volatility estimators into more sophisticated GARCH class or alternative classes of volatility models. For example, Aielli (2011) show that the DCC large system estimator can be inconsistent, and that the traditional interpretation of the DCC correlation parameters can lead to misleading conclusions. The author suggests a more tractable dynamic conditional correlation model (cDCC), which reformulates the correlation driving process of the DCC model. Tested with real data, he finds that the cDCC multi-step-ahead correlation forecasts have been proven to perform equally or significantly better than the corresponding DCC forecasts. With more sophisticated (co)variation models, the economic value of intraday data could be further improved.

## References

- [1] Ahoniemi, K., Fuertes, A.-M., Olmo, J. (2012), Overnight News and Equity Trading Risk Limits. SSRN Working Paper.
- [2] Aielli, G.P. (2011), Dynamic Conditional Correlation: on Properties and Estimation. Seminar working paper.
- [3] Aiolfi, M., Capistran, C., Timmermann, A.G. (2010), Forecast Combinations. CRE-ATES Research Paper No. 2010-21.
- [4] Ait-Sahalia, Y., Fan, J., and Xiu, D. (2010), High Frequency Covariance Estimates with Noisy and Asynchronous Financial Data. *Journal of the American Statistical Association*, 94, 9-51.
- [5] Andersen, T.G., T. Bollerslev (1998), Answering the Sceptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts. *International Economic Review* 39, 885-905.
- [6] Andersen, T.G., Bollerslev, T., Diebold, F.X., Ebens, H. (2001), The Distribution of Realized Stock Return Volatility. *Journal of Financial Economics* 61, 43-76.
- [7] Andersen, T., Bollerslev, T, Diebold, F.X, & Labys, P. (2001), The Distribution of Realised Exchange Rate Volatility. *Journal of the American Statistical Association*, 96, 42-55.
- [8] Andersen, T., Bollerslev, T, Diebold, F.X, & Labys, P. (2003), Modeling and Forecasting Realized Volatility. *Econometrica* 71, 579-625.
- [9] Andersen, T.G., T. Bollerslev, F.X. Diebold. (2008), Parametric and Nonparametric Measurement of Volatility. *Handbook of Financial Econometrics*. Amsterdam: North Holland.

- [10] Areal, N., Taylor, S.J. (2002), The Realized Volatility of FTSE-100 Futures Prices. *Journal of Futures Markets* 22, 627-648.
- [11] Bandi, F.M. and Perron, B. (2006), Long Memory and the Relation Between Implied and Realized Volatility. *Journal of Financial Economics*, 4, 636-670.
- [12] Barndorff-Nielsen, O.E., Shephard, N. (2002). Econometric Analysis of Realised Volatility and its Use in Estimating Stochastic Volatility Models. *Journal of the Royal Statistical Society, Series B*, 64, 253–280.
- [13] Barndorff-Nielsen, O.E., Shephard, N. (2004), Power and Bipower Variation with Stochastic Volatility and Jumps. *Journal of Financial Econometrics*, 2, 1-48.
- [14] Barndorff-Nielsen, O.E., Shephard, N. (2006). Econometrics of Testing for Jumps in Financial Economics using Bipower Variation. *Journal of Financial Econometrics* 4, 1-30.
- [15] Barndorff-Nielsen, O.E., Hansen, P.R., Lunde, A., Shephard, N. (2008), Designing Realized Kernels to Measure the ex post Variation of Equity Prices in the Presence of Noise. 76, 1481-1536.
- [16] Barndorff-Nielsen, O. E., Hansen, P.R., Lunde, A., Shephard, N. (2009), Realised Kernels in Practice: Trades and Quotes. *Econometrics Journal* 12, 1-33.
- [17] Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and Shephard, N. (2011), Multivariate Realized Kernels: Consistent Positive Semi-definite Estimators of the Covariation of Equity Prices with Noise and Non-synchronous Trading, *Journal of Econometrics*, 162, 149-169.
- [18] Bates, J. M., Granger, C.W.J. (1969), The Combination of Forecasts. *Operational Research Quarterly*, 20, 451–468.

- [19] Bauwens, L., Laurent, S., Rombouts, J. (2006), Multivariate GARCH Models: A Survey. *Journal of Applied Econometrics* 21 (1), 79–109.
- [20] Becker, R., Clements, A.E., White, S. (2007), Does Implied Volatility Provide any Information Beyond that Captured in Model-Based Volatility Forecasts?. *Journal of Banking and Finance* 31, 2535-2549.
- [21] Becker, R., Clements, A.E. (2008), Are Combination Forecasts of S&P 500 Volatility Statistically Superior? *International Journal of Forecasting*, 24, 122-133.
- [22] Berkowitz, J. and O'Brien, J. (2002), How Accurate are Value-at-Risk Models at Commercial Banks. *The Journal of Finance* 57 (3), 1093-1111.
- [23] Bessembinder, H., Seguin, P.J. (1993), Price Volatility, Trading Volume and Market Depth: Evidence from Future Markets. *Journal of Financial and Quantitative Analysis*, 28, 21-39.
- [24] Bhardwaj, G. and Swanson, N. (2006), An Empirical Investigation of the Usefulness of ARFIMA Models for Predicting Macroeconomic and Financial Time Series, *Journal of econometrics*, 131, 539-578.
- [25] Bjornland, H.C., Gerdrup, K., Jore, A.S., Smith, C., Thorsrud, L.A. (2012), Does Forecast Combination Improve Norges Bank Inflation Forecasts? *Oxford Bulletin of Economics and Statistics*, 74, 163-179.
- [26] Blair, B.J., Poon, S.H., Taylor, S.J. (2001), Forecasting S&P 100 Volatility: the Incremental Information Content of Implied Volatilities and High-Frequency Index Returns. *Journal of Econometrics*, 105, 5-26.
- [27] Bollerslev, T., Engle, R.F., Nelson, D. (1994), ARCH Models Engle, R.F. and McFadden, D.L. (Eds), *Handbook of Econometrics*, IV Elsevier (Chapter 49).

