University of Kent

Kent Academic Repository

Bersimi, Eirini (2023) *Volatility Forecasting and Asset Allocation In Portfolio Management*. Doctor of Philosophy (PhD) thesis, University of Kent,.

Downloaded from
https://kar.kent.ac.uk/101079/ The University of Kent's Academic Repository KAR

The version of record is available from
https://doi.org/10.22024/UniKent/01.02.101079

This document version
UNSPECIFIED

DOI for this version

Licence for this version
CC BY (Attribution)

Additional information

Versions of research works

Versions of Record
If this version is the version of record, it is the same as the published version available on the publisher's web site. Cite as the published version.

Author Accepted Manuscripts
If this document is identified as the Author Accepted Manuscript it is the version after peer review but before type setting, copy editing or publisher branding. Cite as Surname, Initial. (Year) 'Title of article'. To be published in ***Title of Journal*** , Volume and issue numbers [peer-reviewed accepted version]. Available at: DOI or URL (Accessed: date).

Enquiries
If you have questions about this document contact ResearchSupport@kent.ac.uk. Please include the URL of the record in KAR. If you believe that your, or a third party's rights have been compromised through this document please see our Take Down policy (available from https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies).

Volatility Forecasting and Asset Allocation In Portfolio Management

By

Eirini Bersimi

Submitted to Kent Business School, University of Kent

in fulfillment of the requirements for the degree of

Doctor of Philosophy in Finance

March 2023

Word count: approx. 76807

Declaration

I confirm the work submitted is entirely my own and have fully referenced my sources as appropriate. Moreover, I confirm that Chapter 4 from this thesis was used to draft one paper:

A draft paper of the 4th chapter is submitted to Journal of Operational Research:
Bersimi, E., Costantini, M. and Panopoulou, E. (2021). Directional Volatility Forecasting in the US Market.

The copyright of this thesis rests with the author. Quotations from it are permitted, provided that fully acknowledgment is made.

This thesis was supported by the Economic and Social Research Council [grant number 1938235].

Eirini Bersimi

Abstract

Volatility is a fundamental concept in finance. Referring to the conditional standard deviation of asset returns, volatility represents an important measure of risk and, as such, it has applications in almost all areas of finance. Therefore, it is difficult to overstate the importance of accurate measurement and forecasting of volatility for academics, practitioners and policy makers. The main aim of this thesis is to compare alternative ways of forecasting volatility and to evaluate their performance. This thesis is tackling this research topic in 4 main research questions.

First, Chapter 3 examines how different models can produce efficient forecasts of future volatility. Various approaches are used from simple non-parametric to parametric GARCH models and Realized Volatility (RV) models. Then, the produced forecasts are combined and the extent to which combinations can improve the forecasting accuracy is evaluated. To explore the robustness of the results, single, pairwise and multiple comparisons are considered. The purpose of this study is not just to identify the optimal specification but most importantly it tries to provide an insight interpretation of the models' performance. We use daily data for stock indices returns and 5-min realized variance for the period 2000-2017. Our main finding are that TGARCH and log HAR specifications and the simple mean combination outperform the rest. Our results are robust to various loss functions, stock index, and sample size.

Then, we continue working on univariate volatility forecasting but we move one step further. Chapter 4 aims to evaluate the performance of several GARCH models and forecast combinations for the direction of future volatility. Several alternative forecast combinations are considered based on i) a hierarchical encompassing-MSE procedure; ii) a weighted average scheme based on directional change and iii) weighted schemes based on directional accuracy, along with measures of directional accuracy (DA) and direction forecast value (DV). The empirical analysis is conducted using daily data for the S&P 500 index over the period 2000-2017 and the realized variance is used as a proxy for the true unobservable volatility. Our findings indicate that forecast combinations, in particular those based on directional accuracy and the hierarchical MSE procedure, perform as well as the best single model, thus alleviating uncertainty over the choice of a single predictive model.

Further, a multivariate approach is considered as assets' volatilities have been found to move together. Multivariate volatility models can arise as a direct generalization of univariate GARCH-type models. Alternative approaches include linear and nonlinear combinations of univariate GARCH models, such as generalized orthogonal models, latent factor models, constant and dynamic conditional correlation models (CCC, DCC), the general dynamic covariance model and copula-GARCH models. This study is not another "horse-race" but it is focused on how multivariate models can be used in a portfolio framework, using alternative assets such as Hedge Funds (HF). Since classical portfolio allocation has received a lot of criticism, this study is focusing on the risk-based approach, called Risk Parity (RP). This study aims at constructing a RP portfolio consisting of hedge funds which is competing against the Hedge Funds Research (HFR) RP Indices in an attempt to outperform them. Our empirical results suggest that for both volatility targets (10% and 12%) at least one of the our constructed portfolios (alpha and Sharpe-Ratio based) outperform the HFR Indices in the out-of-sample period.

Finally, given the lack of agreement in the literature as to which approach is superior; volatility forecasts from historical univariate models or forward looking volatility implied from option prices, a comparison of those two is performed. Chapter 6 examines under what circumstances option-Implied Volatility models provide more accurate density forecasts compared to historical volatility. What is more, since densities implied by option prices produce unrealistic risk-neutral densities (RND) calibrated real-world densities (RWD) are also calculated. Future densities are evaluated using Goodness-of-Fit tests. We use options written on European indices for the years 2006-2016. We found that historical densities outperform RND and RWD forecasts.

Overall, the intention is to significantly contribute to the literature by providing methodological solutions and new empirical evidence to important open questions on volatility forecasting. Successfully tackling these issues will result in producing more effective ways of measuring and, importantly, anticipating risk.

Contents

1 Introduction **16**

2 Literature Review and Empirical Framework **19**

2.1 Univariate Volatility models 19

2.1.1 Non-parametric volatility models 20

2.1.2 Parametric volatility models - the GARCH family 22

2.1.3 Implied Volatility (IV) 33

2.1.4 Stochastic Volatility (SV) 35

2.1.5 Realized Volatility (RV) 37

2.2 Multivariate volatility models 43

2.2.1 Multivariate Factor models 52

2.2.2 Realized Covariance 56

2.2.3 Multivariate Stochastic Volatility 60

2.2.4 Correlation Multivariate Stochastic Models 62

2.2.5 Special Methods 62

3 Evaluating univariate volatility forecasts **64**

3.1 Introduction 64

3.2 Literature Review 65

3.2.1 Single model evaluation 66

3.2.2 Pairwise comparisons 69

3.2.3 Comparisons of multiple models 72

3.3 Data and Methodology 73

3.3.1 Data 73

3.3.2 Methodology 76

3.4 Empirical Results 79

3.4.1 Robustness - Sensitivity to the loss function 85

3.4.2 Robustness - Sensitivity to the in-sample estimation period 91

3.4.3 Robustness - Others Stock indices 101

3.5 Conclusion . . . 111
3.6 Appendix . . . 112

4 Directional Volatility Forecasting in the US Market 134
4.1 Introduction . . . 134
4.2 Data and methodology . . . 138
4.2.1 Data . . . 138
4.2.2 Methodology . . . 138
4.3 Empirical results . . . 143
4.4 Conclusion . . . 155
4.5 Appendix. . . . 157

5 Multivariate Volatility Models and Risk Parity Portfolio 158
5.1 Introduction . . . 158
5.2 Literature Review . . . 160
5.2.1 Multivariate volatility models . . . 160
5.2.2 Asset Allocation . . . 162
5.2.3 Risk - based portfolio allocation approaches . . . 165
5.2.4 Hedge Funds . . . 168
5.2.5 Portfolio Performance Evaluation Measures . . . 172
5.3 Data and Methodology . . . 176
5.4 Empirical Results . . . 184
5.5 Conclusion . . . 202

6 Density Volatility Forecasts 204
6.1 Introduction . . . 204
6.2 Literature Review . . . 206
6.2.1 Risk-Neutral & Real World Volatility Density . . . 206
6.2.2 Historical Density . . . 209
6.2.3 Other Methods . . . 210
6.2.4 Density Forecast evaluation . . . 211
6.3 Data & Methodology . . . 215
6.3.1 Data description . . . 215
6.3.2 Methodology . . . 218
6.4 Results . . . 220
6.5 Conclusion . . . 263

7 Conclusion 264

Bibliography **266**

List of Tables

2.1 Univariate volatility models 32

3.1 Sample splits 74
3.2 Descriptive statistics for index returns and realized volatility 75
3.3 Univariate and multivariate volatility models 77
3.4 Pairwise comparisons of unconditional Giacomini and White test for S&P 500 (4-year in-sample) -MSE 82
3.5 Superior Predictive Ability (SPA) test and Model Confidence Set (MCS) results for the S&P 500 (4-year in-sample) - MSE, Robust Patton (b=1), Robust Patton (b=-0.5) . 84
3.6 Pairwise comparisons of unconditional Giacomini and White test for S&P 500 (4-year in-sample) -robust Patton (b=1) 87
3.7 Pairwise comparisons of unconditional Giacomini and White test for S&P 500 (4-year in-sample) - robust Patton (b=-0.5) 89
3.8 Pairwise comparisons of unconditional Giacomini and White test for S&P 500 (1-year in-sample) - MSE 94
3.9 Pairwise comparisons of unconditional Giacomini and White test for S&P 500 (1-year in-sample) - robust Patton (b=1) 96
3.10 Pairwise comparisons of unconditional Giacomini and White test for S&P 500 (1-year in-sample) - robust Patton (b=-0.5) 98
3.11 Superior Predictive Ability (SPA) test and Model Confidence Set (MCS) results for the S&P 500 (1-year in-sample) - MSE, Robust Patton (b=1), Robust Patton (b=-0.5) . 100
3.12 Pairwise comparisons of unconditional Giacomini and White test for CAC 40 (4-year in-sample) - MSE 102
3.13 Superior Predictive Ability (SPA) test and Model Confidence Set (MCS) results for the CAC 40 (4-year in-sample) - MSE, Robust Patton (b=1), Robust Patton (b=-0.5) . . 104
3.14 Pairwise comparisons of unconditional Giacomini and White test for FTSE 100 (4-year in-sample) -MSE 105
3.15 Superior Predictive Ability (SPA) test and Model Confidence Set (MCS) results for the FTSE 100 (4-year in-sample) - MSE, Robust Patton (b=1), Robust Patton (b=-0.5) . 107

3.16 Pairwise comparisons of unconditional Giacomini and White test for DAX (4-year in-sample) - MSE . . . 108
3.17 Superior Predictive Ability (SPA) test and Model Confidence Set (MCS) results for the DAX (4-year in-sample) - MSE, Robust Patton (b=1), Robust Patton (b=-0.5) . . . 110
3.18 Pairwise comparisons of unconditional Giacomini and White test for S&P 500 (4-year in-sample) - all Loss Functions . . . 114
3.19 Pairwise comparisons of unconditional Giacomini and White test for S&P 500 (1-year in-sample) - all Loss Functions . . . 115
3.20 Superior Predictive Ability (SPA) test and Model Confidence Set (MCS) results for the S&P 500 (1-year in-sample) - MSE, Robust Patton (b=1), Robust Patton (b=-0.5) . 116
3.21 Superior Predictive Ability (SPA) test and Model Confidence Set (MCS) results for the S&P 500 (4-year in-sample) - MSE, Robust Patton (b=1), Robust Patton (b=-0.5) . 117
3.22 Pairwise comparisons of unconditional Giacomini and White test for CAC 40 (4-year in-sample) - robust Patton (b=1) . . . 118
3.23 Pairwise comparisons of unconditional Giacomini and White test for CAC 40 (4-year in-sample) - robust Patton (b=-0.5) . . . 120
3.24 Pairwise comparisons of unconditional Giacomini and White test for CAC 40 (4-year in-sample) - all Loss Functions . . . 122
3.25 Pairwise comparisons of unconditional Giacomini and White test for FTSE 100 (4-year in-sample) - robust Patton (b=1) . . . 123
3.26 Pairwise comparisons of unconditional Giacomini and White test for FTSE 100 (4-year in-sample) - robust Patton (b=-0.5) . . . 125
3.27 Pairwise comparisons of unconditional Giacomini and White test for FTSE 100 (4-year in-sample) - all Loss Functions . . . 127
3.28 Pairwise comparisons of unconditional Giacomini and White test for DAX - (4-year in-sample) robust Patton (b=1) . . . 128
3.29 Pairwise comparisons of unconditional Giacomini and White test for DAX (4-year in-sample) - robust Patton (b=-0.5) . . . 130
3.30 Pairwise comparisons of unconditional Giacomini and White test for DAX (4-year in-sample) - all Loss Functions . . . 132
3.31 Pairwise comparisons of unconditional Giacomini and White test for S&P 500 (4-year and 1-year in-sample), CAC 40 (4-year in-sample), FTSE 100 (4-year in-sample) and DAX (4-year in-sample) - MSE Loss Function . . . 133

4.1 Univariate and multivariate volatility models . . . 139
4.2 Directional accuracy (DA) and directional forecast value (DV) of volatility. The holdout period is 500 observations. . . . 145

4.3 Directional accuracy (DA) and directional forecast value (DV) of volatility in bear and bull markets. The holdout period is 500 observations. 147
4.4 Recession and expansion periods directional accuracy (DA) and direction forecast value (DV) of volatility. The holdout period is 500 observations. 148
4.5 Evaluation of strategies. The holdout period is 500 observations. 150
4.6 Directional accuracy (DA) and directional forecast value (DV) of volatility. The holdout period is 250 observations 151
4.7 Bear and bull markets directional accuracy (DA) and directional forecast value (DV) of volatility. The holdout period is 250 observations. 152
4.8 Recession and Expansion periods directional accuracy (DA) and direction forecast value (DV) of volatility. The holdout period is 250 observations. 153
4.9 Evaluation of strategies. The holdout period is 250 observations. 154
4.10 MSE-based model ranking. The holdout period is 250 observations. 157
4.11 MSE-based model ranking. The holdout period is 500 observations. 157

5.1 Mapped Strategies used in the Risk Parity Portfolio 177
5.2 Fund Characteristics 178
5.3 Number of overlapping funds for Alpha-sorted portfolio at each period 192
5.4 Number of overlapping funds for SR-sorted portfolio at each period 193
5.5 Number of overlapping funds for Alpha-sorted portfolio at each period 194
5.6 Number of overlapping funds for SR-sorted portfolio at each period 195
5.7 Evaluation measures for Alpha-sorted, SR-sorted portfolios and HFRI10 and HFRI10I Indices 198
5.8 Evaluation measures for Alpha-sorted, SR-sorted portfolios and HFRI12 and HFRI12I Indices 200

6.1 Summary Statistics for European Options 216
6.2 Summary Statistics for European Options - Cont'd 217
6.3 Summary Statistics for European Options - Cont'd (2) 218
6.4 KS1 test results 222
6.5 LR3 test results 225
6.6 Autocorrelation test results-RND 228
6.7 Autocorrelation test results - RWD 231
6.8 Autocorrelation test results-Historical Simulation GARCH-N 234
6.9 Autocorrelation test results-Historical Simulation GARCH-T 237
6.10 Autocorrelation test results-Historical Simulation GJR-N 240
6.11 Autocorrelation test results-Historical Simulation GJR-T 243

6.12 Hit tests-RND 247
6.13 Hit tests-RWD 249
6.14 Hit tests-Historical Simulation-GARCH-N 251
6.15 Hit tests-Historical Simulation-GARCH-T 253
6.16 Hit tests-Historical Simulation-GJR-N 255
6.17 Hit tests-Historical Simulation-GJR-T 257
6.18 Statistics 260

List of Figures

3.1 Daily returns for S&P 500 returns (2000-2017) . . . 75
3.2 Daily returns for FTSE 100 returns (2000-2017) . . . 75
3.3 Daily returns for DAX returns (2000-2017) . . . 75
3.4 Daily returns for CAC 40 returns (2000-2017) . . . 75
3.5 Autocorrelation function for S&P 500 daily returns (2000-2017) . . . 113
3.6 Autocorrelation function for FTSE 100 daily returns (2000-2017) . . . 113
3.7 Autocorrelation function for DAX daily returns (2000-2017) . . . 113
3.8 Autocorrelation function for CAC 40 daily returns (2000-2017) . . . 113
3.9 Quantile-quantile plot for S&P 500 daily returns (2000-2017) . . . 113
3.10 Quantile-quantile plot for FTSE 100 daily returns (2000-2017) . . . 113
3.11 Quantile-quantile plot for DAX daily returns (2000-2017) . . . 113
3.12 Quantile-quantile plot for CAC 40 daily returns (2000-2017) . . . 113

5.1 HFR RP Indices (closing) monthly prices graph . . . 181
5.2 Relative weights for the Alpha-sorted portfolio. Yearly rebalancing, 10% volatility targeting. . . . 185
5.4 Relative weights for the SR-sorted portfolio. Yearly rebalancing, 10% volatility targeting. 187
5.6 Relative weights for the Alpha-sorted portfolio. Yearly rebalancing, 12% volatility targeting. . . . 189
5.8 Relative weights for the SR-sorted portfolio. Yearly rebalancing, 12% volatility targeting. 191
5.10 Histogram of all funds used in the Alpha-sorted portfolio. Yearly rebalance, 10% volatility targeting . . . 196
5.11 Histogram of all funds used in the SR-sorted portfolio. Yearly rebalance, 10% volatility targeting . . . 196
5.12 Histogram of all funds used in the Alpha-sorted portfolio. Yearly rebalance, 12% volatility targeting . . . 196
5.13 Histogram of all funds used in the SR-sorted portfolio. Yearly rebalance, 12% volatility targeting . . . 197

5.14 Cumulative returns for the two constructed portfolios, Alpha and Sharpe Ratio (SR) as well as the two benchmarks, the HFR Risk Parity Vol 10 Index and the HFR Risk Parity Vol 12 Institutional Index. Initial investment 1000$. 199
5.15 Cumulative returns for the two constructed portfolios, Alpha and Sharpe Ratio (SR) as well as the two benchmarks, the HFR Risk Parity Vol 12 Index and the HFR Risk Parity Vol 12 Institutional Index. Initial investment 1000$. 201

Acknowledgments

I would like to start by thanking my two wonderful supervisors, Prof. Ekaterini Panopoulou and Dr. Nikolaos Voukelatos, without them this thesis would not have been possible. They supported me at each and every step from my master studies to PhD application and submission. They were always available to provide me with invaluable guidance and support for my thesis but also for my professional and personal development. I am deeply grateful for all their encouragement, trust, belief in my abilities, patience and opportunities they offered me. I could not have imagined having two better supervisors for my PhD. I would also like to thank my third supervisor Dr. Vasileios Pappas who has always been helpful and supportive.

I am also grateful to South East Network for Social Sciences (SeNSS) and Kent Business School (KBS) for granting me funds that allowed me to complete my studies and develop my skills. I owe gratitude to the KBS research team for their help and support all those years and for creating a friendly and secure environment that allowed me to focus on my studies. I am also really grateful for all the wonderful people I have met at KBS. Special thanks go to my dear friend Sheri for making my days at the office special and her continuing support and help both practical and emotional. I am also really grateful for my friendship with Mattia and all the other PhD students, namely Enoch, Davide, Rasmi, Omar, Rukiye and Shashitha. I am very happy that I have met them and for all the good times we have shared. I would also like to thank Prof. Mauro Costantini for his warm welcome during my visit at the University of L'Aquila and for his contribution to this thesis. Finally, I would also like to thank Dr. Leonidas Tsiaras for his input to this thesis.

Then I would like to thank my caring partner Ben who provides me with unconditional love and support. He knows all my struggles and he has always been by my side to lift me up and encourage me but also to celebrate good news and happy moments. I truly appreciate his patience the last five years. I am looking forward to many more endeavors together.

Finally, I would like to say special thanks to my parents, Pantelis and Pauline, and my brother, Mark who have always believed in me. I truly appreciate their understanding and support to follow my dreams. I could not forget to thank all of my friends back home and my Glaswegian pals for their

support, hospitality and friendship.

Chapter 1

Introduction

Volatility is defined as the conditional standard deviation of stock returns. Variance (square of standard deviation) is the deviation around the mean of a random variable and therefore it signifies the uncertainty in the financial markets and it can be viewed as a measure to quantify risk. Under this framework, volatility has many applications in finance. It is used as input when pricing derivatives after the seminal model of Black and Scholes (1973) and Merton (1973) [1], and when calculating Value-at-risk (VaR) and Expected Shortfall (ES) which is used to measure market risk for regulatory capital requirements for banks [2]. Finally, many investing and hedging decisions and strategies are designed relying on future volatility levels to evaluate the investment risk.

Accurate volatility forecasts lead to more effective ways of measuring and predicting risk. Thus, accurate measurement and forecast of future variance is crucial for academics, practitioners and policy makers. Therefore, it is not a surprise that the area of forecasting future volatility of asset returns has attracted a lot of attention in the literature over the previous years starting from the seminal papers of Engle (1982) and Bollerslev (1986); (see Poon and Granger (2003) and the references within; Andersen et al. (2005a); Brooks and Persand (2003); Christoffersen and Diebold (2000); Satchell and Knight (2007); Andersen et al. (2011); Poon (2005); Corsi et al. (2012); Taylor (2004); Symitsi et al. (2018); Caporin and McAleer (2012a); Hansen and Lunde (2005); de Almeida et al. (2018); Boudt et al. (2019)).

Despite the fact that various models and methods have been suggested in the literature, there is still not clear consensus on what works best. This thesis aims to contribute to the literature by reviewing and comparing prominent volatility models, methods and categories of volatility models such as univariate, multivariate models, point and density forecasts. Then all these forecasts are evaluated by

[1] The breakthrough came form the fact that volatility is not directly observed in the market as the rest of the parameters.

[2] Since 1996 Basel Capital Accord mandates all commercial banks to set aside regulatory capital to cover exposure to the market risk.

various different criteria such as loss functions, tests of equal predictive accuracy, directional accuracy and other diagnostics. In that way this thesis can provide new insights in the area of volatility and propose under what circumstances forecasting ability can be improved.

Chapter 2 provides an extensive literature review on univariate and multivariate models which aims at establishing the methodological framework for this thesis. For univariate models we discuss various non-parametric, parametric GARCH models, implied volatility, stochastic volatility and realized volatility models. For multivariate models we review multivariate factor models, realized covariance, multivariate stochastic volatility models, correlations multivariate stochastic models and finally some spacial methods that address the issue of dimensionality in large covariance matrices. In each category the existing models and approaches are organized into relevant sub-groups.

Chapter 3 compares the forecasting performance of various univariate volatility models. We consider several specifications from the GARCH family ranging from non-parametric to parametric different distributional assumptions and RV models as well as forecast combinations. Then the produced forecasts are combined and the extent to which combinations can improve the forecasting accuracy and the directional forecasting accuracy is evaluated. This study aims to compare and evaluates alternative volatility forecasting specifications in terms of their out-of-sample performance. To evaluate the forecasts, pairwise comparisons are employed using the Giacomini and White (2006) test of equal predictive ability under alternative loss functions. Additionally, the Model Confidence Set (MCS) of Hansen et al. (2011) and the Superior Predictive Ability (SPA) of Hansen (2005) tests are used with three robust loss functions. An empirical application on the S&P 500 index shows that there is no single specification that outperforms the rest and that simple specifications such as the threshold ARCH (TARCH), the logarithmic heterogeneous autoregressive (log-HAR) model and the mean forecast combination appear superior to the other models considered. Also our results, are robust to the loss function used, the in sample size and the stock index.

Chapter 4 aims to examine the directional forecast accuracy (DA) of GARCH models and forecast combination schemes. Simple forecast combinations are considered but also alternative ones based on a hierarchical-MSE procedure, and past directional accuracy performance. The models are evaluated in terms of DA and directional forecast value (DV) for one and five-steps ahead. This chapter aims to ascertain whether different GARCH models are equally good at predicting the direction using a test for the equivalence of the proportion of successful direction predictions. Further, a sensitivity analysis is performed to explore the robustness of our results and further understand the dynamics of the models and their combinations. The main results show that forecast combinations considered are highly competitive against the best performer GARCH model in terms of both DA and DV. These

results seem to suggest the when there is uncertainty about what model to chose, the methods of combination may be preferable considering directional forecast equivalence among models and forecast combinations.

Chapter 5 is based on the review of the wide set of multivariate volatility models motivated by Chapter 1. In this chapter we construct our own portfolio of hedge funds. Portfolio construction is one of the main applications for multivariate volatility models. As Timmermann (2006) highlighted, forecasts do not have practical value to market participants. They only became useful when interpreting the model performance in terms of gains or loses. Therefore an empirical application in portfolio management is valuable. This chapter briefly explores the classical asset allocation before focusing on the most recent risk-based approaches. Further, it presents the main characteristics of hedge funds and their performance evaluation measures as alternative investments. Classical portfolio allocation has received a lot of criticism (Maillard et al. (2010), Asness et al. (2012a)). Thus, this study focuses on the risk-based approach, called risk parity, where each asset contributes equally to the total portfolio risk resulting to a truly balanced portfolio. Given the lack of an appropriate benchmark for risk parity hedge funds before the recent release of the Risk Parity Indices by Hedge Fund Research, this study aims at constructing a risk parity portfolio consisting of hedge funds which competes against the HFR Risk Parity Indices in terms of portfolio performance evaluation measures. Two portfolios of hedge funds are created, based on the past performance measured by the alpha and Sharpe ratio criteria. Our empirical results suggest that for the out-of-sample period, at least for one of the our constructed portfolios outperform the HFR Indices as confirmed for both volatility targets (10% and 12%) under examination.

At Chapter 6, the density forecasts from historical univariate volatility models are compared against risk transformations of risk-neutral densities (RNDs). Given the lack of agreement in the literature as to which of these two is generally superior to the other, this study examines under what circumstances, risk transformations of risk-neutral densities can outperform historical densities. Further, a non-parametric calibration transforms the risk-neutral implied densities to real-world densities (RWDs). An empirical application on options written on European stock indices for years 2006-2016 is performed. All methods are evaluated using Goodness-of-Fit tests. We find that historical densities outperform RNDs and RWDs.

At the final chapter 7, a conclusion of this thesis is presented, summarizing the main findings of each empirical application and highlighting the limitations and the steps for future work.

Chapter 2

Literature Review and Empirical Framework [1]

2.1 Univariate Volatility models

Volatility of asset returns can be measured in many different ways. There are five main classes of univariate volatility models; non-parametric, GARCH models, Implied Volatility (IV), Stochastic Volatility (SV) and Realized Volatility (RV) models. Traditionally, in finance, volatility is often associated with the standard deviation (or the variance) of asset returns. Non-parametric methods make use of the sample standard deviation and they are also known, as distribution-free methods as there is no assumption on the distribution.

Returns of financial data are heteroskedastic and ARCH effects are present (Engle (1982), French et al. (1987)). So the conditional variance of stock returns needs to be estimated while making an assumption of the error's distribution. This class of models is better known as parametric. These characteristics motivated the classic Autoregressive Conditional Heteroskedastic (ARCH) model and its generalization, the GARCH model. Models of the extended GARCH-family compose the biggest body of parametric models. Extensions of the parsimonious GARCH model are motivated by certain behaviors observed at the stock returns' time series, known as stylized effects. This covers linear, non linear, asymmetric and regime switching.

Alternatively, volatility forecasts can be directly derived as market's expectations. Implied Volatility (IV) is defined as the value that equates the theoretical price derived by an option-pricing model (such as the Black and Scholes (B-S) Model) to the price observed in the market, assuming that the option's prices reflect all the available information and that the option-pricing model is correctly

[1] Please note that not all models and methods presented in this chapter are used in the following empirical chapters. However, we decided to include them as part of the main thesis for completeness.

specified. Since this approach is based on direct option prices it is both market based and forward looking. In that way, IV can link ex-post to ex-ante volatility (Andersen et al. (2003)).

Another popular class of models for the conditional variance are those that treat volatility of stock returns as a stochastic process, known as Stochastic Volatility (SV) models. Strictly speaking, parsimonious GARCH models are not Stochastic models. What is more, SV models attempt to fix some of the problems associated with the Black–Scholes model. Therefore, in SV models the assumption that the variance rate of the return is constant is relaxed. In that way, SV can explain the "volatility smile" and take it into consideration when pricing options resulting in more accurate estimates.

Last but not least, Realized Volatility (RV) is a fairly new approach of modeling the conditional variance of asset returns. RV is an ex-post measure of volatility and it contains almost all the relevant information to predict its future values. Daily squared returns were used prevailing at the past but this measure has received much criticism for being a noisy estimator (Blair et al. (2001a)). Additionally, it has caused a lot of concern among the researchers because they attributed the bad performance of the models to the models themselves and not on the use of the inappropriate estimator on the evaluation stage. Therefore, daily squared returns are not considered trustworthy and later the literature proposed to use intra-daily observations instead, with the 5 minute RV being the dominant estimator. Given that the additive property is valid, the daily RV is calculated as the sum of all the intra daily values. Finally, some alternative ways to estimate and forecast asset returns' volatility can be found in the literature that in some cases can improve the forecasting power such as Bayesian or state space models (Timmer and Weigend (1997), Liu and Maheu (2009)).

2.1.1 Non-parametric volatility models

The simplest way to estimate volatility, is to assume that returns follows a Random Walk without a drift. Variance is expected to follow an unpredictable path. This estimate implies no change and that the past realized values carry no information. Therefore, the best forecast for tomorrow's volatility is today's value:

$$\widehat{\sigma}_{t+1} = \sigma_t \tag{2.1}$$

where σ is the historical standard deviation of asset returns. Another simplistic way to model volatility is to rely on all past realizations included in the sample period. This model is known as Historical Average model and it is given by:

$$\widehat{\sigma}_{t+1} = \frac{\sum_{j=1}^{t} r_{t-j+1}^2}{t} \tag{2.2}$$

where r is the expected return. Please add a sentence to clarify Alternatively, standard deviation can be approximated by a formula provided by Pearson and Tukey (1965). In that case the standard deviation is given as the interval between two quantiles of the returns' distribution divided by a number that corresponds to a known distribution. Pearson and Tukey (1965) suggested a three point approximation that is based on symmetric quantiles based on the 98%, 95% and 90% intervals:

$$\widehat{\sigma}_t = \frac{\widehat{Q}(0.99) - \widehat{Q}(0.01)}{4.65} \tag{2.3}$$

$$\widehat{\sigma}_t = \frac{\widehat{Q}(0.975) - \widehat{Q}(0.025)}{3.92} \tag{2.4}$$

$$\widehat{\sigma}_t = \frac{\widehat{Q}(0.95) - \widehat{Q}(0.05)}{3.25} \tag{2.5}$$

Reflecting on the denominators, they are extracted from 29 distributions with different shapes that were part of the experiment and they represent the central values of the 5%, the 2.5% and the 1% distance, respectively. Under the Gaussian distribution the same expressions would have been divided by 4.653 (=2 x 2.326), 3.920 (=2 x 1.960) and 3.290 (=2 x 1.645) correspondingly (Yeh et al. (2014); Taylor (2005)). The Laplace approximation takes into account the phenomenon of "heavy" tails, which is frequently observed in financial data, and therefore it suggests to divide the third equation by 3.256.

The seminal paper of Pearson and Tukey (1965) motivated a part of the literature that focuses on simple approximations of the standard deviation either by adjusting the existing formulas or by proposing new methods based on various intervals or regressions of moments of the returns' distributions. Moreover, later on, this expertise was transfered to RV estimators; for instance the RV quantile of Christensen et al. (2010). These simple approximations have been used prominently, in many different applications in various scientific fields.

One of the proposed modifications is to consider the median instead of the mean as a measure of the central value. However as Taylor (2005) stated this had unsatisfactory results for financial data as the median of daily observations is approximately close to zero. Mosteller (1946) suggested to employ quantile combinations to achieve better results. What is more, other methods include the 5-point approximation of Pearson and Tukey (1965), the Moder and Rodgers (1968) formulas, or the Keefer (1994) 3-point discrete approximation, the PERT (Program Evaluation and Review Technique) approach of Keefer and Verdini (1993) or the Keefer and Bodily (1983) or the Davidson et al. (1976) or follow Johnson (1998). Additionally, a conservative measure of standard deviation was proposed by Greco et al. (2015). This measure is defined as:

$$\widehat{\sigma}_t = max\left(\frac{\widehat{Q}(0.75) - \widehat{Q}(0.50)}{0.6745}, \frac{\widehat{Q}(0.50) - \widehat{Q}(0.25)}{0.6745}\right) \tag{2.6}$$

where 0.6745 is the 75th percentile of the Normal distribution Greco et al. (2015). Finally, Taylor (2005) proposed to divide the 90% interval with 3.29 and he proposed as an extension a quadratic regression model. However, Johnson (1998) warned about the efficiency of these estimators since they rely on very extreme quantiles. In contrast, the alternative choice of Autoregressive (AR) models can exploit all observations, even though various past observations can be excluded, if appropriate.

2.1.2 Parametric volatility models - the GARCH family

Despite the fact that non-parametric approaches are considered flexible since they are not restricted by impractical assumptions, they fail to provide validation for time-varying volatility, a phenomenon well-known as heteroskedasticity. Starting from the seminal paper of Engle (1982) that introduced the AutoRegressive Conditional Heteroscedasticity (ARCH) model and later its generalization (GARCH) proposed by Bollerslev (1986), conditional variance is a weighted average of past squared forecast errors and following a GARCH model is a weighted average of past squared forecast errors and past variances, as well. The parameters determine the sensitivity of variance to new information and its long run mean. Their representations are given in equations 2.8 and 2.9, respectively, as:

$$r_t = \sigma_t \varepsilon_t \tag{2.7}$$

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} a_i \varepsilon_{t-i}^2 \tag{2.8}$$

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} a_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 \tag{2.9}$$

where r_t is zero mean return series, ϵ is i.i.d and σ_t is the time-varying conditional standard deviation. RiskMetrics is a restricted non-stationary GARCH(1,1) where $\omega = 0$ and $\lambda = \beta = 1 - \alpha$ so $a + \beta = 1$. Under this specification σ_t^2 is assumed to follow an Exponentially Weighted Moving Average (EWMA). The EWMA enforces a smoothing parameter to assign a different weight on each squared return. The weights are given as: $\lambda^j(1 - \lambda)$. Such a model is often called an integrated model. On the same direction, Gallant (1981) proposed another approach where σ_t^2 is approximated by a function of polynomial and trigonometric terms in lagged values of σ_t.

Thresholds and Asymmetries

The basic GARCH model fails to account for a number of well-known characteristics of volatility (stylized facts) such as asymmetric and leverage effects. Black (1976) was the first who examined those effects. On the ground that gains and losses do not have the same effect on future volatility, a different practice is required, therefore the literature has focused on the so called threshold models.

Under threshold models, different states or regimes are defined where certain statistics such as the variance of a time series, is presumed to remain constant in those regimes (Franses and Van Dijk (2000); Priestley (1988)). In threshold models the shifts between regimes are modeled in terms of the time series (X_t or X_{t-1} past realization of an observed variable) which differs from regime switching (RS) models where an unobservable stochastic process (Markov chain) is employed. "This implies that one can never be certain that a particular regime has occurred at a particular point in time, but can only assign probabilities to the occurrence of the different regimes" (Franses and Van Dijk (2000), p.70). Thresholds are related to both the sign and size of shocks. Inside this category lie the open-loop threshold models and the threshold AR (TAR) along with the SETAR model where the regime depends on the time series, as well and it is given as:

$$r_t = \begin{cases} (\phi_{0,1} + \phi_{1,1} y_{t-1} + \varepsilon_t), & \text{if } y_{t-1} \leq c. \\ (\phi_{0,2} + \phi_{1,2} y_{t-1} + \varepsilon_t), & \text{if } y_{t-1} > c \end{cases} \tag{2.10}$$

where r_t is the returns series and c is the threshold (Franses and Van Dijk (2000)). Hagerud (1997) proposed a more smooth transition by its Smooth Transition Autoregressive (STAR) model, where the regimes are determined by "(..) a continuous function $G(r_{t-1}; \gamma, c)$, which changes smoothly from 0 to 1 as r_{t-1} increases" (Franses and Van Dijk (2000), p.72). A popular specification for the transition function $G(r_{t-1}; \gamma, c)$ is the logistic function:
$G(r_{t-1}; \gamma, c) = \frac{1}{1+\frac{1}{e^{\gamma(y_{t-1}-c)}}} = \alpha$ where $\alpha \to 0$ if $\frac{1}{e^{\gamma(r_{t-1}-c)}} \to \infty$ since $e^{\gamma(r_{t-1}-c)} \to \infty$. Then, the model is called Logistic STAR (LSTAR). The parameter c is the threshold between the two regimes such as : $G(y_{r-1}; \gamma, c) = 0$ and $G(y_{r-1}; \gamma, c) = 1$. Hence, the logistic function changes monotonically from 0 to 1 as r_{t-1} increases. The parameter γ determines how smooth the transition between the regimes will be. Finally, van Dijk and Franses (1999) proposed the multiple regime STAR model. Astatkie et al. (1997) proposed the nested TAR (NeTAR) for the case that the time series depend on a second regime as well. This second regime can be nested in the first one, or not (Franses and Van Dijk (2000)).

In opposition to an AR parameterization, there is the moving average (MA) which is simply a linear combination of current and past white noise disturbance terms. Wecker (1981) proposed the asymmetric MA where time series depend differently on positive and on negative innovations.

$$r_t = \varepsilon_t + a_1 \varepsilon_{t-1}^{(1)} + a_2 \varepsilon_{t-1}^{(2)} \tag{2.11}$$

where $\varepsilon_{t-1}^{(1)} = \max(\varepsilon_t, 0)$, for positive innovations and $\varepsilon_{t-1}^{(2)} = \min(\varepsilon_t, 0)$, for negative innovations.
However, the issue of time-varying volatility has not been addressed just yet. To account for the asymmetric effects on time series numerous models have been proposed. The ARCH-Leverage (ARCH-

L) model is one of them.

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} a_i \varepsilon_{t-1}^2 + \gamma \varepsilon_{t-1}^2 I_{t-1} \tag{2.12}$$

where $I_{t-1} = 1$ if $r_{t-i} < 0$ and $I_{t-1} = 0$ otherwise. The most famous model within this classification is the generalized GARCH-L, which is also called GJR by Glosten et al. (1993).

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} a_i \varepsilon_{t-i}^2 + \gamma \varepsilon_{t-1}^2 I_{t-1} + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 \tag{2.13}$$

Similarly, Guégan and Diebolt (1994) proposed the $\beta-$ARCH (Beta ARCH) model which treats asymmetrically positive and negative innovations

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} a_i \varepsilon_{t-i}^{2*\beta} + \gamma \varepsilon_{t-1}^{2*\beta} I_{t-1} \tag{2.14}$$

Caporin and McAleer (2006) identified three possible extensions of the GJR model. For instance, it could accommodate more variables that depend on different thresholds according to the size or the sign. Further, the leverage component could be asymmetric and time-varying.

While GJR can only account for asymmetric effects, the exponential GARCH (EGARCH) by Nelson (1991) is capable of accommodating both leverage effect and asymmetric behavior of the time series. A distribution is characterized as asymmetric when, for zero mean returns, negative values have a higher probability happening than positive ones, a phenomenon that can visually be observed as the distribution having a "heavier" left tail. Whereas the leverage effect is used to describe the situation where negative returns have a greater impact on volatility that the positive ones. An additional advantage of the EGARCH model is that it models the logarithm of the conditional variance to ensure that the positive constraint holds in all cases (Bollerslev (2009)). The EGARCH(1,1) is given as:

$$\log(\sigma_t^2) = \omega + a(|z_{t-1}| - E(|z_{t-1}|) + \gamma z_{t-1} + \beta \log(\sigma_{t-1}^2) \tag{2.15}$$

where $z_t = \varepsilon_t \sigma_t^{-1}$ is the standardized innovations and the γ parameter controls the leverage effect.
In the literature of threshold models there have been introduced more specifications such as the asymmetric threshold GARCH. Further, due to the favorable characteristics of the standard deviation of it being more stable than the variance and it being expressed in the same units as the mean, Zakoian (1994) proposed the Threshold GARCH (TGARCH). Besides, TGARCH models the absolute value of the residuals. This is given as:

$$\sigma_t = \omega + \sum_{i=1}^{p} a_i |\varepsilon_{t-i}| + \gamma \varepsilon_{t-1}^2 I_{t-1} + \sum_{j=1}^{q} \beta_j \sigma_{t-j} \tag{2.16}$$

Without a non-negativity constraint the TGARCH model can be rewritten with the following Marko-

vian representation but then this specification cannot account for cyclical behavior. It can also be viewed as a linear ARCH.

$$\sigma_t = \omega + \alpha_1[z_{t-1}]\sigma_{t-1} + \ldots + \alpha_p[z_{t-p}]\sigma_{t-p} \tag{2.17}$$

where $\varepsilon_t = \sigma_t z_t$. Gonzalez-Rivera (1998) proposed the Smooth Transition GARCH (STGARCH) model, which also motivated the development of the Logistic STGARCH (LSTGARCH) model of Broto and Ruiz (2006). The representation of the latent model fixed for (1,1) lags is:

$$\sigma_t = \omega + \alpha_1\varepsilon_{t-1}^2[1 - F(\varepsilon_{t-1}) + \gamma\varepsilon_{t-1}^2 F(\varepsilon_{t-1})] + \beta_1\sigma_{t-1} \tag{2.18}$$

where $F(\varepsilon_{t-1})$ is the logistic function, where $F(\varepsilon_{t-1}) = (1+\exp(-\vartheta\varepsilon_{t-1}))^{-1} - 1/2$, $\vartheta > 0$ (Khemiri (2011), p.558). Anderson et al. (1999) combined the ideas of Fornari and Mele (1997), Hagerud (1997) and Gonzalez-Rivera (1998) and proposed the Asymmetric Nonlinear Smooth Transition GARCH (ANSTGARCH) model. The model is given by:

$$\sigma_t = \omega + \alpha_1\varepsilon_{t-1}^2 + 1 - F(\varepsilon_{t-1}) + \gamma\varepsilon_{t-1}^2 F(\varepsilon_{t-1}) + \beta_1\sigma_{t-1} \tag{2.19}$$

It is also possible to lose the non-negativity constraints and let ε_t take negative values so that non-cyclical pattens can be accommodated as well. This type of model was introduced from Rabemananjara and Zakoían (1993). Another specification that has been proposed is the Exponential STGARCH (ESTGARCH) model which relies on the assumption that only the shocks' size are responsible for the observed asymmetry (Franses and Van Dijk (2000)). In the same direction ε_{t-1} and $|\varepsilon_{t-1}|$ have been proposed as transition variables instead of the ε_{t-1}^2 that is used in the ESTEGARCH. That is called the smooth transition exponential smoothing (STES) (Taylor (2004)). In an attempt to generalize asymmetric models, Caporin and McAleer (2006) developed a dynamic asymmetric univariate GARCH model (DAGARCH). Their proposal adds flexibility and allows for "time-varying asymmetry with autoregressive structure that is related to both size and the sign of the shocks" in a multiple threshold framework (Caporin and McAleer (2006), p.386). Another characteristic of the DAGARCH models is that it is capable of identifying the source of each effect (Caporin and McAleer (2006)). Finally, Li and Li (1996) proposed the Double Threshold ARCH (DTARCH) model where "parameters in both mean and conditional variance to change across regimes" (Bollerslev (2009), p.10). Audrino and Bühlmann (2001) proposed the Tree Structured GARCH model to describe multiple limiting regimes in volatility. More general models included in this class are the asymmetric ARCH (AGARCH) of Engle (1990b)

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} a_i(\varepsilon_{t-i} - \gamma)^2 + \sum_{j=1}^{q} \beta_j\sigma_{t-j}^2 \tag{2.20}$$

and the Non-linear Asymmetric GARCH (NA-GARCH) of Engle and Ng (1993) given as:

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} a_i(\varepsilon_{t-i}\sigma_{t-i}^2 - \gamma)^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 \tag{2.21}$$

Based on the nonlinear GARCH model, Ding et al. (1993) developed a family of asymmetric GARCH models like the Asymmetric Power GARCH (APGARCH) which nests many other specifications. They tried to exploit the power functions of returns that tend to have a stronger autocorrelation.

$$\sigma_t^\delta = (\omega + a_1|r_{t-1}|^\delta + (\gamma r_{t-1})^\delta + \beta(\sigma_{t-1}^\delta) \tag{2.22}$$

when $\delta = 2$ and $\gamma = 0$ the model collapses to GARCH. and the Quadratic QGARCH which can be extended to accommodate thresholds (QTGARCH) (Bollerslev (2009)).

Recent evidence suggests that the asymmetry depends on the sign of the past shock and its effects on future volatility rather than on the returns (Fornari and Mele (1997)). Fornari and Mele (1997) following this direction proposed two models; the Sign- and Volatility - Switching ARCH models. The former is given as:

$$\sigma_t^2 = w + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 + \sum_{j=1}^{q} \alpha_j \varepsilon_{t-j}^2 + \sum_{x=1}^{j} \phi_x s_{t-x} \tag{2.23}$$

where $s_t = 1$, if $\varepsilon_t > 0$, $s_t = 0$, if $\varepsilon_t = 0$ and $s_t = -1$, if $\varepsilon_t < 0$. Nevertheless, there is evidences that leads to the conclusion that the sign of the past shocks alone was not enough to explain all the asymmetry observed. Therefore, the Volatility Switching ARCH (SV) has been proposed that "(..) captures asymmetries via the impact of past shocks on the level of the volatility, rather than through the unexpected returns" (Fornari and Mele (1997), p.50). The specification of that model is:

$$\sigma_t = w + a\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2 + s_{t-1}v_{t-1} \tag{2.24}$$

$$v_t = \delta_0\varepsilon_t^2 - \delta_1\sigma_t^2 - \delta_2 \tag{2.25}$$

Regime Switching Models

Contrary to the previously presented models, Hamilton (1989) suggested that the different regimes should depend on an unobservable variable that follows a Markov Chain in an attempt to capture the dynamic behavior of the time series during different time periods (Kuan (2002)). This subclass of non linear time series models is know as Markov Switching models. This model assumes that a few regimes happened. However, breaks identification process is not exogenous anymore and it is based on probabilities of the most likely regime at each time. The simplest model of this class is the Markov Switching Model of Conditional Mean:

$$r_t = \begin{cases} a_0 + \beta z_{t-1} + \varepsilon_t, & \text{if } s_t = 0. \\ a_0 + a_1 + \beta z_{t-1} + \varepsilon_t, & \text{if } s_t = 1. \end{cases} \tag{2.26}$$

where $|\beta| < 1$ and ϵ_t are i.i.d. random variables with zero mean and variance σ_ϵ^2. The model it is defined as two AR processes depending on the value of the state variable s_t. This model introduces two dynamic structures at different regimes r_t and it is driven by two distributions (Kuan (2002), p.3). In this specification the variance of ε_t is assumed to be constant in every regime but can vary among regimes (Satchell and Knight (2007)). An illustration of models that allow for this kind of validation is the the Markov-Switching Autoregressive (MS-AR), the Regime Switching AR, the RS-GARCH Regime Switching Models, the SWARCH-L, leverage, the MS-GARCH Markov switching model, multiple regime smooth transition heterogeneous autoregressive HARST by McAleer and Medeiros (2008), the Flexible Coefficient GARCH (FCGARCH) model by Medeiros and Veiga (2003) which is a multiple regime model, the switching-AR (k)-Markov-ARCH and the Markov-ARCH/switching-regime ARCH (SWARCH) model allows to detect which effect is more critical (Cai (1994)).

Persistence in the conditional variance

Shocks in the future conditional variance disappear at a slower rate than the one implied by the classic GARCH model. Models with a long memory process can be used to describe persistence. Previous studies capture this effect by measuring the decay rate of a shock, that is the fractional differencing parameter d, as estimated by the Integrated GARCH (IGARCH) (Nelson (1990)). In IGARCH a shock remains important for all futures horizons, $d = 1$. On the contrary, in the GARCH model a shock fades away at an exponential rate. The former model is given as:

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} a_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 \tag{2.27}$$

where the $a_i + \beta_j = 1$ constraint holds in order to impose a unit root. A special case is the IGARCH with no intercept witch is also called RiskMetrics model. Furthermore, in the Fractional Integrated ARCH (FIARCH) of Ding et al. (1993) and in its extension FIGARCH model of Baillie et al. (1996) a shock will dissipate at a slow hyperbolic rate ($0 < d < 1$), which for financial time series it has been approximated by $d = 0.4$. FIGARCH representation is given respectively:

$$\sigma_t^2 = \omega + [1 - \beta(L) - \phi(L)(1 - L)^d]\varepsilon_t + \beta\sigma_{t-1}^2 \tag{2.28}$$

This methodology can be extended to EGARCH by carrying out the relevant transformations. The resultant model is called FIEGARCH and it was originally proposed by Bollerslev and Mikkelsen (1996):

$$\log(\sigma^2) = \omega + \phi(L)^{-1}(1-L)^{-d}[1 + a(L)]g(z\_\{t-1\}) \tag{2.29}$$

Another specification is the FIAPARCH which was developed by Tse (1998):

$$\sigma_t^{\delta} = w + \{1 - [1 - \beta(L)]^{-1}\Phi(L)(1-L)^{-d}\}(|\varepsilon_t| - \gamma\varepsilon_t)^{\delta} \tag{2.30}$$

Finally, range-based models, such as range-based EGARCH (REGARCH) incorporates long memory and asymmetries and leverage effect. This is possible in a 1-factor or a 2-factor structure. Range-based models use data on the daily range which is the difference between the max and min (log) price for that day (Chou et al. (2010)).

Other specifications - Mixed

Engle et al. (1987) recognized that none of the available models in the literature could explain the risk-premium observed in the financial markets. Therefore, they proposed the GARCH-in-Mean which estimates the parameters in the mean and the variance equations simultaneously. GARCH-M is given by:

$$r_t = \mu + \delta\sigma_{t-1} + \varepsilon_t \tag{2.31}$$

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} a_i\varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j\sigma_{t-j}^2 \tag{2.32}$$

Component models, such as the multiplicative model differ from the rest approaches as they assign a different weight to intra-daily data but they should also not be confused with the RV models; see Gallant et al. (1999), Engle and Lee (1999b) and Chernov et al. (2003). Within this class there is the spline-GARCH of Engle et al. (2008); where "(...) that daily equity volatility is a product of a slowly varying deterministic component and a mean reverting unit GARCH" (Engle et al. (2013), p.777) and the HYBRID GARCH models of Chen et al. (2015). This type of models "(...) addresses the issue of volatility forecasting involving forecast horizons of a different frequency than the information set" (Ghysels (2014), p.28).

Engle et al. (2008), motived by the spline-GARCH and the need to explain the time and market varying nature of unconditional volatility, made used of the existing framework of mixed data sampling (MIDAS) which was originally proposed by Ghysels et al. (2005) to combine data that are sampled at different frequencies. "The weights of the MIDAS estimator implicitly capture the dynamics of the conditional variance. (...) Fortunately, the MIDAS approach can easily be implemented at different frequencies on the left-hand and on the right-hand side" (Ghysels et al. (2005), p.514). "The volatility specification in Ghysels et al. (2005) involves a single polynomial applied to daily data" (Ghysels

(2014), p.28).

However, Engle et al. (2008) were able to blend monthly data for the expected and daily squared returns to estimate the variance. In that way they created a novel component model with direct links to economic activity allowing for that data to be sampled at different frequency than daily returns. Another unique characteristic of this model is that new information release can affect the conditional variance depending on the state of the economy (Engle et al. (2008)). Specifically, it is a GARCH specifications that decomposes the conditional variance in a short-run component which is modeled as a standard mean-reverting GARCH(1,1) and a long-run component which is represented by a long history (monthly, quarterly, or biannually) of the realized volatility or macroeconomic variables weighted by a MIDAS polynomials which allows it to vary smoothly over time (Hang (2017); Engle et al. (2008)). Hence, this model is called GARCH-MIDAS. The model is given by the following series of equations:

$$r_{i,t} = \mu + \sqrt{(}\tau_t g_{i,t})\epsilon_{i,t}, \tag{2.33}$$

$$g_{i,t} = (1 - \alpha - \beta) + \alpha \frac{(r_{i-1,t} - \mu)^2}{\tau_t} + \beta g_{i-1,t}, \tag{2.34}$$

$$\tau_t = m + \theta \sum_{k=1}^{K} \phi_k(w_1, w_2) RV_{t-k}, \tag{2.35}$$

$$RV_t = \sum_{i=1}^{N_t} r_{i,t}^2, \tag{2.36}$$

$$\sigma_{it}^2 = \tau_t g_{i,y}, \tag{2.37}$$

$$\phi_k(\omega) = \frac{(1 - \frac{k}{K})^{w_1 - 1}(1 - \frac{k}{K})^{w_2 - 1}}{\sum_{j=1}^{K}(\frac{j}{K})^{w_1 - 1}(1 - \frac{j}{K})^{w_1 - 1}} \tag{2.38}$$

where $r_{i,t}$ is the return on day i in month t, $g_{i,t}$ is the component for the conditional variance dynamics which follows a (daily) GARCH(1,1) process, τ is defined as smoothed realized volatility in the spirit of MIDAS regression and K is the number of periods over which we smooth the volatility, the total conditional variance is represented by σ and the weighting scheme by ϕ. "The conditional variance is decomposed into the short-run component g_{it} and the long-run $r_{\tau t}$, RV_t is the RV of the month. What is more τ_i is fixed for $i = 1 : N$. Finally, the long-history of $RV_{t-1}, RV_{t-2}, \ldots, RV_{t-K}$, weighted by Beta polynomials, ϕ captures the long-run volatility" (Hang (2017), p.1).

There exist two types of specifications available for the GARCH-MIDAS model. First, the weighting scheme can be fixed over time and this case is suitable for low frequency data, representatively, the component τ_t can vary and under this scenario the GARCH-MIDAS model involves a rolling window

(Engle et al. (2008)). Additionally, Engle et al. (2008) estimated the model using only financial data but it can also be estimated using macroeconomic variables into the long-term component (Ghysels (2014)). Finally, MIDAS filtering can be used in different applications, such as to determine the weights in the case of forecast combinations (Ghysels (2014); Timmermann (2006)). Examples of MIDAS applications can be found on Engle et al. (2013) and Conrad and Loch (2015).

Beyond the predominant practices, there have also suggested in the literature various alternative methods in the area of volatility estimation and forecasting. For instance, artificial network models, focused on the "data themselves" in order to detect various regimes. A distinctive class of models is called State Space Models (SSM) that are mainly used in engineering. They are used in finance in the area of exponential smoothing. A different school of econometric tools and measures is the so called Bayesian statistics. An example of Bayesian model is the Smooth Transition Exponential Smoothing (STES) (see Poon (2005)). Bayesian non-parametric is an alternative approach to the classical parametric Bayesian statistics.

Table 2.1 – Univariate volatility models

Name	Author	Specification
Qualitative ARCH	Gourieroux and Monfort (1992)	-
Nonlinear ARCH	Higgins and Bera (1992)	-
Tobit GARCH or Censored-GARCH	Kodres (1993)	-
Composite GARCH	den Hertog (1994)	-
Modified GARCH	Friedman et al. (1989)	-
Generalized Linear ARCH (GLARCH)	-	-
Partially Nonparametric ARCH (PNP ARCH)	Engle and Ng (1993)	-
Modified ARCH	Friedman et al. (1989)	-
Randomized GARCH (RGARCH)	Nowicka-Zagrajek and Werron (2001)	-
Square-Root GARCH (SQR-GARCH)	-	-
Quadratic ARCH	Sentana (1995)	-
Discrete-time SGARCH	Heston and Nandi (2000)	-
Semiparametric ARCH(SP ARCH)	Engle and Gonzalez-Rivera (1991)	-
Structural ARCH (STARCH)	-	-
RGARCH (Randomized GARCH)	Nowicka-Zagrajek and Werron (2001)	-
Linear ARCH (LARCH)	Robinson (1991)	-
Conditional autoregressive range (CARR)	-	-
Augmented ARCH	Bera et al. (1992)	-
(GAARCH) the hyperbolic GARCH (HYGARCH)	Davidson (2004)	-
HARCH (heterogeneous ARCH)	Muller et al. (1997)	$\sigma_t^2 = \omega + \Sigma_{i=1}^{n} \gamma_i (\Sigma_{j=1}^{i} \varepsilon_{t-j})^2$
H-GARCH	Hentschel (1995)	$\sigma_t^2 = \omega + \alpha\delta\sigma_{t-1}^{\delta}(\left\|\varepsilon_{t-1}\sigma_{t-1}^{-1} - \kappa\right\| - \gamma(\varepsilon_{t-1}\sigma_{t-1}^{-1} - \kappa))^{\nu} + \beta(\sigma_{t-1}^{\delta})$
NGARCH, the non linear GARCH	Higgins and Bera (1992)	-
Power GARCH (PGARCH)	-	$\sigma_t^{\delta} = \omega + \Sigma_{i=1}^{p} a_i \left\|\varepsilon_{t-i}\right\|^{\delta} + \Sigma_{j=1}^{q} \beta_j (\sigma_{t-j})^{\delta}$
Modified ARCH	Friedman et al. (1989)	-
Log-ARCH	Geweke (1986) and Dieobold (1986)	$log(\sigma_t)^2 = \omega + \Sigma_{i-1}^{p} \alpha_i log(\varepsilon_{t-i})^2 + \Sigma_{j=1}^{q} \beta_j log(\sigma_{t-1})^2$
GARCH diffusion	-	$d\sigma^2(t) = (\omega - \theta\sigma^2(t))dt + \sqrt{2}a\sigma^2(t)dW_2(t)$
where are Wiener processes, the absolute residuals models	Taylor (1986) and Schwert (1989)	$\sigma_t^2 = a + \beta\|\varepsilon_{t-1}\|$
PNP-ARCH (Partially NoN-parametric ARCH)	-	$\sigma_t^2 = \omega + \beta\sigma_{t-1}^2 + \Sigma_{i=-m}^{m} \theta_i(\varepsilon_{t-1} - i\sigma)I(\varepsilon_{t-s} < i\sigma)$
Periodic GARCH (PGARCH)	Bollerslev and Ghysels (1996)	$\sigma_t^2 = \omega_{s(t)} + a_{s(t)}\varepsilon_{t-1}^2 + \beta_{s(t)}\sigma_{t-1}^2$
Continuous-time GARCH (COGARCH)	-	$\sigma_t^2 = [\sigma^2(0) - \omega \int_0^t \exp(x(s))ds] \exp(-x(t))$
Exponential Continuous GARCH (ECOGARCH)	Haug and Czado (2007)	-
GARCHX	Hwang and Satchell (2005)	-

Note: This table reports the univariate models.

2.1.3 Implied Volatility (IV)

In contrast to the historical volatility models introduced above, Implied Volatility (IV) is derived from priced options written on stocks. IV represents the market's expectation of the volatility over the remaining duration of the option. Thus, IV is often referred to as the future volatility based on market's expectations (Christensen and Prabhala (1998)). Since IV is both market based and forward looking it is possible to relate ex-ante and ex-post volatility. A favorable characteristic of volatility estimation based on option's prices is that a specification of a time series model is not needed.

Transitionally, IV is the volatility that equates the theoretical price after using a model, such as the Black-Scholes (B-S), to the market price (Bollerslev (2009)). From the Black—Scholes call option-pricing formula it is possible to derive the ex-ante IV. To compute ex-post volatility forecast over the remaining life of the option, the sample standard deviation of the daily index returns over the remaining life can be used (Christensen and Prabhala (1998)). Another way to measure IV is to rely on observations derived from IV indices. A common practice is to consider IV index market prices with various strikes and expiration date and the take the mean, of those market prices of options, with respect to the strike price and the expiration dates in an attempt to minimize the measurement error (Konstantinidi et al. (2008)). The most broadly used index is VIX which now is a Model-Free option IV (MFIV) measure for the S&P 500 index. Initially, it was computed using the B-S IV of eight S&P 500 options (Andersen and Benzoni (2009)). However, since 2003 "the index is computed from concurrent price quotes on a wide range of out-of-the-money" over the coming 30 days (Andersen et al. (2010), p.1). VIX is computed using a linear combination of two different maturities that they are close to 30 days, because of the lack of available option maturities.

$$\sigma = \frac{2}{T}\sum_{i}\frac{\Delta K_i}{K_i^2}e^{RT}Q(K_i) - \frac{1}{T}[\frac{F}{K_0} - 1]^2 \tag{2.39}$$

where $\sigma = 100 * VIX$, T is the time to expiration, F is the forward index level from index option prices, K_0 is the first strike price below F, K_1 strike price for the i^{th} out-of-the-money option; a call if $K - i > K_0$, a put if $K - i, K_0$, a put and a call if $K - i = K_0$, ΔK_i interval between the strike prices - half the distance between the strike on either side of K_i, $\Delta K_i = \frac{K_{i+1} - K_{i-1}}{2}$, R is the risk-free rate to expiration and $Q(K_i)$ is the midpoint of the bid-ask spread for each option with strike K_i (CBOE) (Zhang et al. (2013)). MFIV is considered a substitute to the B-S formula approach. Taylor et al. (2010) argued that MFIV is considered superior from B-S IV as it is model-free, so it makes no (or, at least, very little) assumptions about the dynamics of the volatility and also it relies on multiple strike prices. Jiang and Tian (2005) added price jumps to MFIV model.

In reality, it is very hard to calculate the full MFIV as the number of available strikes is limited.

Therefore, an approximation of the CBOE VIX is used. Hereof, Andersen et al. (2015) suggested a method that makes use of mid-quotes that excludes all out-of-the-money (OTM) options with zero bid prices but accepts all with a positive one, creating in that way a more vigorous measure that corrects for the distortions (eg. bias and breaks) caused by the CBOE stopping rule. The resultant measure is called a model-free Corridor Implied Volatility Index (CX)

$$CIV = 2\int_{B_L}^{B_H} \frac{M_{o(K)}}{K^2} dK \tag{2.40}$$

where $M_{0(K)}$ is the price of the vanilla option with a strike K which is out-of-the-money and $0 \leq B_L \leq B_H \leq \in$ (Andersen et al. (2015)). Illustrations of more indices commonly-used are: the replication indices of VIX, the ultra high frequencies VIX, the end of the day VIX and the daily VIX VXD futures, the CBOE OEX implied volatility which now is called VXO is based on S&P 100 OEX option prices (Christensen and Prabhala (1998)).

Alternatively, IV can be modeled using a regression. One example is to assume that it depends on the natural logarithm of the Black and Scholes implied volatility or at-the-money call options on the S&P 100 index as measured at the beginning of month t and on the natural logarithm of the ex-post daily return volatility of the index, over the life of the option. On the other hand, Konstantinidi et al. (2008) suggested that IV depends on a number of economic variables (p.8). Likewise, a regression model can be used to capture seasonal patterns. It could include dummy variables for Monday and Friday to capture the relevant effect, lags of both, the call and put implied volatility changes, to capture any autocorrelation, interest rate variables and S&P 500 futures contract (Harvey and Whaley (1992), p.54-55). A common assumption when estimating the seasonal component in volatility is that it stays constant over a certain period, for example, a day or a week or a month.

There is a growing interest in the literature about the information content and the predictive ability of IV (Taylor et al. (2010)). Theoretically, option-implied forecasts should dominate forecasts based on historical models since the first are forward looking and already encompass the information included in historical models. In practice, illiquidity of option contracts and other deficiencies affect the performance of IV models. Further, it should be noted that implied volatility is obtained under the risk-neutral measure. So comparing IV models to historical models it can only be performed under the unrealistic assumption that the risk premium is zero. In that case, option-implied forecasts need to be adjusted for the volatility risk premium.

2.1.4 Stochastic Volatility (SV)

Implied Volatility has been criticized since it assumes constant volatility. In fact, the empirically observed volatility smile and skew confirm that this is not accurate, and volatility does change over the option's life. "Stochastic Volatility (SV) models are based on the direct correlation between the innovations in both returns and volatility" (Asai and McAleer (2005), p.317). Therefore in SV models, volatility has its own stochastic process (Bauwens et al. (2012a)).

Following a strict definition, the GARCH models are not genuinely stochastic. SV models can be regarded as an alternative which is often superior to GARCH type models since they are consistent with the economic theory, offer a direct way to move to continuous time framework and add flexibility (Bauwens et al. (2012a)). Moreover, SV models can address in a straightforward way, matters of contention in financial time series modeling, such as the concession of the stylized facts. SV models can be discrete or continuous. Following Taylor (1982) the discrete time SV is :

$$y_t = \exp(h_t/2)u_t \tag{2.41}$$

$$h_t = \mu + \phi(h_{t-1} - \mu) + \eta_t \tag{2.42}$$

where y_t is the log returns and h_t is the log volatility with parameter $|\phi| < 1$.The error terms u_t and η_t are Gaussian white noise processes. An additional advantage of the SV models is that naturally, within their main assumptions can offer a valid justification for the volatility clustering. On the grounds of the Mixture-of-Distributions Hypothesis, in SV models the information set is time-varying and therefore at each time returns will include the normal innovations but also the shocks to the variance (Andersen and Benzoni (2009), p.4). Model representations in 2.41 and 2.42 can be expanded to reflect this hypothesis (Hautsch and Ou (2008)). If in 2.41, the square root of λ_t (an i.i.d random variable), which is assumed to follow the inverse gamma distribution, is added then it reads as:

$$y_t = \exp(h_t/2)u_t\sqrt{\lambda}_t \tag{2.43}$$

Moreover, a simple modification can be made to adjust for jumps.

$$y_t = k_t q_t + \exp(h_t/2)u_t \tag{2.44}$$

$$h_t = \mu + \phi(h_{t-1} - \mu) + \eta_t \tag{2.45}$$

where q_t is a Bernoulli random variable that is 1 in the presence of a jump and 0 otherwise. A serious weakness with this approach however, is that discrete time framework is an unrealistic assumption. Continuous time framework is more consistent in volatility estimation and forecasting. Technological

progress made high frequency returns readily available and supported the spreading of continuous models. The basic framework of models in this class is:

$$dS_t = \mu S_t dt + \sigma(V_t) S_t dW_{1t} \tag{2.46}$$

$$dV_t = \psi V_t dt + \beta V_t dW_{2t} \tag{2.47}$$

where S_t represents logarithmic price and V_t is the latent volatility process, dW_{1t} and dW_{2t} are (possibly correlated) Brownian motions, ψ is the mean of the V_t process and β is the volatility of V_t process. Likewise this class of models can account for jumps following Merton's jump diffusion Model. Finally, Hull and White (1987) introduced the log normal SV model. Often used is the Euler specification. It is possible to have an SV-mix, Euler scheme of a continuous with regime that is called SV-mix. As Durham (2007) pointed out, the SV-mix approach is an alternative to the affine-jump models, where a low-probability, high-variance component corresponds to the jumps.

SV models are mainly used for pricing financial assets, therefore incorporation of leverage effects is critical, since if ignored the estimates may be biased. Leverage effects can be added on the Euler specification following Harvey and Shephard (1996) where the innovations remain independent, or alternatively as Jacquier et al. (2004) suggested the discrete SV model can be adjusted so that the innovations follow a bivariate normal distribution (Yu (2005); Bauwens et al. (2012a), p.26). The general asymmetric SV is given:

$$dS_t = \sigma(t) dW_1(t) \tag{2.48}$$

$$d \ln \sigma^2(t) = a + \beta \ln \sigma^2(t) dt + \sigma_\upsilon dW_2(t) \tag{2.49}$$

where S(t) is the asset price, and $W_1(t), BW_2(t)$ are Brownian motions (Yu (2005)). When the correlation of the two Brownian motions is negative it means that there is a leverage effect. Further an Euler and a discrete specifications are also available.

In the case of stochastic volatility (SV) models, leverage effects can also be specified through the dynamic leverage (DL) model and through the dynamic asymmetric leverage (DAL) model that accommodates both; the sign and the magnitude of threshold effects can be employed (Asai and McAleer (2005)). In addition, the literature suggested the SV model with leverage effects of Omori et al. (2007) and Yu (2005), the threshold SV (TSV) model of So et al. (2002) and the dynamic leverage threshold effects (DLTE) model of Asai and McAleer (2004) (Asai and McAleer (2005), p.318). Harvey and Shephard (1996); So et al. (1998) and Kalimipalli and Susmel (2007) discussed SV models with Markovian regime switches, such as the ARSV(1) stochastic of Taylor (1986). Another suggestion is the Asymmetric, stochastic volatility model (ASV) volatility model with DPM model (ASV-DPM).

Alternatively, the Eigenfunction SV of Meddahi (2001) ESV is another way to specify different performance to different states.

However, the aforementioned models fail to account for the volatility in-mean effect. That motivated Koopman and Uspensky (2002) to propose the the SV in mean (SVM) model which is analogous to the GARCH-in mean model. As discussed before, financial returns exhibit persistence in variance which is mainly observable through the Autocorrelation Function (ACF) plot. Hence, Breidt et al. (1998) and Harvey (1998) proposed the Long Memory SV (LMSV) model that allows for the incorporation of an ARFIMA process.

$$(1 - \phi L)(1 - L)^d \log(\sigma_t^2) = \eta_t \tag{2.50}$$

where d is the long memory parameter. Further, in the context of SV, persistence in volatility can be captured by specifying a RW, squaring the expression and taking logarithm of both sides. This process is not stationary but its first difference is. Another specification can be used instead of the RW it is the doubly integrated RW where it is a white noise.

The literature has focused as well on the effect of the non-trading days and what is the optimal way to include them in a time series analysis. There have identified three ways; the first is to include dummy variable on the volatility level, whereas the second approach suggests to approximate continuous time SV models using a discrete time approach, Finally, it is possible to combine both methods into one (Asai and McAleer (2005)).

Despite their many advantages, SV models are not used extensively due to limitations concerning their estimation. The most common methods for parameter estimation in SV models are the Generalized Method of Moments (GMM), the Efficient Method of Moments (EMM) and the Frequency Domain Quasi-Maximum Likelihood (FDQML) estimation under the assumption that the SV process for the underlying asset price is continuous.

2.1.5 Realized Volatility (RV)

GARCH and SV models have several advantages, yet they come with a critical limitation. They treat the volatility as a latent variable and this renders the evaluation of their performance problematic. Since the true volatility is unobservable and thus unknown, a proxy variable is being used. This makes the evaluation of the models sensitive to the proxy variable and eventually distorts the final inference. Finding the right approximation of the true unobserved realized volatility is imperative.

In the past, due to limited availability of high frequency data, squared daily returns where used as a

proxy for daily volatility. However, when squared daily returns where regressed against the forecasted volatility from GARCH models gave very low R^2 values. Now we know that this implies that this measure it is not the best proxy to evaluate the forecasting performance of models. Additionally, it has been found that the performance of the daily squared returns deteriorates as volatility increases (Liu et al. (2015)). Parkinson (1980) was the first to made an attempt to introduce intra daily observations by introducing the following estimator:

$$\sigma_t = \sqrt{\frac{(\ln H_t - \ln L_t)}{4 \ln 2}} \tag{2.51}$$

where H and L are the high and low intra daily observations. Garman and Klass (1980) extended Parkinson's estimator as:

$$\sigma_t = \sqrt{\frac{1}{2}\left(\ln \frac{H_t}{L_t}\right)^2 - (2\ln 2 - 1)\left(\ln \frac{P_t}{P_{t-1}}\right)^2} \tag{2.52}$$

where P_t is the daily prices. Empirical results confirmed that those two reformed estimators had a better performance than squared daily returns. Logarithmic range estimator is superior to using single logarithmic or absolute returns since it is a closer approximation of the Normal distribution, it is more informative and it can minimize the measurement error (Alizadeh et al. (2002)). Technological progress made high frequency data readily available facilitating the use of continuous time processes like:

$$dp(t) = \mu(t)dt + \sigma(t)dW(t) \tag{2.53}$$

where p is a Poisson process and W_t is a standard Brownian motion. The volatility is represented by the Integrated Volatility (IV) :

$$IV_t^{(d)} = \int_{t-h}^{t} \sigma(s)^2 ds \tag{2.54}$$

The IV(d) can be approximated by the sum of intraday squared returns (Corsi (2009)). Therefore, Realized Volatility (RV) of Andersen et al. (2001a) was used as proxy of the true unobserved volatility.

$$RV_t^{(d)} = \sqrt{\sum_{j=0}^{M-1} r_{t-j\cdot\Delta}^2} \tag{2.55}$$

where $\Delta = 1d/M$ and $r_{t-j\cdot\Delta} = p(t-j)$ is intraday returns sampled at time interval Δ, t is the day, while j is the time within the day (Corsi (2009)). RV is an unbiased estimator under certain conditions (Barndorff-Nielsen and Shephard (2002a)). Finding a better estimator of the true volatility explained partially the bad forecasting performance of GARCH volatility models. The sample frequency is a trade-off between the incorporation of all the relevant information and the reduction of noise. Liu

et al. (2015) showed that nothing beats the 5-min RV. Key papers in the area of Realized Variance (or Volatility) and high frequency data include Andersen et al. (2001a), and Barndorff-Nielsen and Shephard (2002a) and Barndorff-Nielsen and Shephard (2002b), Meddahi (2002), and the references within.

A growing literature recognizes important gains in volatility forecasting from modeling directly the realized variance since "(..) it is the main variable of interest for financial applications (..)" (Corsi et al. (2008), p.48; McAleer and Medeiros (2008)). Therefore, the RV can not only be used as a proxy but also to forecast the volatility as RV time series contain valuable information about the dynamics of volatility (Hansen and Lunde (2004)). In view of that, non-parametric RV models take volatility as an observed variable and tend to perform better than GARCH models (Duong and Swanson (2015)).

The most naive way is to fit RV in an Autoregressive (AR) model. However, it too simplistic and does not comply with the observed characteristics of the financial time series. RV times series exhibit long memory, leverage effects and jumps. Another proposition is to use the ARQ model which allows for a time-varying AR parameter to depend linearly on an estimate of $IQ^{1/2}$ (Bollerslev et al. (2016), p.5-8) and RQ corresponds to RV in the case of data limitations.

$$RV_t = \beta_0 + (\beta_1 + \beta_{1Q} RQ^{1/2}) RV_{t-1} + u_t \tag{2.56}$$

where IQ ($IQ = \int_{t-1}^{t} \sigma_s^4 ds$) is the equivalent of IV, and RQ ($M/3 \sum_{i=t}^{M} r_{t,i}^4$) is the equivalent of RV in case of data limitations (Bollerslev et al. (2016), p.5-8). However, the aforementioned models fail to account for the persistence in the conditional variance. Shocks in the future conditional variance decay at a slow rate. With regard to this phenomenon long memory models were suggested, such FIGARCH model. Andersen et al. (2001a), Barndorff-Nielsen and Shephard (2002a), Andersen et al. (2003) and Martens et al. (2004) introduced long memory, asymmetries and structural models breaks into a realized volatility model, expanding the nonlinear Autoregressive Fractionally Integrated class of models followed by Scharth and Medeiros (2009) who proposed a multiple regime tree structure model.

The main weakness of fractional integration, as highlighted by Corsi (2009), is that it is not an intuitive approach, it does not allow to examine dependently short and long term dynamics, it requires a lot of observations and does not capture "the multiscaling behavior found in the empirical data" (Corsi (2009), p.175). Therefore, Corsi (2009) inspired by the Heterogeneous Market Hypothesis and the Heterogeneous ARCH (HARCH) model of Muller et al. (1997) and Dacorogna et al. (1998) and developed the so called HAR model. This model is a reduced form cascade model that belongs to the reduced form time series forecasting models and despite the fact that is not formally a long memory model it has been proven suitable to capture the observed persistence in variance with only few parameters (Corsi et al. (2012), Corsi (2009)). It is true that RV time series exhibit strong and slowly

decaying persistence (Bollerslev et al. (2016), Corsi et al. (2012)). That is possible as a result of the model's lag structure which even though fixed for most of the cases is capable of accounting for both long memory persistence of volatility (through the 22-day) and for short term dynamics (through the 1-day and 5-day term). Another advantage of the model is that it overcomes measurement errors of RV (Asai et al. (2012)). The basic HAR-RV model is given as:

$$RV_t = \omega + \phi_1 RV_{t-1} + \phi_2 \sum_{i=1}^{5} \frac{1}{5} RV_{t-i} + \phi_3 \sum_{j=1}^{22} \frac{1}{22} RV_{t-j} + u_t \tag{2.57}$$

where RV_{t-i} is the lagged terms of volatility, to capture the 5-day week effect and the 22-day month effect. The models allows for two transformations, the logarithmic and the squared root. The logarithmic one is motivated by the need to approximate the Normal distribution and to ensure that the non negative constraints hold. Whereas, the square root corresponds to the standard deviation which is generally more stable than variance and holds the additional advantage of being measured in the same units as the mean. The two non linear specifications are given respectively.

$$\log(RV_t) = \omega + \phi_1 \log(RV_{t-1}) + \phi_2 \sum_{i=1}^{5} \frac{1}{5} \log(RV_{t-i}) + \phi_3 \sum_{j=1}^{22} \frac{1}{22} \log(RV_{t-j}) + u_t \tag{2.58}$$

$$(RV_t^{1/2}) = \omega + \phi_1 (RV_{t-1}^{1/2}) + \phi_2 \sum_{i=1}^{5} \frac{1}{5} (RV_{t-i}^{1/2}) + \phi_3 \sum_{j=1}^{22} \frac{1}{22} (RV_{t-j}^{1/2}) + u_t \tag{2.59}$$

There are many studies that apply the HAR-RV model; some examples include McAleer and Medeiros (2008), Maheua and McCurdycd (2011), Duong and Swanson (2015), Bollerslev et al. (2016), Corsi et al. (2010) and Andersen et al. (2007). The popularity of the model can by attributed to its simplicity, to the fact that can be estimated by OLS and to its good empirical results. However, the model has attracted negative criticism as well, primarily for producing wrong estimators. Bollerslev et al. (2016) proposed an extension that corrects for the estimated parameters.

$$RV_t = \beta_0 + (\beta_1 + \beta_{1Q} RQ_{t-1}^{1/2}) RV_{t-1} + (\beta_2 + \beta_{2Q} RQ_{t-1|t-5}^{1/2}) RV_{t-1|t-2} + (\beta_3 + \beta_{3Q} RQ_{t-1|t-22}^{1/2}) RV_{t-1|t-22} + u_t \tag{2.60}$$

where $RQ_{t-1|t-k} = \frac{1}{k} \sum_{j=1}^{k} RQ_{t-j}$. This models is then called the full HARQ, but a simpler specification is available that makes it easier to estimate (Bollerslev et al. (2016), p.4). That is:

$$RV_t = \beta_0 + (\beta_1 + \beta_{1Q} RQ^{1/2}) RV_{t-1} + \beta_2 RV_{t-1|t-5} + \beta_3 RV_{t-1|t-22} + u_t \tag{2.61}$$

Following the same direction, the HARQ-h specification has been suggested that adjusted only the term that is related to the forecasting horizon. Therefore, considering a weekly forecasting horizon only the second explanatory variable is being substituted and correspondingly only the third variable

for annual forecasts. However, all the aforementioned HARQ models assume a constant intercept. Therefore, it has been proposed to substitute RQ with $RQ^{1/2}$, RQ^{-1} and log(RQ) (Bollerslev et al. (2016), p.14). As Quaedvlieg (2016) noted "(..) the estimated HARQ models exhibit more persistence in 'normal times' and quicker mean reversion in 'erratic times' compared to the standard HAR model with constant autoregressive parameters" (p.80).

In order to achieve satisfactory forecasting performance it is important to allow for breaks in the time series such as jumps. Corsi and Reno (2009) introduced the extensions of the basic HAR model. The main representative of this class is the HAR-RV-J where an extra explanatory variable is added. It decomposes the total variation into a continuous part and a jump; see also Bollerslev et al. (2016), Andersen et al. (2007).

$$RV_t = \beta_0 + \beta_1 RV_{t-1} + \beta_2 \sum_{i=1}^{5} \frac{1}{5} RV_{t-i} + \beta_3 \sum_{j=1}^{22} \frac{1}{22} RV_{t-j} + \phi_4 J_{t-1} + u_t \tag{2.62}$$

where $J_t = max(RV_t - BVT_t)$. For the same reasons explained above a modification holds:

$$RV_t = \beta_0 + (\beta_1 + \beta_{1Q} RQ^{1/2}) RV_{t-1} + \beta_2 RV_{t-1|t-5} + \beta_3 RV_{t-1|t-22} + \beta_J J_{t-1} + u_t \tag{2.63}$$

More specifications exist in the literature with upper, down, asymmetric and truncated large jumps such as the HAR-TCJ and the HAR-RV-CJ which are extensions of the HAR-RV-J model that can separate continuous variability to jump variation (Andersen et al. (2007)). In the context of HAR model in comes natural to account simultaneously for jumps and the leverage effect which was first discussed on Black (1976). The resultant model is called Leverage Heterogeneous Auto-Regressive with Continuous volatility and Jumps (LHAR-CJ) model and it was introduced by Corsi and Reno (2009) (p.7):

$$\begin{aligned} \log \widehat{V}_{t+h}^{(h)} &= c + a^{(d)} \log(1 + J_t) + a^{(w)} \log(1 + J_t^{(5)}) + a^{(m)} \log(1 + J_t^{(22)}) \\ &+ \beta^{(d)} \log C_t + \beta^{(w)} \log C_t^{(5)} + \beta^{(m)} \log C_t^{(22)} \ + \gamma^{(d)-} r_t^{-} + \gamma^{(w)-} r_t^{(5)-} + \gamma^{(m)-} r_t^{(22)-} + \varepsilon_t^{(h)} \end{aligned} \tag{2.64}$$

At last, models have been proposed that rely on different RV measures such as the bipower variation of Barndorff-Nielsen and Shephard (2004), the Tri-Power or the multiple regime smooth transition extension HAR (HARST) of McAleer and Medeiros (2008) which is a combination of the HAR and the Smooth Transition Autoregressive (STAR) models capturing both the persistence and the asymmetries and the HAR-GARCH of Corsi et al. (2008) which includes a GARCH component to account for volatility clustering of the RV (see also Corsi et al. (2008) p.63).

An alternative to the HAR with jumps class of models is the Continuous HAR (CHAR) model, where continuous variation is inserted on the right hand side of the equation (Bollerslev et al. (2016); Andersen et al. (2007)).

$$RV_t = \omega + \phi_1 BPV_{t-1} + \phi_2 \sum_{i=1}^{5} \frac{1}{5} BPV_{t-i} + \phi_3 \sum_{j=1}^{22} \frac{1}{22} BPV_{t-j} + u_t \tag{2.65}$$

where $BVT_t = \mu_1^2 \sum_{i-1}^{M-1} |r_t, i||r_{t,i=1}|$ with $\mu_1 = \sqrt(2/\pi) = E(|Z|)$ and Z a standard normally distributed random variable (Bollerslev et al. (2016), p.8). In the same direction the semivariance HAR (SHAR) model of Patton and Sheppard (2015) exists that separates positive and negative returns (Bollerslev et al. (2016), p.8). It is given as:

$$RV_t = \beta_0 + \beta^+ RV_{t-1}^+ + \beta^- RV_{t-1}^- + \beta_2 RV_{t-5:t-1} + \beta_3 RV_{t-22:t-1} + u_t \tag{2.66}$$

Chung et al. (2008) added as a depended variable past values of trading volume and the resultant model is called CHAR-RV-cum-volume. RV can also be included in the volatility equation of GARCH model as an exogenous variable as suggested by Engle (2002b) and Koopman et al. (2005). The resultant models is known as GARCH-RV and is given as:

$$h_t = \omega + \alpha \epsilon_{t-1}^2 + \theta h_{t-1} + \lambda RV_{t-1} \tag{2.67}$$

GJR-RV is also another option where daily returns are substituted by RV. See Lu (2005) for a vigorous empirical study. However, GARCH models are estimated through a single volatility process. Alternatively, RV can be estimated directly, using models that operate with multiple volatility processes. On the grounds of this conception, Engle and Gallo (2006) proposed the Multiple Error Model (MEM) which is an extension of the model that Engle (2002b) proposed. Brownlees and Gallo (2009) implemented these ideas in a risk management approach. Further, Shephard and Sheppard (2010) introduced the High-frEquency-bAsed VolatilitY models (HEAVY) model. MEM models introduces a new volatility process for each realized measure so that, the basic MEM model has three latent volatility process whereas the HEAVY model has at least two.

It is worth mentioning that the GARCH and the HEAVY models although closely related they are not considered nested models and it appears that the latter has many advantages over the GARCH model as they "(...) have momentum and mean reversion effects, and they adjust quickly to structural breaks in the level of volatility process" (Shephard and Sheppard (2010), p.197). Finally, HEAVY models are easy to estimate using the quasi-likelihood method. The model is defined by the following system of equations; this system gives the basic representation of the model. Although a common representation of the model is the vector multiplicative form.

$$var(r_t|F_{t-1}^{HF}) = h_t = \omega + \alpha RM_{t-1} + \beta h_{t-1} \tag{2.68}$$

$$E(RM_t|F_{t-1}^{HF}) = \mu_t = \omega_R + \alpha_R RM_{t-1} + \beta_R \mu_{t-1} \tag{2.69}$$

where r corresponds to daily returns while RM to daily realized measures. The first equation is known as HEAVY-r model while the second one as HEAVY-RM model. For 1-step ahead forecasts the HEAVY-r model is sufficient. The HEAVY model can accommodate squared daily returns or intra-daily returns. Under the former case, the performance of the model can be compared with the component GARCH model introduced by Engle and Lee (1999a). "However, HEAVY models are much easier to estimate as they bring two sources of information to identify the longer term component of volatility" (Shephard and Sheppard (2010), p.198). The persistence of the RM is governed by the $\alpha_R + \beta_R$ parameter. Parameter β is called momentum parameter and it has a different function from the GARCH models where it controls the mean reversion of the conditional variance (Shephard and Sheppard (2010), p.203). As expected, various extensions have been proposed. For example, it may be necessary to add on the right hand side of the first equation the r_{t-1}^2 variable to include the lagged values of squared daily returns as a GARCH-X model since it may carry some important information for some assets (Shephard and Sheppard (2010)). More extensions that have been proposed include the targeting parameterization extension that has been developed by Engle and Mezrich (1996) or extensions that account for leverage effects and an alteration of the basic model which can accommodate joint empirical distributions (Shephard and Sheppard (2010)). Finally, similarly to GARCH models, integrated version of the model can be derived by imposing a unit root.

2.2 Multivariate volatility models

Engle and Susmel (1993a) demonstrated on their paper that stock indices' returns and volatilities tend to move together over time due to shocks affecting certain assets simultaneously. In an attempt to exploit and understand those linkages and interactions among markets, and having established the importance of accurate volatility forecasts from the first chapter, it is sensible to consider multivariate models for the volatility, as well. In that case two or more stock markets are modeled jointly.

We can distinguish between non, semi and full parametric models. In most of the cases the multivariate models are extensions of the univariate models. The two main categories, on parametric conditional volatility models, are the VEC and the BEKK GARCH models. Multivariate models have two main issues, too many parameters to estimate and imposing positive conditional covariance matrix. One way to overcome these problems is by restricting the aforementioned models. Another way is to directly model the correlation. The constant Conditional Correlation and the Dynamic Conditional

Correlation model are the two most used models in this category. Another approach is to relay on Multivariate factor models which are trying to identify a common factor in an attempt to minimize the number of parameters that need to be estimated. Finally, in a similar way stochastic and realized volatility models can be extended to the multivariate case.