- [28] Bollerslev, T. (1986), Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics* 31, 307-327.
- [29] Bollerslev, T. (1990), Modeling the Coherence in Short-Run Nominal Exchange Rates: A Multivariate Generalized ARCH Model. *Review of Economics and Statistics* 72, 498-505.
- [30] Bolleslev, T., Engle, R. F., and Wooldridge, J., (1988), Capital Asset Pricing Model with Timevarying Covariances. *Journal of Political Economy* 96, 116–131.
- [31] Breen, W., Glosten, L.R., Jagannatan, R. (1989), Economic Significance of Predictable Variations in Stock Index Returns, *Journal of Finance* 44, 1177-1189.
- [32] Brooks, C. and G. Persand (2003), Volatility Forecasting for Risk Management. *Journal of Forecasting* 22 (1), 1–22.
- [33] Brownlees, C.T., Cipollini, F. and Gallo, G.M. (2011). Intra-daily Volume Modeling and Prediction for Algorithmic Trading, *Journal of Financial Econometrics* 9(3): 489-518.
- [34] Campbell, J.Y., Lo, A.W., and Mackinlay A.G. (1997), *The Econometrics of Financial Markets*. Princeton University Press, Princeton.
- [35] Chen, Z., Knez, P.J. (1996), Portfolio Performance Measurement: Theory and Applications. *Review of Financial Studies* 9, 511-556.
- [36] Christensen, K., & Podolskij, M. (2005), Realised Range-Based Estimation of Integrated Variance. Aarhus School of Business, mimeo.
- [37] Christoffersen, P. (1998), Evaluating Interval Forecasts. *International Economic Review* 39 (4), 841–862.
- [38] Christoffersen, P. (2009), Value-at-Risk Models. In T. Andersen, R. Davis, J.-P. Kreiss, and T. Mikosch (Eds.), *Handbook of Financial Time Series*. Springer Verlag.

- [39] Chou, R.Y., Liu, N. (2008), The Economic Value of Volatility Timing Using a Range-based Volatility Model. Conference Paper, China International Conference in Finance.
- [40] Clark, P.K. (1973), A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices. *Econometrica*, 41, 135-155.
- [41] Clemen, R.T. (1989), Combining Forecasts: A Review and Annotated Bibliography. *International Journal of Forecasting*, 5, 559–583.
- [42] Clements, M. P., Galvao, A.B., & J.H. Kim (2008), Quantile Forecasts of Daily Exchange Returns from Forecasts of Realized Volatility. *Journal of Empirical Finance*, 15, 729-750.
- [43] Clements, A. and Silvennoinen, A. (2012), Volatility Timing and Portfolio Selection: How best to forecast volatility. NCER Working Paper, Queensland University of Technology.
- [44] Copeland M.M. and Copeland, T.E. (1999) Market Timing: Style and Size Rotation using VIX. *Financial Analysts Journal*, 55, 73–81.
- [45] Corrado, C.J., Miller, T.W. (2005). The Forecast Quality of CBOE Implied Volatility Indexes. *Journal of Futures Markets*, 25, 339-373.
- [46] Corsi, F. (2004). A Simple Long Memory Model of Realized Volatility, University of Lugano.
- [47] Della Corte, P.D., Sarno, L., and Tsiakas, I. (2009), Correlation Timing in Asset Allocation: Evidence from the Foreign Exchange Market. Working Paper.
- [48] DeMiguel, V., L. Garlappi, and R. Uppal (2009), Optimal versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy? *Review of Financial Studies* 22 (5), 1915–1953.

- [49] Ding, Z., C.W.J. Granger and R.F. Engle (1993). A Long Memory Property of Stock Market Returns and a New Model. *Journal of Empirical Finance*, 1, 83-106.
- [50] Dumitrescu, E.I., Hurlin, C., Pham, V. (2012), Backtesting Value-at-Risk: From Dynamic Quantile to Dynamic Binary Tests. Working Papers halshs-00671658, HAL.
- [51] Ebens, H. (1999), Realized Stock Volatility. Working Paper, Department of Economics, John Hopkins University.
- [52] Engle, R.F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50, 987-1007.
- [53] Engle, R.F. (2002a) New Frontiers for ARCH Model, *Journal of Applied Econometrics* 17: 425-446.
- [54] Engle, R.F. (2002b), Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models. *Journal of Business and Economic Statistics* 20, 339-350.
- [55] Engle, R.F., Colacito R. (2006), Testing and Valuing Dynamic Correlations for Asset Allocation. *Journal of Business and Economic Statistics*, 24, 238–253.
- [56] Engle, R.F., Gallo, G. (2006). A Multiple Indicators Model for Volatility using Intra-daily Data. *Journal of Econometrics* 131, 3-27.
- [57] Engle, R.F., Gonzalez-Rivera, G. (1991), Semiparametric ARCH Models. *Journal of Business & Economic Statistics* 9, 345-359.
- [58] Engle, R.F., Kroner, K. F. (1995), Multivariate Simultaneous GARCH. *Econometric Theory*, 11, 122-150.

- [59] Engle, R.F., Manganelli, S. (2004), CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles. *Journal of Business & Economic Statistics* 22 (4), 367-382.
- [60] Engle, R.F., Sheppard, K. (2001), Theoretical and Empirical Properties of Dynamic Conditional Correlation MVGARCH. UCSD Working Paper No. 2001-15.
- [61] Engle, F., Sokalska, M.E., Chanda, A. (2006), Forecasting Intraday Volatility in the US Market. Stern School of Business, NYU, mimeo.
- [62] Fama, E., French, K. (1988), Permanent and Temporary Components of Stock Prices, *Journal of Political Economy* 96, 246-273.
- [63] Fleming, J. (1998), The Quality of Market Volatility Forecasts Implied by S&P 100 Index Option Prices. *Journal of Empirical Finance*, 5, 317-345.
- [64] Fleming, J., Kirby, C., Ostdiek, B. (2001), The Economic Value of Volatility Timing. *Journal of Finance* 56, 329-352.
- [65] Fleming, J., Kirby, C., Ostdiek, B. (2003), The Economic Value of Volatility Timing Using “Realized” Volatility. *Journal of Financial Economics* 67, 473-509.
- [66] Forsberg, L., Ghysels, E. (2007), Why Do Absolute Returns Predict Volatility So Well? *Journal of Financial Economics*, 5, 31-67.
- [67] Foster, D.P. and Nelson, D.B. (1996), Continuous Record Asymptotics for Rolling Sample Variance Estimators, *Econometrica* 64, 139-174.
- [68] Franses, P.H., van Dijk, D. (2000), *Nonlinear Time Series Models in Empirical Finance*. Cambridge University Press.
- [69] Fuertes, A.-M., Kalotychou, E., Izzeldin M. (2009), Forecasting Daily Stock Volatility: the Role of Intraday Information and Market Conditions. *International Journal of Forecasting* 25, 259-281.