Non-parametric and Semi-parametric estimation

For the same reasons that have been outlined for univariate volatility models, non-parametric covariance models are present in the literature. The equivalent Moving Average Covariance model is given as:

$$\hat{\Omega}_t = n^{-1} \sum_{i=1}^{n} Y_t Y_t' \tag{2.70}$$

where $Y_t = R_t$ is a $N \times 1$ vector of asset returns (Sheppard (2016)). The limitations of this approach are well-known thus, an Exponentially Weighted Moving Average (EWMA) filter is required. The EWMA covariance model is given as:

$$\hat{\Omega}_t = \gamma Y_t Y_t' + (1-\gamma)\hat{\Omega}_{t-1} \equiv \gamma \sum_{i=1}^{\infty} (1-\gamma)^{i-1} Y_t Y_t' \tag{2.71}$$

and $\gamma = 0.94$ daily data and $\gamma = 0.97$ for monthly data (Morgan et al. (1996); Andersen et al. (2006), p.840; Sheppard (2016)). The recent stream of literature is focused on how long memory effects can be adopted. One way is to assume that weights decay at a hyperbolic rate. The new RiskMetrics 2006 methodology introduces the Student-t distribution for the residuals and two other terms for the scale condition beyond the long memory effect (Zumbach (2007)). Now the covariance matrix is calculated as a weighted sum of EWMAs and that is believed to produce more accurate results (Sheppard (2016)). The original methodology based on a single EWMA process is nested in the 2006 methodology. What is more, the corresponding time-varying correlations are given as: $\hat{p}_{ij,t} = \frac{\{\hat{\Omega}_t\}_{ij}}{\{\hat{\Omega}_t\}_{ii}^{1/2}\{\hat{\Omega}_t\}_{jj}^{1/2}}$ (Andersen et al. (2006), p.840).

Parametric estimation

Vector GARCH (VEC)

Alternatively, trying to bridge the gap between the distribution-free and the parametric designs Hafner and Rombouts (2007) assumes a MGARCH process but there are no assumptions about the distributions of the errors. Respectively, semi-parametric (SP) approaches for the conditional correlations(CC) such as the (SPCC–) GARCH model of Hafner et al. (2006); see also Silvennoinen and Teräsvirta (2009b)). This type of models have some advantages over the parametric models, even though they are very simple. First of all, they have a structure where interpretation of the conditional covariance matrix is straightforward and secondly, they are robust over distributional assumptions (Silvennoinen

and Teräsvirta (2009b)).

Multivariate GARCH class of models were first introduced by Bollerslev (1990). For the ARCH specification see Engle et al. (1984). To begin with, consider a vector of asset's returns, Y_t which can be represented by the model:

$$Y_t = M_{t|t-1} + \Omega_{t|t-1}^{1/2} Z_t \tag{2.72}$$

where $Z_t \sim i.i.d$, Z_t is a vector white noise process with unit variances, $M_{t|t-1} \equiv E(Y_t|F_[t-1])$ and $\Omega_{t|t-1} \equiv Var(Y_t|F_{t-1})$. It is assumed that $E(Z_t) = 0$ and $Var(Z_t) = I$ (Andersen et al. (2006), p.841). $\Omega_{t|t-1}^{1/2}$ can have any parameterization as long as it produces a positive definite matrix, and it is flexible but parsimonious.

The most popular approach is to apply the *vech* operator to the conditional variance. The resulting model is the VEC model which is regarded as the multivariate counterpart of the univariate GARCH(1,1) and it was introduced in the paper of Bollerslev et al. (1988).

$$vech(\Omega_{t|t-1}) = C + Avech(e_{t-1}e'_{t-1}) + Bvech(\Omega_{t-1|t-2}) \tag{2.73}$$

where $e_t \equiv \Omega_{t|t-1}^{1\prime 2} Z_t$, $vech(.)$ denotes the operator that stacks the $\frac{1}{2}N(N+1)$ unique elements in the lower triangular part of a symmetric matrix into a $\frac{1}{2}N(N+1) \times 1$ vector, and the parameter matrices C, A, and B, are of dimensions $\frac{1}{2}N(N+1) \times 1, \frac{1}{2}N(N+1) \times \frac{1}{2}N(N+1)$, and $\frac{1}{2}N(N+1) \times \frac{1}{2}N(N+1)$, respectively (Andersen et al. (2006), p.841, see also Franses and Van Dijk (2000), p.200). Similarly, asymmetries can be accommodated "by including the signed cross-products of the residuals on the right-hand side. The most straightforward generalization would be to simply include $vech(min(e_{t-1}, 0)min(e_{t-1}, 0)')$, but other matrices involving the cross-products of $max(e_{t-1}, 0)$ and/or $min(e_{t-1}, 0)$ have proven important in some empirical applications" (Andersen et al. (2006), p.842). Obviously, more lagged terms and other exogenous variables can be included in the model as in the univariate case. More about non linear specifications can be found in the papers of Hafner and Herwartz (1998a, 1998b) and Kroner and Ng (1998).

VEC model is very flexible and it allows for all elements of covariance matrix $\Omega_{t|t-1}$ to depend on each cross-product of e_{t-1} and all elements of $\Omega_{t-1|t-2}$ (Franses and Van Dijk (2000), p.200). Nevertheless, this entails that the number of parameters that need to be estimated is very large. A total of $\frac{1}{2}N^4 + N^3 + N^2 + \frac{1}{2}N$ unique parameters need to be estimated and therefore another problem that emerges is that is almost impossible to guarantee the positive definiteness of the unconditional covariance matrix $((I - A - B)^{-1}C)$ unless some restrictions are imposed on the C, A, and B parameter matrices (Andersen et al. (2006), p.842). This motivated a stream in the literature that studies

various specifications that aim at ensuring positive definiteness and reducing the number of parameters that need to be estimated. A well-used approach is the diagonal (VEC) model of Bollerslev et al. (1988) which suggests that by restricting the matrices A and B to be diagonal the number of unique elements will be decreased so that "a particular element in the conditional covariance matrix only depends on its own lagged value and the corresponding cross product of the innovations" (Andersen et al. (2006), p.842) However, some find this feature too restrictive (Franses and Van Dijk (2000)).

The specification above can be also written as Hadamard products, or element-by-element multiplication, as:

$$\Omega_{t|t-1} = C + A \circ (e_{t-1}e'_{t-1}) + B \circ \Omega_{t-1|t-2} \tag{2.74}$$

where C, A, and B now denote symmetric positive definite matrices of dimension $N \times N$. Now the number of free parameters is $\frac{3(N^2+N)}{2}$ and this model produces positive definite forecasts of the covariance matrix if C is positive definite and A and B positive semi-definite (see Attanasio (1991); Franses and Van Dijk (2000); Sheppard (2016)). It is a condition which can easily be tested. Moreover, estimation is becoming simpler, since it now it is possible to estimate each equation separately. Although, as any model comes with its drawbacks which in this case is that number of parameters that need to be estimated is still big in large dimensions and restrictions are still needed to enforce the positivity constraint (Sheppard (2016)). Therefore it has been suggested restricting all elements in A and B matrices so that are the same and this can be achieved by imposing the matrices to be scalars.

$$\Omega_{t|t-1} = C + \alpha(e_{t-1}e'_{t-1}) + \beta\Omega_{t-1|t-2} \tag{2.75}$$

That is the scalar MGARCH. It has some similarities to the RiskMetrics exponential smoother under the right conditions (Andersen et al. (2006)). Further, Ledoit and Wolf (2003) suggested to estimate separate bivariate models for all of the possible pairwise combinations of the N elements in Y_t before inserting into the full-dimensional model ensuring however, that the resulting $N \times N$ matrices in are positive definite (Andersen et al. (2006), p.844). Finally, to ensure positive definiteness, it could also be assumed that the autoregressive part of the model is zero.

Further, it is worth mentioning that the multivariate extension of the EGARCH model which was developed by Kawakatsu (2006) offers the additional advantage of being positive definite by default which is guaranteed by taking the logarithmic of the conditional covariance thus, there is no need for redundant parameter constraints (Silvennoinen and Teräsvirta (2009b), p.206-207). The model is defined as:

$$vech(ln\Omega_t - C) = \sum_{i=1}^{q} A_i \eta_{t-i} + \sum_{i=1}^{q} F_i(|\eta_{t-i}| - E|\eta_{t-i}|) + \sum_{i=1}^{p} B_i vech(lnH_{t-i} - C) \tag{2.76}$$

where C is a symmetric N x N matrix. Finally, multivariate extension of the regime-switching models exists including the SETAR, the LSTAR, multivariate MSW (Markov-process), threshold error-correction model (TECM), for more details see Franses and Van Dijk (2000) and the references within.

BEKK

Despite all the efforts to make the VEC model more feasible, it is possible to get a negative define covariance matrix. So, Engle and Kroner (1995) purposed the so called BEKK (Baba, Engle, Kraft and Kroner) model which assigns appropriate constraints on the parameters. The main merit of this specification is that it is intrinsically positive define since quadratic forms are positive semi-definite and so the sum of a positive semi-definite matrix and a positive definite matrix is always positive definite (Sheppard (2016); Silvennoinen and Teräsvirta (2009b)).

$$\Omega_t = CC' + A(e_{t-1}e'_{t-1})A' + B\Omega_{t-1}B' \tag{2.77}$$

where C, A and B are $(N \times N)$ matrices, with C symmetric and positive definite (Franses and Van Dijk (2000), p.202). A very interesting feature is the relationship among the BEKK and the VEC models. Each BEKK model implies a unique VEC model by applying a VEC operator. Correspondingly, each VEC model which is related to a BEKK representation can generate positive definite conditional covariance matrices (Franses and Van Dijk (2000); Silvennoinen and Teräsvirta (2009b)). Engle and Kroner (1995) exposed this preposition and set the conditions that need to hold for the two models to be equal. "Using the VEC operator, the BEKK can be seen as a restricted version of the VEC specification where $A \otimes A$ and $B \otimes B$ control the response to recent news and the smoothing, respectively" (Sheppard (2016), p.537).

Nevertheless, empirically the model is feasible only for five or less assets. In a larger scale the model requires further restrictions so diagonalization could also be applied on the A and B matrices (Sheppard (2016)). Further, the diagonal and the scaler BEKK have been been proposed. The latter was introduced by Ding and Engle (2001) and imposes the constraint that all assets have the same parameter matrix (Sheppard (2016)). The scalar BEKK offers the possibility to be covariance targeted easily. Covariance targeting replaces the intercept (CC') with a consistent estimator, leaving only the other 2 parameters to be estimated. "Scalar BEKK models can be used in large portfolios ($k > 50$), unlike models without targeting" (see Sheppard (2016), p.538). Pedersen and Rahbek (2014) propose the variance-targeting approach for the BEKK model. Targeting can be applied to the standard, to

the diagonal or even to the scalar BEKK specification. (for more see (Caporin and McAleer (2012a)). Finally, a modification of the model that facilitates with the parameters interpretation has been suggested, see also Silvennoinen and Teräsvirta (2009b). The representation of an asymmetric BEKK is given as:

$$\Omega_t = CC' + A\epsilon_{t-1}\epsilon'_{t-1}A' + G\zeta_{t-1}\zeta'_{t-1}G' + B\Omega_{t-1}B' \tag{2.78}$$

where G is a k x k matrix of parameters which control the covariance response the "bad" news, and when $k = 1$ this model reduces to a GJR-GARCH(1,1,1) model for variance (Sheppard (2016), p.542-543).

Matrix GARCH

Matrix GARCH of Ding and Engle (2001) can be viewed as combination of the diagonal VEC and the diagonal BEKK (Sheppard (2016)). Its representation is:

$$\Omega_t = CC' + AA' \odot \epsilon_{t-1}\epsilon_{t-1} + BB' \odot \Omega_{t-1} \tag{2.79}$$

where A, B and C matrices are lower triangular matrices (Sheppard (2016)). The condition that determines whether the positivity constraint is satisfied, it is to have a full rank matrix C. Ding and Engle (2001) also consider a diagonal and a scalar version of the model in their paper. The later is the same as the scalar BEKK model (Sheppard (2016)). Asymmetric dynamics can be introduced into the model, in a similar fashion as in the BEKK model. However, Sheppard mentioned "the asymmetric term has a slightly different interpretation for covariance since it is only relevant when both indicators are one which only occurs if both market experience "bad" news (negative returns)" (Sheppard (2016), p.542-543).

Constant Conditional Correlation (CCC) GARCH

Beyond the multivariate GARCH and its expansion and generalizations, a prominent model for large scale dynamic covariance matrix modeling and forecasting is the Constant Conditional Correlation (CCC) model of Bollerslev (1990) that is also designed to simplify the original model and offer a solution to the curse of dimensionality. The model's value stems from the fact that it can decompose the conditional covariance into conditional standard deviations and correlations and assumes that the conditional correlations remain constant (Sheppard (2016); Sheppard (2016); Silvennoinen and Teräsvirta (2009b); Franses and Van Dijk (2000)). CCC differs from the other models foreasmuch as modeling individual variances instead of the conditional covariance matrix. For $D_{t|t-1}$ denoting the $N \times N$ diagonal matrix with the conditional standard deviations, the CCC model is given as:

$$\Omega_{t|t-1} = D_{t|t-1}\Gamma D_{t|t-1} \tag{2.80}$$

where $\Gamma_{t|t-1}$ denotes the $N \times N$ matrix of conditional correlations (Andersen et al. (2006), p.845; Sheppard (2016), p.540). The conditional variances are usually represented by a GARCH model so the model may also be encountered as CCC-GARCH (Sheppard (2016)). Other specifications of the model include the CCC-GJR of Amado and Teräsvirta (2014) which uses the GJR model for the conditional variances and the Constant Conditional Correlation Smooth-Transition (CCCST) GARCH model where the parameters are also allowed to smoothly change depending on the estimated values of volatility on the previous period. Finally, Silvennoinen and Teräsvirta (2015) proposed the smooth transition (STCC) model.

The main feature of the CCC model is that the conditional correlation matrix is static (time-invariant). It may be seen as an oversimplified assumption but it has the advantage of restricting the number of parameters that need to be estimated. In that way the model can be estimated following the two-step method. "In the first step N individual univariate GARCH models are estimated for each of the series in Y_t, resulting in an estimate for the diagonal matrix, $\hat{D}_{t|t-1}$. Then defining the $N \times 1$ vector of standardized residuals for each of the univariate series, $\hat{\epsilon}_t \equiv \hat{D}^{-1}_{t|t-1}(Y_t - \hat{M}_{t|t-1})$" the elements in Γ may simply be estimated by the corresponding sample analogue (Andersen et al. (2006), p.845). That is:

$$\hat{\Gamma} = \frac{1}{T}\sum_{t=1}^{T} \hat{\epsilon}_t \hat{\epsilon}'_t \tag{2.81}$$

For D_t, it can be assumed that $D_t = S_t G_t$ where $S_t = diag(h_{1t}^{1/2}, ..., h_{Nt}^{1/2})$, $h_{it}^{11/2}$ are the conditional standard deviations and $G_t = diag(g_{1t}^{1/2}, ..., g_{Nt}^{1/2})$, $G_{it}^{1/2}$ are positive-valued deterministic functions of rescaled time (Amado and Teräsvirta (2014)). It follows that:

$$\sigma_{ijt} = p_{ijt}(h_{it}g_{it})^{1/2}(h_{jt}g_{jt})^{1/2}, i \neq j \tag{2.82}$$

$$\sigma_{iit} = h_{it}g_{it}, i = 1, ..., N \tag{2.83}$$

If $G_t \equiv I_N$, and $\Gamma = \Gamma_t$ the model belongs to the family of CCC-GARCH models (Amado and Teräsvirta (2014)). Within the context of the CCC model the positive definiteness is guaranteed as long as the the individual variances are positive. Jeantheau (1998) extended the model in terms of the assumptions about the matrices in the conditional variances. In the CCC they are assumed to be diagonal, but by relaxing this condition it is possible to include ARCH and GARCH effects in each series from all the other series (cross-market) (Silvennoinen and Teräsvirta (2009b)). The resultant model is known as Extended CCC– (ECCC–) GARCH model.

DCC GARCH

Even though the CCC model has a lot of desirable features empirically it was found to be an oversimplified version of reality as it assumes constant conditional correlations. To deal with this limitation Engle (2002a) proposed the Dynamic Conditional Correlation (DCC) model which is also grounded on decomposition, to model directly any temporal dependencies in the conditional correlations. DCC can be viewed as an extension of the CCC model. Both CCC and DCC are easy to estimate even with many assets. The conditional correlations are represented through a scalar BEKK-like model though, noticeably Γ is time dependent (Andersen et al. (2006); Sheppard (2016); Silvennoinen and Teräsvirta (2009b)).

$$\Omega_{t|t-1} = D_{t|t-1}\Gamma_t D_{t|t-1} \tag{2.84}$$

$$\Gamma_t = Q_t^{*-1} Q_t Q_t^{*-1} \tag{2.85}$$

$$Q_{t|t-1} = (1-\alpha-\beta)\hat{\Gamma} + \alpha(\hat{\epsilon}_{t-1}\hat{\epsilon}'_{t-1}) + \beta Q_{t-1|t-2} \tag{2.86}$$

To ensure positive definiteness in DCC model the same conditions, as the one presented for the CCC model, must be satisfied plus that the conditional correlation matrix is positive definite for any t. DCC model guarantees that the conditional covariance matrix will be positive definite yet, to achieve the same for the conditional correlation matrix the individual elements of the covariance matrix need to be standardized (Andersen et al. (2006)). For a GJR specification for the individual conditional variances see Amado and Teräsvirta (2014). Alternatively, another specification for the conditional correlations is an EGARCH or a Varying Correlation (VC-) GARCH model proposed by Tse and Tsui (2002) which models the conditional correlations through its lagged conditional correlations and a set of estimated correlations so that:

$$\Gamma_t = (1-\alpha-b)S + \alpha S_{t-1} + b\Gamma_{t-1} \tag{2.87}$$

where S is a constant, positive definite parameter matrix with ones on the diagonal, α and b are non-negative scalar parameters such that $\alpha + b \leq 1$, and S_{t-1} is a sample correlation matrix of the past standardized residual (Silvennoinen and Teräsvirta (2009a)). The definition of the intercept, $1-\alpha-b$, corresponds to the idea of 'variance targeting' in Engle and Mezrich (1996) (Silvennoinen and Teräsvirta (2009b), p.212; Amado and Teräsvirta (2014); Caporin and McAleer (2012a)). Kwan et al. (2005) allowed for regimes switches into the VC GARCH model. Further specifications of the

model are the Smooth Transition Conditional Correlation (STCC–) GARCH model of of Silvennoinen and Teräsvirta (2005), the Double Smooth Transition Conditional Correlation (DSTCC–) GARCH model by Silvennoinen and Teräsvirta (2009a), the wide class of Volatility Dependent Conditional Correlation (VDCC) models by Bauwens and Otranto (2016) and the Regime Switching Dynamic Correlation (RSDC–) GARCH model introduced by Pelletier (2006). What is more, similarly to the univariate case exogenous variables may be added as well. "This variable may be either an observable variable, a combination of observable variables, or a latent variable that represents factors that are difficult to quantify" (Silvennoinen and Teräsvirta (2009b), p.212). Finally, the Cappiello et al. (2006), extended the DCC–GARCH model following the recent stream in literature that indicated the addition of asymmetric effects forming the Asymmetric Generalized DCC (AG–DCC) (Andersen et al. (2006)). Finally, extensions are available that allow for more flexibility and for regime switches.

Despite its seeming dominance, the DCC has a quite a few parameters that need to be estimated and that convey complications. First of all, the estimation is a three-step procedure (Sheppard (2016), p.541). Secondly, there is need for further parameter restriction by utilizing the appropriate methods. Billio and Caporin (2009) utilized a BEKK specification for the conditional correlations developing in that way the Quadratic Flexible DCC (GFDCC) GARCH model(for more see Silvennoinen and Teräsvirta (2009b), p.212). Alternatively, the scalar version or the variance targeting can be used. Caporin and McAleer (2012a) highlighted that the Scalar BEKK and the scalar DCC specifications are almost identical. However, Aielli (2013) claimed that since the sample correlation is an inconsistent estimator of the covariance matrix, the variance targeting is not suitable anymore (Caporin and McAleer (2012a)). Therefore Aielli (2013) suggested the corrected or consistent DCC (cDCC) model.

Other specifications of the model include, the Quadratic Flexible DCC of Billio et al. (2006) the Generalized DCC suggested by Cappiello et al. (2006), the Time-Varying Correlation (TVC) of Tse and Tsui (2002) or the Hadamard product representation of Hafner and Franses (2009), the Non Linear AutoRegressive Correlation (NLARC) and the Flexible RSDC (FRSDC) models (see Bauwens and Otranto (2020)), the Sequential Conditional Correlations (SCC) decomposition of Palandri (2009), the asymmetric DCC model as proposed in Cappiello et al. (2006) (Samkharadze (2013), Sheppard (2016)). Finally, Aielli and Caporin (2015) proposed the Dynamic Principal Component (DPC) GARCH model which through orthonormal dynamic eigenvectors adds flexibility.

Multivariate Components models: DCC-MIDAS

The incentive for using components models has already been highlighted in the first part of the chapter when MIDAS approach was introduced. Components models with their ability to decompose short and long memory dynamics are attractive in the multivariate framework as well. Dynamic correlations

are replaced with component specifications so that long-run correlation component can be extracted through mixed data sampling and the resultant model was proposed by Colacito et al. (2011) (Engle et al. (2008)). Audrino (2014) has also discussed DCC-MIDAS and CCC-MIDAS specifications. The short term is represented through a GARCH model using daily (squared) returns, and it is assumed that it fluctuates around a long-run component which is described by the RV using monthly or quarterly returns. That is possible by the MIDAS weighting scheme (Colacito et al. (2011)). Likewise, "short-lived effects to correlations will be captured by the autoregressive dynamic structure of DCC, with the intercept of the latter being a slowly moving process that reflects the fundamental or long-run causes of time variation in correlation" (Colacito et al. (2011), p.45).

Equicorrelation

Focusing on the limitations of the DCC model, caused by the number of parameters that need to be estimated, the literature has also suggested an alternative way to impose restrictions. That is is known as Equicorrelation and it is defined as $p\sigma_i\sigma_j$ where the σ is the volatility of the individual series. "This estimator is clearly mis-specified for any $k > 2$, and is generally only appropriate for assets where the majority of the pairwise correlations are positive and the correlations are fairly homogeneous" (Sheppard (2016), p.532). On the same track Elton and Gruber (1973) proposed a model where every pair of assets has the same correlation. In this category also falls the Dynamic Equicorrelation (DECO) model proposed by Engle and Kelly (2012).

Dynamic Conditional Beta

Further, Engle (2016) motivated by the work of Bollerslev et al. (1988) proposed a new model called Dynamic Conditional Beta (DCB) to obtain time-varying slope coefficients. The model is based on extracting the time-dependent parameters from the estimated conditional covariance (Darolles et al. (2018)). The conditional betas are superior than the conditional correlations in view of the fact that they are not restricted between the minus and plus one. Finally, Darolles et al. (2018) proposed a novel approach "based on a natural orthogonalization of the observed time series, initially proposed by Pourahmadi (1999) in the context of longitudinal data" (Darolles et al. (2018), p.223-224).

2.2.1 Multivariate Factor models

Multivariate factor GARCH and stochastic volatility models were first analyzed by Diebold and Nerlove (1989) and Engle et al. (1990). "Factor models are motivated by economic theory" (Silvennoinen and Teräsvirta (2009b), p.207). Various assets tend to move together as a reaction to a new information. Therefore, identifying common underlying factors in the series under study that

cause this behavior is crucial for volatility estimation and forecasting. Factor models are successful at decreasing the number of parameters as long as the number of factors is relative smaller than the number of assets (Silvennoinen and Teräsvirta (2009b)).

Engle et al. (1990) were the first to analyze multivariate factor GARCH models. They assumed that the conditional variance is modeled through K number of factors, not necessarily uncorrelated. Factors can be observable or unobservable.

$$\Omega_t = C + \sum_{k=1}^{K} w_k w_k' f_{k,t} \tag{2.88}$$

where C is an $N \times N$ positive semi-definite matrix, w_k, $k = 1, ..., K$, are linearly independent $N \times 1$ vectors of factor weights, and the $f_{k,t}$ are the factors which follow a first-order GARCH (Silvennoinen and Teräsvirta (2009b), p.207). The model can be estimated applying the two-step procedure. Running along side Diebold and Nerlove (1989) were also contributed to the literature by proposing a factor model which however was a stochastic volatility model rather than a GARCH one (Silvennoinen and Teräsvirta (2009b)).

However, imposing uncorrelated factors is more appropriate for financial data. An additional gain from adopting this assumption is that it is possible to describe the dynamics of the series using less number of factors. The intuition behind this argument is that if the engaged factors are are uncorrelated then they are truly different components that drive the returns. Within this framework, the original observed series contained in r_t are assumed to be linked to unobserved, uncorrelated variables, or factors, z_t, which usually follow a GARCH process, through a linear, invertible transformation W: $r_t = W z_t$. There various specifications proposed for the W $(N \times N)$ matrix. One simple case is to consider an $r - factor$ multivariate GARCH model:

$$\epsilon_t = B f_t + u_t \tag{2.89}$$

where B is a $k \times r$ matrix of full-column rank of so-called factor loadings, f_t is a $r \times 1$ vector containing the common factors and u_t is a $k \times 1$ vector of idiosyncratic noise. It is assumed that f_t and u_t have conditional mean zero and conditional variance matrices given by Λ_t and Σ_t, respectively. Both Λ_t and Σ_t are diagonal, while it is also common to assume that Σ_t is constant over time. Finally, f_t and u_t are assumed to be uncorrelated. The preceding assumptions imply that the conditional covariance matrix of ϵ_t is given by (Franses and Van Dijk (2000), p.203): :

$$\Omega_t = B \Lambda_t B' + \Sigma \tag{2.90}$$

The above model exists with or without a constant term. A famous illustration is the CAPM model

where a portfolio's return is represented by the following regression model

$$r_t^P = \alpha + \beta R_t + \epsilon_t \tag{2.91}$$

In that case the *beta* coefficient is given as $\beta = \dfrac{Cov(r_t^P, R_t)}{Var(R_t)}$. Extensions of the basic models have been forward so the return of the portfolio depends on multiple risk factors like:

$$r_t^P = \alpha + \beta_1 R_{1t} + .. + \beta_n R_{nt} + \epsilon t \tag{2.92}$$

where R_i with $i = 1, .., n$. The betas assigned to each individual risk factor can be derived from the covariances between the portfolio return and the risk factors and the risk factors' variances. That represents the simplest type of factor models. In the literature it has been suggested the idea of sparse factor model that is based on the assumption of a few weak factors that are allowed to affect just a subset of the available data. It is worth mentioning that the factors considered could be the market model which is basically the single-index covariance matrix of Sharpe (1963), or a common ARCH-Feature or industry factors (for example as defined by Fama and French (1997)), or other factors that are not industry-based (Engle and Susmel (1993a)). Nevertheless, factors can be conditionally heteroskedastic, with individual conditional variances denoted by:

$$\sigma^2_{t|t-1} \equiv Var(f_t|F_{t-1}) \tag{2.93}$$

Under this case the conditional covariance matrix for R_t takes the form (Andersen et al. (2006), p.845):

$$\Omega_{t|t-1} \equiv Var(R_t|F_{t-1}) = bb'\sigma^2_{t|t-1} + \Lambda \tag{2.94}$$

The main gain achieved is that the conditional covariance is inherently positive definite. Another advantage is that multi-period ahead forecasts are easy to be calculated. However, forecasts are easily constructed as long as the factors are observable conditional at the information set. For latent or unobservable observations/factors the procedure could be somewhat more complicated.

Factor models have been used to transform models to accommodate many assets as in the case of the Markov-Switching Multifractal model by Calvet et al. (2006). Alternatively, the Dynamic Factor model was introduced by Aguilar and West (2000) and Christodoulakis and Satchell (2000) to accomplish better forecasting performance. These models are different in the way that the factor loadings matrix is time-varying. Finally, factor models have been extend to component models with the Pure Variance Common Features model which "(...) decomposes the conditional variance into a short-run idiosyncratic component (a low-order ARCH process) and a long-run component (the variance factors)" (Engle and Marcucci (2006), p.37).

Orthogonal GARCH (OGARCH) model

In the same category lies the Orthogonal GARCH (OGARCH) model of Alexander and Chibumba (1997). "Orthogonal GARCH uses the first p principle components to model the conditional covariance (...)" (Sheppard (2016), p.541). In OGARCH model the factors are unobservable.

$$\Omega_t = \beta \Omega_t^f \beta' + \Sigma \tag{2.95}$$

where β is the k x p matrix of factor loadings corresponding to the p factors with the highest total R^2, Σ is the conditional covariance of the residuals and the conditional covariance of the factors (Σ_t^f) is assumed diagonal (Sheppard (2016), p.541). Van der Weide (2002) extended the orthogonal GARCH model in a way so that Σ is not required to be orthogonal, only invertible, and in that model the uncorrelated factors are standardized to have unit unconditional variances (Silvennoinen and Teräsvirta (2009b)).

Various specifications of the conditional variances are possible in both models, O-GARCH and its generalization (GO– GARCH). However, unlike the simple factor models presented above, orthogonal models are not genuine associated with less factors than assets. In the O-GARCH model that is feasible but it leads to reduced rank conditional covariance matrix but for GO-GARCH that is impossible (Silvennoinen and Teräsvirta (2009b)).

What is more,ther suggestions are given by, Vrontos et al. (2003) and their Full Factor (FF-) GARCH, Lanne and Saikkonen (2007) and their Generalized Orthogonal Factor (GOF–) GARCH model. GOF models can be seen as a combination of factor models orthogonal models as they are easier to estimate than factor models and have less heteroskedastic factors than orthogonal factors (Silvennoinen and Teräsvirta (2009b), p.209-210). Finally, other specifications include the O-EWMA specification which is an alternative to the O-GARCH model. In addition, just as in the simple factor models, it is possible to consider only a reduced factor structure.

Principal Component Analysis (PCA)

An alternative approach to reduce the dimension on the conditional covariance or correlation matrix is the principal components. This approach was motivated by the statistical Principal Component Analysis (PCA) which aims at decomposing a matrix into orthogonal factors and factor loadings (see Sheppard (2016), p.529). Thus using this method is possible to extract factors within the matrix. The main limitations of this approach is that is a statistical technique which aims at reconstructing the data so that few components can explain a significant part of the risk, as a result, these latent

factors are lacking economic interpretation (Ledoit and Wolf (2003)).

Several variations have been prosed such as the Principal Volatility Components (PVC) of Hu and Tsay (2014) and Li et al. (2016) where "this methodology produces two types of components. The first type corresponds to components with conditional covariance matrix evolving over time whilst the other type corresponds to components with constant conditional covariance matrix" (Trucíos et al. (2019), p.202). Further, conditionally uncorrelated components (CUC) can be used which is similar to the Independent Component Analysis (ICA) of Back and Weigend (1997) but it relies only on the assumption hat the component series are conditionally uncorrelated (Fan et al. (2012)). In the ICA the components are statistically independent (Wu and Philip (2005)).

Factor model come in many forms and shapes and Santos and Moura (2014) provide a fair synoptic overview of the literature. In more details, factors can be directly derived from the asset returns like in Chan et al. (1999) or by PCA as in Alexander (2001). Engle (1990b), Alexander (2001) and Vrontos et al. (2003) consider GARCH models for the factors, whereas Aguilar and West (2000) and Han (2006) assume SV models (Santos and Moura (2014)). Moreover, Van der Weide (2002) assumes not mutually orthogonal components while Alexander (2001) work is based orthogonal matrix of eigenvectors.

2.2.2 Realized Covariance

High-frequency data are having a strong impact on multivariate models as well. The framework is that prices follow the $N - variate$ continuous time diffusion:

$$dP_t = M(t)dt + \Sigma(t)dW(t), t \in [0, T] \tag{2.96}$$

where $M(t)$ denotes the $N \times 1$ instantaneous drifts, $\Sigma(t)$ refer to the $N \times N$ instantaneous diffusion matrix, and $W(t)$ now denotes an $N \times 1$-dimensional vector of independent standard Brownian motions. (Andersen et al. (2006); Sheppard (2016)). Intuitively, "(...) for small values of $\Delta > 0$,

$$R(t, \Delta) \equiv P(t) - P(t-1) \simeq M(t-\Delta)\Delta + \Sigma(t-\Delta)\Delta W(t) \tag{2.97}$$

where $\Delta W(t) \equiv W(t) - W(t-\Delta) \sim N(0, \Delta I_N)$" (Andersen et al. (2006), p.831). As $\Delta \to 0$ and under the assumption that price process does not contain any jumps the realized covariance is an estimate of the integrated covariance over some period, for instance one day using intra daily returns (Brito et al. (2018); Andersen et al. (2006)).

$$RCOV(t, \Delta) \to \int_{t-1}^{t} \Sigma(s)\Sigma(s)'ds \equiv ICOV(t) \tag{2.98}$$

The realized covariance is then formally defined as (Andersen et al. (2006)):

$$RCOV(t,\Delta) = \sum_{j=1}^{1/\Delta} R(t-1+j\cdot\Delta,\Delta)R(t-1+j\cdot\Delta,\Delta)' \tag{2.99}$$

A common problem encountered when working with high-frequency returns on multiple time series is the noisy and a-synchronized data (Sheppard (2016)). Two methods have been mainly proposed in order to manage a-synchronized data pairwise-refresh and all-refresh. So to overcome these problems is important to identify the optimal sample frequency. As the $\Delta \to 0$, the multivariate realized volatility errors are decreasing (Barndorff-Nielsen and Shephard (2004); Andersen et al. (2006); Sheppard (2016)). Empirically it has been found that the best sample frequency is every 5-minutes. Alternatively, a modified realized covariance estimator which uses sub sampling can be used. This allows to exploit all the data on the available frequency not just non overlapping ones. The resultant realized covariance estimator is:

$$RC_t^{ss} = \frac{m}{n(m-n+1)} \sum_{t=1}^{m-n+1} \sum_{j=1}^{n} r_{i+j-1,t} r'_{i+j-1,t} = \frac{1}{n} \sum_{j=1}^{n} \widetilde{RC}_t \tag{2.100}$$

where there are m high-frequency returns available and the selected sampling sampling time is based on n returns (Sheppard (2016)). Finally, the counterpart of realized kernel for the multivariate case has been proposed by Barndorff-Nielsen et al. (2011). It is a weighted version of the realized covariance that relies on a Newey and West (1987) covariance estimator (Sheppard (2016)). Finally, Martens and Poon (2001) proposed a general model which accounts for the problem of a-synchronicity among the financial markets:

$$r_t = \mu + \epsilon_t + M\epsilon_{t-1}, \epsilon_t \sim N(0, H_t) \tag{2.101}$$

$$h_{iit} = \theta_{iit} \tag{2.102}$$

$$h_{ijt} = p_{ijt}\sqrt{h_{iit}}\sqrt{h_{jjt}} + \phi_{ij}\theta_{ijt} \tag{2.103}$$

$$\theta ijt = \omega + b_i' H_{t-1} b_j + \alpha_i' \epsilon_{t-1}\epsilon_{t-1}'\alpha_j + g_i'\eta_{t-1}\eta_{t-1}'g_j \tag{2.104}$$

A-synchronicity is controlled through the M matrix. An additional advantage of this model is that it adds flexibility since it includes asymmetries to both the spillover effect and the last term in θ_{ijt}. This model nests the following specifications the restricted VECH of Bollerslev et al. (1988), the CCC model of Bollerslev (1990), the BEKK model of Engle and Kroner (1995) and the Factor ARCH of Engle et al. (1990) (Poon (2005)).

Realized covariances have a dual purpose; they can be used as a proxy of the true unobserved covariance when evaluating forecasted valued derived from relevant models whereas its past values are also used directly in a model to forecast the conditional covariance matrix. For the purposes of estimation the positive definiteness of the realized covariance is guaranteed however, this might not hold when future values need to be forecasted thus the various matrix restriction need to be introduced again. Realized covariance and realized kernels can be incorporated in the basic multivariate GARCH models presented on the previous section. This is possible under the assumption that "the mean of the realized covariance, conditional on the time $t-1$ information, is Ω_t" (Sheppard (2016), p.553). Then treating the Realized Covariance as a "(...) direct observation (with uncorrelated measurement errors) on the unique elements in the covariation matrix of interest (...)" (Andersen et al. (2006), p.850). For instance the BEKK model presented on the previous section can be modified as:

$$\Omega_t = CC' + ARC_{t-1}A' + B\Omega_{t-1}B' \tag{2.105}$$

Alternatively, Fleming et al. (2003) proposed a weighted average of past valued of realized covariance and of conditional covariances.

$$\Omega_t = \exp(-\alpha)\Omega_{t-1} + \alpha \exp(-\alpha)RCV_{t-1} \tag{2.106}$$

It is based on a rolling scheme which makes the model parsimonious and it guarantees that the conditional covariance matrix will be positive definite as long as $\alpha > 0$. Further, a model which has attracted some attention is the Conditional Autoregressive Wishart (CAW) model of Golosnoy et al. (2012) which assumes that: $RCV_t|I^{t-1} \sim W_k(v, S_t/v)$ where v is the degrees of freedom and the S_t/v is the positive definite matrix. It is motivated by the Wishart Autoregressive (WAR) model of Gouriéroux et al. (2009) which however was proven to be very hard to estimate because the high number of parameters involved in large dimensions. The CAW model however manages to reduce the number of parameters to be estimated through its structure where the conditional covariance depends directly on the realized covariance. The diagonal CAW with covariance targeting specification reduces further the number of parameters that need to be estimated. Finally, more specifications of the model include a BEKK structure for the model, the Realized DCC (Re-DCC) CAW model of Bauwens et al. (2012b), the VHAR model, the O-CAW (orthogonal) specification, the diagonal d-CAW HAR, and the Dynamic Principal Component CAW (DPC-CAW) model for Realized covariances which is an adaptation of the DPC-GARCH model of Aielli and Caporin (2015) which allows for long memory dynamics by construction.

Realized Covariance matrix and factor analysis

A few researchers have worked on decomposing the realized covariance matrix to facilitate its estimation and forecasting. Brito et al. (2018) describe a process where the matrix is primarily decomposed into a factor covariance matrix and residual covariance matrix. With the following decomposition it is easier to forecast each input separately and plug them all in the final equation. Let Ω_t denote the realized covariance matrix of returns at time t, that is, $\Omega_t = cov(r_t)$ and relying on the assumption that $E(\epsilon_t|f_t) = 0$, the resulting model is:

$$\Omega_t = cov(B_t' f_t) + cov(\epsilon_t) = B_t' \Sigma_{f,t} B_t + \Sigma_{\epsilon,t} \tag{2.107}$$

where B_t is a $K \times N$ matrix of loadings of N assets on K factors,$\Sigma_{f,t}$ is the $K \times K$ factor covariance matrix and $\Sigma_{\epsilon,t}$ is the $N \times N$ residual covariance matrix all in time t.

$$\Sigma_{f,t} = cov(f_t) = cov(W' r_t) = W' \Omega_t W \tag{2.108}$$

Realized Betas

Starting from the previous decomposition, assume that the corresponding realized covariance matrix is:

$$RC = \begin{bmatrix} RV_t & RC_{fi}' \\ RC_{fti} & RC_{f,f} \end{bmatrix} \tag{2.109}$$

where $RV_{i,i}$ is the realized variance of the asset being studied, RC_{if} is the k by 1 vector of realized covariance between the asset and the factors, and RC_{ff} is the k by k covariance of the factors. In this case the realized beta is defined as (Sheppard (2016), p.552):

$$R\beta = RC_{ff}^{-1} RC_{fi} \tag{2.110}$$

As Sheppard (2016) mentioned "realized Betas are similar to other realized measures in that they are model free and, as long as prices can be sampled frequently and have little market microstructure noise, is an accurate measure of the current exposure to changes in the market" (Sheppard (2016), p.552).

Realized Correlation

Realized Correlation can be calculated in a straightforward way from the realized covariance:

$$RCorr = \frac{RC_{ij}}{\sqrt{RC_{ii} RC_{jj}}} \tag{2.111}$$

where RC_{ij} is the realized covariance between assets i and j and RC_{ii} and RC_{jj} are the realized variances of assets i and j, respectively (Sheppard (2016)). See also Andersen et al. (2001a) and Andersen et al. (2001b). To ensure that the correlations will fit into the [-1, 1] interval some transformations might be necessary such as the Fisher transformation (Andersen et al. (2006)).

2.2.3 Multivariate Stochastic Volatility

Following Chib et al. (2009) Multivariate Stochastic Volatility (MSV) models fall into three broad categories the classic MSV models which are the counterpart of the univariate ones, the factor MSV models and the dynamic correlation ones. The basic framework is consistent with the univariate SV structure.

$$y_t = exp(h_t/2)\epsilon_t,\ t = 1, ..., n \tag{2.112}$$

$$h_{t+1} = \mu + \phi(h_t - \mu) + \eta_t,\ t = 1, ..., n-1 \tag{2.113}$$

$$h_1 \sim N(\mu, \sigma_\eta^2/(1-\phi^2)),\ \begin{pmatrix}\epsilon_t \\ \eta_t\end{pmatrix} | h_t \sim N_2(\mathbf{0}, \mathbf{\Sigma}),\ \mathbf{\Sigma} = \begin{pmatrix}1 & 0 \\ 0 & \sigma_\eta^2\end{pmatrix} \tag{2.114}$$

where y_t is a univariate outcome, h_t is a univariate latent variable and $N(\mu, \sigma^2)$ and $N_m(\mu, \mathbf{\Sigma})$ denote respectively a univariate Normal distribution with mean μ and variance σ^2, and an m-variate Normal distribution with mean vector μ and variance-covariance matrix $\Sigma(...)$. The SV model is thus a state-space model, with a linear evolution of the state variable h_t but with a nonlinear measurement equation (because h_t enters the outcome model nonlinearly) (Ghysels et al. (1996); Broto and Ruiz (2004); Shephard (2005) in Chib et al. (2009), p.366). Estimation of this type of models is complex thus, various methods have been proposed such as Bayesian Markov chain Monte Carlo methods or non-parametric ones. The basic MSV is described by the following series of equations:
$y_t = V_t^{1/2}\epsilon_t, V_t^{1/2} = diag(\exp(h_{1t}/2), ..., \exp(h_{1pt}/2)), t = 1, ..., n$ with the additional assumptions that $\begin{pmatrix}\epsilon_t \\ \eta_t\end{pmatrix} | h_t \sim N_{2p}(0, \Sigma), \Sigma = \begin{pmatrix}\Sigma_{\epsilon\epsilon} & 0 \\ 0 & \Sigma_{\eta\eta}\end{pmatrix}$ and $\Sigma_{\epsilon\epsilon}$ is a matrix in correlation (with units in the main diagonal). The individual log volatilities are assumed to follow an AR(1) process :

$$h_{t-1} = \mu + \Phi(h_t - \mu) + \eta_t,\ t = 1, ..., n-1 \tag{2.115}$$

Φ and $\Sigma_{\eta\eta}$ matrices might be assumed to be diagonal so that the log volatilities will be conditionally independent reducing the final number of parameters that needs to be estimated as in Harvey et al. (1994). Whereas So et al. (1997) relaxed the assumption that the non diagonal elements of Φ are

equal zero (Chib et al. (2009)). For further reduction a Cholesky decomposition might be employed as well, resulting in the Cholesky MSV model of Chiriac and Voev (2011). Then depending on the specification given for the y_t and on the assumption about the underlying distribution various specifications can be developed (Chib et al. (2009)). Further contributions where made by Danielsson (1998), Ishihara et al. (2016), Chan et al. (2006) and Smith and Pitts (2006) (for more see Chib et al. (2009), p.373-374).

Factor Multivariate Stochastic Models

The need to derive time-varying conditional correlation from MSV motivated the development of Factor Multivariate Stochastic Models where factors are assumed to follow a SV process: $y_t = Bf_t + V_t^{1/2}\epsilon_t, V_t^{1/2} = diag(\exp(h_{1t}/2), ..., \exp(h_{pt}/2))$, where B is a $p \times q$ matrix called the loading matrix, and $f_t = (f_{1t}, ..., f_{qt})$ is a $q \times 1$ latent factor at time t (Chib et al. (2009)). Extensions of the model were provided by Harvey et al. (1994), Calvet et al. (2006) and Calvet and Fisher (2001). Another well-used model is the mean factor model. The model is described by the following set of equation:

$$y_t = Bf_t + V_t^{1/2}\epsilon_t \tag{2.116}$$

$$f_t = D_t^{1/2}\gamma_t \tag{2.117}$$

$$h_{t+1} = \mu + \Phi(h_t - \mu) + \eta_t \tag{2.118}$$

where

$$V_t = diag(\exp(h_{1t}), ..., \exp(h_{pt})) \tag{2.119}$$

$$D_t = diag(\exp(h_{p+1,t}), ..., \exp(h_{p+q,t})) \tag{2.120}$$

$$\Phi = diag(\phi_1, ..., \phi_{p+q}) \tag{2.121}$$

$$\Sigma_{\eta\eta} = diag(\sigma_{1,\eta\eta}, ..., \phi_{p+q,\eta\eta}) \tag{2.122}$$

and $h_t = (h_{1t}, ..., h_{pt}, h_{p+1,t}, ..., h_{p+q,t})$ (Chib et al. (2009)). This models allows for all the factors and the errors to follow univariate SV models (Chib et al. (2009)). Various extensions of the models are also available. Chib et al. (2006) suggested a factor MSV model which is flexible enough to include series-specific jumps $\forall t$, and heavy tailed innovations whereas, Asai and McAleer (2015) proposed a factor MSV model that can model realized covariance measures and can accommodate asymmetric and long memory effects. Kroner and Ng (1998) and Asai and McAleer (2015) also worked on how to include asymmetric effects.

2.2.4 Correlation Multivariate Stochastic Models

To model time-varying correlations, Correlation Multivariate Stochastic Models have been proposed. Yu and Meyer (2006) suggested a model, however this approach can be successfully applied for only one pair of assets. To speed up the estimation and the forecasting a Cholesky decomposition and a decomposition based on the matrix exponential were proposed. More specifications were proposed by Asai and McAleer (2007) and Gouriéroux et al. (2009).

2.2.5 Special Methods

Now it follows a remark of alternative methods that the literature has proposed in order to address the issue of dimensionality in large covariance matrices. First of all, a simplistic tool is to use the identity matrix so that the covariance matrix is a scalar multiple of the identity matrix. Similarly, Fan et al. (2011) combined a factor structure with sparsity of the variance matrix estimation based on bivariate pairs. Sparsity, means that (asymptotically) the zero elements in the matrix are estimated as zero, and it can be achieved by enforcing strong assumptions on the individual volatility series (for example, thresholds). However, this approach does not work well in very large dimensions and empirically it has been observed that heavily constrained models are experiencing higher specification error and may harder to be estimated (Caporin and McAleer (2012a); Briner and Connor (2008)). Finally, it is also possible to get covariance matrix estimates directly from univariate volatility estimates (Andersen et al. (2006)).

This review on multivariate volatility models and shrinkage of covariance matrices could not be complete without mentioning the shrinkage estimators and the Penalized Maximum Likelihood estimation. Shrinkage estimators suggested by Ledoit and Wolf (2004b), Ledoit and Wolf (2004a) and Ledoit and Wolf (2003), and effectively shrink the eigenvalues of the sample covariance matrix toward some fixed target. As targets could be used the identity matrix (Ledoit and Wolf (2004b)), a single factor model/shrinkage to the market (Ledoit and Wolf (2003)), and Equicorrelation (Ledoit and Wolf (2004a)) where the extreme values are shrunk to the center. To reduce the bias Hafner and Reznikova (2012) proposed to apply the shrinkage methods of Ledoit and Wolf to the DCC model or to the cDCC model (Ledoit and Wolf (2004b); Ledoit and Wolf (2004a); Ledoit and Wolf (2003)). "Intuitively, the role of the shrinkage parameter is to balance the estimation error coming from the ill-conditioned variance matrix and the specification error associated with the target matrix" (Abadir et al. (2014), p.165).

On the other hand, penalized maximum likelihood are applied on the factor loadings matrix. Callot et al. (2017) advocate the use of the Least Absolute Shrinkage and Selection Operator (LASSO) of Tibshirani (1996) and the adaptive LASSO of Zou (2006) to model the dynamics of the realized covariance. "LASSO minimizes residual sum of squares subject to the sum of the absolute value of

the coefficients than a constant (...)" (Tibshirani (1996), p.267). This approach gives coefficients that are easy to interpret but it still remains unfeasible for hundreds of assets without imposing some restrictions on the data before the factors.

Chapter 3

Evaluating univariate volatility forecasts

3.1 Introduction

In the previous chapter an extensive review of volatility models, was provided. All of those volatility models can be used to derive forecasts of future volatility. Others in a more straightforward way like GARCH-type models or others in a less direct one like stochastic volatility models. Gneiting (2011) highlighted, that although forecasts ought to be probabilistic in nature that aim to predict the likelihood of an outcome, they can be quite vague, and thus not suitable for decision-making process. This makes density forecast less common in the literature, compared to the more practical, singled-valued, point forecasts. Nevertheless, this chapter is focusing on point forecast leaving the probabilistic predictions for the last empirical chapter of this thesis. The various models and approaches can be used to produce forecast for one-step ahead or multiple-step ahead (Patton (2011); Fuertes et al. (2009a); Marcucci (2005)). More studies focus on 1-step ahead. Figlewski (1997) provided a review on multi-steps ahead forecast. Some studies focus to define ex-ante which forecast horizon will give the best results. Finally, when constructing a forecast there are two different schemes that can be used; a recursive or a rolling window (Stock and Watson (2002); Rapach and Strauss (2008)). Although, Clark and McCracken (2009) suggested to use a combination of the two schemes to achieve better results.

In economics and finance, there has been been a lot of research done on forecasts, especially of asset returns, of exchange rates and of macroeconomics variables. However, for volatility forecasts there are still many research questions open. That is because due to some unique characteristics of volatility. What makes volatility forecasting special is that true realized volatility is unobserved even after estimation. As a result, forecast evaluation is performed with respect to an ex-ante volatility estimator, a volatility proxy, and this leads to some form of distortion (Violante and Laurent (2012)). Or as

Patton (2011) warns, "it is not true that using a conditionally unbiased proxy will always lead to the same outcome as if the true latent variable were used (...)" (p.1).

The large amount of alternative forecasting models and methods advocates the need for tools to evaluate and rank the alternative specifications as for their forecasting power and performance in an attempt to find the optimal method. The importance of producing accurate and trustworthy forecasts of future volatility cannot be overstated. As Timmermann (2006) emphasized, "forecasts do not intrinsically have direct value to decision making. Rather, they become valuable only to the extent that they can be used to improve decision makers"(p.141). Forecast evaluation can be exercised for the in-sample period. However, the critical objective is to test the out-of-sample performance of the competitive methods. "A good forecasting model should be the one that can withstand the robustness of out-of-sample test-a-test design that is closer to reality" (Poon (2005), p. 30).

The aim of this chapter is to evaluate and rank alternative univariate volatility forecasting specifications in terms of their out-of-sample performance. Numerous approaches for volatility estimation and forecasting are considered from simple, non-parametric, to parametric GARCH models and Realized Volatility (RV) models. Including non-parametric approaches, which are commonly overlooked in volatility forecasting, to a big pool of models is our first contribution to the literature. Then the produced forecasts are combined and the extent to which combinations can improve the forecasting accuracy is evaluated. For the purposes of evaluation pairwise tests of equal predictive accuracy are performed. In line with the common practice this study focuses on comparing the forecast accuracy in terms of the forecast error. Following the relevant literature, since our model set is comprised of more than 10 models, we perform bootstrapping tests as well (for example, the MCS or the SPA). For the empirical application data from S&P 500 and 3 more stock indices (FTSE 100, DAX and CAC 40) are used. Lastly, this chapter evaluates the sensitivity of the results in terms of the underlying loss function focusing on robust loss functions, the in-sample estimation period and the stock index.

This chapter is organized as follows. A literature review on forecast evaluation methods and techniques is presented followed by data and methodology and empirical results sections. Finally, a summary of the main points is given in the conclusion.

3.2 Literature Review

Forecast evaluation can be performed with economic or statistical methods. This chapter will focus on statistical forecast evaluation which concerns single models, pairwise comparisons between models or comparisons across multiple models. For evaluation of a single model the most popular approach to rank the models based on a loss functions to measure the forecast error. Alternatively, an indication of

the model's accuracy can be given by examining the correlation between the forecast and the realization (for example, Mincer and Zarnowitz (1969) regression). To perform pairwise comparisons tests of equal predictive accuracy can be used. Other common measures of forecast evaluation aim to compare the correlation of the prediction errors for two models or of the errors' autocorrelation (Violante and Laurent (2012)). Finally, bootstrapping approaches allow for multiple model evaluation. In contrast, the success of a forecast can be examined through its applications. One way is to use backtesting to assess whether the various models are adequate enough to predict the Value-at-Risk (VaR) measure at the chosen significance level. Another approach suggests to measure the gains in a portfolio using as inputs the predicted values (return and covariance)(West (1996)). Alternatively, the forecasts can be used to construct naive volatility strategies. Then those investments can be evaluated for their profitability by means of simple returns or risk-adjusted returns (for example, the Sharpe Ratio of Sharpe (1966)) (Costantini et al. (2016a)). Finally, a utility function or option-pricing errors can also be used (González-Rivera et al. (2004)). In addition to being practical, approaches based on economic evaluation, offer the advantage that a volatility proxy is not required (Andersen et al. (2009)).

3.2.1 Single model evaluation

The literature has suggested various loss functions to measure the forecast error but, all of them seem to suffer from some sort of bias. Patton (2011) identified the upward and downward bias of the various (non-robust) loss functions, as the reason for the mixed findings reported in the literature as to what the best forecasting model is. So, it is important to elaborate further on loss functions. In point forecasts, the objective is to find the forecast that minimizes the expected loss, conditional on time information. This can be expressed as follows:

$$\widehat{Y}^{*}_{t+h,t} \equiv \arg_{\widehat{y}\epsilon y} \min E[L(Y_{t+h}, \widehat{y})|F_t] \tag{3.1}$$

where L is the forecaster's loss function, $\hat{y}$ is the set of possible forecasts and F_t is the time t information set. Hence, the optimal volatility forecast given a loss function will be:

$$h^{*}_{t} = \arg_{\widehat{y}\epsilon y} \min E[L(\widehat{\sigma^2_t}, h_t)|F_{t-1}] \tag{3.2}$$

where $\widehat{\sigma}_t$ is the realized volatility measure and h_t is the the conditional volatility. However, it is not guaranteed that for any choice of loss function it will the true conditional variance equal to the optimal forecast (Patton (2011)). Two of the most commonly-used loss functions are the Mean Squared Error (MSE) and the Quasi-Likelihood (QLIKE) given as:

$$MSE : L(\widehat{\sigma}_t, h) = \frac{1}{m}\sum_{t=1}^{m}((\widehat{\sigma}_t - h_t)^2) \tag{3.3}$$

$$Quasi - Likelihood - QLIKE : L(\widehat{\sigma}_t, h) = \frac{1}{m}\sum_{t=1}^{m}(\log h_t + \frac{\widehat{\sigma}_t}{h_t}) \tag{3.4}$$

where m is the number of forecasts. Other popular loss functions are listed in the Appendix (Caporin and McAleer (2012b), p.206-207; Hansen and Lunde (2005), p.877; Fuertes et al. (2009a); Patton (2011), p.248). In the literature of forecast evaluation more loss functions can be found such as the Gaussian Maximum Likelihood Error (GMLE) and the heteroskedasticity adjusted version of MSE and MAE[1] (HMSE and HMAE). Loss functions can be classified as symmetric or asymmetric. Symmetric loss functions such as the MAE, the MSE, the HMSE penalize equally overpredictions and underpredictions. On the other hand the GMLE is asymmetric and it penalizes under-predictions more heavily than over-predictions (Fuertes et al. (2009a)). MAE is performing very well in the presence of many large errors while MSE is more robust to small errors (Fuertes et al. (2009a)). On the contrary, Vilhelmsson (2006) argues that MAE is more robust to outliers and Koehler (2009) suggested that in times that high volatility is observed the HMAE loss function should be "trusted" more (Näsholm and Bunjaku (2010)).

Volatility proxies are inherently associated with noise. Therefore, robust loss functions are needed to address this issue. A robust loss function should rank the models as the true volatility was available. What is more, it should be less sensitive to outliers (Patton (2018); Patton (2011)). Patton (2011) elaborated on loss functions and illustrated under what circumstances the loss functions can be robust. Thus, the author examined combinations of nine, most commonly-used, loss functions and of three realized volatility measure and found that only the MSE and QLIKE loss functions are robust. This finding indicates that when the MSE or the QLIKE loss functions are used then the true conditional variance is the optimal forecast (Patton (2011)). Though, that is not true for the non robust loss functions. The MSE-prop and the MAE-prop generated upwards bias while the rests of the loss functions were under-estimating the true variance. Additionally, after testing with three proxies Patton (2011) concluded that the bias is independent of the accuracy of the used proxy. However, that was only an exercise using nine loss functions that resulted in identifying only two as robust following a pre-specified condition. Patton (2011) proved that under a more general set of conditions it is possible to draw "a class of loss functions" that are robust (p.251). "It is worth noting that although the ranking obtained by a robust loss function will be invariant to noise in the proxy the actual level of expected loss obtained using a proxy will be larger than that would have obtained when using the true conditional variance" (Patton (2011), p.248).

$$L(\widehat{\sigma}_t, h_t, b) = \frac{1}{(b+1)(b+2)}(\widehat{\sigma}_t^{2(b+2)} - h_t^{(b+2)}) - \frac{1}{b+1}h_t^{(b+1)}(\widehat{\sigma}_t^2 - h_t), b \notin \{1, -2\} \tag{3.5}$$

[1]Mean Absolute Error

where $L(\widehat{\sigma}_t, h_t, b) = h_t - \widehat{\sigma}_t^2 + \widehat{\sigma}_t^2 \log \frac{\widehat{\sigma}_t^2}{h_t}, b = -1$ and $L(\widehat{\sigma}_t, h_t, b) = \frac{\widehat{\sigma}_t^2}{h_t} - \log \frac{\widehat{\sigma}^2}{h} - 1, b = -2$. The first case corresponds to the QLIKE while the second does to the MSE (Patton (2011)).

Prediction accuracy can also be tested through the correlation between prediction and realization or the proxy (Violante and Laurent (2012)). The standard way is to rely on the Mincer and Zarnowitz (1969) (MZ) regression where the actual value or its proxy is regressed on the forecasted value. It has been used a lot for out-of-sample forecast comparisons. A simple representation is:

$$\widehat{\sigma}_t^2 = a + \beta h_t + \epsilon_t \tag{3.6}$$

where $\widehat{\sigma}_t^2$ is the proxy and h_t is the forecast. The condition for an unbiased prediction is $a = 0$ and $\beta = 1$, and for an accurate prediction is R^2 close to one and a small $var(\epsilon_t)$ (Poon and Granger (2003)). Rankings in terms of the MZ regression and the MSE loss should be alike if the forecasts are unbiased (Patton (2011)). Blair et al. (2001b) proposed a ratio to measure the proportion of variance explained by the forecasts:

$$P = 1 - \frac{\sum(\sigma_t^2 - \widehat{\sigma}_t^2)^2}{\sum(\sigma_t^2 - \mu_t)^2} \tag{3.7}$$

Under this criterion the forecasts are accurate when P is close to 1 (Poon and Granger (2003)). The standard MZ regression is not consistent against all possible deviations so an augmented version was proposed which is given as:

$$\widehat{\sigma}_t^2 = a + \beta h_t + \gamma z_{t-1} + \epsilon_t \tag{3.8}$$

where variable z includes any other measurement errors that could be considered relevant. This version allows for other variables to be included in order to test whether other stylized facts have been omitted (Violante and Laurent (2012)). Other extensions of the regression model include various variables that have been found to add explanatory power and/or lagged observations of the dependent variable. Another alternative is to change the proxy in order to avoid any asymptotic size distortions. The standardized volatility proxy could be used so the above MZ regression becomes:

$$\frac{\widehat{\sigma}_t^2}{h_t} = \delta + \vartheta \frac{\widehat{\sigma}_{t-1}^2}{h_{t-1}} + u_t \tag{3.9}$$

What is more, other transformations of the proxy and the forecasts that are less sensitive to outliers are also possible (Andersen et al. (2009), p.807). For instance:

$$|r_t| = a + \beta \sqrt{h_t} + \varepsilon_t \tag{3.10}$$

$$\log(r_t^2) = a + \beta \log(h_t) + \varepsilon_t \tag{3.11}$$

However, these transformations of the forecasts are not guaranteed to produce unbiased estimators of

the latent variable (Andersen et al. (2005b) referenced by Violante and Laurent (2012), p.467). The above regressions can be estimated using Least-Squares (LS) method but, in some case utilizing the Generalized LS (GLS) method could improve performance (Patton and Sheppard (2009a)).

At last, a recent stream of literature recognizes the benefits of relying on non-direct ways to perform out-of-sample comparisons. A popular approach is to employ Value-at-Risk (VaR) measure in order to evaluate forecasts of future volatility. This approach is intuitive since it suggests to examine the performance of competing models with regards to risk analysis. The cornerstone of this approach is that VaR estimates, obtained from various models and methods, are input into relevant loss functions. Under this framework, the loss function requires as inputs, VaR estimates and daily returns. Since the loss function penalizes the violations, the greater the number of unexpected violations the worse the model's performance. There are many loss functions designed for this purposes , such as, for example the the Violation Ratio (VR):

$$VR = \frac{1}{T}\sum_{t=1}^{T} I_t \tag{3.12}$$

where $I_t = 1$ if $r_t < VaR_t$ and equal to 0 otherwise. An issue that has not attracted much attention but could potentially improved the evaluation techniques is to account for different market conditions when calculating the forecasts errors. One approach could be to filter the data for high and low volume and then calculate the mean of the forecast errors (Fuertes et al. (2009a)).

3.2.2 Pairwise comparisons

Although, the measures presented above are useful in model evaluation and ranking, in order to conduct a robust study it is essential to test the statistical significance of the forecast errors across the models by carrying out pairwise comparisons. As highlighted by Diebold and Mariano (1995) "comparison of forecast accuracy is important for discriminating among competing models" (1995, p.1). The null hypothesis of these tests is that the compared models have equal forecast accuracy. Diebold and Mariano (1995) proposed the namesake, DM, test which tests this hypothesis "while allowing one to assess the significance of apparent predictive superiority" (Diebold (2015), p.1). The DM test was also used to test the significant difference in MSE and R^2 (Patton and Sheppard (2009a)). The DM test relaxed the assumptions of the previous tests so that any loss function to be used and the forecast errors can have a non Normal distribution with mean values of zero and also no condition was specified about the autocorrelation and the contemporaneous correlation (Diebold and Mariano (1995)). Diebold and Mariano (1995) and West (2006) proposed three asymptotic and exact finite sample tests. The DM test can be used for comparisons of many loss functions. However, Diebold and Mariano (1995) performed the test on the MSE loss function (Poon and Granger (2003)). The null hypothesis of equal accuracy will be $E[g(e_{it})] = E[g(e_{jt})]$ or $E[d_t] = 0$, where $d_t = [g(e_{it}) - g(e_{jt})]$ where the associated forecast errors bee_{it} and e_{jt}. In large samples, under the null hypothesis of zero

population mean of the loss differential, $\overline{d}$, has a standard normal distribution (Forte and Manera (2006), p.8-9).

$$DM = \frac{\overline{d}}{\sqrt{\widehat{f}}} \sim N(0,1) \tag{3.13}$$

However, in order to account for the presence of autocorrelation and correct for its bias in small samples improved version of the DM test were proposed by Forte and Manera (2006)and Harvey et al. (1997).

On the same direction, two powerful non-parametric tests proposed are the sign test which is based on the observed loss differentials and the Wilcoxon's signed-rank test which is based on the loss differentials' ranks. The null hypothesis of the first test is a zero median loss differential. Wilcoxon's Signed-rank test related distribution-free procedure that requires symmetry of the loss differential (but can be more powerful than the sign test in that case) (Diebold and Mariano (1995), p.7). The test statistic is (Forte and Manera (2006), p.8-9):

$$S_3 = \sum_{t=1}^{T} I_+(d_t) Rank(|d_t|) \tag{3.14}$$

Additionally, there are three more tests in this category. One of them is the simple F-test which is considered as a naive benchmark. "If loss is quadratic and the forecast errors are zero mean, Gaussian, serially correlated or contemporaneously uncorrelated then the null hypothesis of equal forecast accuracy corresponds to equal forecast error variances, the ratio of sample variances has the usual F distribution under the null hypothesis" (Diebold and Mariano (1995), p.7).

$$F = \frac{e_i' e_i}{e_j' e_j} \tag{3.15}$$

However, F-test's conditions are considered too strict and thus it is of little use. The second is the Moegan-Granger-Newbold test where it tests for zero correlation and finally, the Meese and Rogoff (1988) test which is used when forecast errors are serially and contemporaneously correlated. Their approach considers a squared error loss. Harvey et al. (1997) built the test in a regression form (Diebold and Mariano (1995)).

However, for nested models the DM statistic is non normal resulting in undersized tests with low power. Clark and West (2007) suggest a test for equal MSE for nested models that is approximately standard normal. Further, DM tests fails to account for forecast uncertainty due to parameter estimation. West (1996) proposed to perform tests accounting for the estimation error and making in that way possible to evaluate the model's performance instead of the forecast [2]. One limitation of

[2]For more information see Patton (2018)

the test is that it cannot used to perform nested models and even though it is designed for multiple forecast evaluation, it works better when only two models are compared (Patton (2018)). Therefore, a modified version of the DM test is proposed that corrects the autocorrelation of the errors in multi-period forecasts. West (1996) provides the asymptotic theory for testing the predictive ability of model. "The machinery allows for non-nested and non-linear models and dependence of predictions on estimated regression parameters"(Clements and Hendry (2008), p.293). This new modified test is called Diebold-Mariano and West (DMW) test. The DMW can be used for pairwise comparisons.

Giacomini and White (2006) continued West's work by providing a theoretical motivation for performing tests on estimated parameters. Another major contribution that they proposed is about conditional expectations of forecasts and forecasts errors [3] with the unconditional expectations being nested (Giacomini and White (2006)). In the case where there is no condition the test corresponds to the DM test. The null hypothesis following the West and GW approach respectively is given respectively as: $\mathbf{E}[f(Y_{t+1}, \hat{Y}(X_{t+1}, \beta^*), Z_{t+1})] = 0$, $\mathbf{E}[f(Y_{t+1}, \hat{Y}(X_{t+1}, \hat{\beta_{t+1}}), Z_{t+1})|G_t] = 0$ where, $\hat{Y}$ is a function mapping the predictor variables X_t, β^* is the the pseudo-true parameter, $\hat{\beta}_{t+1}$ is is an estimator of β^* using data up until date $t+1$, Z_{t+1} is a vector of other observable data, G_t is the information set. The t-statistic under this case follows the Normal distribution. The GW test can be regarded as superior since it can compare nested and non nested models and forecasting methods, not just models. Forecasting methods include the model, the window and the method, for example parametric or not. Finally, it is relative easy to be conducted (Patton (2018)). However, this test is only suitable for a rolling window and a finite amount of data is required (Patton (2018)). According to Giacomini and White (2006) the test was built for a rolling window, rather than a recursive one, for three main reasons. Firstly, the popularity of the scheme, secondly the rolling window method is not comparatively specific and thirdly it is more straightforward with better results. The test is given as:

$$unconditional\ Giacomini - White(GW) = \frac{\tilde{d}}{\sqrt{\frac{2\pi \hat{f}_d(0)}{T}}} \tag{3.16}$$

where T is the out-of-sample, d is the average loss differential and $\hat{f}_d(0)$ is the consistent estimator of $f_d(0)$ which is the population mean of the loss differential and is given as:

$$\hat{f}_d(0) = \frac{1}{2\pi}\hat{\gamma}_d(0)$$

[3]The conditional specification tests for serial correlation in the errors of the regression and it is suitable for comparing and evaluating forecasting models at a specific future point in time (Giacomini and White (2006)).

3.2.3 Comparisons of multiple models

An important limitation of pairwise comparisons is that they fail to account for data mining (snooping) (Caporin and McAleer (2012b)). Also with the tests of equal predictive accuracy test we cannot "(...) exclude a priori the possibility of having different model ranking associated with different robust loss functions" (Caporin and McAleer (2012b), p.208). Therefore, it is essential to test that the predictive accuracy is not random or due to luck (Costantini et al. (2016a)). The literature has suggested to employ a bootstrap approach to test the predictive superiority. Bootstrapping technique allows to deal with the problem of dimensionality when many models (> 10 models) are compared.

Multiple comparisons tests can address the issues presented above through multiple hypothesis testing. What is more when the aim is to compare a large set of models then the multiple comparison is superior compared to pairwise tests as over-rejection of null arising from the process of sequentially searching for the best model across several alternatives. However, in some multiple comparisons tests more than one model can be the best (Hansen et al. (2011)). Originally, White (2000) proposed the reality check. In this tests, a benchmark is chosen ex-ante to test the null hypothesis that a benchmark is inferior to all the other models. Later Hansen et al. (2005) proposed the test of superior predictive ability (SPA) which is believed to be more powerful and less sensitive to the alternatives (Hansen et al. (2005)). They used a different test statistic and employed a sample dependent distribution under the null hypothesis (Hansen et al. (2005), p.1). SPA test uses the Hansen's P-value and the White's Upper P-value (Original RC P-values). It is valid for any loss function, however Hansen et al. (2005) used the squared forecast deviation from the proxy and accounts for the parameter estimation uncertainty, and can compare nested and non nested models and with less strict assumptions. This test can be used to test the superiority of a trading strategy as well (White (2000)).

However, the SPA test from the way it is defined it can only demonstrate the superiority of a certain model which is the benchmark. So, Hansen et al. (2011) proposed the Model Confidence set (MCS) which is designed to compare multiple models but without the need to specify a benchmark and thus, it is considered superior to the SPA test. Also it is defined as a single hypothesis test. The way MCS works is by performing pairwise comparisons of all the models and then, at some pre-specified level of confidence it discards the worse performing models to create a set of the best performing models, M(minimum expected loss) from the an initial set of models, M^0 (Symitsi et al. (2018); Liu et al. (2015); Wang et al. (2015); Violante and Laurent (2012)). The test sequentially performs tests of equal predictive accuracy and the worst performing models are discarded. The null hypothesis is given as: $E[d_{ij,t}] = 0$, where $d_{ij,t} = [L_{it} - L_{jt}]\ \forall i, j = 1, ..., M$ where L is the loss function. At each step the tests statistics are given as:

$$T_{SQ} = \sum_{i<j} \frac{(\bar{d}_{ij,t})^2}{\widehat{var}(\bar{d}_{ij})} \tag{3.17}$$

$$T_R = \max_{i,j \in M} \frac{(|\bar{d}_{ij})|}{\sqrt{\widehat{var}(\bar{d}_{ij})}} \tag{3.18}$$

where $\bar{d}_{ij} = \frac{1}{T}\sum_{t=1}^{T} d_{ij,t}$ is the average sample loss between forecast i and j, $\widehat{var}(\bar{d}_{ij})$ is the estimate of the asymptotic variance of $\bar{d}_{ij}$ computed using a block bootstrap procedure, T_{SQ} is the semi-quadratic statistic, and T_R is the range statistic. If the null hypothesis is rejected at a significance level, then the model with the largest statistic $t_i = \frac{(\bar{d}_i)}{\sqrt{\widehat{var}(\bar{d}_i)}}$ is eliminated. Where $\bar{d}_i = \frac{1}{m}\sum_{j=1}^{M} \bar{d}_{ij,t}$ is the sample loss of model i relative to the average loss across of all other m models that are currently in the set, M and $\widehat{var}(\bar{d}_i)$ is the estimate of the asymptotic variance of $\bar{d}_i$. The resulting set can be used to form forecast combinations as well. However, the decision of the loss function plays a significant role to the identification of the optimal set of forecasting models (Luo et al. (2017); Charles and Darné (2019)). MCS can be applied using a VaR-based loss function too (Tian and Hamori (2015)).

3.3 Data and Methodology

3.3.1 Data

The data used for the empirical application are daily open-to-close returns for S&P 500 (Standard and Poor's 500), the CAC 40, the FTSE 100, and the DAX stock indices. The data were downloaded directly from the Oxford-Man Institute's Realized Library (library version 0.2). The full sample for S&P 500 index extends from January, 3 2000 to December, 29 2017 (4,518 daily observations). Two different sample split schemes were considered following the recent stream of literature on forecasting which suggests a small number of observations which is adaptive to changes and up-to-date- in order to produce satisfactory out-of-sample results. The first in-sample period for S&P 500 is from 03/01/2000 to 31/12/2003 (997 observations) so of the rest, 3521 observations, are used as the out-of-sample. While, the second sample examined for the same index, consists of an in-sample period form 03/01/2000 to 29/12/2000 (251 observations) and and out-of-sample period from 02/01/2001 to 29/12/2017 (4267 observations). We are using a rolling window to generate one-step ahead forecasts and we rely on calendar time to choose the window size as specified above.

Respectively, the sample period for CAC 40 is from 03/01/2000 to 29/12/2017 (4568 observations). The in-sample period extends from 03/01/2000 to 31/12/2003 (1005 observations) and the rest 3585 observations were the out of sample period. As for the FTSE index the sample ranges from 04/01/2000 to 29/12/2017 (4537 observations). The in-sample period is from beginning of the sample until the

31/12/2003 (1001 observations) and from 02/01/2004 until the end of the sample (3536 observations) is the out-of-sample period. Correspondingly, for the DAX the sample includes 4569 observations from 03/01/2000 until 29/12/2017 and the in-sample period covers 4 years; 1010 daily returns, leaving 3559 observations for the out-of-sample to be used for evaluation purposes.

The sample period considered includes some important events such as the dot-com bubble of 2000 which caused the S&P 500 index to an all-time intraday high on March 24, 2000, the record high closing price on October 9, 2007, the financial crisis of 2008 which caused the highest daily percentage gain and its largest single-day point increase on October 13, 2008 but also the year (2008) with the highest yearly percentage loss.

The information presented above about the sample periods considered for estimation and evaluation for each stock index is summarized in the Table 3.1.

Table 3.1 – Sample splits

	In-sample period (number of observations)	Out-of-sample period (number of observations)
S&P 500 -sample 1	from 03/01/2000 to 31/12/2003 (997 obs)	from 02/01/2004 to 29/12/2017 (3521 obs)
S&P 500 -sample 2	from 03/01/2000 to 29/12/2000 (251 obs)	from 02/01/2001 to 29/12/2017 (4267 obs)
CAC 40	from 03/01/2000 to 31/12/2003 (1005 obs)	from 02/01/2004 to 29/12/2017 (3585 obs)
FTSE 100	from 04/01/2000 to 31/12/2003 (1001 obs)	from 02/01/2004 to 29/12/2017 (3536 obs)
DAX	from 03/01/2000 to 30/12/2003 (1010 obs)	from 02/01/2004 to 29/12/2017 (3559 obs)

Note: This table reports the in-sample and out-of-sample period used for the empirical application as well as the number of observations included in the relevant period for all stock indices.

Figure 3.1 displays the plots of the stock indices' returns for the whole sample period. It seems like the returns' mean is constant and around zero. Another characteristic is that high (low) volatility periods are followed by high (low) volatility periods, a phenomenon called volatility clustering. From the graph it can be seen that the most highly volatile periods for the S&P 500 index are 2003, 2009, 2011 and 2016. For the rest of the indices similar observations can be made. The CAC 40 is the as seen by more volatile, followed by the DAX index.

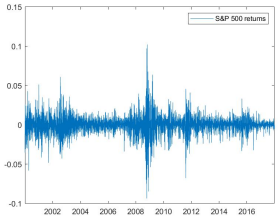


Figure 3.1 – Daily returns for S&P 500 returns (2000-2017)

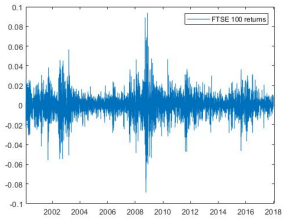


Figure 3.2 – Daily returns for FTSE 100 returns (2000-2017)

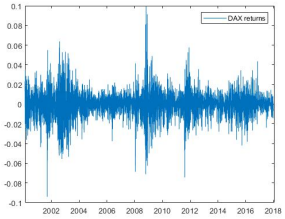


Figure 3.3 – Daily returns for DAX returns (2000-2017)

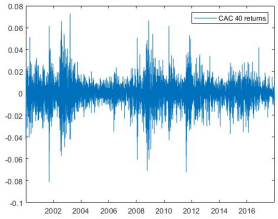


Figure 3.4 – Daily returns for CAC 40 returns (2000-2017)

Table 3.2 – Descriptive statistics for index returns and realized volatility

	Panel A: Returns				Panel B: Realized Variance (RV)			
	S&P 500	CAC 40	FTSE 100	DAX	S&P 500	CAC 40	FTSE 100	DAX
Mean	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Volatility	0.01	0.01	0.01	0.01	0.00	0.00	0.00	0.00
Kurtosis	11.52	7.4	9.36	8.07	243.54	133.74	202.90	98.39
Skewness	-0.18	-0.14	-0.2	-0.10	11.63	8.85	10.69	7.57
Max	0.10	0.07	0.09	0.10	0.01	0.01	0.00	0.01
Min	-0.09	-0.08	-0.09	-0.09	0.00	0.00	0.00	0.00
Count	4518	4568	4537	4569	4518	4568	4537	4569

Note: This table reports the main descriptive statistics for the S&P 500, CAC 40, the FTSE 100, the DAX for the daily returns and the 5-min RV for the period from January 2000 to December 2017 respectively for each index. The last row simply reports the number of observations in each sample.

Table 3.2 presents summary statistics for the returns of all stock indices. Kurtosis and skewness statistics, indicate that the distribution of S&P 500 returns has fatter tails than the Normal dis-

tribution and it is asymmetric. To provide more information about the returns' distribution and persistence (long memory) we plot the Quantile-Quantile (qq) plot and the correlogram (also known as Autocorrelation Function ACF plot) of all assets for the whole sample period (see Figures 3.5 - 3.12 in the appendix) have been plotted. GARCH-type models treat volatility as a latent variable therefore it remains unobservable even after estimation. Thus, a measure of the true volatility needs to be specified. For the purposes of this study, the 5 min RV (realized variance) will be the proxy taken from Oxford-Man Institute's Realized Library (library version 0.2) for each stock index in line with previous studies (see for example Liu et al. (2015)). The summary statistics of the RV estimators are also reported in Table 3.2 where very high values of kurtosis are observed.

3.3.2 Methodology

From the data analysis it can be confirmed the presence of the stylized facts namely, the leverage effects, the leptokurtosis, the volatility clustering and the long memory. So GARCH models are suitable to estimate and to forecast the conditional variance of returns. Univariate volatility models have been defined in detailed in Chapter 2. In this study models from the GARCH family the following models are included: the ARCH, the GARCH, the GJR-GARCH, the TGARCH, the IGARCH, AGARCH, the NAGARCH, the EGARCH, the APGARCH and the FIGARCH models. Additionally, for comparison reasons the GARCH-MIDAS and Realized Volatility (RV) forecasting models were included in the study. The latter include the HAR model and its logarithmic transformation, the HEAVY model and the GARCH-RV model. Finally, a couple of simple standard deviation's (SD) approximations are evaluated, that is the Laplace approximation and the conservative SD measures [4].

Below Table 3.3 reports the models' specifications. In addition, Engle et al. (2008) GARCH-MIDAS model can be scripted by the following series of equations:

$$r_{i,t} = \mu + \sqrt{(}\tau_t g_{i,t})\epsilon_{i,t}, \tag{3.19}$$

$$g_{i,t} = (1 - \alpha - \beta) + \alpha \frac{(r_{i-1,t} - \mu)^2}{\tau_t} + \beta g_{i-1,t}, \tag{3.20}$$

$$\tau_t = m + \theta \sum_{k=1}^{K} \phi_k(w_1, w_2) RV_{t-k}, \tag{3.21}$$

$$RV_t = \sum_{i=1}^{N_t} r_{i,t}^2, \tag{3.22}$$

$$\sigma_{it}^2 = \tau_t g_{i,y}, \tag{3.23}$$

[4] All models and specifications were estimated using the codes provided by the Oxford MFE Toolbox of Sheppard (2009) apart from the GARCH-RV which was estimated using the rugarch package developed by Ghalanos (2018).

$$\phi_k(\omega) = \frac{(1-\frac{k}{K})^{w_1-1}(1-\frac{k}{K})^{w_2-1}}{\sum_{j=1}^{K}(\frac{j}{K})^{w_1-1}(1-\frac{j}{K})^{w_1-1}} \tag{3.24}$$

where $r_{i,t}$ is the return on day i in month t, $g_{i,t}$ is the component for the conditional variance dynamics which follows a (daily) GARCH(1,1) process, τ is defined as smoothed realized volatility in the spirit of MIDAS regression and K is the number of periods over which we smooth the volatility, the total conditional variance is represented by σ and the weighting scheme by ϕ. For the empirical application of the GARCH-MIDAS the period was selected at the 22; the day-month aggregation and the number of MIDAS lags was chosen to ten months.

Table 3.3 – Univariate and multivariate volatility models

Name	Model specification	Author
ARCH	$\sigma_t^2 = \omega + \sum_{i=1}^{p} a_i \varepsilon_{t-i}^2$	Engle (1982)
GARCH	$\sigma_t^2 = \omega + \sum_{i=1}^{p} a_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2$	Bollerslev (1986)
GJR-GARCH	$\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \gamma\varepsilon_{t-1}^2 I_{t-1} + \beta_1\sigma_{t-1}^2$	Glosten et al. (1993)
	($I_{t-1} = 1$ if $r_{t-1} < 0$ and $I_{t-1} = 0$ otherwise)	
TGARCH	$\sigma_t^2 = \omega + \alpha\lvert\varepsilon_{t-1}\rvert + \gamma\lvert\varepsilon_{t-1}\rvert I_{t-1} + \beta_1\sigma_{t-1}$	Zakoian (1994)
	($I_{t-1} = 1$ if $r_{t-1} < 0$ and $I_{t-1} = 0$ otherwise)	
IGARCH	$\sigma_t^2 = \omega + \sum_{i=1}^{p} a_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2$	Nelson (1990)
	$(a_i + \beta_j = 1)$	
AGARCH	$\sigma_t^2 = \omega + a(r_{t-1} - \gamma)^2 + \beta\sigma_{t-1}^2$	Engle (1990b)
NAGARCH	$\sigma_t^2 = \omega + a(\varepsilon_{t-1}\sigma_{t-1}^{-1} + \gamma)^2 + \beta\sigma_{t-1}^2$	Engle and Ng (1993)
EGARCH	$\log(\sigma_t^2) = \omega + a(\lvert z_{t-1}\rvert - E(\lvert z_{t-1}\rvert) + \gamma z_{t-1} + \beta\log(\sigma_{t-1}^2)$	Nelson (1991)
	$(z_t = \varepsilon_t\sigma_t^{-1})$	
APGARCH	$\sigma_t^2 = (\omega + a_1\lvert r_{t-1}\rvert + \gamma r_{t-1}^{\delta} + \beta(\sigma_{t-1}^2)^{\delta/2})^{2/\delta}$	Ding et al. (1993)
FIGARCH	$\sigma_t^2 = \omega + [1 - \beta(L) - \phi(L)(1-L)^d]\varepsilon_t + \beta\sigma_{t-1}^2 (0 < d < 1)$	Baillie et al. (1996)
HAR	$RV_t = \omega + \phi_1 RV_{t-1} + \phi_2 \sum_{i=1}^{5} \frac{1}{5} RV_{t-i} + \phi_3 \sum_{j=1}^{22} \frac{1}{22} RV_{t-j} + u_t$	Corsi et al. (2012)
HEAVY	$var(r_t \lvert F_{t-1}^{HF}) = h_t = \omega + \alpha RM_{t-1} + \beta h_{t-1}$	
	$E(RM_t \lvert F_{t-1}^{HF}) = \mu_t = \omega_R + \alpha_R RM_{t-1} + \beta_R \mu_{t-1}$	Shephard and Sheppard (2010)
GARCH-RV	$h_t = \omega + \alpha\epsilon_{t-1}^2 + \theta h_{t-1} + \lambda RV_{t-1}$	Engle (2002b) and Koopman et al. (2005)
$quant_1$	$\widehat{\sigma}_t = \frac{\widehat{Q}(0.95) - \widehat{Q}(0.05)}{3.256}$	Taylor (2005)
$quant_2$	$\widehat{\sigma}_t = max(\frac{\widehat{Q}(0.75) - \widehat{Q}(0.50)}{0.6745}, \frac{\widehat{Q}(0.50) - \widehat{Q}(0.25)}{0.6745})$	Greco et al. (2015)

Notes: This table reports the univariate models used in this chapter.

The analysis is based on 1-step ahead rolling forecasts of the future conditional volatility (Rapach and Strauss (2008); Day and Lewis (1993); Brooks and Persand (2003)). For the GARCH models at each step the optimal number of lags was selected recursively using the Bayesian Information Criterion (BIC) since finding the best specification is crucial for accurate forecasting. The BIC of Schwarz et al. (1978) is usually preferred over the AIC in forecasting exercises, since it selects the more parsimonious model which leads to better forecasts. Following the recent stream of literature,

the process was restricted to only two lags and for the asymmetric models, the number of asymmetries was controlled to one. Each of the aforementioned GARCH models was estimated depending on three different error distribution to robustify the results and explore any forecasting gains. The examined error distributions were the Normal, the Student-t and the Skew-t.

There is a growing stream in the literature supporting the combination of individual forecasts, on the ground that any single model will fail at some level due to misspecification bias or structural breaks and that one forecast could capture information another has missed (Timmermann (2006)); Pesaran and Timmerman (2005)). What also adds to that complexity is the fact that different loss function select different single model. As a result, multiple loss functions need to be consider to achieve robust results. Thus, Timmermann (2006) suggested to associate forecast combinations with diversification gains in order to alleviate the importance of the chosen loss function. Since the seminal paper of Bates and Granger (1969), the literature on forecasting has widely investigated the performance of the combination of forecasts against single forecasting models. Evidence of the superiority of the combination of forecasts is found in several studies (see, for example Clemen (1989); Stock and Watson (2004); Timmermann (2006); Becker and Clements (2008); Costantini and Pappalardo (2010); Aiolfi et al. (2011); Costantini et al. (2016a); Costantini et al. (2017); Samuels and Sekkel (2017)). So, in this study forecast combinations are included too.

The literature on forecast combination expands to quantiles, intervals and probability distributions combinations but also to combinations of various realized measures (Patton and Sheppard (2009b); Timmermann (2006)). There are many approaches for combinations such as "variance-covariance methods" or methods that aim to achieve a lower MSE and the "regression-based" methods (Diebold and Lopez (1996), Granger and Ramanathan (1984) and Patton and Sheppard (2009b)). What distinguishes most methods is the weighting scheme that is being adopted. Regardless that, the most naive way is to rely on equal weights which is more successful than more sophisticated schemes (Patton and Sheppard (2009b); Palm and Zellner (1992); Makridakis and Winkler (1983) and Armstrong (1984)). Finally, it is often optimal not to include all the models but to perform some kind of model selection (Armstrong (1989); Timmermann (2006)). In that case loss functions can help to rank the models and select the optimal ones. Further, encompassing tests can be used which directly compare two sets of forecasts to evaluate whether the informational content of one is such that it dominates the other. In that case, a forecast should be included only " (...) if the expected loss of a combination of that forecast and a rival forecast is not significantly less than the expected loss of the original forecast alone" (Clements and Hendry (1998), p.228). For the purposes of this study, the optimal set for combination was recursively selected using one year as an evaluation window. Then at each evaluation window the optimal model was selected based on the MSE loss function. Different loss functions can

be tested as well but this out-of the scope of this study. To from the forecast combinations that are used in the empirical application, the mean, the median and the trimmed mean of the optimal set of forecasts were calculated, forming the first, the second and the third combinations, respectively. To calculate our third forecast combination (the trimmed mean) the mean of the optimal models was calculated after the 10% of the highest and the 10% of the lowest observations were discarded.