- [70] Fuertes, A.-M., Olmo, J. (2012), The Role of High-Frequency Prices, Long Memory and Jumps for Value-at-Risk Prediction. SSRN Working Paper.
- [71] Ghysels, E., Harvey, A., Renault, E. (1996), Stochastic volatility, pp119-191, in Maddala G.S. and C.R. Rao eds. Handbook of Statistics: Statistical Methods in Finance, Vol.14, Amsterdam: Elsevier Science.
- [72] Glosten, L.R., Jagannathan, R., Runkle, D.E. (1992), On the Relation Between the Expected Value and Volatility of the Nominal Excess Return on Stocks. Journal of Finance, 46, 1779-1801.
- [73] Gouriéroux, C., Jasiak, J. (2001), Financial Econometrics. Princeton University Press: Princeton.
- [74] Graham, J.R., Harvey, C.R. (1996), Market Timing Ability and Volatility Implied in Investment Newsletters' Asset Allocation Recommendations. Journal of Financial Economics, 42(3), 397-421.
- [75] Grant, D. (1977), Portfolio Performance and the "Cost" of Timing Decisions. Journal of Finance 32, 837-846.
- [76] Hansen, P.R., Lunde, A. (2005), A Forecast Comparison of Volatility Models: Does Anything Beat a GARCH(1,1)? Journal of Applied Econometrics, 20, 873-889.
- [77] Hansen, P.R., Lunde, A. (2006), Realised Variance and Market Microstructure Noise. Journal of Business and Economic Statistics, 24, 127-161.
- [78] Hardle, W.K., Hautsch, N., Mihoci, A. (2012), Local Adaptive Multiplicative Error Models for High-Frequency Forecasts, SFB 649 Discussion Paper 2012-031.
- [79] Harvey, C.R., Whaley, R.E. (1992), Market Volatility Prediction and the Efficiency of the S&P 100 Index Option Market. Journal of Financial Economics 31, 43-73.

- [80] Hautsch, N., Malec, P. and Schienle, M. (2011), Capturing the zero: A New Class of Zero-Augmented Distribution and Multiplicative Error Processes, Submitted to *J. Financial Econometrics*.
- [81] Hendry, D.F., Clements, M.P. (2004), Pooling of Forecasts. *Econometrics Journal* 7, 1–31.
- [82] Hsieh, D.A. (1991), Chaos and Nonlinear Dynamics: Application to Financial Markets. *Journal of Finance* 46, 1839-1877.
- [83] Huang, H.Y., Lee, T.H. (2010), To Combine Forecasts or to Combine Information? *Econometric Reviews*, 29, 534-570.
- [84] Hull, J., White, A. (1987), The Pricing of Options on Assets with Stochastic Volatilities. *Journal of Finance* 42, 281-300.
- [85] Hendry, D.F., Clements, M.P. (2004), Pooling of Forecasts. *Econometrics Journal* 7, 1–31.
- [86] Hurlin, C., Gilbert, C., Tokpavi, S., Candelon, B. (2008), Backtesting Value-at-Risk: A GMM Duration-Based Test. Working Papers halshs-00329495, HAL.
- [87] Jorion P. (1995), Predicting Volatility in the Foreign Exchange Market. *Journal of Finance* 50, 507-528.
- [88] Jorion, P. (2000), Risk Management Lessons from Long-Term Capital Management.” *European Financial Management*. 6, 277–300.
- [89] JPMorgan, (1996), RiskMetrics Technical Document. 4th ed.
- [90] Kalotychou, E., Staikouras, S.K., Zhao, G. (2009), On the Dynamics of Asset Correlations: What Matters to Portfolio Managers, Cass Business School.

- [91] Kapetanios, G., Labhard, V., Price, Simon. (2008), Forecast Combination and the Bank of England's Suite of Statistical Forecasting Models. *Economic Modelling* 25, 772-792.
- [92] Kinlay, J. (2005), Forecasting and Trading Volatility in the S&P 100 index - An Empirical Test of Options Market Efficiency. Working Paper.
- [93] Koopman, S.J., Jungbacker, B., Hol, E. (2005), Forecasting Daily Variability of the S&P 100 Stock Index Using Historical, Realized and Implied Volatility Measurements. *Journal of Empirical Finance*, 12, 445-475.
- [94] Korn, R., Trautmann, S. (1995), Continuous Time Portfolio Optimization under Terminal Wealth Constraints. *Math. Methods Oper. Res.* 42, 69-92.
- [95] Korn, R. (1998), Portfolio Optimization with Strictly Positive Transaction Costs and Impulse Control. *Finance and Stochastics*, 2, 85-114.
- [96] Kuester, K., Mittnik, S., Paoletta, M. (2006), Value-at-Risk Prediction: A Comparison of Alternative Strategies. *Journal of Financial Econometrics*, 4, 53-89.
- [97] Laux, P.A., Ng, L.K. (1993), The Sources of GARCH: Empirical Evidence from an Intraday Returns Model Incorporating Systematic and Unique Risks. *Journal of International Money and Finance*, 12, 543-560.
- [98] Leippold, M., Trojani, F. and Vanini, P. (2004), A Geometric Approach to Multi-period Mean Variance Optimization of Assets and Liabilities. *Journal of Economic Dynamics & Control*, 28, 1079-1113.
- [99] Li, D., Ng, W.L. (2000), Optimal Dynamic Portfolio Selection: Multi-period Mean-Variance Formulation. *Math. Fin* 10, 387-406.
- [100] Liu, C., Maheu, J. (2005), Modeling and Forecasting Realised Volatility: The Role of Power Variation. University of Toronto, Department of Economics, Mimeo.