Even though measuring forecast accuracy is important, in order to evaluate and rank the models, what is also often interesting is to test the hypothesis that two forecasts have equal forecasting accuracy. So, for the forecast evaluation the unconditional test of Giacomini and White (2006), hereafter GW, was performed at 5% significance level. The GW test is considered the most appropriate for this study among the other equally predictive tests since it is suitable for comparisons of rolling forecasts, it is capable of capturing the effect of uncertainty in the estimation and can operate with nested and non nested models. What is more the sign of the t-statistic can indicate which specification performs better. Under this test, a positive t-statistic indicates that model A forecast produces larger average loss than the model B forecast and hence model B is superior at a given level of significance. To test the sensitivity of our results to the loss function used, the GW test was performed under three different robust loss functions following Patton (2011)'s robust loss function family. Those are the MSE which corresponds to the Patton's robust family when $b = 0$, and the specifications derived when b is selected equal to 1 and to -0.5. What is more, the test was repeated for more popular loss functions logMSE, sdMSE, propMSE, RMSE, propRMSE, MAE, logMAE, sdMAE, MAPE and GMLE loss functions. To test the in-sample sensitivity the test was performed for both in-sample periods for S&P 500. Finally, to asses the sensitivity of the data the analysis was replicated for three additional stock indices, DAX, CAC 40 and FTSE 100.

3.4 Empirical Results

Table 3.4 reports the results for S&P 500 using the GW and the MSE loss function for the 4 years in-sample period. ARCH models are performing significantly worse than almost all the other models and various specifications. This is somewhat expected since ARCH model is restrictive, symmetric and it is unable to capture the persistence in volatility unless the number of lags is large (Liu et al. (2015)). What is more, the assumption about the distribution of model does not seem to affect. Similarly, the two simple SD approximations specifications have each, at least 39 cases where the hypothesis of equal predictive accuracy could not be rejected at the 5% significant level and all of them demonstrate cases where they were outperformed by the competing specifications. They only manage to beat the HAR model. Finally, HAR seems to be the worst performing model.Following, all GARCH, IGARCH, FIGARCH and GARCH-RV specifications, the GARCH-MIDAS and the HEAVY

models can be grouped together, as they have almost the same ratio of number of "wins" over the overall number of comparisons and they, only, manage to outperform the worse performing specifications, namely the ARCH models, the two simple SD approximations and the HAR model. However, GARCH specifications in some cases seem to outperform IGARCH and FIGARCH specifications as well.

On the contrary, the three TGARCH approximations, along with the first and the third combinations outperform all of the other specifications. Remarkably, the first and the third combinations have zero loses. Yet, there are cases where the TGARCH models or the 1st or the 3rd combinations compete against certain models and/or specifications and the null hypothesis of equal forecasting accuracy could not be rejected. More specifically, the TGARCH-N did not seem to perform significantly better or worse against all three specifications of the NAGARCH and EGARCH models, the three combinations, the two more TGARCH specifications, the logHAR and the APGARCH-St. The TGARCH-t seems to perform equally well with the TGARCH-N, TGARCH-st, the three combinations, the logHAR, the three EGARCH specifications, the AGARCH-N, the NAGARCH-N and the NAGARCH-St. Further, for the TGARCH-st the models towards which is not significantly better or worse are the TGARCH-N, the TGARCH-t, the three combinations, the three EGARCH specifications and the NAGARCH-St. Finally, for the first and the third combinations the null hypothesis of equal predictive accuracy could not be rejected against all the specifications of the TGARCH and the NAGARCH models, the EGARCH-t, the APGARCH-t, the GJR-st, the EGARCH-st, the APGARCH-st. Finally, the third combination has equal predictive accuracy with the HAR, the GJR, the EGARCH and the APGARCH-N and the GJR-t specifications.

From that list it can be seen that for the best performing models, namely the three TGARCH approximations, the first and the third combinations, the null hypothesis of the GW test cannot be rejected at 5% significant level for: the NAGARCH, the EGARCH, the logHAR, and the APGARCH. More specifically, the NAGARCH and the EGARCH under the Skew-t specifications beat 18 and 22 models, respectively and they are indifferent from the rest (zero lose cases). NAGARCH-St is superior to its Student-t and indifferent to its normal, while EGARCH is indifferent to its two other specifications. Following, logHAR logarithmic transformation has 20 "wins" and only 2 "loses" that is the first and the third combinations. Finally, APGARCH, is outperformed by the NAGARCH-N. APGARCH-N is superior to its Student-t and inferior to its Skew-t specifications.

Finally, GJR and AGARCH are beaten by the rest specifications at 5% significant level; both NAGARCH-N and EGARCH-t outperforms AGARCH-t, the AGARCH-st. However, AGARCH-N is superior to its Student-t transformation and indifferent to its Skew-t. Then the AGARCH-N is outper-

formed by the TGARCH-N, the EGARCH-St and the first and the second combination. Nevertheless, the AGARCH-St (and the APGARCH-St) are superior than the three ARCH, the three GARCH, the three IGARCH and the three FIGARCH specifications, as well as the HAR, and the GARCH-MIDAS and the three GARCH-RV. In addition the APGARCH-St is beating the GJR Student-t and Skew-t. Nevertheless, the GJR-St is outperforming its Student-t transformation but both are indifferent to the Normal specification. GJR-N is beaten by the first combination and TGARCH-t. Principally, the GJR is winning the three ARCH, the three GARCH, GARCH-RV, and the three FIGARCH specifications, the HAR, the HEAVY, the GARCH-MIDAS and the IGARCH Student-t and Skew-t while it is losing just by four specifications. The GJR-t is outperformed by the three TGARCH specifications, and the GJR-St, the first combination and the APGARCH-St as already mentioned. Thus, the GJR-St is losing also from the three TGARCH specifications plus the NAGARCH-N and the APGARCH-St.

Furthermore, in order to compare all the models at the same time and because many models and specifications are being tested, the Superior Predictive Ability (SPA) of Hansen et al. (2005) and the model confidence set (MCS) of Hansen et al. (2011) test was conducted at 5% significant level, using the three main robust losses from this study, since the tests of equally predictive ability can only perform pairwise comparisons. The test was conducted using the bootstrap approach and the average block size was selected to be 10 following the papers of Liu et al. (2015) and Politis and Romano (1994). The SPA test requires a benchmark, which in tables 3.11 and 3.5 is stated by the specification in the rows. With respect to the Table 3.5, under the MSE when any of the ARCH or the GARCH-N or the IGARCH-N or the AGARCH-N or the IGARCH-St or the HAR was used as a benchmark then the null hypothesis that the benchmark is inferior to all the other models was rejected when all the models were included into the set at the 5% level which confirms our results. However, the MCS dictates that all models should be included. As for the 4-years in-sample the results based on the SQ (semi-quadratic statistic) criterion are presented in the Appendix (see table 3.21). At last, 3.31 in the Appendix contains information on how many times each specific model beats the rest of the models for S&P 500 using the GW test and the MSE loss function for the 4 years in-sample period.

Table 3.4 – Pairwise comparisons of unconditional Giacomini and White test for S&P 500 (4-year in-sample) -MSE

	Spec 1	Spec 2	Spec 3	Spec 4	Spec 5	Spec 6	Spec 7	Spec 8	Spec 9	Spec 10	Spec 11	Spec 12	Spec 13	Spec 14	Spec 15	Spec 16	Spec 17	Spec 18	Spec 19	Spec 20	Spec 21
Spec 1:A_N	-	2.43*	3.80*	4.30*	2.03*	4.03*	3.54*	4.44*	3.89*	1.78	-4.50*	2.19*	3.16*	4.01*	2.16*	3.79*	2.84*	4.38*	3.31*	1.72	-4.22*
Spec 2:GA_N	-	-	3.61*	4.37*	-3.87*	6.22*	2.82*	3.73*	3.82*	-1.84	-3.50*	-1.21	2.30*	3.61*	-1.35	5.89*	1.77	3.94*	2.54*	-2.14*	-3.54*
Spec 3:GJR_M	-	-	-	2.55*	-4.49*	-1.20	1.00	0.13	0.92	-3.29*	-4.53*	-4.51*	-1.59	2.57*	-4.63*	-1.75	-0.37	0.69	-0.93	-3.49*	-4.60*
Spec 4:TGA_N	-	-	-	-	-5.11*	-2.34*	-0.33	-1.05	-2.70*	-3.82*	-5.01*	-5.13*	-2.41*	0.61	-5.21*	-2.91*	-1.20	-0.80	-2.29*	-3.99*	-5.10*
Spec 5:IGA_N	-	-	-	-	-	6.64*	3.38*	4.17*	4.68*	-0.56	-3.08*	2.38*	2.98*	4.27*	2.16*	6.93*	2.21*	4.48*	3.23*	-0.78	-3.11*
Spec 6:AGA_N	-	-	-	-	-	-	1.21	1.55	1.41	-4.61*	-4.97*	-6.06*	0.26	1.89	-6.15*	-1.99*	0.33	1.97*	0.50	-4.79*	-5.08*
Spec 7:$NAGA_N$	-	-	-	-	-	-	-	-0.42	-0.82	-2.73*	-4.16*	-3.38*	-2.59*	1.21	-3.43*	-1.51	-2.54*	-0.18	-2.51*	-2.88*	-4.21*
Spec 8:EGA_N	-	-	-	-	-	-	-	-	0.02	-3.69*	-5.34*	-3.86*	-0.61	0.97	-3.91*	-1.83	-0.29	1.33	-0.41	-3.70*	-5.48*
Spec 9:$APGA_N$	-	-	-	-	-	-	-	-	-	-3.42*	-4.62*	-4.75*	-1.70	2.42*	-4.84*	-1.98*	-0.51	0.51	-1.27	-3.63*	-4.70*
Spec 10:$FIGA_N$	-	-	-	-	-	-	-	-	-	-	-2.90*	0.98	2.33*	3.31*	0.90	4.15*	1.88	3.69*	2.50*	-0.61	-2.91*
Spec 11:A_t	-	-	-	-	-	-	-	-	-	-	-	3.20*	3.85*	4.66*	3.17*	4.73*	3.44*	5.20*	3.99*	2.81*	1.19
Spec 12:GA_t	-	-	-	-	-	-	-	-	-	-	-	-	2.92*	4.29*	-0.94	6.60*	2.14*	4.27*	3.21*	-1.24	-3.23*
Spec 13:GJR_t	-	-	-	-	-	-	-	-	-	-	-	-	-	3.79*	-3.00*	-0.56	0.36	1.09	1.37	-2.50*	-3.89*
Spec 14:TGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-4.35*	-2.30*	-2.02*	-0.81	-4.06*	-3.47*	-4.72*
Spec 15:IGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	6.78*	2.18*	4.34*	3.28*	-1.15	-3.20*
Spec 16:AGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.53	2.32*	0.82	-4.42*	-4.82*
Spec 17:$NAGA_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.59	-0.00	-2.00*	-3.47*
Spec 18:EGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-0.87	-3.76*	-5.31*
Spec 19:$APGA_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-2.68*	-4.03*
Spec 20:$FIGA_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-2.82*
Spec 21:A_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 22:GA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 23:GJR_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 24:TGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 25:IGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 26:AGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 27:$NAGA_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 28:EGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 29:$APGA_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 30:$FIGA_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 31:$comb_1$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 32:$comb_2$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 33:$comb_3$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 34:HAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 35:logHAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 36:HEAVY	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 37:$GA-MIDAS$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 38:$GA-RV_N$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 39:$GA-RV_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 40:$GA-RV_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 41:$quant_1$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 42:$quant_2$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Note: This table reports the out-of-sample t-statistic of the Unconditional Giacomini and White (2006) test under the MSE loss function for S&P 500. The period is from 02/01/2004 to 29/12/2017. A positive t-statistic indicates that average loss of model A (in row) is larger than of model B (in column) and hence model B is superior. *denotes significance at the 5% level

Table 3.4 cont. - Pairwise comparisons of unconditional Giacomini and White test for S&P 500 (4-year in-sample) - MSE

	Spec 22	Spec 23	Spec 24	Spec 25	Spec 26	Spec 27	Spec 28	Spec 29	Spec 30	Spec 31	Spec 32	Spec 33	Spec 34	Spec 35	Spec 36	Spec 37	Spec 38	Spec 39	Spec 40	Spec 41	Spec 42
Spec 1:A_N	2.21*	3.27*	3.97*	2.18*	3.79*	2.93*	4.44*	3.58*	1.72	4.93*	4.55*	4.97*	-5.77*	4.55*	2.11*	2.01*	1.08	0.92	0.04	-3.05*	-3.10*
Spec 2:GA_N	-1.01	2.45*	3.48*	-1.13	5.89*	1.92	3.79*	2.96*	-2.16*	5.82*	3.69*	5.92*	-6.36*	3.69*	-0.37	-1.81	-1.56	-1.88	-1.89	-4.05*	-4.08*
Spec 3:GJR_M	-4.51*	-1.01	2.53*	-4.65*	-1.75	0.16	0.59	0.52	-3.54*	2.21*	-0.37	1.75	-6.50*	-0.37	-2.10*	-3.86*	-2.55*	-2.75*	-2.63*	-4.28*	-4.31*
Spec 4:TGA_N	-5.12*	-2.04*	0.88	-5.22*	-2.91*	-0.59	-0.68	-1.36	-4.04*	0.91	-1.13	0.53	-6.69*	-1.13	-2.62*	-4.51*	-2.96*	-3.16*	-2.97*	-4.55*	-4.58*
Spec 5:IGA_N	2.48*	3.13*	4.09*	2.16*	6.93*	2.32*	4.26*	3.65*	-0.78	6.36*	3.90*	6.37*	-6.17*	3.90*	0.06	-0.12	-1.03	-1.30	-1.52	-3.75*	-3.79*
Spec 6:AGA_N	-6.00*	0.47	1.88	-6.07*	-1.99*	0.62	1.87	1.03	-4.85*	4.40*	0.71	4.49*	-6.82*	0.71	-2.41*	-5.10*	-3.09*	-3.37*	-3.00*	-4.75*	-4.78*
Spec 7:$NAGA_N$	-3.37*	-2.11*	2.13*	-3.43*	-1.51	-0.88	-0.17	-1.25	-2.92*	0.64	-0.65	0.46	-6.25*	-0.65	-1.96*	-3.18*	-2.31*	-2.47*	-2.44*	-3.99*	-4.02*
Spec 8:EGA_N	-3.82*	-0.42	1.06	-3.88*	-1.83	0.04	1.93	0.08	-3.71*	2.13*	-0.69	1.91	-7.01*	-0.69	-2.98*	-3.68*	-3.09*	-3.28*	-3.30*	-5.05*	-5.08*
Spec 9:$APGA_N$	-4.75*	-1.17	2.35*	-4.86*	-1.98*	0.04	0.43	0.19	-3.69*	2.22*	-0.49	1.70	-6.54*	-0.49	-2.21*	-4.06*	-2.64*	-2.84*	-2.69*	-4.34*	-4.36*
Spec 10:$FIGA_N$	1.03	2.46*	3.23*	0.94	4.15*	2.00*	3.65*	2.82*	-0.43	4.93*	4.28*	5.04*	-6.40*	4.28*	0.33	0.50	-1.16	-1.58	-1.69	-4.11*	-4.15*
Spec 11:A_t	3.22*	3.95*	4.60*	3.19*	4.73*	3.51*	5.28*	4.25*	2.82*	5.70*	5.61*	5.77*	-5.39*	5.61*	3.34*	3.07*	2.25*	2.08*	1.07	-2.38*	-2.45*
Spec 12:GA_t	1.78	3.09*	4.11*	-0.24	6.60*	2.27*	4.02*	3.66*	-1.27	6.34*	3.47*	6.28*	-6.17*	3.47*	-0.14	-0.98	-1.18	-1.44	-1.60	-3.76*	-3.79*
Spec 13:GJR_t	-2.92*	4.59*	4.19*	-3.01*	-0.56	0.98	0.98	3.46*	-2.53*	2.05*	0.10	1.75	-6.15*	0.10	-1.51	-2.68*	-1.98*	-2.15*	-2.16*	-3.83*	-3.86*
Spec 14:TGA_t	-4.29*	-3.34*	1.67	-4.37*	-2.30*	-1.06	-0.72	-3.27*	-3.51*	0.29	-1.04	0.09	-6.49*	-1.04	-2.37*	-3.87*	-2.68*	-2.86*	-2.75*	-4.30*	-4.33*
Spec 15:IGA_t	1.79	3.17*	4.17*	1.75	6.78*	2.30*	4.09*	3.73*	-1.17	6.37*	3.47*	6.30*	-6.15*	3.47*	-0.11	-0.83	-1.14	-1.40	-1.58	-3.73*	-3.77*
Spec 16:AGA_t	-6.57*	0.79	2.26*	-6.72*	1.23	0.82	2.16*	1.40	-4.50*	4.89*	0.99	4.84*	-6.69*	0.99	-2.04*	-4.98*	-2.78*	-3.06*	-2.76*	-4.56*	-4.59*
Spec 17:$NAGA_t$	-2.13*	-0.03	2.59*	-2.17*	-0.53	3.16*	0.55	0.89	-2.02*	1.21	-0.04	1.05	-5.87*	-0.04	-1.36	-2.12*	-1.74	-1.87	-1.96*	-3.54*	-3.57*
Spec 18:EGA_t	-4.24*	-0.87	0.95	-4.31*	-2.32*	-0.18	-0.06	-0.28	-3.78*	1.73	-0.94	1.35	-6.88*	-0.94	-2.81*	-3.90*	-3.00*	-3.19*	-3.14*	-4.83*	-4.85*
Spec 19:$APGA_t$	-3.21*	-0.08	4.46*	-3.29*	-0.82	0.70	0.77	5.06*	-2.72*	1.88	-0.05	1.58	-6.20*	-0.05	-1.64	-2.93*	-2.11*	-2.28*	-2.26*	-3.91*	-3.94*
Spec 20:$FIGA_t$	1.3	2.63*	3.38*	1.19	4.42*	2.11*	3.69*	3.00*	0.35	5.10*	4.19*	5.20*	-6.30*	4.19*	0.37	0.75	-0.97	-1.34	-1.54	-3.95*	-3.99*
Spec 21:A_{st}	3.25*	3.99*	4.66*	3.22*	4.82*	3.53*	5.40*	4.30*	2.83*	5.83*	5.75*	5.90*	-5.42*	5.75*	3.38*	3.09*	2.22*	2.06*	1.03	-2.43*	-2.49*
Spec 22:GA_{st}	-	3.08*	4.11*	-1.05	6.57*	2.26*	4.00*	3.66*	-1.32	6.32*	3.42*	6.26*	-6.17*	3.42*	-0.17	-1.07	-1.19	-1.46	-1.61	-3.76*	-3.80*
Spec 23:GJR_{st}	-	-	3.78*	-3.18*	-0.79	0.69	0.78	2.47*	-2.66*	1.81	-0.04	1.53	-6.18*	-0.04	-1.62	-2.82*	-2.07*	-2.24*	-2.23*	-3.88*	-3.91*
Spec 24:TGA_{st}	-	-	-	-4.18*	-2.26*	-1.45	-0.85	-3.90*	-3.42*	0.04	-1.10	-0.11	-6.45*	-1.10	-2.35*	-3.74*	-2.65*	-2.82*	-2.73*	-4.25*	-4.28*
Spec 25:IGA_{st}	-	-	-	-	6.72*	2.30*	4.06*	3.74*	-1.22	6.37*	3.42*	6.29*	-6.15*	3.42*	-0.13	-0.91	-1.15	-1.41	-1.58	-3.73*	-3.77*
Spec 26:AGA_{st}	-	-	-	-	-	0.82	2.16*	1.40	-4.50*	4.89*	0.99	4.84*	-6.69*	0.99	-2.04*	-4.98*	-2.78*	-3.06*	-2.76*	-4.56*	-4.59*
Spec 27:$NAGA_{st}$	-	-	-	-	-	-	0.16	0.02	-2.13*	0.74	-0.29	0.61	-5.88*	-0.29	-1.50	-2.23*	-1.84	-1.97*	-2.04*	-3.57*	-3.61*
Spec 28:EGA_{st}	-	-	-	-	-	-	-	-0.24	-3.71*	1.57	-0.96	1.27	-6.95*	-0.96	-2.92*	-3.77*	-3.03*	-3.21*	-3.22*	-4.94*	-4.96*
Spec 29:$APGA_{st}$	-	-	-	-	-	-	-	-	-3.04*	1.28	-0.43	1.02	-6.30*	-0.43	-1.91	-3.31*	-2.33*	-2.50*	-2.44*	-4.05*	-4.07*
Spec 30:$FIGA_{st}$	-	-	-	-	-	-	-	-	-	5.16*	4.17*	5.25*	-6.29*	4.17*	0.36	0.75	-0.96	-1.32	-1.53	-3.93*	-3.97*
Spec 31:$comb_1$	-	-	-	-	-	-	-	-	-	-	-2.35*	-1.50	-7.00*	-2.35*	-3.47*	-5.68*	-3.67*	-3.89*	-3.52*	-5.02*	-5.04*
Spec 32:$comb_2$	-	-	-	-	-	-	-	-	-	-	-	2.46*	-7.20*	-2.42*	-3.75*	-3.81*	-3.92*	-4.17*	-3.82*	-5.43*	-5.45*
Spec 33:$comb_3$	-	-	-	-	-	-	-	-	-	-	-	-	-7.04*	-2.46*	-3.56*	-5.73*	-3.75*	-3.98*	-3.58*	-5.08*	-5.10*
Spec 34:HAR	-	-	-	-	-	-	-	-	-	-	-	-	-	7.20*	7.03*	6.19*	6.69*	6.57*	6.85*	4.95*	4.90*
Spec 35:logHAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-3.75*	-3.81*	-3.92*	-4.17*	-3.82*	-5.43*	-5.45*
Spec 36:HEAVY	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-0.09	-1.42	-1.62	-2.81*	-5.47*	-5.50*
Spec 37:$GA-MIDAS$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-1.04	-1.33	-1.53	-3.79*	-3.82*
Spec 38:$GA-RV_N$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-2.57*	-1.88	-4.84*	-4.88*
Spec 39:$GA-RV_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-1.44	-4.55*	-4.59*
Spec 40:$GA-RV_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-6.51*	-6.52*
Spec 41:$quant_1$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-3.39*
Spec 42:$quant_2$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Note: This table reports the out-of-sample t-statistic of the Unconditional Giacomini and White (2006) test under the MSE loss function for S&P 500. The period covers is 02/01/2004 to 29/12/2017. A positive t-statistic indicates that average loss of model A (in row) is larger than of model B (in column) and hence model B is superior. *denotes significance at the 5% level

Table 3.5 – Superior Predictive Ability (SPA) test and Model Confidence Set (MCS) results for the S&P 500 (4-year in-sample) - MSE, Robust Patton (b=1), Robust Patton (b=-0.5)

	MSE - Robust Patton (b=1) - Robust Patton (b=-0.5)					
	P-value (Hansen)	R-statistic	P-value (Hansen)	R-statistic	P-value (Hansen)	R-statistic
Spec 1:A_N	0.03*	✓	0.09	✓	≤ 0.00*	
Spec 2:GA_N	0.04*	✓	0.11	✓	0.01*	
Spec 3:GJR_M	0.23	✓	0.28	✓	0.33	✓
Spec 4:TGA_N	0.84	✓	0.87	✓	0.27	✓
Spec 5:IGA_N	0.03*	✓	0.12	✓	0.01*	
Spec 6:AGA_N	0.04*	✓	0.11	✓	0.01*	
Spec 7:$NAGA_N$	0.28	✓	0.60	✓	0.24	✓
Spec 8:EGA_N	0.24	✓	0.37	✓	0.04*	✓
Spec 9:$APGA_N$	0.25	✓	0.33	✓	0.25	✓
Spec 10:$FIGA_N$	0.05	✓	0.12	✓	0.02*	
Spec 11:A_t	≤ 0.00*	✓	0.09	✓	≤ 0.00*	
Spec 12:GA_t	0.06	✓	0.15	✓	0.01*	
Spec 13:GJR_t	0.16	✓	0.17	✓	0.18	✓
Spec 14:TGA_t	0.41	✓	0.61	✓	0.43	✓
Spec 15:IGA_t	0.05	✓	0.11	✓	0.01*	
Spec 16:AGA_t	0.06	✓	0.13	✓	0.01*	
Spec 17:$NAGA_t$	0.14	✓	0.18	✓	0.12	✓
Spec 18:EGA_t	0.23	✓	0.84	✓	0.02*	✓
Spec 19:$APGA_t$	0.18	✓	0.20	✓	0.19	✓
Spec 20:$FIGA_t$	0.05	✓	0.17	✓	0.01*	
Spec 21:A_{st}	0.01*	✓	0.06	✓	≤ 0.00*	
Spec 22:GA_{st}	0.05	✓	0.11	✓	0.01*	
Spec 23:GJR_{st}	0.18	✓	0.19	✓	0.20	✓
Spec 24:TGA_{st}	0.94	✓	1.00	✓	0.78	✓
Spec 25:IGA_{st}	0.04*	✓	0.12	✓	≤ 0.00*	
Spec 26:AGA_{st}	0.06	✓	0.10	✓	0.02*	
Spec 27:$NAGA_{st}$	0.40	✓	0.67	✓	0.24	✓
Spec 28:EGA_{st}	0.33	✓	0.79	✓	0.04*	✓
Spec 29:$APGA_{st}$	0.19	✓	0.21	✓	0.25	✓
Spec 30:$FIGA_{st}$	0.06	✓	0.14	✓	0.01*	
Spec 31:$comb_1$	1.00	✓	0.94	✓	1.00	✓
Spec 32:$comb_2$	0.13	✓	0.38	✓	0.06	✓
Spec 33:$comb_3$	0.46	✓	0.59	✓	0.23	✓
Spec 34:HAR	0.01*	✓	0.06	✓	≤ 0.00*	
Spec 35:logHAR	0.14	✓	0.37	✓	0.05	✓
Spec 36:HEAVY	0.05	✓	0.19	✓	≤ 0.00*	
Spec 37:$GA-MIDAS$	0.06	✓	0.14	✓	0.01*	
Spec 38:$GA-RV_N$	0.05	✓	0.11	✓	0.01*	
Spec 39:$GA-RV_t$	0.06	✓	0.16	✓	0.02*	
Spec 40:$GA-RV_{st}$	0.07	✓	0.16	✓	0.02*	
Spec 41:$quant_1$	0.07	✓	0.15	✓	0.03*	
Spec 42:$quant_2$	0.06	✓	0.17	✓	0.03*	

Note: This table reports the out-of-sample P-values of the Superior Predictive Ability test of Hansen (2005) under the MSE loss function, the robust loss of Patton (2011) for b=1 and for b=-0.5 for S&P 500. The period covers is 02/01/2004 to 29/12/2017. Each row name indicates the benchmark which is compared to each of the alternative models. In the same table the models that are included in the Model Confidence set (MCS) of Hansen (2011) using the range statistic are indicated by the "tick" symbol. The White's Upper P-value (Original RC P-values) of the SPA test and the SQ-method of the MCS are reported in the Appendix. *significant at 5% level

3.4.1 Robustness - Sensitivity to the loss function

Table 3.6 reports the out-of-sample results for the GW test under the loss function for the robust family of Patton (2011) for b=1 at 5% significance level. One general observation is that there are fewer significant cases overall. Once more the specifications of the ARCH models and the SD approximations are among the worst performing, the latter have 1 win and 19 loses and the all the ARCH specification have zero wins apart from the Skew-t one which outperforms its student specification. Similarly, the HAR model has no cases where it outperforms any competing specification and 34 cases where it is outperformed. In a similar way, as with the MSE loss function, all the IGARCH, the FIGARCH, the GARCH, the GARCH-RV specifications and the GARCH-MIDAS, the HEAVY models are also among the worst performing. Table 3.7 reports the out-of-sample results for GW under the loss function from Patton (2011)'s robust family for b=-0.5. It correspondingly verifies that the same specifications are performing better than the three ARCH, the three GARCH, the three IGARCH and the three FIGARCH specifications, the HAR, the GARCH-MIDAS, the HEAVY. However, the number of cases that the three GARCH-RV and the two SD approximation were not preferred over is almost double times for when b=-0.5 compared to b=1. The reason for that is that there are more significant cases.

Following the table 3.6 the three TGARCH specifications again they have no loses. The same holds for all three forecast combination schemes. Except for the second combination scheme that is outperformed by the first and the third scheme. Similar results are obtained for logHAR as well, since it is only outperformed by the 1st and the 2nd combination. The NAGARCH-St and the EGARCH-St tested for the robust loss function of Patton (b=1) have zero cases where they are outperformed however the NAGARCH-St is beaten by the TGARCH-St and the EGARCH-St, by the first and the second combinations for the robust loss of Patton (b=-0.5). An important difference is than the TGARCH-N under the robust loss of Patton (b=-0.5), it is outperformed by the first combination.

The rest of the models are also almost the same. One difference is that the APGARCH-N when evaluated using the robust loss b=1 it is significantly better than the AGARCH-t. AGARCH-st model it seems that only when evaluated under the MSE or the robust loss of Patton (b=1) outperforms the EGARCH-St, in terms of the rest of the specifications it seems that all the examined loss function deduce the same outcome. At last, the APGARCH-St for the robust loss of Patton (b=1) is "winning" in 20 cases but is losing in the cases where it is compared to the TGARCH-t and the TGARCH-St, along with the first combination for the robust loss of Patton (b=-0.5). Similarly, only the MSE and the robust loss (b=-0.5) suggest that the NAGARCH-N outperforms its Student-t specification, then the set of the specification the certain model "wins" is almost identical. What is more, the Table 3.6 reports that as in the case of the evaluation under the MSE the EGARCH-t is superior than both the

AGARCH-t and the AGARCH-St, while under the robust loss of Patton (b=-0.5) that relationship applies only for the AGARCH-St model. The GJR-N is "losing" in 3 cases and for the robust loss of Patton (b=1) and in 5 for the robust loss of Patton (b=-0.5). But GJR-ST (b=-0.5) is also "losing" to the best models. Finally for b=1, GJR-t outperform GJR-St but it is indifferent to the Normal specification, while, for b=-0.5 GJR-t outperforms both of it's other specifications.

To conclude, the MSE and the robust loss of Patton (b=1), of the second sample (from 02/01/2001 to 29/12/2017) agree with the robust loss of Patton (b=1) of the first sample (from 02/01/2004 to 29/12/2017), that all the specifications must be included in the MCS. On the contrary, the robust loss of Patton (b=-0.5) reports that only the three GJR, the three TGARCH, the three NAGARCH, the three EGARCH, the three APGARCH specifications and the three combinations together with the HAR model should be contained. For both samples [5], the SPA of the b=-0.5 agrees more with the MSE but it adds more models to the list that the null hypothesis could not be rejected. These are almost all the models that are not included in the MCS. Surprisingly when b=1, there are no models that reject the null of the SPA test at 5% significance level.

The results for the MSE loss function and the robust loss functions of Patton when $b = 1$ and $b = -0.5$ are displayed exhaustively. In the Appendix, Table 3.18 reports only the aggregated results of the total number of "wins" and "losses" for each specification for the second sample of the S&P 500 index for a number of non-robust loss functions (namely the logMSE, the sdMSE, the propMSE, the RMSE, the propRMSE, the MAE, the logMAE, the sdMAE, the MAPE and the GMLE). We count as a "win" ("loss") the case where a certain specification outperformed (was outperformed) significantly (by) a competing specification. It should be noted that the number of the specifications compared is 42 thus the overall number of comparisons is 41. It can be observed that most functions indicate that the logHAR, and the combinations have the highest number of wins. However, unexpectedly, HEAVY and GARCH-RV were the best-performing ones, under RMSE and GMLE while GARCH-RV-St had the highest win score under prop-MSE. From table 3.19 [6] it is possible to observe that for all the other non-robust loss-functions the results are uniformed and not far from the ones derived from the robust ones. It seems like the logHAR and the TGARCH specifications have a good performance under any loss but APGARCH have also a high ratio of "wins".

[5] The first sample being from 02/01/2004 to 29/12/2017 and the second one from from 02/01/2001 to 29/12/2017.
[6] Table 3.19 can also be found the Appendix.

Table 3.6 – Pairwise comparisons of unconditional Giacomini and White test for S&P 500 (4-year in-sample) -robust Patton (b=1)

	Spec 1	Spec 2	Spec 3	Spec 4	Spec 5	Spec 6	Spec 7	Spec 8	Spec 9	Spec 10	Spec 11	Spec 12	Spec 13	Spec 14	Spec 15	Spec 16	Spec 17	Spec 18	Spec 19	Spec 20	Spec 21
Spec 1:A_N	-	1.87	2.61*	2.91*	1.67	2.55*	2.65*	2.78*	2.67*	1.49	-1.83	1.73	2.31*	2.86*	1.72	2.46*	2.19*	2.87*	2.40*	1.39	-1.07
Spec 2:GA_N	-	-	2.80*	3.57*	-3.77*	5.06*	2.18*	3.15*	2.98*	-1.81	-2.37*	-1.52	1.57	2.87*	-1.62	4.86*	1.22	3.40*	1.76	-2.30*	-2.35*
Spec 3:GJR_M	-	-	-	2.49*	-3.47*	-0.87	1.17	0.18	1.89	-2.78*	-3.05*	-3.57*	-1.13	2.39*	-3.61*	-1.35	-0.03	0.88	-0.66	-3.07*	-3.06*
Spec 4:TGA_N	-	-	-	-	-4.15*	-1.95	0.08	-0.99	-2.41*	-3.30*	-3.36*	-4.21*	-2.01*	0.73	-4.24*	-2.46*	-0.66	-0.56	-1.85	-3.52*	-3.39*
Spec 5:IGA_N	-	-	-	-	-	5.25*	2.59*	3.47*	3.65*	-0.65	-2.17*	1.43	2.06*	3.36*	1.10	5.49*	1.52	3.80*	2.27*	-1.23	-2.13*
Spec 6:AGA_N	-	-	-	-	-	-	1.07	1.37	1.07	-4.03*	-3.06*	-4.77*	0.10	1.59	-4.77*	-2.17*	0.31	1.87	0.29	-4.16*	-3.09*
Spec 7:$NAGA_N$	-	-	-	-	-	-	-	-0.53	-0.99	-2.27*	-3.03*	-2.62*	-2.78*	0.55	-2.65*	-1.32	-1.69	-0.28	-2.59*	-2.48*	-3.02*
Spec 8:EGA_N	-	-	-	-	-	-	-	-	-0.01	-3.27*	-3.29*	-3.25*	-0.55	0.95	-3.26*	-1.66	-0.10	1.67	-0.37	-3.32*	-3.33*
Spec 9:$APGA_N$	-	-	-	-	-	-	-	-	-	-2.91*	-3.11*	-3.76*	-1.35	2.17*	-3.80*	-1.58	-0.14	0.69	-0.94	-3.19*	-3.13*
Spec 10:$FIGA_N$	-	-	-	-	-	-	-	-	-	-	-1.99*	0.86	1.81	2.80*	0.79	3.62*	1.43	3.33*	1.95	-1.77	-1.94
Spec 11:A_t	-	-	-	-	-	-	-	-	-	-	-	2.23*	2.71*	3.27*	2.21*	2.96*	2.53*	3.36*	2.80*	1.88	2.39*
Spec 12:GA_t	-	-	-	-	-	-	-	-	-	-	-	-	2.08*	3.42*	-2.82*	5.22*	1.50	3.66*	2.30*	-1.43	-2.19*
Spec 13:GJR_t	-	-	-	-	-	-	-	-	-	-	-	-	-	3.43*	-2.12*	-0.35	0.55	1.04	1.54	-2.06*	-2.69*
Spec 14:TGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-3.45*	-1.94	-1.24	-0.71	-3.56*	-3.02*	-3.29*
Spec 15:IGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	5.23*	1.52	3.68*	2.34*	-1.36	-2.17*
Spec 16:AGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.47	2.21*	0.56	-3.90*	-2.98*
Spec 17:$NAGA_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.35	-0.30	-1.59	-2.49*
Spec 18:EGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-0.85	-3.43*	-3.41*
Spec 19:$APGA_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-2.20*	-2.79*
Spec 20:$FIGA_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-1.82
Spec 21:A_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 22:GA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 23:GJR_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 24:TGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 25:IGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 26:AGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 27:$NAGA_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 28:EGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 29:$APGA_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 30:$FIGA_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 31:$comb_1$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 32:$comb_2$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 33:$comb_3$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 34:HAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 35:logHAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 36:HEAVY	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 37:$GA-MIDAS$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 38:$GA-RV_N$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 39:$GA-RV_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 40:$GA-RV_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 41:$quant_1$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 42:$quant_2$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Note: This table reports the out-of-sample t-statistic of the Unconditional Giacomini and White (2006) test under the robust loss of Patton (2011) for b=1 for S&P 500. The period is from 02/01/2004 to 29/12/2017. A positive t-statistic indicates that average loss of model A (in row) is larger than of model B (in column) and hence model B is superior. *denotes significance at the 5% level

Table 3.6 cont. - Pairwise comparisons of unconditional Giacomini and White test for S&P 500 (4-year in-sample) - robust Patton (b=1)

	Spec 22	Spec 23	Spec 24	Spec 25	Spec 26	Spec 27	Spec 28	Spec 29	Spec 30	Spec 31	Spec 32	Spec 33	Spec 34	Spec 35	Spec 36	Spec 37	Spec 38	Spec 39	Spec 40	Spec 41	Spec 42
Spec 1:A_N	1.74	2.37*	2.86*	1.73	2.46*	2.26*	2.86*	2.57*	1.38	3.02*	2.67*	3.00*	-1.88	2.67*	2.04*	1.58	1.49	1.36	1.45	0.82	0.82
Spec 2:GA_N	-1.36	1.71	2.73*	-1.43	4.86*	1.33	3.24*	2.15*	-2.34*	4.57*	2.85*	4.63*	-2.56*	2.85*	0.60	-2.00*	-1.01	-1.48	-0.66	-1.69	-1.69
Spec 3:GJR_M	-3.56*	-0.72	2.25*	-3.62*	-1.35	0.33	0.66	0.45	-3.12*	1.48	-0.47	1.08	-2.83*	-0.47	-1.22	-3.19*	-2.08*	-2.32*	-1.74	-2.41*	-2.41*
Spec 4:TGA_N	-4.20*	-1.73	0.89	-4.24*	-2.46*	-0.22	-0.61	-1.10	-3.56*	-0.03	-1.26	-0.29	-2.93*	-1.26	-1.78	-3.79*	-2.57*	-2.81*	-2.20*	-2.85*	-2.85*
Spec 5:IGA_N	1.43	2.21*	3.18*	1.13	5.49*	1.60	3.60*	2.66*	-1.28	5.01*	3.10*	5.00*	-2.50*	3.10*	0.98	-0.81	-0.41	-0.81	-0.27	-1.29	-1.29
Spec 6:AGA_N	-4.69*	0.26	1.57	-4.69*	-2.17*	0.53	1.71	0.74	-4.18*	3.23*	0.25	3.26*	-2.79*	0.25	-1.36	-4.11*	-2.78*	-3.15*	-2.08*	-2.92*	-2.91*
Spec 7:$NAGA_N$	-2.62*	-2.45*	1.33	-2.65*	-1.32	-0.53	-0.33	-1.69	-2.51*	-0.07	-0.80	-0.18	-2.87*	-0.80	-1.27	-2.61*	-1.86	-2.03*	-1.65	-2.19*	-2.19*
Spec 8:EGA_N	-3.21*	-0.39	1.00	-3.23*	-1.66	0.16	1.99*	0.05	-3.32*	1.42	-1.01	1.15	-2.89*	-1.01	-2.03*	-3.19*	-2.83*	-3.04*	-2.61*	-3.47*	-3.47*
Spec 9:$APGA_N$	-3.76*	-0.97	2.04*	-3.81*	-1.58	0.23	0.48	0.16	-3.24*	1.29	-0.60	0.89	-2.84*	-0.60	-1.32	-3.34*	-2.18*	-2.42*	-1.82	-2.49*	-2.49*
Spec 10:$FIGA_N$	0.88	1.92	2.71*	0.82	3.62*	1.51	3.27*	2.25*	-1.64	4.14*	3.50*	4.21*	-2.47*	3.50*	1.41	0.06	-0.11	-0.69	-0.06	-1.29	-1.29
Spec 11:A_t	2.23*	2.77*	3.26*	2.22*	2.96*	2.57*	3.35*	2.97*	1.87	3.51*	3.22*	3.50*	-1.75	3.22*	2.58*	2.08*	2.01*	1.88	1.97*	1.32	1.31
Spec 12:GA_t	1.10	2.23*	3.24*	-0.41	5.22*	1.59	3.44*	2.72*	-1.50	4.94*	2.74*	4.87*	-2.51*	2.74*	0.79	-1.24	-0.53	-0.90	-0.36	-1.32	-1.32
Spec 13:GJR_t	-2.08*	4.98*	3.79*	-2.12*	-0.35	0.93	0.89	3.31*	-2.10*	1.36	0.00	1.13	-2.73*	0.00	-0.69	-2.08*	-1.42	-1.62	-1.21	-1.82	-1.82
Spec 14:TGA_t	-3.42*	-3.10*	1.30	-3.45*	-1.94	-0.55	-0.73	-2.96*	-3.05*	-0.38	-1.15	-0.51	-2.93*	-1.15	-1.61	-3.22*	-2.26*	-2.46*	-1.98*	-2.57*	-2.57*
Spec 15:IGA_t	3.03*	2.27*	3.27*	1.18	5.23*	1.61	3.46*	2.75*	-1.43	4.95*	2.75*	4.88*	-2.51*	2.75*	0.81	-1.14	-0.50	-0.86	-0.34	-1.30	-1.30
Spec 16:AGA_t	-5.14*	0.52	1.88	-5.14*	1.45	0.67	2.01*	1.04	-3.95*	3.77*	0.63	3.71*	-2.76*	0.63	-0.97	-4.02*	-2.38*	-2.75*	-1.76	-2.60*	-2.60*
Spec 17:$NAGA_t$	-1.50	-0.33	1.65	-1.52	-0.47	2.11*	0.29	0.30	-1.61	0.51	-0.21	0.40	-2.71*	-0.21	-0.70	-1.61	-1.22	-1.36	-1.10	-1.60	-1.60
Spec 18:EGA_t	-3.63*	-0.86	0.81	-3.66*	-2.21*	-0.03	-0.76	-0.35	-3.45*	0.63	-1.31	0.25	-2.92*	-1.31	-2.01*	-3.45*	-2.76*	-2.98*	-2.49*	-3.25*	-3.25*
Spec 19:$APGA_t$	-2.30*	-0.17	3.95*	-2.34*	-0.56	0.73	0.70	4.49*	-2.24*	1.16	-0.13	0.94	-2.75*	-0.13	-0.81	-2.27*	-1.53	-1.73	-1.31	-1.92	-1.92
Spec 20:$FIGA_t$	1.46	2.17*	2.91*	1.39	3.90*	1.66	3.35*	2.50*	-0.84	4.29*	3.45*	4.34*	-2.43*	3.45*	1.52	0.66	0.25	-0.20	0.14	-1.01	-1.01
Spec 21:A_{st}	2.19*	2.76*	3.27*	2.18*	2.98*	2.53*	3.40*	2.97*	1.81	3.56*	3.26*	3.56*	-1.81	3.26*	2.57*	2.03*	1.96*	1.82	1.91	1.22	1.22
Spec 22:GA_{st}	-	2.23*	3.24*	-2.26*	5.14*	1.59	3.41*	2.72*	-1.53	4.91*	2.69*	4.84*	-2.51*	2.69*	0.77	-1.29	-0.54	-0.91	-0.37	-1.33	-1.33
Spec 23:GJR_{st}	-	-	3.49*	-2.27*	-0.52	0.75	0.72	2.51*	-2.22*	1.17	-0.11	0.95	-2.75*	-0.11	-0.79	-2.20*	-1.51	-1.71	-1.29	-1.90	-1.90
Spec 24:TGA_{st}	-	-	-	-3.27*	-1.88	-0.80	-0.82	-3.51*	-2.95*	-0.53	-1.17	-0.62	-2.94*	-1.17	-1.60	-3.09*	-2.21*	-2.4*	-1.95	-2.51*	-2.51*
Spec 25:IGA_{st}	-	-	-	-	5.14*	1.61	3.43*	2.76*	-1.46	4.93*	2.70*	4.85*	-2.51*	2.70*	0.79	-1.20	-0.51	-0.87	-0.35	-1.30	-1.30
Spec 26:AGA_{st}	-	-	-	-	-	0.67	2.01*	1.04	-3.95*	3.77*	0.63	3.71*	-2.76*	0.63	-0.97	-4.02*	-2.38*	-2.75*	-1.76	-2.60*	-2.60*
Spec 27:$NAGA_{st}$	-	-	-	-	-	-	-0.00	-0.25	-1.68	0.17	-0.41	0.09	-2.75*	-0.41	-0.83	-1.68	-1.31	-1.44	-1.19	-1.67	-1.67
Spec 28:EGA_{st}	-	-	-	-	-	-	-	-0.24	-3.36*	0.76	-1.27	0.42	-2.92*	-1.27	-2.06*	-3.31*	-2.78*	-2.99*	-2.55*	-3.35*	-3.35*
Spec 29:$APGA_{st}$	-	-	-	-	-	-	-	-	-2.54*	0.64	-0.48	0.45	-2.81*	-0.48	-1.08	-2.62*	-1.78	-1.98*	-1.53	-2.13*	-2.13*
Spec 30:$FIGA_{st}$	-	-	-	-	-	-	-	-	-	4.32*	3.43*	4.36*	-2.43*	3.43*	1.52	0.73	0.27	-0.16	0.16	-0.98	-0.98
Spec 31:$comb_1$	-	-	-	-	-	-	-	-	-	-	-2.07*	-1.60	-2.95*	-2.07*	-2.32*	-4.50*	-3.21*	-3.47*	-2.67*	-3.35*	-3.35*
Spec 32:$comb_2$	-	-	-	-	-	-	-	-	-	-	-	2.14*	-2.83*	1.21	-2.02*	-3.07*	-3.48*	-3.74*	-2.81*	-3.69*	-3.69*
Spec 33:$comb_3$	-	-	-	-	-	-	-	-	-	-	-	-	-2.94*	-2.14*	-2.34*	-4.50*	-3.29*	-3.56*	-2.71*	-3.41*	-3.41*
Spec 34:HAR	-	-	-	-	-	-	-	-	-	-	-	-	-	2.83*	2.67*	2.46*	2.48*	2.44*	2.50*	2.28*	2.28*
Spec 35:logHAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-2.02*	-3.07*	-3.48*	-3.74*	-2.81*	-3.69*	-3.69*
Spec 36:HEAVY	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-1.14	-2.05*	-2.30*	-2.34*	-3.85*	-3.84*
Spec 37:$GA-MIDAS$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-0.11	-0.49	-0.07	-1.09	-1.09
Spec 38:$GA-RV_N$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-3.09*	-0.00	-1.85	-1.85
Spec 39:$GA-RV_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.43	-1.36	-1.36
Spec 40:$GA-RV_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-5.20*	-5.19*
Spec 41:$quant_1$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-0.50
Spec 42:$quant_2$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Note: This table reports the out-of-sample t-statistic of the Unconditional Giacomini and White (2006) test under the robust loss of Patton (2011) for b=1 for S&P 500.for S&P 500. The period is from 02/01/2004 to 29/12/2017. A positive t-statistic indicates that average loss of model A (in row) is larger than of model B (in column) and hence model B is superior. *denotes significance at the 5% level

Table 3.7 – Pairwise comparisons of unconditional Giacomini and White test for S&P 500 (4-year in-sample) - robust Patton (b=-0.5)

	Spec 1	Spec 2	Spec 3	Spec 4	Spec 5	Spec 6	Spec 7	Spec 8	Spec 9	Spec 10	Spec 11	Spec12	Spec 13	Spec 14	Spec 15	Spec 16	Spec 17	Spec 18	Spec 19	Spec 20	Spec 21
Spec 1:A_N	-	4.18*	6.02*	6.58*	3.64*	6.58*	5.45*	6.99*	6.12*	3.38*	-7.21*	3.87*	5.13*	6.09*	3.82*	6.26*	4.59*	6.67*	5.32*	3.38*	-7.61*
Spec 2:GA_N	-	-	4.60*	5.34*	-3.30*	7.32*	3.71*	4.31*	4.94*	-1.91	-5.65*	-0.90	3.20*	4.54*	-1.15	6.74*	2.51*	4.48*	3.56*	-1.98*	-5.81*
Spec 3:GJR_M	-	-	-	1.96*	-5.83*	-1.70	0.59	-0.30	0.23	-4.10*	-6.98*	-5.50*	-1.98*	2.28*	-5.93*	-2.28*	-0.82	0.01	-1.21	-4.23*	-7.18*
Spec 4:TGA_N	-	-	-	-	-6.35*	-2.78*	-0.65	-1.39	-2.51*	-4.60*	-7.50*	-6.12*	-2.31*	0.57	-6.42*	-3.35*	-1.66	-1.40	-2.40*	-4.72*	-7.71*
Spec 5:IGA_N	-	-	-	-	-	8.26*	4.46*	4.92*	6.08*	-0.49	-5.10*	2.64*	4.16*	5.44*	2.98*	8.48*	3.13*	5.25*	4.51*	-0.42	-5.26*
Spec 6:AGA_N	-	-	-	-	-	-	1.44	1.51	1.89	-5.48*	-7.79*	-7.22*	0.59	2.26*	-7.75*	-1.37	0.43	1.86	0.85	-5.60*	-8.03*
Spec 7:$NAGA_N$	-	-	-	-	-	-	-	-0.53	-0.57	-3.42*	-6.31*	-4.36*	-1.90	1.58	-4.49*	-1.75	-3.25*	-0.37	-2.08*	-3.53*	-6.47*
Spec 8:EGA_N	-	-	-	-	-	-	-	-	0.38	-4.24*	-8.15*	-4.41*	-0.35	1.24	-4.60*	-1.72	-0.29	1.01	-0.17	-4.18*	-8.42*
Spec 9:$APGA_N$	-	-	-	-	-	-	-	-	-	-4.26*	-7.07*	-5.96*	-1.61	2.42*	-6.26*	-2.49*	-0.92	-0.06	-1.48	-4.42*	-7.26*
Spec 10:$FIGA_N$	-	-	-	-	-	-	-	-	-	-	-4.99*	1.15	3.08*	4.04*	0.99	4.99*	2.50*	4.19*	3.30*	0.43	-5.12*
Spec 11:A_t	-	-	-	-	-	-	-	-	-	-	-	5.26*	6.07*	6.96*	5.21*	7.47*	5.45*	7.73*	6.24*	4.94*	-2.30*
Spec 12:GA_t	-	-	-	-	-	-	-	-	-	-	-	-	3.89*	5.28*	-0.84	7.46*	2.98*	4.79*	4.35*	-1.20	-5.41*
Spec 13:GJR_t	-	-	-	-	-	-	-	-	-	-	-	-	-	3.49*	-4.16*	-0.90	-0.01	0.72	0.87	-3.18*	-6.23*
Spec 14:TGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-5.52*	-2.70*	-2.65*	-1.24	-4.10*	-4.16*	-7.15*
Spec 15:IGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	8.32*	3.08*	5.04*	4.57*	-1.00	-5.37*
Spec 16:AGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.63	2.15*	1.19	-5.17*	-7.70*
Spec 17:$NAGA_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.57	0.43	-2.58*	-5.58*
Spec 18:EGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-0.51	-4.19*	-7.97*
Spec 19:$APGA_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-3.43*	-6.40*
Spec 20:$FIGA_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-5.07*
Spec 21:A_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 22:GA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 23:GJR_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 24:TGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 25:IGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 26:AGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 27:$NAGA_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 28:EGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 29:$APGA_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 30:$FIGA_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 31:$comb_1$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 32:$comb_2$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 33:$comb_3$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 34:HAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 35:logHAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 36:HEAVY	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 37:$GA-MIDAS$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 38:$GA-RV_N$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 39:$GA-RV_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 40:$GA-RV_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 41:$quant_1$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 42:$quant_2$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Note: This table reports the out-of-sample t-statistic of the Unconditional Giacomini and White (2006) test under the robust loss of Patton (2011) for b=-0.5 for S&P 500. The period is from 02/01/2004 to 29/12/2017. A positive t-statistic indicates that average loss of model A (in row) is larger than of model B (in column) and hence model B is superior. *denotes significance at the 5% level

Table 3.7 cont. -Pairwise comparisons of unconditional Giacomini and White test for S&P 500 (4-year in-sample) - robust Patton (b=-0.5)

	Spec 22	Spec 23	Spec 24	Spec 25	Spec 26	Spec 27	Spec 28	Spec 29	Spec 30	Spec 31	Spec 32	Spec 33	Spec 34	Spec 35	Spec 36	Spec 37	Spec 38	Spec 39	Spec 40	Spec 41	Spec 42
Spec 1:A_N	3.90*	5.24*	6.02*	3.84*	6.26*	4.68*	6.82*	5.64*	3.40*	7.80*	7.01*	7.88*	-8.95*	7.01*	2.51*	3.76*	1.36	1.29	-0.95	-5.00*	-5.11*
Spec 2:GA_N	-0.67	3.31*	4.38*	-0.95	6.74*	2.66*	4.28*	3.97*	-1.97*	7.58*	4.21*	7.68*	-9.32*	4.21*	-1.51	-1.43	-2.26*	-2.47*	-3.32*	-5.77*	-5.84*
Spec 3:GJR_M	-5.54*	-1.30	2.39*	-5.96*	-2.28*	-0.16	0.04	0.38	-4.28*	2.68*	-0.44	2.20*	-9.23*	-0.44	-3.23*	-4.74*	-3.25*	-3.39*	-3.87*	-5.92*	-5.98*
Spec 4:TGA_N	-6.15*	-1.91	0.89	-6.43*	-3.35*	-0.94	-1.12	-1.35	-4.77*	2.15*	-1.04	1.59	-9.39*	-1.04	-3.73*	-5.42*	-3.61*	-3.76*	-4.16*	-6.10*	-6.15*
Spec 5:IGA_N	2.97*	4.27*	5.22*	2.91*	8.48*	3.23*	4.97*	4.91*	-0.37	8.17*	4.30*	8.11*	-9.09*	4.30*	-0.98	0.44	-1.70	-1.88	-2.93*	-5.54*	-5.61*
Spec 6:AGA_N	-7.26*	0.80	2.25*	-7.67*	-1.37	0.75	1.78	1.44	-5.69*	5.76*	0.76	5.69*	-9.69*	0.76	-3.63*	-5.99*	-3.70*	-3.90*	-4.29*	-6.26*	-6.31*
Spec 7:$NAGA_N$	-4.36*	-1.42	2.25*	-4.50*	-1.75	-1.28	-0.31	-0.63	-3.57*	1.62	-0.60	1.33	-8.94*	-0.60	-3.04*	-4.01*	-3.02*	-3.16*	-3.65*	-5.69*	-5.76*
Spec 8:EGA_N	-4.39*	-0.16	1.33	-4.56*	-1.72	0.07	1.50	0.37	-4.19*	3.22*	-0.34	2.95*	-9.83*	-0.34	-3.99*	-4.16*	-3.60*	-3.73*	-4.41*	-6.43*	-6.47*
Spec 9:$APGA_N$	-5.99*	-1.05	2.39*	-6.30*	-2.49*	-0.23	-0.02	0.25	-4.49*	3.23*	-0.52	2.58*	-9.28*	-0.52	-3.39*	-5.09*	-3.37*	-3.53*	-3.96*	-5.96*	-6.02*
Spec 10:$FIGA_N$	1.21	3.18*	3.95*	1.03	4.99*	2.62*	4.13*	3.62*	0.56	6.22*	4.68*	6.33*	-9.55*	4.68*	-0.92	0.81	-2.07*	-2.32*	-3.35*	-5.92*	-5.98*
Spec 11:A_t	5.28*	6.16*	6.86*	5.22*	7.47*	5.50*	7.90*	6.53*	4.96*	8.71*	8.36*	8.83*	-8.59*	8.36*	4.22*	5.22*	2.87*	2.79*	0.20	-4.48*	-4.60*
Spec 12:GA_t	1.59	3.98*	5.07*	-0.49	7.46*	3.10*	4.50*	4.76*	-1.19	8.15*	3.97*	8.06*	-9.09*	3.97*	-1.25	-0.78	-1.92	-2.11*	-3.04*	-5.56*	-5.64*
Spec 13:GJR_t	-3.90*	3.46*	3.92*	-4.17*	-0.90	0.70	0.68	2.86*	-3.22*	2.59*	0.02	2.25*	-8.88*	0.02	-2.60*	-3.47*	-2.73*	-2.86*	-3.42*	-5.60*	-5.67*
Spec 14:TGA_t	-5.30*	-2.95*	1.80	-5.54*	-2.70*	-1.55	-1.06	-3.13*	-4.21*	1.24	-0.99	0.92	-9.16*	-0.99	-3.43*	-4.71*	-3.35*	-3.49*	-3.93*	-5.90*	-5.96*
Spec 15:IGA_t	1.42	4.27*	5.29*	1.90	8.32*	3.20*	4.72*	5.01*	-0.99	8.23*	3.90*	8.08*	-9.05*	3.90*	-1.17	-0.47	-1.83	-2.00*	-2.97*	-5.53*	-5.60*
Spec 16:AGA_t	-7.57*	1.13	2.66*	-8.28*	0.45	0.96	2.02*	1.85	-5.27*	6.06*	0.93	5.81*	-9.56*	0.93	-3.27*	-5.88*	-3.44*	-3.63*	-4.08*	-6.14*	-6.19*
Spec 17:$NAGA_t$	-2.96*	0.36	3.22*	-3.08*	-0.63	3.67*	0.55	1.53	-2.60*	2.10*	0.02	1.85	-8.61*	0.02	-2.38*	-2.82*	-2.47*	-2.59*	-3.20*	-5.38*	-5.45*
Spec 18:EGA_t	-4.79*	-0.48	1.37	-5.01*	-2.15*	-0.12	0.19	0.15	-4.21*	3.07*	-0.53	2.62*	-9.60*	-0.53	-3.74*	-4.34*	-3.49*	-3.62*	-4.23*	-6.24*	-6.29*
Spec 19:$APGA_t$	-4.36*	0.01	4.43*	-4.60*	-1.19	0.45	0.48	4.95*	-3.47*	2.68*	-0.11	2.28*	-8.95*	-0.11	-2.79*	-3.85*	-2.88*	-3.02*	-3.54*	-5.66*	-5.73*
Spec 20:$FIGA_t$	1.27	3.28*	4.06*	1.06	5.17*	2.70*	4.11*	3.75*	1.02	6.41*	4.64*	6.51*	-9.45*	4.64*	-0.93	0.79	-2.02*	-2.26*	-3.26*	-5.83*	-5.90*
Spec 21:A_{st}	5.43*	6.33*	7.04*	5.38*	7.70*	5.64*	8.15*	6.70*	5.08*	8.95*	8.56*	9.07*	-8.53*	8.56*	4.35*	5.37*	2.93*	2.85*	0.26	-4.42*	-4.55*
Spec 22:GA_{st}	-	4.00*	5.08*	-1.02	7.57*	3.09*	4.49*	4.78*	-1.26	8.15*	3.90*	8.05*	-9.09*	3.90*	-1.27	-0.88	-1.94	-2.12*	-3.04*	-5.56*	-5.64*
Spec 23:GJR_{st}	-	-	3.41*	-4.28*	-1.13	0.38	0.46	1.93	-3.32*	2.33*	-0.10	2.00*	-8.90*	-0.10	-2.68*	-3.57*	-2.78*	-2.91*	-3.47*	-5.63*	-5.69*
Spec 24:TGA_{st}	-	-	-	-5.31*	-2.66*	-2.00*	-1.19	-3.76*	-4.10*	0.90	-1.05	0.64	-9.11*	-1.05	-3.39*	-4.56*	-3.31*	-3.44*	-3.89*	-5.86*	-5.92*
Spec 25:IGA_{st}	-	-	-	-	8.28*	3.20*	4.69*	5.03*	-1.04	8.25*	3.85*	8.09*	-9.05*	3.85*	-1.18	-0.56	-1.84	-2.01*	-2.97*	-5.52*	-5.60*
Spec 26:AGA_{st}	-	-	-	-	-	0.96	2.02*	1.85	-5.27*	6.06*	0.93	5.81*	-9.56*	0.93	-3.27*	-5.88*	-3.44*	-3.63*	-4.08*	-6.14*	-6.19*
Spec 27:$NAGA_{st}$	-	-	-	-	-	-	0.14	0.46	-2.72*	1.55	-0.23	1.35	-8.60*	-0.23	-2.51*	-2.95*	-2.56*	-2.67*	-3.26*	-5.39*	-5.47*
Spec 28:EGA_{st}	-	-	-	-	-	-	-	0.10	-4.12*	2.75*	-0.56	2.40*	-9.70*	-0.56	-3.85*	-4.19*	-3.51*	-3.64*	-4.30*	-6.33*	-6.37*
Spec 29:$APGA_{st}$	-	-	-	-	-	-	-	-	-3.80*	2.05*	-0.48	1.69	-9.02*	-0.48	-3.04*	-4.22*	-3.07*	-3.21*	-3.69*	-5.75*	-5.82*
Spec 30:$FIGA_{st}$	-	-	-	-	-	-	-	-	-	6.49*	4.61*	6.59*	-9.43*	4.61*	-0.95	0.74	-2.00*	-2.24*	-3.23*	-5.81*	-5.88*
Spec 31:$comb_1$	-	-	-	-	-	-	-	-	-	-	-2.84*	-1.51	-9.83*	-2.84*	-5.05*	-7.38*	-4.56*	-4.74*	-4.89*	-6.52*	-6.56*
Spec 32:$comb_2$	-	-	-	-	-	-	-	-	-	-	-	2.99*	-10.27*	0.67	-5.75*	-4.35*	-4.77*	-4.95*	-5.27*	-6.81*	-6.84*
Spec 33:$comb_3$	-	-	-	-	-	-	-	-	-	-	-	-	-9.90*	-2.99*	-5.20*	-7.44*	-4.65*	-4.84*	-4.97*	-6.57*	-6.61*
Spec 34:HAR	-	-	-	-	-	-	-	-	-	-	-	-	-	10.27*	10.42*	9.22*	10.29*	10.14*	11.14*	9.82*	9.52*
Spec 35:logHAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-5.75*	-4.35*	-4.77*	-4.95*	-5.27*	-6.81*	-6.84*
Spec 36:HEAVY	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.11	-1.03	-1.13	-3.80*	-6.58*	-6.64*
Spec 37:$GA-MIDAS$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-1.92	-2.12*	-3.09*	-5.66*	-5.73*
Spec 38:$GA-RV_N$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-1.38	-4.03*	-6.43*	-6.48*
Spec 39:$GA-RV_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-3.60*	-6.27*	-6.33*
Spec 40:$GA-RV_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-6.97*	-7.03*
Spec 41:$quant_1$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-4.72*
Spec 42:$quant_2$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Note: This table reports the out-of-sample t-statistic of the Unconditional Giacomini and White (2006) test under the robust loss of Patton (2011) for b=-0.5 for S&P 500. The period is from 02/01/2004 to 29/12/2017. A positive t-statistic indicates that average loss of model A (in row) is larger than of model B (in column) and hence model B is superior. *denotes significance at the 5% level

3.4.2 Robustness - Sensitivity to the in-sample estimation period

To test how robust our findings to the in-sample estimation period employed, we repeat our analysis using a 1-year in-sample period keeping the rest of the observations for the out-of-sample evaluation. Table 3.8 reports results for the 1-year in-sample of S&P 500 under the MSE loss function. Similar to the previous results, ARCH models, non-parametric specifications, GARCH, FIGARCH, and GARCH-RV specifications, GARCH-MIDAS, HEAVY are among the worst performing ones.

Six specifications that have zero "losses" that is TGARCH using any of the examined error distributions, the NAGARCH-St and the two HAR specifications. Even though, HAR only wins the poor performing models. The two HAR specifications are also indifferent to each other. NAGARCH-St outperforms its other two specifications. Explicitly, the TGARCH-N and the TGARCH-St are preferred over the counter specifications in 36 cases leaving only 5 cases were the there is no statistically significant difference among the alternative specifications. However, the NAGARCH-St is only preferred in 30 cases. One difference is that forecast combinations do not perform that well and are outperformed by the TGARCH specifications. The 1st forecast combination outperforms the 3rd and, it is indifferent to the 2nd, for all robust loss functions for the 4-year in-sample. For 1-year in-sample case the rest of the loss function confirm that the 1st beats the 3rd. However, the 1st forecast combination is the best among all combinations.

The GJR-N is outperformed by the three TGARCH specifications while the two other GJR specifications are outperformed by the same models plus the GJR-N and the NAGARCH-St. What is more, for all the three GJR specifications the null hypothesis of equal predictive accuracy could not be rejected against the AGARCH-N, the NAGARCH-N, the NAGARCH-t, the three EGARCH specifications, the APGARCH-St and the three combinations and the HAR and its logarithmic transformation. In addition, the Normal GJR is indifferent to EGARCH, NAGARCH and APGARCH under the Skew-t at the 5% significant level while the Student-t and the Skew-T GJR specifications are additionally indifferent to the APGARCH-N, AGARCH-St and to each other and only the Student-t GJR is indifferent to HEAVY model. All of the NAGARCH specifications are outperformed by the TGARCH Student-t and Skew-t, IGARCH-N and logHAR where the Student-t and the Skew-t NAGARCH are outperformed by the NAGARCH, EGARCH and APGARCH under the Skew-t specification. One exception as mentioned before is that the TGARCH-t is indifferent to the NAGARCH-t. The three AGARCH specifications are outperformed by TGARCH Student-t and Skew-t, by the IGARCH-N and the logHAR whil,e only the AGARCH Student-t and Skew-t are also outperformed by the NAGARCH, EGARCH and APGARCH under the Skew-t.

The results from the robust loss of Patton (2011) for $b = 1$ at 5% significant level are a bit different

to the ones derived from the 4-year in-sample in the sense that there are less cases where the null hypothesis of equal accuracy was rejected. Moving on to Table 3.9 it can be reported that overall it is worth mentioning that the following specifications had no wins with regard to their pairwise comparisons, namely, all the ARCH, FIGARCH, GARCH-RV, the IGARCH-N, the NAGARCH-N, the NAGARCH-t, the APGARCH-t, the GARCH-MIDAS, the HAR and the quantiles. The only exception is that the second quantile is winning against the first while for the first sample(4-years) it was indifferent. TGARCH specifications, log HAR and the forecast combinations still perform very well. The same holds for NAGARCH-St. NAGARCH-St outperforms its other two specifications. The NAGARCH-N is not preferred over the three TGARCH specifications and the NAGARCH-St. For the 4-years in-sample it was seen that the NAGARCH-St outperforming just its student-t specifications while now its both specifications. Each one of the three AGARCH specifications are outperformed by the three TGARCH specifications and the EGARCH Normal and Skew-t and the NAGARCH-St. Now focusing on the results presented on Table 3.10 it can be seen that the results using this specification of the robust losses of Patton (2011) are more in accordance with the ones derived using the MSE. However, now it can been seen that both of the HAR specifications are not outperformed by any model but they only outperform the worse performing models whereas for 4-years in-sample (b=-0.5) that was hardly the case for HAR which was outperformed by almost all models and the logHAR which was outperformed by just the 1st combination.

Table 3.19 in the Appendix reports the total number of "wins" along with the total number of "losses" for each specification for the 1-year in-sample of the S&P 500 index at 5% significant level for the non robust loss functions. From the table it is easy to observe that for the non robust loss functions, the results are uniformed and not far from the ones derived from the robust ones.

According to the table 3.11 and following with the MSE the models that should be included in the set agree with the ones from 4-years (b=-0.5). The models that should definitely be included following all loss functions are: three GJR and the three TGARCH specifications, the EGARCH-St, the APGARCH-St, the HAR and logHAR. Additionally the MSE indicated that the second combination, the AGARCH-N, the NAGARCH-N, the EGARCH-N, the APGARCH-N, the AGARCH-St and the EGARCH-St. Comparing the two samples, it can been seen that for b=1, the results are identical. As expected under the MSE and the robust loss of Patton (b=-0.5) when any of the ARCH, IGARCH, FIGARCH, GARCH-RV (except from GARCH-RV-St for MSE) or the GARCH-MIDAS, or the two combinations (except from 2nd one for MSE) was used as a benchmark then the null hypothesis that the benchmark is inferior to all the other models was rejected when all the models were included into the set at the 5% level. Surprisingly, that list also contained the the EGARCH-t, the EGARCH-St and the forecast combinations. The SQ (semi-quadratic statistic) criterion is presented on the Ap-

pendix. At last, table 3.31 in the Appendix reports how many times a specific model beats the rest of the models in pairwise comparisons for S&P 500 using the GW test and the MSE loss function for the 1 year in-sample period.

Table 3.8 – Pairwise comparisons of unconditional Giacomini and White test for S&P 500 (1-year in-sample) - MSE

	Spec 1	Spec 2	Spec 3	Spec 4	Spec 5	Spec 6	Spec 7	Spec 8	Spec 9	Spec 10	Spec 11	Spec 12	Spec 13	Spec 14	Spec 15	Spec 16	Spec 17	Spec 18	Spec 19	Spec 20	Spec 21
Spec 1:A_N	-	3.67*	4.31*	4.95*	3.35*	4.84*	4.23*	5.18*	4.02*	3.21*	-2.30*	3.68*	4.00*	4.67*	3.41*	4.44*	3.83*	4.06*	3.63*	3.19*	-2.42*
Spec 2:GA_N	-	-	3.68*	4.55*	-1.12	1.45	2.92*	2.07*	3.27*	-2.24*	-4.55*	0.29	2.94*	4.01*	-0.80	3.23*	1.98*	2.74*	2.07*	-2.23*	-4.78*
Spec 3:GJR_M	-	-	-	3.43*	-4.92*	-1.73	-1.07	-0.96	-2.37*	-3.35*	-5.18*	-3.7*	-3.21*	3.09*	-4.72*	-2.29*	-1.39	0.19	-3.46*	-3.35*	-5.41*
Spec 4:TGA_N	-	-	-	-	-5.44*	-2.69*	-2.90*	-2.10*	-3.74*	-3.93*	-5.64*	-4.55*	-3.93*	0.13	-5.33*	-3.58*	-3.14*	-4.15*	-4.34*	-3.93*	-5.91*
Spec 5:IGA_N	-	-	-	-	-	1.57	3.59*	2.08*	4.60*	-1.55	-4.31*	1.40	4.00*	4.80*	1.39	3.81*	2.50*	3.32*	2.93*	-1.55	-4.51*
Spec 6:AGA_N	-	-	-	-	-	-	1.29	1.32	1.20	-5.30*	-5.23*	-1.41	1.27	2.42*	-1.46	0.66	0.79	1.43	0.56	-5.18*	-5.55*
Spec 7:$NAGA_N$	-	-	-	-	-	-	-	-0.36	0.09	-3.15*	-5.05*	-2.88*	0.44	2.60*	-3.45*	-1.53	-1.01	1.02	-0.81	-3.15*	-5.27*
Spec 8:EGA_N	-	-	-	-	-	-	-	-	0.39	-3.83*	-5.69*	-2.06*	0.56	1.87	-2.01*	-0.54	0.07	0.89	-0.13	-3.81*	-6.04*
Spec 9:$APGA_N$	-	-	-	-	-	-	-	-	-	-3.06*	-4.96*	-3.28*	0.88	3.38*	-4.36*	-1.45	-0.49	1.28	-2.17*	-3.05*	-5.16*
Spec 10:$FIGA_N$	-	-	-	-	-	-	-	-	-	-	-3.90*	2.26*	2.90*	3.59*	1.68	3.69*	2.46*	2.72*	2.38*	-0.14	-4.08*
Spec 11:A_t	-	-	-	-	-	-	-	-	-	-	-	4.55*	4.92*	5.43*	4.35*	5.07*	4.74*	5.04*	4.72*	3.89*	1.29
Spec 12:GA_t	-	-	-	-	-	-	-	-	-	-	-	-	2.95*	4.02*	-1.09	3.18*	1.95	2.73*	2.06*	-2.26*	-4.78*
Spec 13:GJR_t	-	-	-	-	-	-	-	-	-	-	-	-	-	3.88*	-3.82*	-1.55	-0.81	0.95	-2.50*	-2.91*	-5.12*
Spec 14:TGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-4.70*	-3.12*	-2.99*	-4.21*	-4.07*	-3.59*	-5.67*
Spec 15:IGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3.62*	2.37*	3.23*	2.76*	-1.69	-4.55*
Spec 16:AGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.74	1.72	0.40	-3.67*	-5.36*
Spec 17:$NAGA_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.38	-0.36	-2.46*	-4.93*
Spec 18:EGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-2.28*	-2.72*	-5.24*
Spec 19:$APGA_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-2.38*	-4.89*
Spec 20:$FIGA_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-4.07*
Spec 21:A_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 22:GA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 23:GJR_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 24:TGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 25:IGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 26:AGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 27:$NAGA_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 28:EGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 29:$APGA_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 30:$FIGA_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 31:$comb_1$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 32:$comb_2$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 33:$comb_3$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 34:HAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 35:logHAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 36:HEAVY	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 37:$GA-MIDAS$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 38:$GA-RV_N$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 39:$GA-RV_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 40:$GA-RV_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 41:$quant_1$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 42:$quant_2$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Note: This table reports the out-of-sample t-statistic of the Unconditional Giacomini and White (2006) test under the MSE loss function for S&P 500. The period is from 02/01/2001 to 29/12/2017. A positive t-statistic indicates that average loss of model A (in row) is larger than of model B (in column) and hence model B is superior. *denotes significance at the 5% level

Table 3.8 cont. - Pairwise comparisons of unconditional Giacomini and White test for S&P 500 (1-year in-sample) - MSE

	Spec 22	Spec 23	Spec 24	Spec 25	Spec 26	Spec 27	Spec 28	Spec 29	Spec 30	Spec 31	Spec 32	Spec 33	Spec 34	Spec 35	Spec 36	Spec 37	Spec 38	Spec 39	Spec 40	Spec 41	Spec 42
Spec 1:A_N	3.72*	4.12*	4.78*	3.44*	4.44*	4.21*	4.44*	4.18*	3.22*	6.73*	5.21*	6.76*	3.40*	7.12*	4.21*	2.57*	3.19*	3.15*	2.07*	-2.54*	-2.71*
Spec 2:GA_N	0.11	3.24*	4.21*	-0.95	3.23*	3.24*	3.34*	3.50*	-2.16*	-0.25	-0.33	-0.43	0.52	2.46*	-1.28	-1.94	-1.59	-1.61	-1.76	-3.60*	-3.67*
Spec 3:GJR_M	-3.61*	-3.19*	3.16*	-4.63*	-2.29*	1.50	1.04	0.19	-3.32*	-1.85	-1.45	-1.89	-0.78	-0.03	-2.18*	-5.23*	-2.41*	-2.43*	-2.40*	-3.93*	-3.98*
Spec 4:TGA_N	-4.46*	-3.99*	0.66	-5.26*	-3.58*	-1.05	-3.54*	-3.43*	-3.89*	-2.73*	-2.05*	-2.73*	-1.50	-1.10	-2.76*	-6.07*	-2.94*	-2.96*	-2.86*	-4.30*	-4.31*
Spec 5:IGA_N	1.18	4.49*	5.03*	0.92	3.81*	3.87*	3.96*	4.50*	-1.50	-0.05	-0.17	-0.22	0.64	2.48*	-1.09	-1.68	-1.33	-1.33	-1.56	-3.39*	-3.47*
Spec 6:AGA_N	-1.51	1.35	2.53*	-1.53	0.66	1.88	1.85	1.65	-5.11*	-1.19	-1.01	-1.36	0.11	1.99*	-1.98*	-2.00*	-2.64*	-2.70*	-2.48*	-4.43*	-4.44*
Spec 7:$NAGA_N$	-2.82*	0.46	2.79*	-3.42*	-1.53	2.00*	1.64	1.22	-3.10*	-1.49	-1.23	-1.57	-0.47	0.55	-1.98*	-4.44*	-2.27*	-2.29*	-2.29*	-3.92*	-3.96*
Spec 8:EGA_N	-2.15*	0.59	1.98*	-2.08*	-0.54	1.36	1.30	0.95	-3.78*	-1.99*	-1.55	-2.11*	-0.26	1.08	-2.34*	-2.42*	-3.11*	-3.17*	-2.83*	-4.60*	-4.60*
Spec 9:$APGA_N$	-3.15*	1.04	3.59*	-4.23*	-1.45	2.34*	2.00*	2.41*	-3.01*	-1.45	-1.18	-1.52	-0.46	0.46	-1.92	-5.16*	-2.18*	-2.20*	-2.20*	-3.78*	-3.83*
Spec 10:$FIGA_N$	2.35*	3.07*	3.71*	1.71	3.69*	3.06*	3.14*	3.17*	0.35	1.05	0.61	0.77	1.18	3.88*	-0.66	0.19	-1.00	-0.99	-1.40	-3.60*	-3.68*
Spec 11:A_t	4.57*	5.01*	5.51*	4.36*	5.07*	5.08*	5.30*	5.10*	3.91*	5.05*	4.53*	4.94*	4.07*	5.97*	3.78*	3.61*	3.54*	3.56*	2.77*	0.00	-0.32
Spec 12:GA_t	-0.31	3.26*	4.22*	-1.29	3.18*	3.23*	3.33*	3.49*	-2.19*	-0.28	-0.34	-0.45	0.50	2.43*	-1.30	-1.99*	-1.61	-1.62	-1.77	-3.61*	-3.68*
Spec 13:GJR_t	-2.88*	-0.03	3.88*	-3.74*	-1.55	2.18*	1.71	1.57	-2.87*	-1.52	-1.22	-1.58	-0.58	0.25	-1.95	-4.63*	-2.16*	-2.18*	-2.20*	-3.74*	-3.80*
Spec 14:TGA_t	-3.94*	-3.72*	0.88	-4.63*	-3.12*	-1.41	-3.02*	-3.27*	-3.56*	-2.50*	-1.93	-2.51*	-1.47	-1.01	-2.62*	-5.59*	-2.76*	-2.78*	-2.73*	-4.13*	-4.16*
Spec 15:IGA_t	0.88	4.28*	4.92*	-0.44	3.62*	3.78*	3.88*	4.34*	-1.64	-0.11	-0.22	-0.28	0.59	2.43*	-1.14	-1.83	-1.38	-1.39	-1.60	-3.42*	-3.50*
Spec 16:AGA_t	-3.16*	1.69	3.29*	-3.70*	2.36*	2.31*	2.18*	2.13*	-3.59*	-1.21	-0.98	-1.33	-0.05	1.43	-1.87	-3.41*	-2.22*	-2.24*	-2.25*	-4.06*	-4.10*
Spec 17:$NAGA_t$	-1.91	0.83	3.17*	-2.35*	-0.74	2.50*	1.96*	1.62	-2.42*	-1.17	-1.00	-1.26	-0.32	0.69	-1.74	-3.62*	-1.95	-1.96*	-2.05*	-3.67*	-3.73*
Spec 18:EGA_t	-2.67*	-0.94	4.55*	-3.18*	-1.72	2.20*	2.39*	-0.10	-2.69*	-1.70	-1.36	-1.75	-0.80	-0.09	-2.06*	-4.36*	-2.21*	-2.22*	-2.27*	-3.75*	-3.81*
Spec 19:$APGA_t$	-1.99*	2.53*	4.19*	-2.69*	-0.40	3.06*	2.81*	3.96*	-2.34*	-1.01	-0.87	-1.10	-0.17	0.81	-1.61	-4.01*	-1.83	-1.85	-1.93	-3.53*	-3.60*
Spec 20:$FIGA_t$	2.35*	3.07*	3.71*	1.72	3.67*	3.06*	3.13*	3.17*	1.45	1.05	0.61	0.78	1.19	3.87*	-0.65	0.20	-0.99	-0.98	-1.39	-3.59*	-3.67*
Spec 21:A_{st}	4.81*	5.23*	5.76*	4.57*	5.36*	5.28*	5.54*	5.31*	4.08*	5.48*	4.77*	5.35*	4.21*	6.45*	4.00*	3.73*	3.67*	3.68*	2.81*	-0.14	-0.48
Spec 22:GA_{st}	-	3.17*	4.13*	-1.07	3.16*	3.15*	3.26*	3.37*	-2.27*	-0.27	-0.34	-0.45	0.51	2.48*	-1.31	-1.87	-1.63	-1.65	-1.78	-3.64*	-3.70*
Spec 23:GJR_{st}	-	-	3.88*	-4.18*	-1.69	2.14*	1.75	1.66	-3.04*	-1.59	-1.27	-1.65	-0.58	0.26	-2.01*	-4.94*	-2.24*	-2.25*	-2.26*	-3.82*	-3.87*
Spec 24:TGA_{st}	-	-	-	-4.85*	-3.29*	-1.83	-3.83*	-3.62*	-3.68*	-2.63*	-2.01*	-2.63*	-1.54	-1.12	-2.70*	-5.89*	-2.85*	-2.86*	-2.80*	-4.20*	-4.22*
Spec 25:IGA_{st}	-	-	-	-	3.70*	3.72*	3.82*	4.23*	-1.67	-0.10	-0.21	-0.27	0.60	2.48*	-1.14	-1.76	-1.40	-1.40	-1.61	-3.44*	-3.52*
Spec 26:AGA_{st}	-	-	-	-	-	2.36*	2.31*	2.18*	-3.59*	-1.21	-0.98	-1.33	-0.05	1.43	-1.87	-3.41*	-2.22*	-2.24*	-2.25*	-4.06*	-4.10*
Spec 27:$NAGA_{st}$	-	-	-	-	-	-	-1.04	-1.94	-3.03*	-2.05*	-1.64	-2.08*	-1.21	-0.64	-2.30*	-5.14*	-2.43*	-2.44*	-2.45*	-3.86*	-3.90*
Spec 28:EGA_{st}	-	-	-	-	-	-	-	-1.21	-3.10*	-2.09*	-1.62	-2.12*	-1.04	-0.42	-2.33*	-4.94*	-2.49*	-2.50*	-2.50*	-3.98*	-4.02*
Spec 29:$APGA_{st}$	-	-	-	-	-	-	-	-	-3.13*	-1.78	-1.42	-1.83	-0.80	-0.07	-2.13*	-5.69*	-2.34*	-2.36*	-2.35*	-3.85*	-3.90*
Spec 30:$FIGA_{st}$	-	-	-	-	-	-	-	-	-	1.03	0.59	0.75	1.17	3.84*	-0.67	0.17	-1.02	-1.02	-1.41	-3.61*	-3.69*
Spec 31:$comb_1$	-	-	-	-	-	-	-	-	-	-	-0.32	-2.48*	0.66	4.16*	-2.13*	-0.56	-2.51*	-2.39*	-2.86*	-5.24*	-5.09*
Spec 32:$comb_2$	-	-	-	-	-	-	-	-	-	-	-	0.01	0.68	2.72*	-2.06*	-0.29	-4.39*	-3.67*	-4.27*	-6.12*	-5.89*
Spec 33:$comb_3$	-	-	-	-	-	-	-	-	-	-	-	-	0.77	4.23*	-2.02*	-0.38	-2.38*	-2.24*	-2.87*	-5.35*	-5.18*
Spec 34:HAR	-	-	-	-	-	-	-	-	-	-	-	-	-	1.03	-1.48	-1.08	-1.54	-1.51	-1.90	-3.69*	-3.78*
Spec 35:logHAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-3.79*	-2.83*	-4.08*	-4.04*	-3.84*	-5.56*	-5.46*
Spec 36:HEAVY	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.58	-0.10	-0.02	-1.87	-5.96*	-5.86*
Spec 37:$GA-MIDAS$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-0.71	-0.69	-1.09	-2.89*	-3.00*
Spec 38:$GA-RV_N$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.58	-1.61	-4.66*	-4.65*
Spec 39:$GA-RV_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-1.54	-4.51*	-4.52*
Spec 40:$GA-RV_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-7.01*	-6.64*
Spec 41:$quant_1$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-2.73*
Spec 42:$quant_2$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Note: This table reports the out-of-sample t-statistic of the Unconditional Giacomini and White (2006) test under the MSE loss function for S&P 500. The period is from 02/01/2001 to 29/12/2017. A positive t-statistic indicates that average loss of model A (in row) is larger than of model B (in column) and hence model B is superior. *denotes significance at the 5% level

Table 3.9 – Pairwise comparisons of unconditional Giacomini and White test for S&P 500 (1-year in-sample) - robust Patton (b=1)

	Spec 1	Spec 2	Spec 3	Spec 4	Spec 5	Spec 6	Spec 7	Spec 8	Spec 9	Spec 10	Spec 11	Spec12	Spec 13	Spec 14	Spec 15	Spec 16	Spec 17	Spec 18	Spec 19	Spec 20	Spec 21
Spec 1:A_N	-	0.73	1.63	2.43*	0.68	0.72	1.02	1.83	1.23	-0.22	-1.49	0.75	1.34	2.27*	0.71	0.80	0.68	1.71	0.73	-0.24	-1.53
Spec 2:GA_N	-	-	2.54*	3.42*	-0.10	-0.33	0.87	1.72	1.81	-2.20*	-1.66	0.49	1.86	3.02*	0.20	0.14	0.29	2.18*	0.45	-2.23*	-1.71
Spec 3:GJR_M	-	-	-	3.23*	-3.02*	-1.86	-1.06	-0.59	-2.64*	-2.65*	-2.09*	-2.52*	-2.97*	2.85*	-2.92*	-2.27*	-1.22	1.16	-3.54*	-2.67*	-2.23*
Spec 4:TGA_N	-	-	-	-	-3.80*	-2.70*	-2.31*	-1.81	-3.29*	-3.22*	-2.39*	-3.42*	-3.70*	0.52	-3.75*	-3.24*	-2.30*	-2.63*	-3.85*	-3.24*	-2.59*
Spec 5:IGA_N	-	-	-	-	-	-0.25	0.92	1.53	2.19*	-1.90	-1.65	0.25	2.22*	3.35*	1.38	0.18	0.31	2.42*	0.54	-1.93	-1.69
Spec 6:AGA_N	-	-	-	-	-	-	1.02	2.08*	1.33	-2.61*	-1.64	0.37	1.41	2.44*	0.33	0.56	0.42	1.78	0.47	-2.65*	-1.70
Spec 7:$NAGA_N$	-	-	-	-	-	-	-	0.33	0.37	-1.82	-1.82	-0.83	0.64	2.22*	-0.86	-1.06	-1.15	1.42	-0.47	-1.84	-1.89
Spec 8:EGA_N	-	-	-	-	-	-	-	-	-0.04	-2.91*	-2.03*	-1.71	0.21	1.63	-1.50	-1.36	-0.55	0.91	-0.76	-2.94*	-2.19*
Spec 9:$APGA_N$	-	-	-	-	-	-	-	-	-	-2.27*	-1.94	-1.78	1.07	3.00*	-2.08*	-1.52	-0.67	1.89	-2.55*	-2.29*	-2.04*
Spec 10:$FIGA_N$	-	-	-	-	-	-	-	-	-	-	-1.28	2.23*	2.24*	2.96*	1.98*	2.18*	1.15	2.34*	1.40	-0.86	-1.25
Spec 11:A_t	-	-	-	-	-	-	-	-	-	-	-	1.66	1.98*	2.37*	1.66	1.68	1.67	2.21*	1.78	1.27	1.33
Spec 12:GA_t	-	-	-	-	-	-	-	-	-	-	-	-	1.84	3.01*	0.04	0.09	0.27	2.16*	0.42	-2.26*	-1.72
Spec 13:GJR_t	-	-	-	-	-	-	-	-	-	-	-	-	-	3.50*	-2.13*	-1.66	-0.90	1.91	-2.87*	-2.25*	-2.09*
Spec 14:TGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-3.30*	-2.91*	-2.30*	-3.26*	-3.65*	-2.97*	-2.55*
Spec 15:IGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.07	0.27	2.37*	0.46	-2.01*	-1.71
Spec 16:AGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.29	2.08*	0.33	-2.21*	-1.74
Spec 17:$NAGA_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.61	-0.01	-1.17	-1.69
Spec 18:EGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-3.02*	-2.35*	-2.35*
Spec 19:$APGA_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-1.42	-1.84
Spec 20:$FIGA_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-1.24
Spec 21:A_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 22:GA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 23:GJR_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 24:TGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 25:IGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 26:AGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 27:$NAGA_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 28:EGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 29:$APGA_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 30:$FIGA_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 31:$comb_1$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 32:$comb_2$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 33:$comb_3$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 34:HAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 35:logHAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 36:HEAVY	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 37:$GA-MIDAS$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 38:$GA-RV_N$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 39:$GA-RV_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 40:$GA-RV_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 41:$quant_1$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 42:$quant_2$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Note: This table reports the out-of-sample t-statistic of the Unconditional Giacomini and White (2006) test under the robust loss of Patton (2011) for b=1 for S&P 500. The period is from 02/01/2001 to 29/12/2017. A positive t-statistic indicates that average loss of model A (in row) is larger than of model B (in column) and hence model B is superior. *denotes significance at the 5% level