- [101] Liu, Q. (2004), On Portfolio Optimization: How Do We Benefit from High-Frequency Data. working paper, University of Hawaii.
- [102] Ling, S., McAleer, M. (2003), Asymptotic theory for a vector ARMA-GARCH model, *Econometric Theory* 19, 278-308.
- [103] Lopez, J.A. (1998), Methods for Evaluating value-at-Risk Estimates. *Economic Policy Review*, Vol. 4, 3.
- [104] Makridakis, S. (1993), Accuracy Measures: Theoretical and Practical Concerns. *International Journal of Forecasting*, 9, 527-529.
- [105] Makridakis, S., Hibon, M. (2000), The M3-Competition: Results, Conclusions, and Implications. *International Journal of Forecasting*, 16, 451-476.
- [106] Mancino, M.E., Sanfelici, S. (2008), Robustness of Fourier Estimator of Integrated Volatility in the Presence of Microstructure noise. *Computational Statistics & Data Analysis* 52, 2966-2989.
- [107] Manganelli, S. (2005), Duration, Volume and Volatility Impact of Trades, *Journal of Financial Markets* 8: 377-399.
- [108] Markowitz, H. (1952), The Utility of Wealth. *Journal of Political Economy*, 60, 151-158.
- [109] Martens, M. (2001), Forecasting Daily Exchange Return Volatility using Intraday Returns. *Journal of International Money and Finance*, 20, 1-23.
- [110] Martens, M., D. van Dijk, de Pooter, M. (2008), Predicting the Daily Covariance Matrix for S&P 100 Stocks Using Intraday Data – But Which Frequency to Use? *Econometric Reviews* 27, 199-229.

- [111] Martens, M. and Zein, J. (2004), Predicting Financial Volatility: High-frequency Time-series Forecasts vis-a-vis Implied Volatility. *Journal of Futures Markets* 24, 1005-1028
- [112] Martens, M., van Dijk, D. (2006), Measuring Volatility with the Realised Range. *Journal of Econometrics*, 138, 181-207.
- [113] McAleer, M. (2009), The Ten Commandments for Optimizing Value-at-Risk and Daily Capital Charges. *Journal of Economic Surveys* 23 (5), 831–849.
- [114] McAleer, M., da Veiga, B. (2008), Single-Index and Portfolio Models for Forecasting Value-at-Risk Thresholds. *Journal of Forecasting*, John Wiley & Sons, Ltd., vol. 27(3), 217-235.
- [115] McNeil, A.J., Frey, R. (2001), Estimation of Tail-Related Risk Measures for Heteroscedastic Financial Time Series: An Extreme Value Approach. *Journal of Empirical Finance*, 7, 271-300.
- [116] Merton, R.C. (1980), On Estimating the Expected Return on the Market: An Exploratory Investigation. *Journal of Financial Economics* 8, 323-361.
- [117] Nelson, D. B. (1991), Conditional heteroskedasticity in Asset Returns: A New Approach. *Econometrica* 59: 347-370.
- [118] Newbold, P., Harvey, D.I. (2002), Forecast Combination and Encompassing. In: Clements, M., Hendry, D.F. (Eds.), *A Companion to Economic Forecasting*. Basil Blackwell, Oxford (Chap. 12).
- [119] Noh, J., Engel, R.F., Kane, A. (1994), Forecasting Volatility and Option Prices of the S&P 500 Index. *Journal of Derivatives* 2(1), 17-30.
- [120] Oomen, R.C.A. (2006), Properties of Realized Variance under Alternative Sampling Schemes. *Journal of Business and Economic Statistics*, 24, 219-237.

- [121] Opdyke, J.D. (2008), Comparing Sharpe Ratios: So Where are the p-values? *Journal of Asset Management*, 8(5).
- [122] Park, S., Linton, O. (2012), Realized Volatility: Theory and Applications, in *Handbook of Volatility Models and Their Applications* (eds L. Bauwens, C. Hafner and S. Laurent). John Wiley & Sons.
- [123] Patton, A.J., Sheppard, K. (2009), Optimal Combinations of Realised Volatility Estimators. *International Journal of Forecasting* 25, 218-238.
- [124] Pagan, A.R. and W. Schwert (1990), Alternative Models for Conditional Stock Volatility. *Journal of Econometrics*, 50, 267-290.
- [125] Parkinson, M. (1980), The Extreme Value Method for Estimating the Variance of the Rate of Return. *Journal of Business*, 53, 61-65.
- [126] Pesaran, M.H., Timmermann, A. (1995), Predictability of Stock Returns: Robustness and Economic Significance. *The Journal of Finance*, 50, 4, 1201-1228.
- [127] Petkova, R. (2006), Do the Fama-French Factors Proxy for Innovations in Predictive Variables? *Journal of Finance* 61, 581-612.
- [128] Pong, E., Shackleton, M.B., Taylor, S.J., Xu, X. (2004), Forecasting Currency Volatility: A Comparison of Implied Volatilities and AR(FI)MA Models. *Journal of Banking and Finance* 28, 10, p. 2541-2563.
- [129] Poon S-H. and Granger C.W.J. (2003), Forecasting Volatility in Financial Markets: a Review. *Journal of Economic Literature*, 41, 478-539.
- [130] Poon S-H. and Granger C.W.J. (2005), Practical Issues in Forecasting Volatility. *Financial Analysts Journal*, 61, 45-56.
- [131] Poterba, J., Summers, L. (1988), Mean Reversion in Stock Prices: Evidence and Implications. *Journal of Financial Economics* 22, 27-59.