Table 3.10 cont. - Pairwise comparisons of unconditional Giacomini and White test for S&P 500 (1-year in-sample) - robust Patton (b=1)

	Spec 22	Spec 23	Spec 24	Spec 25	Spec 26	Spec 27	Spec 28	Spec 29	Spec 30	Spec 31	Spec 32	Spec 33	Spec 34	Spec 35	Spec 36	Spec 37	Spec 38	Spec 39	Spec 40	Spec 41	Spec 42
Spec 1:A_N	0.76	1.44	2.38*	0.73	0.80	1.95	2.12*	1.57	-0.23	4.10*	2.87*	3.99*	-0.45	2.74*	1.93	0.44	0.75	0.33	0.96	-1.66	-1.41
Spec 2:GA_N	0.25	2.15*	3.19*	0.34	0.14	2.51*	2.83*	2.32*	-2.20*	0.98	0.64	0.82	-0.69	1.44	0.07	-0.41	-0.38	-0.64	-0.28	-1.20	-1.12
Spec 3:GJR_M	-2.44*	-3.49*	3.10*	-2.82*	-2.27*	1.98*	2.29*	0.60	-2.65*	-0.42	-0.52	-0.50	-1.09	-0.08	-0.89	-3.29*	-1.32	-1.52	-1.11	-1.83	-1.76
Spec 4:TGA_N	-3.35*	-3.62*	1.28	-3.68*	-3.24*	-0.19	-2.20*	-2.91*	-3.22*	-1.41	-1.33	-1.45	-1.43	-1.16	-1.66	-3.93*	-2.06*	-2.24*	-1.81	-2.47*	-2.39*
Spec 5:IGA_N	0.18	2.58*	3.53*	1.05	0.18	2.77*	3.14*	2.69*	-1.90	0.93	0.61	0.78	-0.69	1.4	0.08	-0.44	-0.33	-0.57	-0.26	-1.14	-1.06
Spec 6:AGA_N	0.37	1.57	2.57*	0.36	0.56	2.08*	2.28*	1.78	-2.61*	1.36	0.97	1.17	-0.65	1.78	0.22	-0.07	-0.29	-0.62	-0.21	-1.33	-1.23
Spec 7:$NAGA_N$	-0.82	0.72	2.35*	-0.83	-1.06	2.00*	1.92	1.19	-1.82	0.22	0.04	0.13	-0.89	0.62	-0.34	-1.17	-0.72	-0.91	-0.61	-1.36	-1.29
Spec 8:EGA_N	-1.78	0.28	1.77	-1.52	-1.36	1.28	1.42	0.66	-2.93*	-0.02	-0.27	-0.17	-0.95	0.45	-0.78	-1.33	-1.55	-1.88	-1.15	-2.12*	-2.01*
Spec 9:$APGA_N$	-1.70	1.44	3.18*	-1.98*	-1.52	2.42*	2.67*	2.41*	-2.27*	0.01	-0.14	-0.08	-0.94	0.37	-0.53	-2.56*	-0.95	-1.15	-0.78	-1.52	-1.44
Spec 10:$FIGA_N$	2.28*	2.42*	3.07*	2.05*	2.18*	2.60*	2.81*	2.50*	-0.25	2.32*	1.94	2.13*	-0.38	2.57*	1.08	1.19	0.89	0.62	0.62	-0.52	-0.43
Spec 11:A_t	1.66	2.01*	2.41*	1.66	1.68	2.30*	2.33*	2.11*	1.28	1.93	1.84	1.89	0.62	2.01*	1.67	1.53	1.51	1.44	1.52	1.11	1.14
Spec 12:GA_t	-0.12	2.13*	3.19*	0.17	0.09	2.49*	2.82*	2.29*	-2.23*	0.97	0.63	0.81	-0.70	1.44	0.06	-0.44	-0.40	-0.66	-0.30	-1.22	-1.14
Spec 13:GJR_t	-1.78	0.64	3.70*	-2.04*	-1.66	2.60*	2.90*	2.40*	-2.24*	-0.16	-0.29	-0.24	-1.00	0.17	-0.66	-2.65*	-1.05	-1.24	-0.89	-1.59	-1.52
Spec 14:TGA_t	-2.95*	-3.25*	2.04*	-3.23*	-2.91*	-0.64	-2.17*	-2.86*	-2.95*	-1.32	-1.26	-1.36	-1.44	-1.11	-1.58	-3.64*	-1.92	-2.09*	-1.71	-2.33*	-2.26*
Spec 15:IGA_t	-0.06	2.47*	3.48*	0.43	0.07	2.72*	3.09*	2.60*	-1.98*	0.91	0.58	0.76	-0.70	1.37	0.05	-0.52	-0.37	-0.62	-0.28	-1.17	-1.09
Spec 16:AGA_t	-0.10	1.87	3.06*	-0.04	2.48*	2.70*	2.17*	2.45*	-2.17*	0.93	0.59	0.77	-0.71	1.41	0.03	-0.45	-0.41	-0.66	-0.32	-1.28	-1.20
Spec 17:$NAGA_t$	-0.27	0.95	2.41*	-0.26	-0.29	2.18*	2.00*	1.36	-1.15	0.41	0.25	0.33	-0.80	0.78	-0.10	-0.53	-0.40	-0.55	-0.35	-1.04	-0.98
Spec 18:EGA_t	-2.11*	-1.66	3.50*	-2.31*	-2.08*	1.96*	2.57*	-1.10	-2.33*	-0.69	-0.73	-0.75	-1.20	-0.44	-1.05	-2.73*	-1.39	-1.55	-1.23	-1.86	-1.79
Spec 19:$APGA_t$	-0.41	2.89*	3.78*	-0.43	-0.33	3.19*	3.50*	3.97*	-1.40	0.48	0.28	0.38	-0.75	0.83	-0.10	-0.94	-0.44	-0.62	-0.37	-1.09	-1.02
Spec 20:$FIGA_t$	2.32*	2.44*	3.08*	2.08*	2.21*	2.61*	2.82*	2.52*	1.12	2.35*	1.96*	2.15*	-0.38	2.59*	1.10	1.21	0.92	0.65	0.64	-0.51	-0.42
Spec 21:A_{st}	1.72	2.13*	2.60*	1.71	1.74	2.45*	2.51*	2.24*	1.25	2.07*	1.95	2.02*	0.48	2.16*	1.74	1.54	1.55	1.45	1.55	1.04	1.08
Spec 22:GA_{st}	-	2.05*	3.13*	0.17	0.10	2.44*	2.77*	2.22*	-2.29*	1.00	0.65	0.84	-0.70	1.47	0.06	-0.41	-0.40	-0.68	-0.30	-1.24	-1.15
Spec 23:GJR_{st}	-	-	3.49*	-2.38*	-1.87	2.34*	2.73*	1.80	-2.42*	-0.21	-0.34	-0.30	-1.01	0.13	-0.72	-2.85*	-1.14	-1.34	-0.96	-1.69	-1.61
Spec 24:TGA_{st}	-	-	-	-3.42*	-3.06*	-0.95	-2.94*	-3.08*	-3.07*	-1.44	-1.35	-1.47	-1.48	-1.23	-1.67	-3.78*	-2.02*	-2.19*	-1.80	-2.42*	-2.35*
Spec 25:IGA_{st}	-	-	-	-	0.04	2.66*	3.03*	2.52*	-2.05*	0.91	0.58	0.76	-0.70	1.38	0.04	-0.53	-0.39	-0.64	-0.30	-1.19	-1.11
Spec 26:AGA_{st}	-	-	-	-	-	2.48*	2.70*	2.17*	-2.17*	0.93	0.59	0.77	-0.71	1.41	0.03	-0.45	-0.41	-0.66	-0.32	-1.28	-1.20
Spec 27:$NAGA_{st}$	-	-	-	-	-	-	-0.64	-2.16*	-2.59*	-1.05	-1.04	-1.10	-1.38	-0.87	-1.33	-3.26*	-1.64	-1.78	-1.47	-2.05*	-1.99*
Spec 28:EGA_{st}	-	-	-	-	-	-	-	-2.23*	-2.80*	-1.08	-1.06	-1.13	-1.33	-0.83	-1.39	-3.34*	-1.76	-1.93	-1.55	-2.20*	-2.13*
Spec 29:$APGA_{st}$	-	-	-	-	-	-	-	-	-2.50*	-0.47	-0.55	-0.54	-1.11	-0.18	-0.90	-3.21*	-1.29	-1.47	-1.10	-1.77	-1.70
Spec 30:$FIGA_{st}$	-	-	-	-	-	-	-	-	-	2.35*	1.96*	2.15*	-0.38	2.58*	1.10	1.18	0.91	0.64	0.63	-0.52	-0.43
Spec 31:$comb_1$	-	-	-	-	-	-	-	-	-	-	-0.64	-2.41*	-0.95	0.64	-1.87	-0.92	-2.88*	-3.09*	-2.65*	-4.21*	-3.92*
Spec 32:$comb_2$	-	-	-	-	-	-	-	-	-	-	-	0.32	-0.88	0.86	-1.43	-0.65	-3.60*	-3.59*	-2.86*	-4.30*	-4.00*
Spec 33:$comb_3$	-	-	-	-	-	-	-	-	-	-	-	-	-0.92	0.81	-1.73	-0.80	-2.70*	-2.91*	-2.66*	-4.33*	-4.01*
Spec 34:HAR	-	-	-	-	-	-	-	-	-	-	-	-	-	1.26	0.71	0.63	0.57	0.51	0.57	0.19	0.22
Spec 35:logHAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-1.49	-1.32	-2.14*	-2.38*	-1.87	-3.00*	-2.86*
Spec 36:HEAVY	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-0.19	-0.88	-1.18	-1.44	-4.34*	-3.99*
Spec 37:$GA-MIDAS$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-0.13	-0.33	-0.11	-0.89	-0.83
Spec 38:$GA-RV_N$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-2.97*	0.00	-2.08*	-1.88
Spec 39:$GA-RV_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.42	-1.56	-1.39
Spec 40:$GA-RV_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-6.19*	-5.38*
Spec 41:$quant_1$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2.38*
Spec 42:$quant_2$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Note: This table reports the out-of-sample t-statistic of the Unconditional Giacomini and White (2006) test under the robust loss of Patton (2011) for b=1 for S&P 500. The period is from 02/01/2001 to 29/12/2017. A positive t-statistic indicates that average loss of model A (in row) is larger than of model B (in column) and hence model B is superior. *denotes significance at the 5% level

Table 3.10 – Pairwise comparisons of unconditional Giacomini and White test for S&P 500 (1-year in-sample) - robust Patton (b=-0.5)

	Spec 1	Spec 2	Spec 3	Spec 4	Spec 5	Spec 6	Spec 7	Spec 8	Spec 9	Spec 10	Spec 11	Spec12	Spec 13	Spec 14	Spec 15	Spec 16	Spec 17	Spec 18	Spec 19	Spec 20	Spec 21
Spec 1:A_N	-	6.59*	7.22*	7.86*	5.96*	8.37*	7.34*	8.25*	7.00*	6.41*	-3.05*	6.58*	6.90*	7.50*	6.04*	7.78*	6.91*	6.70*	6.66*	6.39*	-3.41*
Spec 2:GA_N	-	-	4.98*	5.99*	-1.98*	3.47*	4.84*	2.48*	4.98*	-2.23*	-8.10*	0.68	4.27*	5.30*	-1.54	6.11*	3.68*	3.14*	3.85*	-2.10*	-8.40*
Spec 3:GJR_M	-	-	-	3.28*	-7.14*	-1.90	-1.33	-1.59	-2.17*	-4.43*	-8.62*	-4.99*	-2.84*	3.06*	-6.87*	-2.52*	-1.69	-1.38	-3.26*	-4.38*	-8.90*
Spec 4:TGA_N	-	-	-	-	-7.59*	-2.93*	-3.43*	-2.71*	-4.01*	-5.02*	-9.22*	-5.99*	-3.78*	-0.05	-7.45*	-4.05*	-3.76*	-5.31*	-4.81*	-4.96*	-9.54*
Spec 5:IGA_N	-	-	-	-	-	3.81*	6.33*	2.82*	7.37*	-1.07	-7.43*	2.63*	6.25*	6.77*	1.27	7.63*	4.98*	4.04*	5.74*	-1.00	-7.70*
Spec 6:AGA_N	-	-	-	-	-	-	1.27	0.14	1.21	-7.33*	-9.63*	-3.35*	1.41	2.61*	-3.67*	0.69	0.80	0.85	0.64	-6.97*	-9.96*
Spec 7:$NAGA_N$	-	-	-	-	-	-	-	-0.96	0.16	-4.35*	-8.73*	-4.84*	0.68	3.00*	-6.24*	-1.46	-0.92	0.02	-0.75	-4.29*	-9.02*
Spec 8:EGA_N	-	-	-	-	-	-	-	-	0.95	-4.33*	-9.72*	-2.41*	1.17	2.41*	-2.72*	0.32	0.62	0.76	0.48	-4.23*	-10.02*
Spec 9:$APGA_N$	-	-	-	-	-	-	-	-	-	-4.16*	-8.43*	-5.01*	1.06	3.65*	-7.12*	-1.48	-0.62	-0.11	-1.89	-4.10*	-8.69*
Spec 10:$FIGA_N$	-	-	-	-	-	-	-	-	-	-	-7.54*	2.35*	3.96*	4.59*	1.28	5.22*	3.58*	3.08*	3.52*	0.45	-7.80*
Spec 11:A_t	-	-	-	-	-	-	-	-	-	-	-	8.08*	8.29*	8.84*	7.51*	9.14*	8.30*	8.18*	8.17*	7.52*	-0.50
Spec 12:GA_t	-	-	-	-	-	-	-	-	-	-	-	-	4.31*	5.33*	-2.35*	6.22*	3.67*	3.10*	3.86*	-2.25*	-8.38*
Spec 13:GJR_t	-	-	-	-	-	-	-	-	-	-	-	-	-	3.97*	-6.06*	-1.74	-1.09	-0.72	-2.40*	-3.93*	-8.56*
Spec 14:TGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-6.67*	-3.50*	-3.55*	-5.15*	-4.55*	-4.56*	-9.14*
Spec 15:IGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	7.56*	4.87*	3.94*	5.54*	-1.22	-7.79*
Spec 16:AGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.74	0.84	0.49	-5.10*	-9.48*
Spec 17:$NAGA_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.38	-0.30	-3.55*	-8.57*
Spec 18:EGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-0.74	-3.04*	-8.45*
Spec 19:$APGA_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-3.48*	-8.41*
Spec 20:$FIGA_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-7.78*
Spec 21:A_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 22:GA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 23:GJR_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 24:TGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 25:IGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 26:AGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 27:$NAGA_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 28:EGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 29:$APGA_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 30:$FIGA_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 31:$comb_1$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 32:$comb_2$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 33:$comb_3$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 34:HAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 35:logHAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 36:HEAVY	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 37:$GA-MIDAS$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 38:$GA-RV_N$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 39:$GA-RV_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 40:$GA-RV_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 41:$quant_1$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 42:$quant_2$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Note: This table reports the out-of-sample t-statistic of the Unconditional Giacomini and White (2006) test under the robust loss of Patton (2011) for b=-0.5 for S&P 500. The period is from 02/01/2001 to 29/12/2017. A positive t-statistic indicates that average loss of model A (in row) is larger than of model B (in column) and hence model B is superior. *denotes significance at the 5% level

Table 3.10 cont. - Pairwise comparisons of unconditional Giacomini and White test for S&P 500 (1-year in-sample) - robust Patton (b=-0.5)

	Spec 22	Spec 23	Spec 24	Spec 25	Spec 26	Spec 27	Spec 28	Spec 29	Spec 30	Spec 31	Spec 32	Spec 33	Spec 34	Spec 35	Spec 36	Spec 37	Spec 38	Spec 39	Spec 40	Spec 41	Spec 42
Spec 1:A_N	6.63*	7.02*	7.61*	6.05*	7.78*	6.93*	7.03*	7.06*	6.42*	8.71*	7.05*	8.77*	7.96*	10.24*	5.20*	4.59*	3.32*	4.65*	1.88	-3.06*	-3.76*
Spec 2:GA_N	0.59	4.61*	5.65*	-1.79	6.11*	4.27*	3.60*	4.95*	-1.98*	-2.01*	-1.23	-2.12*	2.27*	3.30*	-2.70*	-3.16*	-3.85*	-2.83*	-3.46*	-5.32*	-5.54*
Spec 3:GJR_M	-5.01*	-2.90*	2.81*	-6.89*	-2.52*	0.58	-0.78	-0.55	-4.33*	-3.86*	-2.49*	-3.80*	-0.45	-0.21	-3.64*	-5.50*	-4.44*	-3.67*	-4.03*	-5.64*	-5.82*
Spec 4:TGA_N	-5.94*	-4.03*	0.07	-7.44*	-4.05*	-1.83	-4.53*	-3.79*	-4.91*	-4.77*	-3.01*	-4.64*	-1.34	-1.16	-4.16*	-6.66*	-4.87*	-4.12*	-4.39*	-5.91*	-6.04*
Spec 5:IGA_N	2.59*	6.91*	7.17*	0.67	7.63*	5.40*	4.49*	6.74*	-0.91	-1.37	-0.85	-1.50	2.58*	3.40*	-2.30*	-2.37*	-3.27*	-2.30*	-3.08*	-5.00*	-5.26*
Spec 6:AGA_N	-3.57*	1.41	2.70*	-3.87*	0.69	1.72	1.16	1.58	-6.92*	-4.63*	-2.64*	-4.56*	0.82	1.58	-3.90*	-3.86*	-5.45*	-4.40*	-4.58*	-6.29*	-6.38*
Spec 7:$NAGA_N$	-4.72*	0.60	3.28*	-6.24*	-1.46	1.87	0.51	1.22	-4.20*	-3.78*	-2.41*	-3.74*	0.09	0.44	-3.61*	-5.92*	-4.59*	-3.75*	-4.08*	-5.72*	-5.88*
Spec 8:EGA_N	-2.49*	1.15	2.48*	-2.84*	0.32	1.56	1.07	1.35	-4.17*	-4.62*	-2.68*	-4.58*	0.71	1.39	-3.85*	-3.48*	-5.42*	-4.37*	-4.62*	-6.27*	-6.36*
Spec 9:$APGA_N$	-4.87*	0.99	3.93*	-7.03*	-1.48	1.92	0.42	1.90	-4.03*	-3.55*	-2.29*	-3.51*	0.03	0.33	-3.49*	-5.81*	-4.36*	-3.56*	-3.92*	-5.56*	-5.74*
Spec 10:$FIGA_N$	2.46*	4.12*	4.74*	1.23	5.22*	3.78*	3.41*	4.12*	1.01	-1.04	-0.65	-1.28	3.03*	4.65*	-2.34*	-1.11	-3.92*	-2.66*	-3.45*	-5.49*	-5.71*
Spec 11:A_t	8.12*	8.42*	8.96*	7.52*	9.14*	8.29*	8.48*	8.50*	7.56*	8.90*	7.39*	8.78*	8.83*	10.55*	5.67*	5.93*	4.42*	5.58*	2.95*	-1.32	-2.27*
Spec 12:GA_t	-0.06	4.68*	5.68*	-2.68*	6.22*	4.25*	3.57*	4.97*	-2.13*	-2.09*	-1.26	-2.19*	2.20*	3.18*	-2.73*	-3.27*	-3.87*	-2.86*	-3.47*	-5.33*	-5.55*
Spec 13:GJR_t	-4.31*	-0.74	3.62*	-6.06*	-1.74	1.19	-0.16	0.41	-3.88*	-3.47*	-2.28*	-3.44*	-0.20	0.05	-3.43*	-5.16*	-4.21*	-3.45*	-3.85*	-5.49*	-5.69*
Spec 14:TGA_t	-5.28*	-3.99*	0.15	-6.65*	-3.50*	-1.97*	-4.10*	-3.61*	-4.51*	-4.38*	-2.85*	-4.29*	-1.27	-1.06	-3.97*	-6.26*	-4.64*	-3.91*	-4.23*	-5.75*	-5.90*
Spec 15:IGA_t	2.33*	6.73*	7.08*	-1.45	7.56*	5.30*	4.40*	6.58*	-1.13	-1.49	-0.93	-1.62	2.50*	3.34*	-2.37*	-2.55*	-3.35*	-2.39*	-3.14*	-5.04*	-5.30*
Spec 16:AGA_t	-6.34*	1.79	3.66*	-7.79*	2.16*	1.28	2.11*	2.40*	-5.00*	-3.94*	-2.32*	-3.88*	0.59	1.11	-3.63*	-5.03*	-4.72*	-3.79*	-4.15*	-5.89*	-6.03*
Spec 17:$NAGA_t$	-3.57*	1.02	3.88*	-4.86*	-0.74	2.60*	0.88	1.79	-3.47*	-3.28*	-2.16*	-3.28*	0.26	0.58	-3.36*	-5.42*	-4.23*	-3.41*	-3.84*	-5.52*	-5.70*
Spec 18:EGA_t	-3.03*	0.58	5.50*	-3.94*	-0.84	2.59*	1.56	1.43	-2.99*	-3.17*	-2.07*	-3.16*	0.07	0.34	-3.27*	-4.83*	-3.96*	-3.19*	-3.69*	-5.34*	-5.55*
Spec 19:$APGA_t$	-3.77*	2.32*	4.66*	-5.50*	-0.49	2.61*	1.21	3.56*	-3.42*	-3.07*	-2.02*	-3.07*	0.34	0.65	-3.22*	-5.14*	-4.07*	-3.27*	-3.71*	-5.39*	-5.60*
Spec 20:$FIGA_t$	2.37*	4.09*	4.70*	1.17	5.10*	3.75*	3.37*	4.07*	1.66	-1.07	-0.68	-1.31	2.99*	4.58*	-2.35*	-1.14	-3.94*	-2.68*	-3.46*	-5.50*	-5.71*
Spec 21:A_{st}	8.42*	8.69*	9.27*	7.80*	9.48*	8.56*	8.76*	8.77*	7.81*	9.34*	7.53*	9.21*	9.20*	11.02*	5.87*	6.13*	4.50*	5.67*	3.00*	-1.29	-2.23*
Spec 22:GA_{st}	-	4.66*	5.59*	-2.78*	6.34*	4.13*	3.48*	4.81*	-2.25*	-2.11*	-1.27	-2.21*	2.20*	3.18*	-2.74*	-3.13*	-3.90*	-2.88*	-3.49*	-5.35*	-5.57*
Spec 23:GJR_{st}	-	-	3.96*	-6.71*	-1.79	1.38	-0.03	0.72	-4.04*	-3.59*	-2.32*	-3.55*	-0.15	0.11	-3.50*	-5.42*	-4.31*	-3.53*	-3.92*	-5.55*	-5.74*
Spec 24:TGA_{st}	-	-	-	-7.03*	-3.66*	-2.58*	-4.79*	-4.25*	-4.64*	-4.56*	-2.93*	-4.45*	-1.34	-1.14	-4.08*	-6.83*	-4.76*	-4.03*	-4.31*	-5.81*	-5.95*
Spec 25:IGA_{st}	-	-	-	-	7.79*	5.25*	4.40*	6.48*	-1.08	-1.46	-0.91	-1.59	2.55*	3.41*	-2.35*	-2.45*	-3.35*	-2.38*	-3.13*	-5.05*	-5.30*
Spec 26:AGA_{st}	-	-	-	-	-	2.16*	1.28	2.11*	-5.00*	-3.94*	-2.32*	-3.88*	0.59	1.11	-3.63*	-5.03*	-4.72*	-3.79*	-4.15*	-5.89*	-6.03*
Spec 27:$NAGA_{st}$	-	-	-	-	-	-	-1.91	-1.48	-3.69*	-3.71*	-2.48*	-3.68*	-0.72	-0.47	-3.65*	-6.45*	-4.30*	-3.57*	-3.95*	-5.50*	-5.68*
Spec 28:EGA_{st}	-	-	-	-	-	-	-	0.62	-3.31*	-3.55*	-2.28*	-3.53*	-0.13	0.13	-3.49*	-5.27*	-4.20*	-3.42*	-3.88*	-5.50*	-5.68*
Spec 29:$APGA_{st}$	-	-	-	-	-	-	-	-	-4.00*	-3.72*	-2.42*	-3.68*	-0.32	-0.06	-3.60*	-6.35*	-4.40*	-3.62*	-3.98*	-5.57*	-5.75*
Spec 30:$FIGA_{st}$	-	-	-	-	-	-	-	-	-	-1.13	-0.71	-1.36	2.93*	4.48*	-2.37*	-1.17	-3.97*	-2.71*	-3.48*	-5.52*	-5.73*
Spec 31:$comb_1$	-	-	-	-	-	-	-	-	-	-	0.00	-2.56*	3.75*	7.68*	-2.55*	-0.16	-3.93*	-2.31*	-3.84*	-6.12*	-6.15*
Spec 32:$comb_2$	-	-	-	-	-	-	-	-	-	-	-	-0.33	2.71*	4.22*	-2.93*	-0.11	-7.64*	-4.63*	-6.43*	-7.36*	-7.20*
Spec 33:$comb_3$	-	-	-	-	-	-	-	-	-	-	-	-	3.86*	7.70*	-2.47*	0.04	-3.92*	-2.19*	-3.90*	-6.23*	-6.24*
Spec 34:HAR	-	-	-	-	-	-	-	-	-	-	-	-	-	0.41	-4.24*	-3.86*	-4.96*	-4.00*	-4.60*	-6.27*	-6.39*
Spec 35:logHAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-5.94*	-4.69*	-6.98*	-5.77*	-5.84*	-7.15*	-7.09*
Spec 36:HEAVY	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.40	-1.08	0.53	-2.98*	-6.66*	-6.85*
Spec 37:$GA-MIDAS$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-2.26*	-1.25	-2.36*	-4.39*	-4.73*
Spec 38:$GA-RV_N$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	5.72*	-1.91	-5.12*	-5.48*
Spec 39:$GA-RV_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-3.57*	-5.88*	-6.09*
Spec 40:$GA-RV_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-6.70*	-6.91*
Spec 41:$quant_1$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-5.12*
Spec 42:$quant_2$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Note: This table reports the out-of-sample t-statistic of the Unconditional Giacomini and White (2006) test under the robust loss of Patton (2011) for b=-0.5 for S&P 500.. The period is from 02/01/2001 to 29/12/2017. A positive t-statistic indicates that average loss of model A (in row) is larger than of model B (in column) and hence model B is superior. *denotes significance at the 5% level

Table 3.11 – Superior Predictive Ability (SPA) test and Model Confidence Set (MCS) results for the S&P 500 (1-year in-sample) - MSE, Robust Patton (b=1), Robust Patton (b=-0.5)

	MSE - Robust Patton (b=1) - Robust Patton (b=-0.5)					
	P-value (Hansen)	R-statistic	P-value (Hansen)	R-statistic	P-value (Hansen)	R-statistic
Spec 1:A_N	≤ 0.00*		0.05	✓	≤ 0.00*	
Spec 2:GA_N	0.08		0.19	✓	≤ 0.00*	
Spec 3:GJR_M	0.21	✓	0.19	✓	0.17	✓
Spec 4:TGA_N	0.91	✓	0.77	✓	0.97	✓
Spec 5:IGA_N	0.03*		0.15	✓	≤ 0.00*	
Spec 6:AGA_N	0.26	✓	0.36	✓	0.17	
Spec 7:$NAGA_N$	0.27	✓	0.47	✓	0.12	
Spec 8:EGA_N	0.24	✓	0.29	✓	0.11	
Spec 9:$APGA_N$	0.12	✓	0.16	✓	0.09	
Spec 10:$FIGA_N$	≤ 0.00*		0.08	✓	≤ 0.00*	
Spec 11:A_t	≤ 0.00*		0.06	✓	≤ 0.00*	
Spec 12:GA_t	0.08		0.19	✓	≤ 0.00*	
Spec 13:GJR_t	0.15	✓	0.17	✓	0.13	✓
Spec 14:TGA_t	0.74	✓	0.43	✓	0.93	✓
Spec 15:IGA_t	0.03*		0.14	✓	≤ 0.00*	
Spec 16:AGA_t	0.23	✓	0.34	✓	0.12	
Spec 17:$NAGA_t$	0.19	✓	0.33	✓	0.08	
Spec 18:EGA_t	0.04*	✓	0.23	✓	≤ 0.00*	
Spec 19:$APGA_t$	0.12	✓	0.15	✓	0.07	
Spec 20:$FIGA_t$	0.01*		0.09	✓	≤ 0.00*	
Spec 21:A_{st}	≤ 0.00*		0.05	✓	≤ 0.00*	
Spec 22:GA_{st}	0.08		0.18	✓	≤ 0.00*	
Spec 23:GJR_{st}	0.12	✓	0.16	✓	0.12	✓
Spec 24:TGA_{st}	1.00	✓	1.00	✓	1.00	✓
Spec 25:IGA_{st}	0.03*		0.17	✓	≤ 0.00*	
Spec 26:AGA_{st}	0.21	✓	0.32	✓	0.11	
Spec 27:$NAGA_{st}$	0.20	✓	0.57	✓	0.04*	✓
Spec 28:EGA_{st}	0.02*	✓	0.24	✓	≤ 0.00*	
Spec 29:$APGA_{st}$	0.12	✓	0.21	✓	0.05	✓
Spec 30:$FIGA_{st}$	≤ 0.00*		0.10	✓	≤ 0.00*	
Spec 31:$comb_1$	≤ 0.00*		0.50	✓	≤ 0.00*	
Spec 32:$comb_2$	0.02*	✓	0.13	✓	0.01*	
Spec 33:$comb_3$	≤ 0.00*		0.20	✓	≤ 0.00*	
Spec 34:HAR	0.26	✓	0.16	✓	0.39	✓
Spec 35:logHAR	0.64	✓	0.80	✓	0.56	✓
Spec 36:HEAVY	0.07		0.24	✓	0.01*	
Spec 37:$GA-MIDAS$	0.01*		0.13	✓	≤ 0.00*	
Spec 38:$GA-RV_N$	0.03*		0.19	✓	≤ 0.00*	
Spec 39:$GA-RV_t$	0.04*		0.18	✓	≤ 0.00*	
Spec 40:$GA-RV_{st}$	0.06		0.19	✓	≤ 0.00*	
Spec 41:$quant_1$	0.04*		0.08	✓	0.02*	
Spec 42:$quant_2$	0.05		0.16	✓	0.02*	

Note: This table reports the out-of-sample P-values of the Superior Predictive Ability test of Hansen (2005) under the MSE loss function, the robust loss of Patton (2011) for b=1 and for b=-0.5 for S¶500. The period is from 02/01/2001 to 29/12/2017. Each row name indicates the benchmark which is compared to each of the alternative models. In the same table the models that are included in the Model Confidence set (MCS) of Hansen (2011) using the range statistic are indicated by the "tick" symbol. The White's Upper P-value (Original RC P-values) of the SPA test and the SQ-method of the MCS are reported in the Appendix. *significant at 5% level

3.4.3 Robustness - Others Stock indices

Tables 3.12 - 3.17 report the results for the three additional European stock indices under examination for GW (MSE) and for MCS and SPA at 5% significant level.[7] CAC 40 index seems to support the results derived for the S&P 500 index. Even though, the second combination does not seem to perform that well. More specifically the first combination outperforms the second and it is indifferent to the 3rd (for MSE) for the other two loss functions all combinations perform quite well. What is more, logHAR performs much better than HAR for all loss functions. The MCS under all the three losses reports that the TGARCH-N, and the first and the third combination should be included. The MSE includes more specifications and the complete set is similar to the one for the S&P 500 and the robust loss (b=1) once again suggests that all specifications should be included. FTSE 100 follows the same direction with the two other indices, though it has less number of significant cases. The TGARCH, the combinations and the log HAR specifications seems to be competing for the model with the best performance. The MCS is more full that the one for the CAC 40 index and it reports that almost all models should be included but not any of the ARCH specifications, or the HAR or the GARCH-MIDAS or the HEAVY models. Moving on, the performance of the tested specifications employing DAX stock index is parallel to the one derived for the FTSE 100. However, it supplementary suggests that all IGARCH and FIGARCH specifications should not be included in the MCS. At last, table 3.31 in the Appendix reports how many times a specific model beats the rest of the models in pairwise comparisons for DAX 40, FTSE 100 and DAX using the GW test and the MSE loss function for the 4 years in-sample period.

[7] In the Appendix we report the Tables: 3.22 - 3.30 for the out-of-sample results of GW test under the robust loss function of Patton (2011) for b=1 and for b=-0.5 as well as for the rest of the non-robust loss function for each stock index respectively.

Table 3.12 – Pairwise comparisons of unconditional Giacomini and White test for CAC 40 (4-year in-sample) - MSE

	Spec 1	Spec 2	Spec 3	Spec 4	Spec 5	Spec 6	Spec 7	Spec 8	Spec 9	Spec 10	Spec 11	Spec 12	Spec 13	Spec 14	Spec 15	Spec 16	Spec 17	Spec 18	Spec 19	Spec 20	Spec 21
Spec 1:A_N	-	5.25*	5.27*	6.07*	4.71*	5.86*	5.38*	6.25*	6.00*	4.37*	-0.93	4.91*	5.31*	6.05*	4.67*	5.95*	5.31*	6.21*	5.96*	4.23*	-0.50
Spec 2:GA_N	-	-	2.38*	3.34*	-0.89	3.84*	2.77*	3.45*	3.12*	-3.19*	-6.10*	-1.82	2.31*	3.13*	-1.04	3.30*	2.61*	3.03*	2.87*	-3.08*	-6.10*
Spec 3:GJR_M	-	-	-	2.51*	-3.30*	-1.18	2.05*	0.11	1.99*	-3.02*	-5.73*	-2.43*	0.53	2.25*	-3.04*	-1.21	2.01*	-0.17	1.73	-2.93*	-5.64*
Spec 4:TGA_N	-	-	-	-	-4.22*	-2.42*	-1.48	-2.16*	-2.22*	-3.62*	-6.37*	-3.28*	-2.56*	-0.05	-4.00*	-2.44*	-0.77	-2.19*	-1.66	-3.52*	-6.26*
Spec 5:IGA_N	-	-	-	-	-	2.54*	3.63*	4.01*	3.92*	-1.88	-5.32*	-0.12	3.15*	3.88*	-0.28	2.32*	3.30*	3.40*	3.56*	-1.73	-5.26*
Spec 6:AGA_N	-	-	-	-	-	-	1.85	2.20*	2.20*	-5.09*	-6.90*	-4.69*	1.18	2.27*	-2.79*	-0.52	1.80	1.84	2.02*	-5.19*	-6.95*
Spec 7:$NAGA_N$	-	-	-	-	-	-	-	-1.19	0.77	-3.13*	-5.69*	-2.80*	-1.99*	1.22	-3.45*	-1.86	1.18	-1.31	0.55	-3.08*	-5.60*
Spec 8:EGA_N	-	-	-	-	-	-	-	-	1.82	-4.20*	-7.06*	-3.59*	-0.02	1.99*	-4.03*	-2.31*	1.28	-1.89	1.62	-4.09*	-7.03*
Spec 9:$APGA_N$	-	-	-	-	-	-	-	-	-	-3.46*	-6.30*	-3.09*	-1.99*	1.54	-3.73*	-2.23*	-0.18	-1.89	-0.48	-3.36*	-6.19*
Spec 10:$FIGA_N$	-	-	-	-	-	-	-	-	-	-	-5.32*	2.96*	2.95*	3.46*	1.98*	4.93*	2.98*	4.11*	3.28*	0.84	-5.37*
Spec 11:A_t	-	-	-	-	-	-	-	-	-	-	-	5.80*	5.76*	6.33*	5.30*	7.05*	5.58*	7.11*	6.23*	5.16*	1.85
Spec 12:GA_t	-	-	-	-	-	-	-	-	-	-	-	-	2.35*	3.08*	0.05	3.95*	2.63*	3.32*	2.86*	-3.18*	-5.82*
Spec 13:GJR_t	-	-	-	-	-	-	-	-	-	-	-	-	-	2.32*	-2.90*	-1.22	2.05*	-0.24	1.75	-2.84*	-5.67*
Spec 14:TGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-3.69*	-2.31*	-0.70	-2.05*	-2.48*	-3.35*	-6.22*
Spec 15:IGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2.50*	3.14*	3.50*	3.40*	-1.87	-5.26*
Spec 16:AGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.82	1.97*	2.05*	-4.89*	-7.11*
Spec 17:$NAGA_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-1.37	0.05	-2.91*	-5.48*
Spec 18:EGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.71	-4.00*	-7.11*
Spec 19:$APGA_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-3.17*	-6.11*
Spec 20:$FIGA_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-5.20*
Spec 21:A_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 22:GA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 23:GJR_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 24:TGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 25:IGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 26:AGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 27:$NAGA_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 28:EGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 29:$APGA_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 30:$FIGA_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 31:$comb_1$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 32:$comb_2$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 33:$comb_3$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 34:HAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 35:logHAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 36:HEAVY	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 37:$GA-MIDAS$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 38:$GA-RV_N$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 39:$GA-RV_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 40:$GA-RV_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 41:$quant_1$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 42:$quant_2$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Note: This table reports the out-of-sample t-statistic of the Unconditional Giacomini and White (2006) test under the MSE loss function for CAC 40. The period is from 02/01/2004 to 29/12/2017. A positive t-statistic indicates that average loss of model A (in row) is larger than of model B (in column) and hence model B is superior. *denotes significance at the 5% level

Table 3.12 cont. - Pairwise comparisons of unconditional Giacomini and White test for CAC 40 (4-year in-sample) - MSE