- [132] Protter, P. (2004), *Stochastic Integration and Differential Equations*. Springer-Verlag, New York.
- [133] Rose C. and Smith, M. (2002), *Mathematical Statistics with Mathematica*. Springer-Verlag, New York.
- [134] Santos, A.P., Nogales, F.J. and Ruiz, E. (2012), Comparing Univariate and Multivariate Models to Forecast Portfolio Value-at-Risk. Working Paper, Universidad Carlos III de Madrid.
- [135] Schwert, G.W. (1989a), Stock Volatility and Crash of '87". *Review of Financial Studies*, 3, 77-102.
- [136] Schwert, G.W. (1989b), Does Stock Market Volatility Change Over Time? *Journal of Finance* 44, 1115-1153.
- [137] Shephard, N. (2007), Variation, Jumps and High Frequency Data in Financial Econometrics. *Advances in Economics and Econometrics*, 328-372.
- [138] Smith, J., Wallis, K.F. (2009), A Simple Explanation of the Forecast Combination Puzzle. *Oxford Bulletin of Economics and Statistics*, 74, 331-355.
- [139] Spanos, A. (1994), On Modeling Heteroskedasticity: The Student's  $t$  and Elliptical Linear Regression Models, *Econometric Theory*, 10, 286-315
- [140] Stock, J.H., Watson, M.W. (1999), Forecasting Inflation. *Journal of Monetary Economics* 44, 293-335.
- [141] Stock, J.H., Watson, M.W. (2001), Forecasting Output and Inflation: The Role of Asset Prices. NBER Working Paper No. 8180
- [142] Stock, J.H., Watson, M.W. (2004), Combination Forecasts of Output Growth in a Seven-Country Data Set. *Journal of Forecasting* 23, 405-430.

- [143] Taylor, S.J. (1986), *Modelling Financial Time Series*. New York, Wiley.
- [144] Taylor, S.J. (1987), Forecasting the Volatility of Currency Exchange Rates. *International Journal of Forecasting*, 3, 159-170.
- [145] Taylor, S.J., Xu, G. (1997), The Incremental Volatility Information in One Million Foreign Exchange Quotations. *Journal of Empirical Finance* 4:4, pp. 317–40.
- [146] Timmermann, A. (2006), *Forecast Combinations*. Handbook of Economic Forecasting. Amsterdam: North Holland Press.
- [147] Williamson J. (1972), Measuring Mutual Fund Performance. *Financial Analysts Journal*. November/December.
- [148] West, K.D., Edison, H.J. and Cho, D (1993), A Utility-Based Comparison of Some Models of Exchange Rate Volatility. *Journal of International Economics*, 35, 23–45.
- [149] Zaffaroni, P. (2007), Contemporaneous Aggregation of GARCH Processes. *Journal of Time Series Analysis* 28 (4), 521–544.
- [150] Zhang, L. (2010), Estimating Covariation: Epps effect, Microstructure Noise, *Journal of Econometrics*, 160, 33-47.
- [151] Zumbach, G. O. (2007), *The Riskmetrics 2006 Methodology*, Working Paper.

## Appendix 3.1 Index Explanation

### Nasdaq-100 Index

The index began in 1985 and consists of 100 of the largest US and international non-financial companies listed on the NASDAQ stock exchange. The components are weighted based on their market capitalization with certain rules capping the influence of the largest components. The index is rebalanced in December every a year.

### Russell 2000 Index

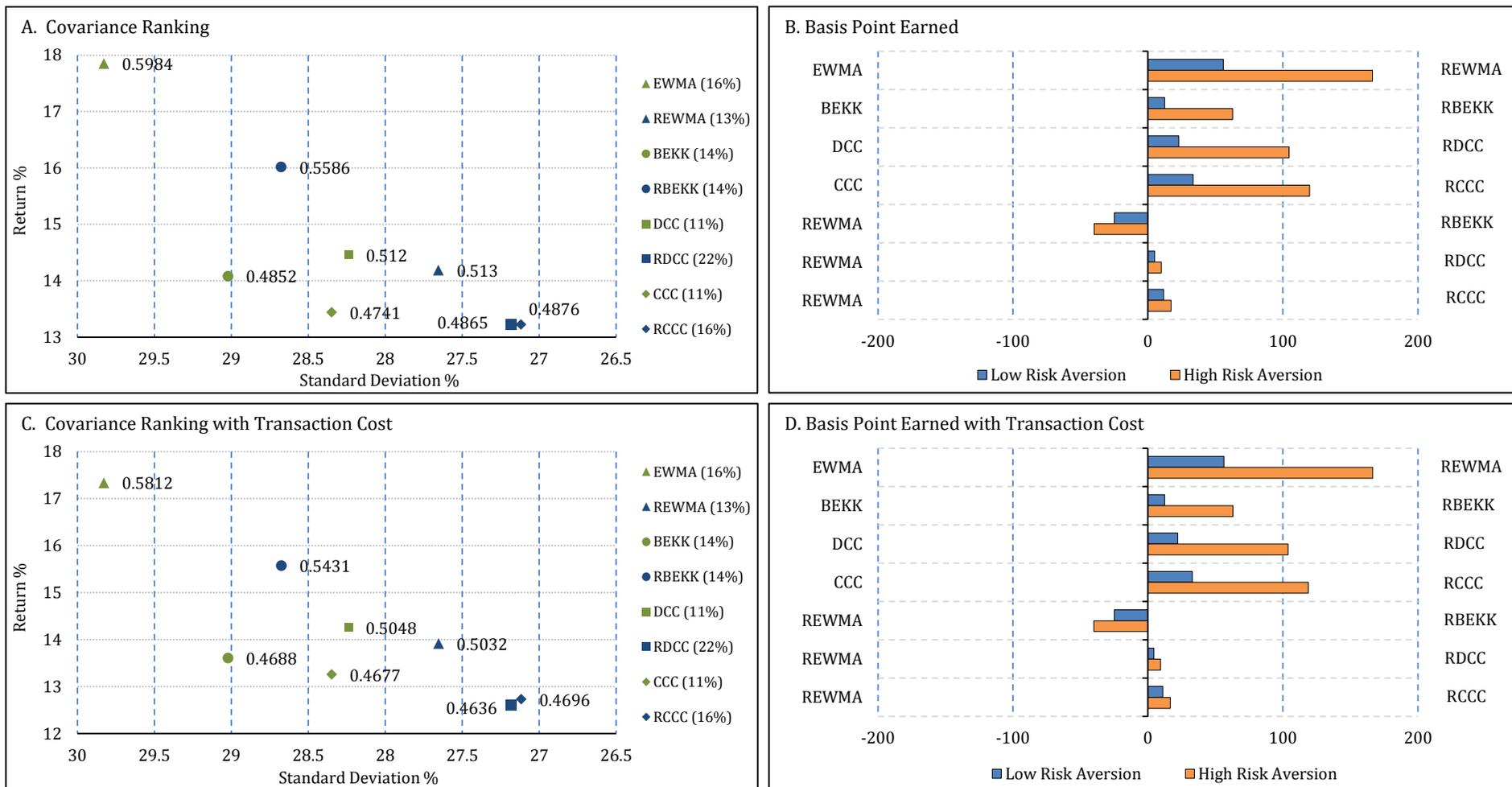
The index is commonly regarded as the benchmark for "small-cap" mutual funds. It tracks the performance of the small-cap segment of the U.S. equity universe. The index includes about 2000 of the smallest securities based on a combination of market cap and current index membership. The index is rebalanced annually to ensure larger stocks do not distort the representation of the true small-cap opportunity set.

### Reuters-CRB Index

The Reuters-CRB Index (CCI) is a commodity price index established by Commodity Research Bureau in 1957. It currently consists of 19 commodities as quoted on the NYMEX, CBOT, LME, CME and COMEX exchanges. The commodities are classified into 4 groups, each with different weightings. These groups are petroleum based products, liquid assets, highly liquid assets and diverse commodities.

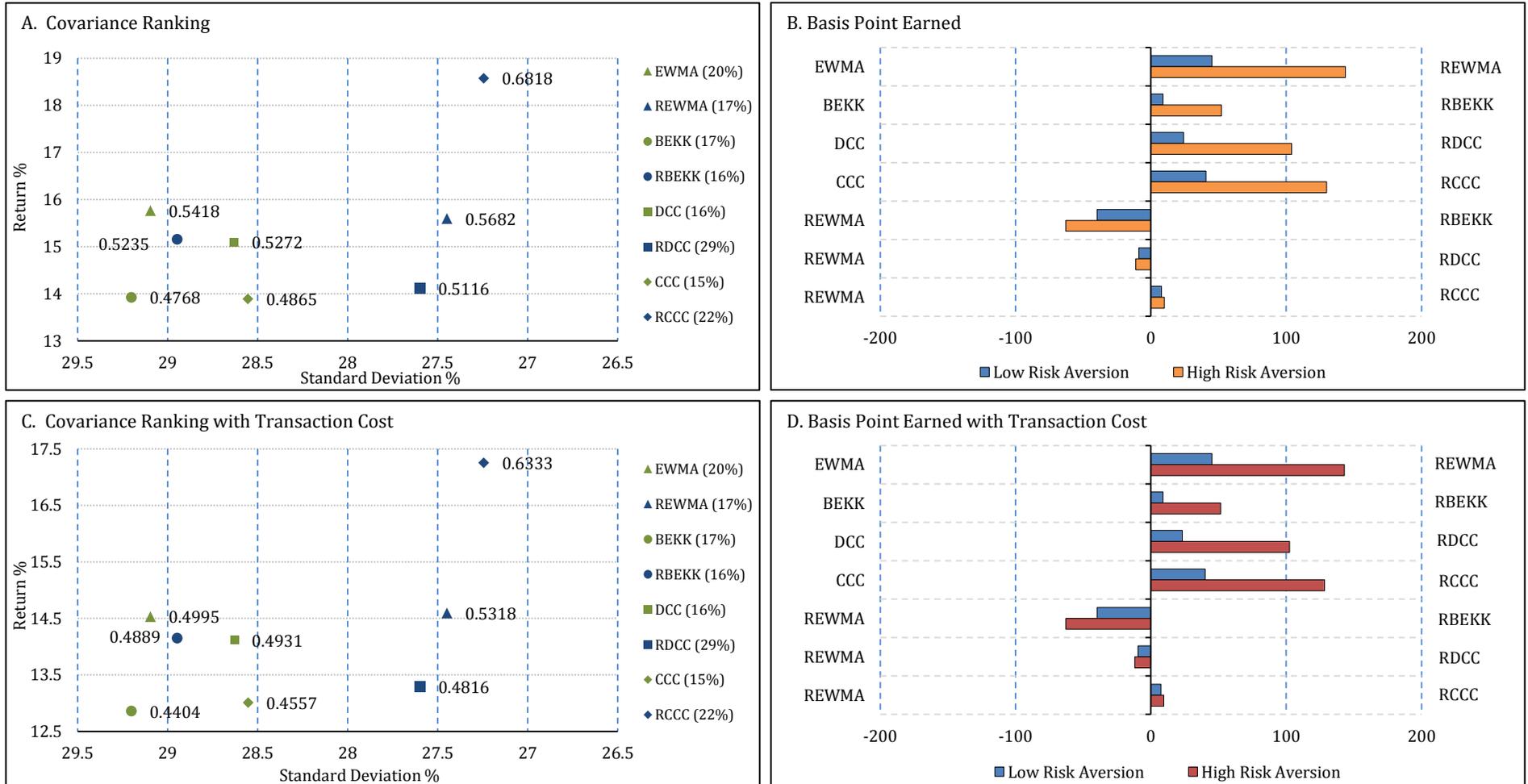
# Appendix 3.2 Additional Covariance Ranking Figures

FIGURE A.3.2.1  
Covariance Ranking • 12% Target Return, Return-Driven Rebalancing with a 20% Threshold



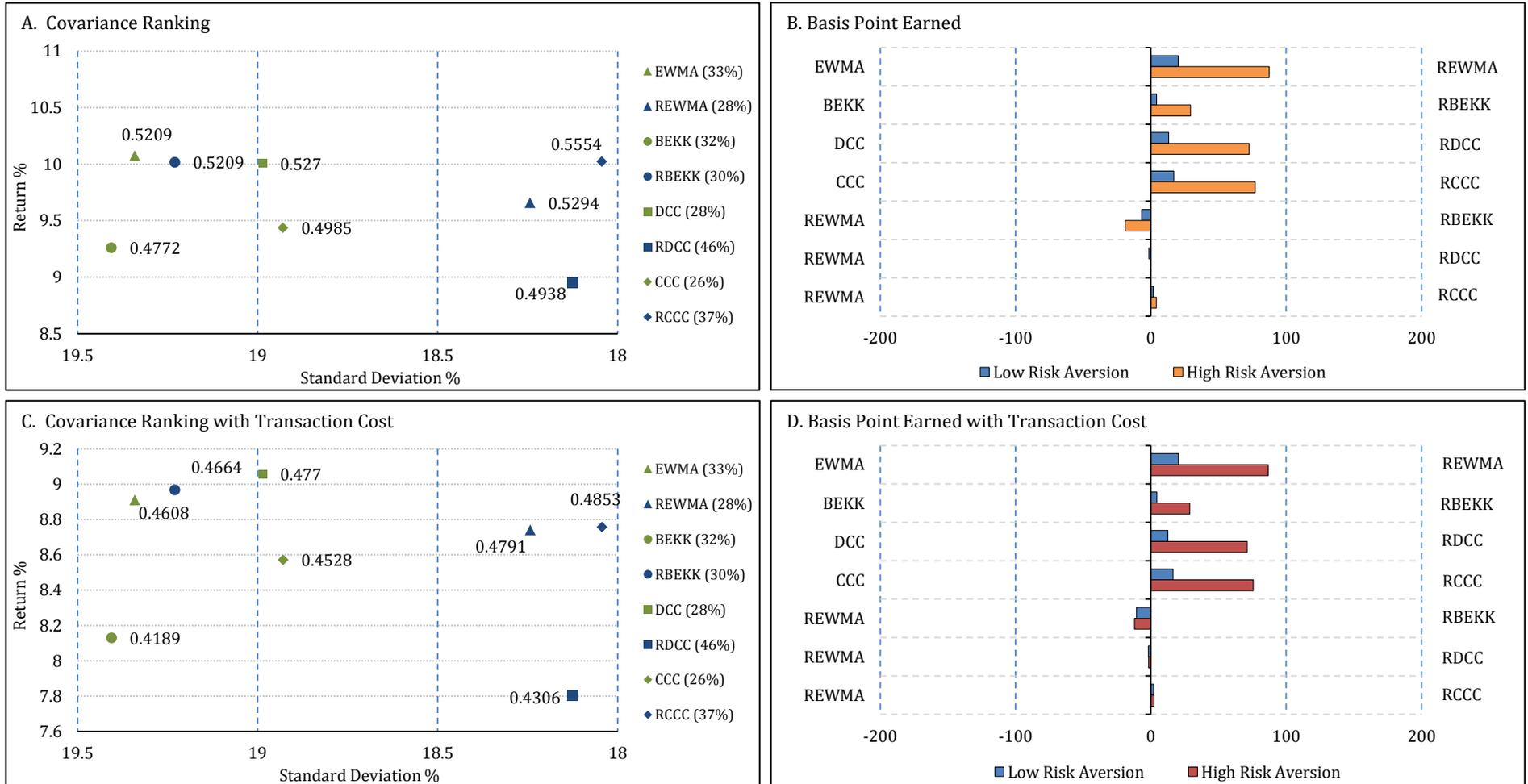
Notes: We rebalance the portfolio on day  $t$  if the overall percentage change in returns exceeds 20% on day  $t - 1$ . See note to Figure 3.3.

FIGURE A.3.2.2  
Covariance Ranking • 12% Target Return, Volume-Driven Rebalancing with a 50<sup>th</sup> Percentile



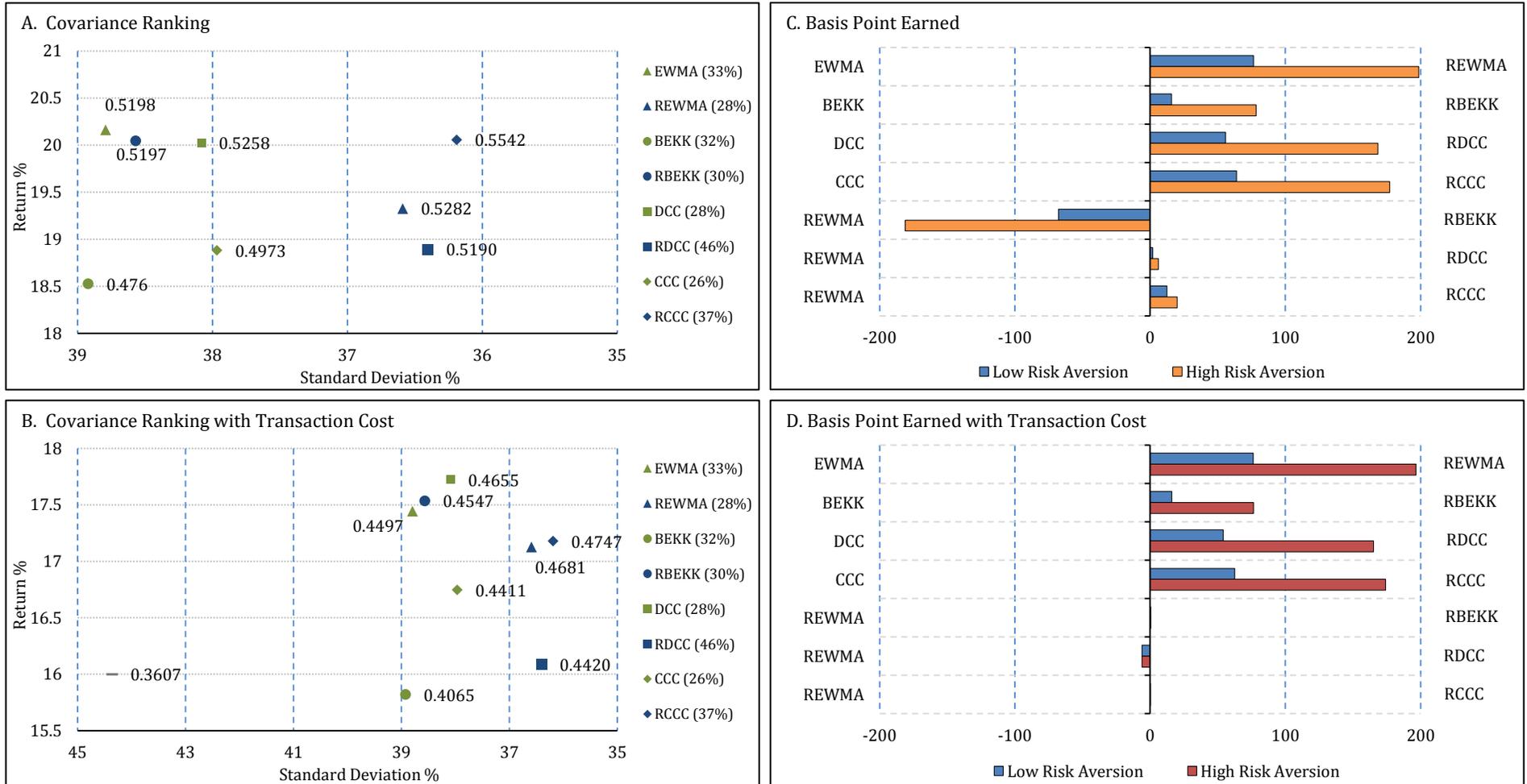
Notes: We rebalance on the day  $t$  if transaction volume on day  $t - 1$  exceeds the average of the in-sample period. S&P500 volume is used as a volume proxy since it is more representative of the market than the three indices we have. See note to Figure 3.3.

FIGURE A.3.2.3  
Covariance Ranking · 8% Target Return, Daily Rebalancing



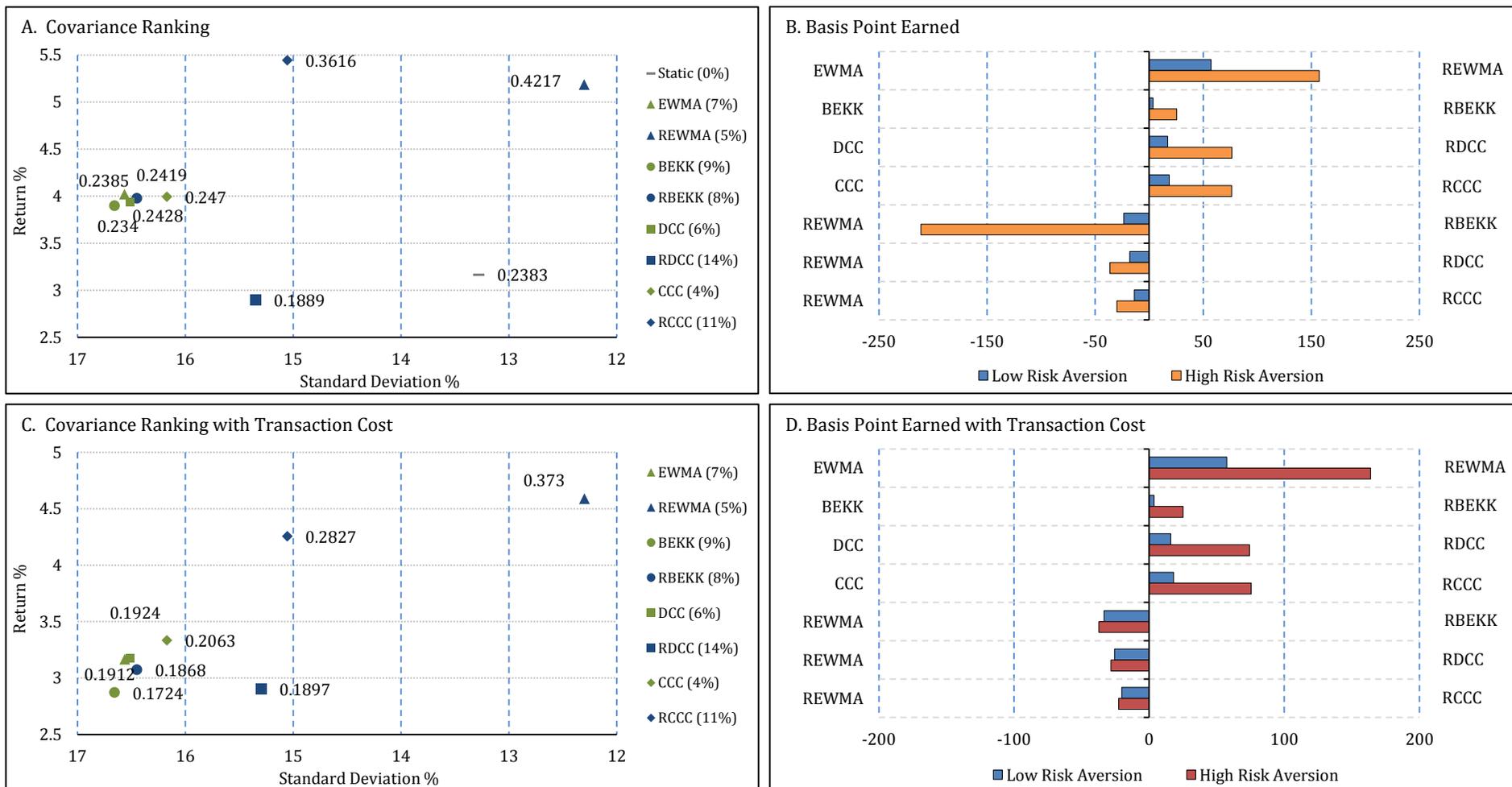
Notes: Each day the optimal portfolio weights are allocated to minimize conditional volatility subject to a target return of 8%. See note to Figure 3.3.

FIGURE A.3.2.4  
Covariance Ranking · 16% Target Return, Daily Rebalancing



Notes: Each day the optimal portfolio weights are allocated to minimize conditional volatility subject to a target return of 16%. See note to Figure 3.3.

FIGURE A.3.2.5  
Covariance Ranking · 12% Target Return, No Short-Selling, Daily Rebalancing



Notes: The figure summarizes the model performances under a target return strategy with a no short-selling constraint. See note to Figure 3.3.