	Spec 22	Spec 23	Spec 24	Spec 25	Spec 26	Spec 27	Spec 28	Spec 29	Spec 30	Spec 31	Spec 32	Spec 33	Spec 34	Spec 35	Spec 36	Spec 37	Spec 38	Spec 39	Spec 40	Spec 41	Spec 42
Spec 1:A_N	4.86*	5.38*	6.16*	4.68*	5.95*	5.39*	6.11*	6.09*	4.20*	6.45*	6.15*	6.40*	-5.53*	5.45*	0.36	4.38*	1.06	0.80	0.03	-3.45*	-3.63*
Spec 2:GA_N	-1.87	2.41*	3.18*	-1.03	3.30*	2.71*	2.52*	2.89*	-3.02*	4.83*	4.56*	4.67*	-6.73*	2.23*	-3.71*	-3.85*	-1.55	-1.80	-1.90	-5.75*	-5.87*
Spec 3:GJR_M	-2.41*	0.30	2.14*	-3.01*	-1.21	2.16*	-0.55	1.57	-2.89*	2.30*	1.27	2.06*	-6.71*	0.07	-3.97*	-3.26*	-2.56*	-2.76*	-2.64*	-5.38*	-5.49*
Spec 4:TGA_N	-3.23*	-2.55*	-0.85	-3.96*	-2.44*	-0.71	-2.31*	-1.90	-3.48*	-0.15	-0.88	-0.24	-6.76*	-1.36	-4.80*	-4.09*	-3.01*	-2.96*	-3.16*	-5.46*	-5.56*
Spec 5:IGA_N	-0.30	3.26*	3.92*	-0.16	2.32*	3.40*	2.73*	3.56*	-1.72	5.84*	5.28*	5.57*	-6.59*	2.55*	-3.36*	-2.29*	-1.01	-1.30	-1.42	-5.36*	-5.48*
Spec 6:AGA_N	-4.85*	1.23	2.30*	-2.81*	-0.52	1.89	1.25	2.02*	-5.08*	3.54*	3.08*	3.40*	-6.89*	1.21	-4.35*	-5.88*	-3.01*	-3.19*	-3.21*	-6.14*	-6.25*
Spec 7:$NAGA_N$	-2.76*	-1.93	0.88	-3.41*	-1.86	1.45	-1.51	0.27	-3.05*	0.62	-0.22	0.49	-6.68*	-0.86	-4.31*	-3.49*	-2.31*	-2.37*	-2.47*	-5.31*	-5.42*
Spec 8:EGA_N	-3.57*	-0.06	2.03*	-3.99*	-2.31*	1.37	-2.66*	1.61	-4.04*	2.99*	1.54	2.66*	-6.88*	0.00	-4.94*	-4.98*	-3.07*	-3.12*	-3.32*	-5.92*	-6.02*
Spec 9:$APGA_N$	-3.05*	-2.00*	0.69	-3.69*	-2.23*	-0.09	-2.05*	-1.13	-3.33*	0.18	-0.60	0.07	-6.74*	-1.17	-4.76*	-3.95*	-2.61*	-2.64*	-2.66*	-5.42*	-5.53*
Spec 10:$FIGA_N$	2.91*	3.04*	3.52*	2.02*	4.93*	3.06*	3.96*	3.33*	0.58	5.06*	4.98*	5.02*	-6.74*	3.66*	-3.06*	-0.49	-1.19	-1.45	-1.54	-5.86*	-5.99*
Spec 11:A_t	5.79*	5.86*	6.48*	5.31*	7.05*	5.68*	7.10*	6.39*	5.14*	7.03*	6.86*	6.99*	-5.56*	6.04*	0.63	5.42*	2.23*	2.05*	1.06	-3.55*	-3.74*
Spec 12:GA_t	-1.60	2.44*	3.13*	0.09	3.95*	2.72*	2.94*	2.89*	-3.14*	4.86*	4.81*	4.76*	-6.75*	2.66*	-3.64*	-3.43*	-1.99*	2.02	2.10	-5.88*	-6.00*
Spec 13:GJR_t	-2.33*	-0.39	2.20*	-2.87*	-1.22	2.21*	-0.60	1.58	-2.81*	2.09*	1.11	1.88	-6.69*	0.02	-3.96*	-3.21*	1.99*	1.98*	2.14*	-5.32*	-5.43*
Spec 14:TGA_t	-3.05*	-2.31*	-1.27	-3.65*	-2.31*	-0.64	-2.18*	-2.37*	-3.32*	-0.12	-0.79	-0.21	-6.73*	-1.27	-4.76*	-3.92*	-2.66*	-2.87*	-2.72*	-5.39*	-5.49*
Spec 15:IGA_t	-0.27	3.00*	3.72*	0.85	2.50*	3.24*	2.86*	3.40*	-1.85	5.79*	5.41*	5.58*	-6.63*	2.68*	-3.41*	-2.40*	1.13	1.40	1.33	-5.46*	-5.58*
Spec 16:AGA_t	-4.14*	1.27	2.34*	-2.50*	-1.76	1.91	1.39	2.07*	-4.80*	3.49*	3.00*	3.35*	-6.89*	1.21	-4.36*	-5.95*	-3.01*	-2.71*	-2.78*	-6.12*	-6.23*
Spec 17:$NAGA_t$	-2.60*	-1.97*	0.38	-3.11*	-1.82	0.60	-1.54	-0.16	-2.89*	0.23	-0.42	0.14	-6.64*	-0.94	-4.27*	-3.31*	1.70	-1.88	-1.96*	-5.19*	-5.30*
Spec 18:EGA_t	-3.35*	0.22	2.08*	-3.46*	-1.97*	1.47	-3.15*	1.71	-3.96*	3.05*	1.81	2.78*	-6.89*	0.24	-4.96*	-4.90*	-3.01*	-3.12*	-3.19*	-5.98*	-6.08*
Spec 19:$APGA_t$	-2.84*	-1.76	1.21	-3.37*	-2.05*	0.02	-1.88	-1.14	-3.14*	0.24	-0.49	0.13	-6.71*	-1.05	-4.71*	-3.75*	-2.11*	-2.28*	-2.21	-5.35*	-5.46*
Spec 20:$FIGA_t$	3.23*	2.93*	3.40*	1.91	4.89*	3.00*	3.85*	3.21*	-1.13	4.98*	4.97*	4.95*	-6.77*	3.57*	-3.17*	-0.82	-0.95	-1.32	-1.52	-5.96*	-6.09*
Spec 21:A_{st}	5.82*	5.78*	6.37*	5.27*	7.11*	5.58*	7.12*	6.28*	5.19*	6.98*	6.84*	6.95*	-5.62*	6.02*	0.55	5.48*	0.56	-0.96	1.21	-3.68*	-3.87*
Spec 22:GA_{st}	-	2.42*	3.09*	0.31	4.14*	2.69*	3.02*	2.86*	-3.22*	4.75*	4.72*	4.67*	-6.77*	2.72*	-3.63*	-3.36*	-1.17	-1.45	-1.60	-5.94*	-6.06*
Spec 23:GJR_{st}	-	-	2.21*	-2.96*	-1.27	2.14*	-0.61	1.60	-2.89*	2.20*	1.19	1.98*	-6.70*	0.04	-4.01*	-3.32*	-2.08*	-2.34*	-2.32*	-5.38*	-5.49*
Spec 24:TGA_{st}	-	-	-	-3.68*	-2.34*	-0.32	-2.23*	-2.31*	-3.37*	0.02	-0.70	-0.07	-6.75*	-1.21	-4.83*	-4.02*	-2.62*	-3.01*	-2.75*	-5.44*	-5.54*
Spec 25:IGA_{st}	-	-	-	-	2.50*	3.21*	2.84*	3.37*	-1.90	5.76*	5.40*	5.55*	-6.63*	2.67*	-3.42*	-2.44*	-1.10	-1.23	-1.52	-5.47*	-5.59*
Spec 26:AGA_{st}	-	-	-	-	-	1.91	1.39	2.07*	-4.80*	3.49*	3.00*	3.35*	-6.89*	1.21	-4.36*	-5.95*	-2.83*	-2.96*	-2.71*	-6.12*	-6.23*
Spec 27:$NAGA_{st}$	-	-	-	-	-	-	-1.64	-0.25	-2.97*	0.19	-0.48	0.10	-6.66*	-0.99	-4.35*	-3.42*	-1.62	-1.97*	-2.14*	-5.24*	-5.35*
Spec 28:EGA_{st}	-	-	-	-	-	-	-	1.90	-3.82*	3.27*	2.24*	3.05*	-6.89*	0.61	-4.92*	-4.75*	-2.98**	-3.12*	-3.21*	-6.04*	-6.14*
Spec 29:$APGA_{st}$	-	-	-	-	-	-	-	-	-3.18*	0.42	-0.37	0.30	-6.73*	-0.99	-4.81*	-3.85*	-2.73*	-2.30*	-2.84*	-5.40*	-5.51*
Spec 30:$FIGA_{st}$	-	-	-	-	-	-	-	-	-	4.91*	4.90*	4.88*	-6.78*	3.57*	-3.16*	-0.73	-0.86	-1.28	-1.51	-6.00*	-6.13*
Spec 31:$comb_1$	-	-	-	-	-	-	-	-	-	-	-2.34*	-1.21	-7.00*	-2.69*	-5.30*	-5.41*	-3.42*	-3.81*	-3.30*	-6.04*	-6.15*
Spec 32:$comb_2$	-	-	-	-	-	-	-	-	-	-	-	2.26*	-6.99*	-1.58	-4.95*	-5.26*	-3.72*	-4.02*	-3.28*	-6.09*	-6.19*
Spec 33:$comb_3$	-	-	-	-	-	-	-	-	-	-	-	-	-7.01*	-2.82*	-5.28*	-5.34*	-3.72*	-3.97*	-3.85*	-6.06*	-6.16*
Spec 34:HAR	-	-	-	-	-	-	-	-	-	-	-	-	-	6.97*	5.97*	6.60*	7.01*	6.53*	6.81*	4.48*	4.34*
Spec 35:logHAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-4.74*	-3.62*	-4.02*	-4.17*	-3.96*	-6.09*	-6.20*
Spec 36:HEAVY	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2.99*	-1.41	-1.52	-2.78*	-4.33*	-4.54*
Spec 37:$GA-MIDAS$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-1.13	-1.34	-1.55	-5.63*	-5.76*
Spec 38:$GA-RV_N$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-2.75*	-1.65	-4.88*	-4.23*
Spec 39:$GA-RV_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-1.41	-3.78*	-3.97*
Spec 40:$GA-RV_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-6.97*	-5.24*
Spec 41:$quant_1$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-5.84*
Spec 42:$quant_2$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Note: This table reports the out-of-sample t-statistic of the Unconditional Giacomini and White (2006) test under the MSE loss function for CAC 40. The period is from 02/01/2004 to 29/12/2017. A positive t-statistic indicates that average loss of model A (in row) is larger than of model B (in column) and hence model B is superior. *denotes significance at the 5% level

Table 3.13 – Superior Predictive Ability (SPA) test and Model Confidence Set (MCS) results for the CAC 40 (4-year in-sample) - MSE, Robust Patton (b=1), Robust Patton (b=-0.5)

	MSE - Robust Patton (b=1) - Robust Patton (b=-0.5)					
	P-value (Hansen)	R-statistic	P-value (Hansen)	R-statistic	P-value (Hansen)	R-statistic
Spec 1:A_N	≤ 0.00*		0.03*	✓	≤ 0.00*	
Spec 2:GA_N	≤ 0.00*		0.05	✓	≤ 0.00*	
Spec 3:GJR_M	0.06	✓	0.15	✓	≤ 0.00*	
Spec 4:TGA_N	1.00	✓	0.89	✓	0.48	✓
Spec 5:IGA_N	≤ 0.00*		0.02*	✓	≤ 0.00*	
Spec 6:AGA_N	0.05	✓	0.16	✓	≤ 0.00*	
Spec 7:$NAGA_N$	0.29	✓	0.33	✓	0.08	
Spec 8:EGA_N	0.04*	✓	0.55	✓	≤ 0.00*	
Spec 9:$APGA_N$	0.31	✓	0.47	✓	0.13	
Spec 10:$FIGA_N$	0.02*	✓	0.07	✓	≤ 0.00*	
Spec 11:A_t	≤ 0.00*		0.01*	✓	≤ 0.00*	
Spec 12:GA_t	0.01*		0.04*	✓	≤ 0.00*	
Spec 13:GJR_t	0.08	✓	0.11	✓	≤ 0.00*	
Spec 14:TGA_t	0.96	✓	1.00	✓	0.39	
Spec 15:IGA_t	≤ 0.00*		0.02*	✓	≤ 0.00*	
Spec 16:AGA_t	0.06	✓	0.16	✓	0.01*	
Spec 17:$NAGA_t$	0.62	✓	0.80	✓	0.16	
Spec 18:EGA_t	0.07	✓	0.53	✓	≤ 0.00*	
Spec 19:$APGA_t$	0.16	✓	0.42	✓	0.04*	
Spec 20:$FIGA_t$	0.01*	✓	0.06	✓	≤ 0.00*	
Spec 21:A_{st}	≤ 0.00*		0.02*	✓	≤ 0.00*	
Spec 22:GA_{st}	0.01*		0.04*	✓	≤ 0.00*	
Spec 23:GJR_{st}	0.06	✓	0.17	✓	≤ 0.00*	
Spec 24:TGA_{st}	0.69	✓	0.74	✓	0.256	
Spec 25:IGA_{st}	≤ 0.00*		0.03*	✓	≤ 0.00*	
Spec 26:AGA_{st}	0.06	✓	0.18	✓	0.01*	
Spec 27:$NAGA_{st}$	0.72	✓	0.93	✓	0.17	
Spec 28:EGA_{st}	0.07	✓	0.33	✓	0.005*"≤ 0.00*	
Spec 29:$APGA_{st}$	0.11	✓	0.32	✓	0.04*	
Spec 30:$FIGA_{st}$	0.01*	✓	0.09	✓	≤ 0.00*	
Spec 31:$comb_1$	0.97	✓	0.77	✓	1.00	✓
Spec 32:$comb_2$	0.03*	✓	0.22	✓	≤ 0.00*	
Spec 33:$comb_3$	0.35	✓	0.53	✓	0.30	✓
Spec 34:HAR	0.01*		0.06	✓	≤ 0.00*	
Spec 35:logHAR	0.10	✓	0.29	✓	0.07	
Spec 36:HEAVY	≤ 0.00*		0.08	✓	≤ 0.00*	
Spec 37:$GA - MIDAS$	0.02*		0.05	✓	≤ 0.00*	
Spec 38:$GA - RV_N$	0.05	✓	0.10	✓	0.02*	
Spec 39:$GA - RV_t$	0.07	✓	0.15	✓	0.02*	
Spec 40:$GA - RV_{st}$	0.08	✓	0.15	✓	0.01*	
Spec 41:$quant_1$	0.02*		0.10	✓	≤ 0.00*	
Spec 42:$quant_2$	0.03*		0.08	✓	0.01*	

Note: This table reports the out-of-sample P-values of the Superior Predictive Ability test of Hansen (2005) under the MSE loss function, the robust loss of Patton (2011) for b=1 and for b=-0.5 for CAC 40. The period is from 02/01/2004 to 29/12/2017. Each row name indicates the benchmark which is compared to each of the alternative models. In the same table the models that are included in the Model Confidence set (MCS) of Hansen (2011) using the range statistic are indicated by the "tick" symbol. *significant at 5% level

Table 3.14 – Pairwise comparisons of unconditional Giacomini and White test for FTSE 100 (4-year in-sample) -MSE

	Spec 1	Spec 2	Spec 3	Spec 4	Spec 5	Spec 6	Spec 7	Spec 8	Spec 9	Spec 10	Spec 11	Spec 12	Spec 13	Spec 14	Spec 15	Spec 16	Spec 17	Spec 18	Spec 19	Spec 20	Spec 21
Spec 1:A_N	-	2.65*	3.04*	3.23*	2.58*	2.99*	2.98*	3.41*	3.12*	2.68*	0.69	2.63*	3.05*	3.23*	2.58*	3.02*	2.94*	3.45*	3.15*	2.67*	0.72
Spec 2:GA_N	-	-	1.62	4.02*	-3.03*	3.80*	1.48	4.54*	3.22*	0.58	-3.22*	-0.36	1.56	3.60*	-2.94*	3.59*	1.29	4.42*	3.11*	0.58	-3.20*
Spec 3:GJR_M	-	-	-	0.57	-1.83	-0.42	-0.69	0.94	0.22	-1.23	-3.28*	-1.56	0.67	0.70	-1.82	-0.30	-0.84	1.02	0.43	-1.13	-3.25*
Spec 4:TGA_N	-	-	-	-	-4.29*	-1.99*	-1.18	1.67	-1.95	-3.33*	-3.85*	-3.91*	-0.45	0.13	-4.28*	-1.62	-1.11	1.83	-0.90	-3.06*	-3.83*
Spec 5:IGA_N	-	-	-	-	-	4.09*	1.70	4.76*	3.52*	1.21	-3.12*	1.89	1.76	3.88*	0.13	3.91*	1.50	4.63*	3.40*	1.15	-3.10*
Spec 6:AGA_N	-	-	-	-	-	-	0.19	4.01*	1.10	-3.82*	-3.75*	-4.31*	0.44	1.64	-4.16*	1.84	0.13	3.91*	1.16	-3.90*	-3.74*
Spec 7:$NAGA_N$	-	-	-	-	-	-	-	1.42	0.82	-1.10	-3.26*	-1.42	0.73	1.45	-1.69	-0.07	-0.49	1.52	1.11	-1.00	-3.23*
Spec 8:EGA_N	-	-	-	-	-	-	-	-	-1.78	-4.85*	-4.23*	-4.64*	-0.83	-1.20	-4.78*	-3.78*	-1.35	2.08*	-1.40	-4.79*	-4.22*
Spec 9:$APGA_N$	-	-	-	-	-	-	-	-	-	-2.45*	-3.65*	-3.10*	-0.11	3.64*	-3.50*	-0.83	-0.79	1.88	1.35	-2.23*	-3.63*
Spec 10:$FIGA_N$	-	-	-	-	-	-	-	-	-	-	-3.27*	-0.74	1.21	2.90*	-1.23	4.14*	0.97	4.75*	2.41*	0.24	-3.25*
Spec 11:A_t	-	-	-	-	-	-	-	-	-	-	-	3.22*	3.26*	3.80*	3.12*	3.82*	3.18*	4.29*	3.67*	3.30*	1.01
Spec 12:GA_t	-	-	-	-	-	-	-	-	-	-	-	-	1.51	3.46*	-2.00*	4.07*	1.25	4.50*	2.99*	0.76	-3.20*
Spec 13:GJR_t	-	-	-	-	-	-	-	-	-	-	-	-	-	0.55	-1.74	-0.33	-0.90	0.91	0.29	-1.12	-3.23*
Spec 14:TGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-3.86*	-1.35	-1.35	1.37	-2.18*	-2.65*	-3.78*
Spec 15:IGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3.97*	1.49	4.65*	3.38*	1.17	-3.10*
Spec 16:AGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.02	3.74*	0.92	-4.60*	-3.81*
Spec 17:$NAGA_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.44	1.06	-0.89	-3.15*
Spec 18:EGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-1.53	-4.70*	-4.28*
Spec 19:$APGA_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-2.20*	-3.64*
Spec 20:$FIGA_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-3.28*
Spec 21:A_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 22:GA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 23:GJR_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 24:TGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 25:IGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 26:AGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 27:$NAGA_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 28:EGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 29:$APGA_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 30:$FIGA_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 31:$comb_1$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 32:$comb_2$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 33:$comb_3$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 34:HAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 35:logHAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 36:HEAVY	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 37:$GA-MIDAS$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 38:$GA-RV_N$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 39:$GA-RV_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 40:$GA-RV_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 41:$quant_1$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 42:$quant_2$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Note: This table reports the out-of-sample t-statistic of the Unconditional Giacomini and White (2006) test under the MSE loss function for FTSE 100. The period is from 02/01/2004 to 29/12/2017. A positive t-statistic indicates that average loss of model A (in row) is larger than of model B (in column) and hence model B is superior. *denotes significance at the 5% level

Table 3.14 cont. - Pairwise comparisons of unconditional Giacomini and White test for FTSE 100 (4-year in-sample) - MSE

	Spec 22	Spec 23	Spec 24	Spec 25	Spec 26	Spec 27	Spec 28	Spec 29	Spec 30	Spec 31	Spec 32	Spec 33	Spec 34	Spec 35	Spec 36	Spec 37	Spec 38	Spec 39	Spec 40	Spec 41	Spec 42
Spec 1:A_N	2.65*	3.05*	3.25*	2.59*	3.06*	2.95*	3.52*	3.23*	2.70*	3.44*	3.30*	3.43*	-3.57*	2.79*	-1.52	-5.26*	1.06	0.95	0.02	-0.58	-0.64
Spec 2:GA_N	0.38	1.48	3.31*	-2.50*	3.90*	1.24	4.44*	3.29*	0.72	5.22*	5.51*	5.22*	-5.58*	0.92	-4.02*	-8.13*	-1.49	-1.71	-1.81	-3.04*	-3.09*
Spec 3:GJR_M	-1.51	0.60	0.92	-1.83	-0.22	-0.65	1.21	0.88	-1.13	1.04	0.74	1.05	-5.33*	-0.22	-3.70*	-7.42*	-2.47*	-2.62*	-2.52*	-2.88*	-2.92*
Spec 4:TGA_N	-3.82*	-0.36	0.32	-4.26*	-1.45	-0.89	2.22*	0.21	-3.09*	1.99*	0.92	2.14*	-5.88*	-0.69	-4.44*	-8.42*	-3.12*	-2.96*	-2.94*	-3.52*	-3.56*
Spec 5:IGA_N	2.52*	1.66	3.58*	0.72	4.21*	1.43	4.65*	3.56*	1.32	5.48*	5.77*	5.47*	-5.53*	1.07	-3.95*	-8.07*	-1.01	-1.28	-1.45	-2.97*	-3.02*
Spec 6:AGA_N	-4.19*	0.45	1.50	-3.98*	2.82*	0.18	3.88*	1.46	-3.83*	3.81*	3.39*	3.81*	-5.92*	0.02	-4.54*	-8.66*	-3.23*	-4.37*	-3.01*	-3.54*	-3.59*
Spec 7:$NAGA_N$	-1.36	0.74	1.85	-1.69	0.02	0.10	1.75	1.71	-1.00	1.63	1.23	1.66	-5.36*	-0.08	-3.73*	-7.51*	-2.31*	-2.47*	-3.44*	-2.64*	-2.93*
Spec 8:EGA_N	-4.59*	-0.73	-0.87	-4.71*	-3.61*	-1.17	2.53*	-0.91	-4.79*	-0.24	-1.26	-0.19	-6.16*	-1.45	-4.88*	-8.92*	-2.07*	-3.21*	-3.23*	-3.92*	-3.96*
Spec 9:$APGA_N$	-3.00*	-0.04	2.54*	-3.50*	-0.67	-0.58	2.17*	2.45*	-2.24*	2.12*	1.64	2.22*	-5.74*	-0.41	-4.25*	-8.19*	-3.64*	-2.21*	-2.85*	-3.34*	-3.38*
Spec 10:$FIGA_N$	-0.56	1.16	2.64*	-1.14	4.43*	0.95	4.70*	2.62*	0.73	4.80*	4.49*	4.77*	-5.75*	0.90	-4.28*	-8.46*	-1.07	-1.43	-1.61	-3.27*	-3.31*
Spec 11:A_t	3.25*	3.22*	3.77*	3.13*	3.89*	3.14*	4.37*	3.75*	3.34*	4.22*	4.05*	4.20*	-4.24*	3.61*	-2.10*	-6.47*	3.25*	2.47*	1.07	-1.03	-1.10
Spec 12:GA_t	2.69*	1.44	3.17*	-1.58	4.35*	1.20	4.48*	3.15*	0.92	5.15*	5.44*	5.14*	-5.61*	0.97	-4.07*	-8.20*	-1.07	-1.21	-1.13	-3.08*	-3.13*
Spec 13:GJR_t	-1.46	0.51	0.73	-1.75	-0.25	-0.76	1.09	0.69	-1.12	0.90	0.61	0.92	-5.30*	-0.24	-3.66*	-7.36*	-1.98*	-3.15*	-2.62*	-3.01*	-2.89*
Spec 14:TGA_t	-3.39*	-0.43	0.55	-3.86*	-1.21	-1.06	1.74	0.29	-2.68*	1.41	0.58	1.49	-5.80*	-0.65	-4.32*	-8.26*	-2.78*	-3.16*	-2.56*	-3.42*	-3.46*
Spec 15:IGA_t	2.67*	1.65	3.56*	0.60	4.27*	1.42	4.66*	3.54*	1.34	5.49*	5.78*	5.47*	-5.54*	1.08	-3.96*	-8.08*	-1.04	-1.51	-1.38	-2.98*	-3.03*
Spec 16:AGA_t	-3.98*	0.35	1.25	-3.82*	2.41*	0.07	3.71*	1.21	-4.44*	3.35*	2.79*	3.34*	-5.98*	-0.09	-4.63*	-8.77*	-3.78*	-2.16*	-3.76*	-3.63*	-3.68*
Spec 17:$NAGA_t$	-1.20	0.90	1.70	-1.49	0.06	0.63	1.66	1.59	-0.88	1.53	1.16	1.55	-5.30*	-0.05	-3.66*	-7.41*	-1.74	-1.54	-1.96*	-2.83*	-2.87*
Spec 18:EGA_t	-4.46*	-0.81	-1.04	-4.59*	-3.60*	-1.26	2.60*	-1.07	-4.71*	-0.70	-1.45	-0.66	-6.19*	-1.58	-4.92*	-8.95*	-2.87*	-3.00*	-4.14*	-3.96*	-4.00*
Spec 19:$APGA_t$	-2.91*	-0.19	3.03*	-3.38*	-0.78	-0.80	1.83	3.41*	-2.22*	1.63	1.07	1.70	-5.72*	-0.47	-4.22*	-8.14*	-3.15*	-2.38*	-2.75*	-3.32*	-3.36*
Spec 20:$FIGA_t$	-0.59	1.08	2.42*	-1.08	4.79*	0.87	4.62*	2.40*	0.68	4.51*	4.19*	4.49*	-5.81*	0.94	-4.37*	-8.58*	-0.92	-1.21	-1.34	-3.35*	-3.39*
Spec 21:A_{st}	3.23*	3.19*	3.75*	3.11*	3.88*	3.11*	4.36*	3.73*	3.32*	4.20*	4.03*	4.19*	-4.27*	3.60*	-2.14*	-6.52*	3.22*	2.61*	1.01	-1.07	-1.14
Spec 22:GA_{st}	-	1.39	3.10*	-2.18*	4.30*	1.16	4.46*	3.08*	0.75	5.11*	5.36*	5.09*	-5.62*	0.93	-4.08*	-8.21*	-1.11	-1.23	-1.61	-3.10*	-3.15*
Spec 23:GJR_{st}	-	-	0.58	-1.66	-0.27	-0.81	0.97	0.54	-1.08	0.78	0.51	0.80	-5.26*	-0.25	-3.61*	-7.27*	-2.21*	-4.24*	-2.75*	-2.81*	-2.86*
Spec 24:TGA_{st}	-	-	-	-3.57*	-1.13	-1.33	1.38	-0.61	-2.45*	0.99	0.29	1.03	-5.74*	-0.65	-4.24*	-8.14*	-2.71*	-3.21*	-2.45*	-3.36*	-3.40*
Spec 25:IGA_{st}	-	-	-	-	4.13*	1.42	4.62*	3.55*	1.25	5.45*	5.74*	5.43*	-5.53*	1.06	-3.95*	-8.06*	-1.12	-1.37	-1.45	-2.97*	-3.02*
Spec 26:AGA_{st}	-	-	-	-	-	-0.00	3.62*	1.09	-4.76*	3.22*	2.59*	3.21*	-5.99*	-0.19	-4.64*	-8.77*	-3.23*	-4.05*	-3.26*	-3.65*	-3.69*
Spec 27:$NAGA_{st}$	-	-	-	-	-	-	1.46	1.25	-0.87	1.30	0.97	1.32	-5.26*	-0.09	-3.61*	-7.32*	-1.99	-2.97*	-2.14*	-2.79*	-2.83*
Spec 28:EGA_{st}	-	-	-	-	-	-	-	-1.39	-4.65*	-1.46	-1.83	-1.41	-6.20*	-1.74	-4.93*	-8.94*	-3.12*	-3.12*	-3.07*	-3.98*	-4.02*
Spec 29:$APGA_{st}$	-	-	-	-	-	-	-	-	-2.43*	1.02	0.35	1.07	-5.74*	-0.63	-4.24*	-8.14*	-2.98*	-4.50*	-2.31*	-3.35*	-3.39*
Spec 30:$FIGA_{st}$	-	-	-	-	-	-	-	-	-	4.62*	4.25*	4.58*	-5.79*	0.89	-4.35*	-8.54*	-0.95	-1.31	-1.41	-3.33*	-3.37*
Spec 31:$comb_1$	-	-	-	-	-	-	-	-	-	-	-1.47	0.44	-6.05*	-1.26	-4.69*	-8.69*	-4.67*	-4.01*	-3.37*	-3.76*	-3.80*
Spec 32:$comb_2$	-	-	-	-	-	-	-	-	-	-	-	1.59	-5.97*	-0.92	-4.58*	-8.58*	-3.97*	-3.42*	-3.58*	-3.65*	-3.69*
Spec 33:$comb_3$	-	-	-	-	-	-	-	-	-	-	-	-	-6.05*	-1.26	-4.69*	-8.68*	-4.01*	-4.12*	-4.21*	-3.76*	-3.80*
Spec 34:HAR	-	-	-	-	-	-	-	-	-	-	-	-	-	6.54*	3.95*	-1.58	5.42*	4.57*	5.81*	4.86*	4.79*
Spec 35:logHAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-5.81*	-10.07*	-4.12*	-5.87*	-4.59*	-4.63*	-4.67*
Spec 36:HEAVY	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-8.44*	-1.42	-1.53	-3.81*	7.34*	6.97*
Spec 37:$GA-MIDAS$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-1.12	-1.31	-1.56	9.46*	9.36*
Spec 38:$GA-RV_N$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-3.87*	-1.89	-5.16*	-3.23*
Spec 39:$GA-RV_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-1.44	-4.23*	-4.23*
Spec 40:$GA-RV_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-5.94*	-6.50*
Spec 41:$quant_1$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-5.03*
Spec 42:$quant_2$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Note: This table reports the out-of-sample t-statistic of the Unconditional Giacomini and White (2006) test under the MSE loss function for S&P 500. The period is from 02/01/2004 to 29/12/2017. A positive t-statistic indicates that average loss of model A (in row) is larger than of model B (in column) and hence model B is superior. *denotes significance at the 5% level

Table 3.15 – Superior Predictive Ability (SPA) test and Model Confidence Set (MCS) results for the FTSE 100 (4-year in-sample) - MSE, Robust Patton (b=1), Robust Patton (b=-0.5)

	MSE - Robust Patton (b=1)		- Robust Patton (b=-0.5)			
	P-value (Hansen)	R-statistic	P-value (Hansen)	R-statistic	P-value (Hansen)	R-statistic
Spec 1:A_N	0.03*		0.08	✓	≤ 0.00*	
Spec 2:GA_N	0.04*	✓	0.12	✓	0.01*	✓
Spec 3:GJR_M	0.66	✓	0.60	✓	0.53	✓
Spec 4:TGA_N	0.22	✓	0.43	✓	0.08	✓
Spec 5:IGA_N	0.04*	✓	0.06	✓	0.01*	
Spec 6:AGA_N	0.06	✓	0.14	✓	0.02*	✓
Spec 7:$NAGA_N$	0.40	✓	0.50	✓	0.24	✓
Spec 8:EGA_N	0.20	✓	0.26	✓	0.35	✓
Spec 9:$APGA_N$	0.23	✓	0.31	✓	0.09	✓
Spec 10:$FIGA_N$	0.03*	✓	0.15	✓	0.01*	✓
Spec 11:A_t	0.02*		0.09	✓	≤ 0.00*	
Spec 12:GA_t	0.06	✓	0.14	✓	0.01*	✓
Spec 13:GJR_t	0.71	✓	0.58	✓	0.42	✓
Spec 14:TGA_t	0.45	✓	0.58	✓	0.12	✓
Spec 15:IGA_t	0.04*	✓	0.05	✓	≤ 0.00*	
Spec 16:AGA_t	0.08	✓	0.16	✓	0.01*	✓
Spec 17:$NAGA_t$	0.44	✓	0.53	✓	0.22	✓
Spec 18:EGA_t	0.13	✓	0.22	✓	0.14	✓
Spec 19:$APGA_t$	0.20	✓	0.25	✓	0.13	✓
Spec 20:$FIGA_t$	0.02*	✓	0.13	✓	0.01*	✓
Spec 21:A_{st}	0.01*		0.09	✓	≤ 0.00*	
Spec 22:GA_{st}	0.06	✓	0.11	✓	0.02*	✓
Spec 23:GJR_{st}	0.58	✓	0.51	✓	0.54	✓
Spec 24:TGA_{st}	0.64	✓	0.67	✓	0.33	✓
Spec 25:IGA_{st}	0.03*	✓	0.06	✓	≤ 0.00*	
Spec 26:AGA_{st}	0.08	✓	0.16	✓	0.02*	✓
Spec 27:$NAGA_{st}$	0.44	✓	0.46	✓	0.27	✓
Spec 28:EGA_{st}	1.00	✓	1.00	✓	0.81	✓
Spec 29:$APGA_{st}$	0.63	✓	0.64	✓	0.28	✓
Spec 30:$FIGA_{st}$	0.03*	✓	0.11	✓	0.01*	✓
Spec 31:$comb_1$	0.74	✓	0.58	✓	1.00	✓
Spec 32:$comb_2$	0.38	✓	0.49	✓	0.22	✓
Spec 33:$comb_3$	0.70	✓	0.55	✓	0.99	✓
Spec 34:HAR	0.01*		0.09	✓	≤ 0.00*	
Spec 35:logHAR	0.20	✓	0.38	✓	0.16	✓
Spec 36:HEAVY	≤ 0.00*		0.12	✓	≤ ≤ 0.00*	
Spec 37:$GA-MIDAS$	0.01*		0.05	✓	≤ 0.00*	
Spec 38:$GA-RV_N$	0.06	✓	0.21	✓	0.01*	
Spec 39:$GA-RV_t$	0.05	✓	0.12	✓	0.02*	
Spec 40:$GA-RV_{st}$	0.07	✓	0.14	✓	0.01*	
Spec 41:$quant_1$	0.07	✓	0.24	✓	0.02*	
Spec 42:$quant_2$	0.07	✓	0.23	✓	0.03*	

Note: This tale reports the out-of-sample P-values of the Superior Predictive Ability test of Hansen (2005) under the MSE loss functions, the robust loss of Patton (2011) for b=1 and for b=-0.5 for FTSE 100. The period is from 02/01/2004 to 29/12/2017. Each row name indicates the benchmark which is compared to each of the alternative models. In the same table the models that are included in the Model Confidence set (MCS) of Hansen (2011) using the range statistic are indicated by the "tick" symbol. *significant at 5% level

Table 3.16 – Pairwise comparisons of unconditional Giacomini and White test for DAX (4-year in-sample) - MSE

	Spec 1	Spec 2	Spec 3	Spec 4	Spec 5	Spec 6	Spec 7	Spec 8	Spec 9	Spec 10	Spec 11	Spec 12	Spec 13	Spec 14	Spec 15	Spec 16	Spec 17	Spec 18	Spec 19	Spec 20	Spec 21
Spec 1:A_N	-	3.52*	2.28*	2.17*	3.09*	3.74*	2.59*	1.40	2.02*	1.88	-0.14	2.57*	2.19*	2.25*	2.48*	3.52*	2.43*	1.36	2.13*	1.51	0.24
Spec 2:GA_N	-	-	0.55	0.49	-1.51	1.82	0.98	-0.42	0.30	-1.56	-4.28*	-2.69*	0.34	0.56	-4.27*	1.40	0.80	-0.41	0.45	-2.05*	-4.01*
Spec 3:GJR_M	-	-	-	-0.11	-0.73	0.11	1.01	-2.51*	-0.78	-1.87	-2.54*	-1.16	-1.68	0.11	-1.21	-0.05	0.60	-2.45*	-0.23	-2.36*	-2.37*
Spec 4:TGA_N	-	-	-	-	-0.67	0.14	1.23	-2.72*	-2.76*	-1.67	-2.41*	-1.07	-0.90	0.80	-1.13	-0.01	0.79	-2.82*	-0.62	-2.16*	-2.25*
Spec 5:IGA_N	-	-	-	-	-	1.86	1.19	-0.20	0.49	-0.92	-3.53*	-0.93	0.55	0.75	-1.82	1.54	1.01	-0.20	0.64	-1.36	-3.33*
Spec 6:AGA_N	-	-	-	-	-	-	0.42	-1.26	-0.38	-3.81*	-4.65*	-3.90*	-0.48	-0.07	-3.60*	-2.31*	0.21	-1.22	-0.19	-4.50*	-4.31*
Spec 7:$NAGA_N$	-	-	-	-	-	-	-	-2.13*	-1.64	-2.25*	-2.83*	-1.63	-1.61	-1.27	-1.72	-0.60	-2.46*	-2.14*	-1.46	-2.68*	-2.67*
Spec 8:EGA_N	-	-	-	-	-	-	-	-	2.33*	-0.29	-1.59	-0.01	2.18*	2.54*	-0.09	1.16	1.91	-0.02	2.38*	-0.72	-1.46
Spec 9:$APGA_N$	-	-	-	-	-	-	-	-	-	-1.39	-2.25*	-0.87	-0.02	2.36*	-0.92	0.24	1.24	-2.49*	2.17*	-1.88	-2.10*
Spec 10:$FIGA_N$	-	-	-	-	-	-	-	-	-	-	-2.27*	0.56	1.68	1.80	0.31	3.61*	1.97*	0.28	1.61	-1.98*	-2.04*
Spec 11:A_t	-	-	-	-	-	-	-	-	-	-	-	3.27*	2.47*	2.50*	3.04*	4.40*	2.65*	1.54	2.37*	1.85	1.80
Spec 12:GA_t	-	-	-	-	-	-	-	-	-	-	-	-	1.00	1.18	-0.81	3.60*	1.42	0.01	1.04	-1.38	-3.00*
Spec 13:GJR_t	-	-	-	-	-	-	-	-	-	-	-	-	-	1.13	-1.05	0.30	1.18	-2.11*	0.62	-2.29*	-2.30*
Spec 14:TGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-1.23	-0.09	0.74	-2.61*	-1.85	-2.30*	-2.34*
Spec 15:IGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3.24*	1.51	0.09	1.10	-1.01	-2.80*
Spec 16:AGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.37	-1.13	-0.04	-4.55*	-4.07*
Spec 17:$NAGA_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-1.93	-0.99	-2.38*	-2.50*
Spec 18:EGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2.49*	-0.69	-1.42
Spec 19:$APGA_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-2.10*	-2.22*
Spec 20:$FIGA_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-1.65
Spec 21:A_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 22:GA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 23:GJR_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 24:TGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 25:IGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 26:AGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 27:$NAGA_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 28:EGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 29:$APGA_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 30:$FIGA_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 31:$comb_1$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 32:$comb_2$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 33:$comb_3$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 34:HAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 35:logHAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 36:HEAVY	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 37:$GA-MIDAS$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 38:$GA-RV_N$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 39:$GA-RV_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 40:$GA-RV_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 41:$quant_1$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 42:$quant_2$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Note: This table reports the out-of-sample t-statistic of the Unconditional Giacomini and White (2006) test under the MSE loss function for DAX. The period is from 02/01/2004 to 29/12/2017. A positive t-statistic indicates that average loss of model A (in row) is larger than of model B (in column) and hence model B is superior. *denotes significance at the 5% level

Table 3.16 cont. - Pairwise comparisons of unconditional Giacomini and White test for DAX (4-year in-sample) - MSE

	Spec 22	Spec 23	Spec 24	Spec 25	Spec 26	Spec 27	Spec 28	Spec 29	Spec 30	Spec 31	Spec 32	Spec 33	Spec 34	Spec 35	Spec 36	Spec 37	Spec 38	Spec 39	Spec 40	Spec 41	Spec 42
Spec 1:A_N	2.65*	2.25*	2.30*	2.59*	3.73*	2.43*	1.39	2.00*	1.60	5.33*	5.28*	5.28*	-5.47*	5.28*	-3.60*	-7.23*	0.98	0.79	0.02	-3.17*	-3.04*
Spec 2:GA_N	-3.21*	0.42	0.65	-4.44*	1.87	0.82	-0.36	0.34	-2.34*	3.81*	3.73*	3.73*	-6.76*	3.73*	-5.29*	-8.76*	-1.41	-1.81	-1.89	-4.91*	-4.74*
Spec 3:GJR_M	-1.10	-1.41	0.49	-1.12	0.09	0.64	-2.34*	-0.51	-2.17*	2.51*	2.46*	2.46*	-7.66*	2.46*	-6.79*	-10.27*	-2.83*	-3.75*	-3.13*	-6.37*	-6.14*
Spec 4:TGA_N	-1.02	-0.59	1.71	-1.04	0.11	0.80	-2.68*	-1.32	-2.00*	2.54*	2.49*	2.49*	-7.67*	2.49*	-6.85*	-10.24*	-3.42*	-3.14*	-2.73*	-6.44*	-6.21*
Spec 5:IGA_N	-0.99	0.62	0.84	-1.78	1.90	1.03	-0.16	0.52	-1.53	4.28*	4.18*	4.18*	-6.49*	4.18*	-4.92*	-8.39*	-1.01	-1.25	-1.26	-4.54*	-4.38*
Spec 6:AGA_N	-3.60*	-0.37	0.05	-3.27*	-0.65	0.24	-1.15	-0.29	-4.81*	3.10*	3.02*	3.02*	-7.35*	3.02*	-6.20*	-9.62*	-4.09*	-4.01*	-2.97*	-5.82*	-5.62*
Spec 7:$NAGA_N$	-1.57	-1.52	-1.00	-1.62	-0.44	-1.21	-2.05*	-1.56	-2.59*	2.75*	2.68*	2.68*	-7.37*	2.68*	-6.21*	-9.67*	-2.45*	-3.47*	-2.48*	-5.82*	-5.63*
Spec 8:EGA_N	0.00	2.21*	2.58*	-0.03	1.23	1.85	0.43	2.14*	-0.63	2.68*	2.64*	2.64*	-7.87*	2.64*	-7.67*	-10.88*	-4.09*	-4.21*	-3.89*	-7.19*	-6.88*
Spec 9:$APGA_N$	-0.82	0.23	2.63*	-0.84	0.36	1.21	-2.32*	0.63	-1.73	2.62*	2.57*	2.57*	-7.67*	2.57*	-6.88*	-10.25*	-2.58*	-3.84*	-3.00*	-6.48*	-6.25*
Spec 10:$FIGA_N$	0.61	1.79	1.90	0.46	3.82*	1.94	0.34	1.38	-1.82	3.76*	3.67*	3.67*	-6.98*	3.67*	-5.69*	-9.26*	-1.12	-1.32	-1.69	-5.27*	-5.07*
Spec 11:A_t	3.34*	2.53*	2.55*	3.14*	4.63*	2.64*	1.57	2.22*	1.97*	5.24*	5.17*	5.17*	-5.74*	5.17*	-3.92*	-7.66*	4.53*	3.08*	0.89	-3.46*	-3.31*
Spec 12:GA_t	0.58	1.08	1.26	-0.13	4.08*	1.41	0.06	0.88	-1.58	3.87*	3.78*	3.78*	-6.84*	3.78*	-5.44*	-8.97*	-1.18	-1.12	-1.60	-5.04*	-4.86*
Spec 13:GJR_t	-0.94	1.88	1.42	-0.95	0.45	1.15	-1.95	0.19	-2.08*	2.67*	2.62*	2.62*	-7.69*	2.62*	-6.88*	-10.36*	-2.98*	-4.78*	-2.86*	-6.45*	-6.21*
Spec 14:TGA_t	-1.12	-0.83	2.28*	-1.14	0.05	0.75	-2.46*	-1.64	-2.14*	2.59*	2.54*	2.54*	-7.62*	2.54*	-6.73*	-10.15*	-2.78*	-3.34*	-5.95*	-6.33*	-6.11*
Spec 15:IGA_t	1.56	1.14	1.32	1.95	3.77*	1.51	0.14	0.94	-1.15	4.18*	4.08*	4.08*	-6.65*	4.08*	-5.16*	-8.70*	-1.09	-1.43	-1.48	-4.76*	-4.59*
Spec 16:AGA_t	-3.24*	-0.19	0.22	-2.88*	1.85	0.40	-1.05	-0.16	-4.82*	3.11*	3.04*	3.04*	-7.37*	3.04*	-6.25*	-9.68*	-3.42*	-2.64*	-4.76*	-5.87*	-5.67*
Spec 17:$NAGA_t$	-1.37	-1.06	-0.43	-1.41	-0.23	0.68	-1.84	-1.17	-2.30*	2.82*	2.75*	2.75*	-7.32*	2.75*	-6.12*	-9.62*	-1.96*	-2.12*	-2.01*	-5.73*	-5.54*
Spec 18:EGA_t	0.00	2.14*	2.64*	-0.03	1.19	1.87	1.91	2.28*	-0.60	2.65*	2.61*	2.61*	-7.86*	2.61*	-7.67*	-10.84*	-4.50*	-2.79*	-5.52*	-7.21*	-6.90*
Spec 19:$APGA_t$	-0.99	-0.35	2.73*	-1.01	0.17	0.98	-2.35*	-1.30	-1.94	2.61*	2.56*	2.56*	-7.61*	2.56*	-6.74*	-10.14*	-3.52*	-3.12*	-3.09*	-6.34*	-6.11*
Spec 20:$FIGA_t$	1.36	2.38*	2.38*	1.11	4.51*	2.33*	0.75	1.82	0.34	3.78*	3.69*	3.69*	-7.03*	3.69*	-5.83*	-9.41*	-0.86	-1.14	-1.52	-5.40*	-5.18*
Spec 21:A_{st}	3.07*	2.36*	2.39*	2.91*	4.29*	2.49*	1.44	2.07*	1.76	5.14*	5.07*	5.07*	-5.72*	5.07*	-3.91*	-7.62*	3.78*	5.06*	1.23	-3.46*	-3.31*
Spec 22:GA_{st}	-	1.02	1.20	-0.45	3.76*	1.37	0.03	0.83	-1.57	3.92*	3.84*	3.84*	-6.78*	3.84*	-5.35*	-8.88*	-1.12	-1.44	-1.12	-4.95*	-4.78*
Spec 23:GJR_{st}	-	-	1.17	-1.04	0.34	1.04	-1.99*	-0.02	-2.17*	2.67*	2.61*	2.61*	-7.66*	2.61*	-6.81*	-10.29*	-3.81*	-3.44*	-3.12*	-6.38*	-6.14*
Spec 24:TGA_{st}	-	-	-	-1.23	-0.07	0.49	-2.52*	-2.36*	-2.22*	2.57*	2.51*	2.51*	-7.60*	2.51*	-6.66*	-10.08*	-2.65*	-3.22*	-4.01*	-6.26*	-6.04*
Spec 25:IGA_{st}	-	-	-	-	3.41*	1.42	0.07	0.86	-1.27	4.19*	4.09*	4.09*	-6.63*	4.09*	-5.13*	-8.65*	-1.15	-1.15	-1.68	-4.73*	-4.56*
Spec 26:AGA_{st}	-	-	-	-	-	0.26	-1.11	-0.27	-4.89*	3.14*	3.06*	3.06*	-7.33*	3.06*	-6.15*	-9.58*	-5.58*	-5.34*	-5.71*	-5.78*	-5.58*
Spec 27:$NAGA_{st}$	-	-	-	-	-	-	-1.79	-1.16	-2.26*	2.84*	2.77*	2.77*	-7.26*	2.77*	-6.02*	-9.51*	-1.61	-2.04*	-2.05*	-5.63*	-5.44*
Spec 28:EGA_{st}	-	-	-	-	-	-	-	2.16*	-0.65	2.61*	2.57*	2.57*	-7.86*	2.57*	-7.65*	-10.83*	-5.00*	-5.04*	-5.37*	-7.19*	-6.89*
Spec 29:$APGA_{st}$	-	-	-	-	-	-	-	-	-1.69	2.62*	2.57*	2.57*	-7.58*	2.57*	-6.69*	-10.10*	-3.47*	-2.99*	-3.01*	-6.29*	-6.06*
Spec 30:$FIGA_{st}$	-	-	-	-	-	-	-	-	-	3.93*	3.83*	3.83*	-6.92*	3.83*	-5.62*	-9.22*	-0.96	-1.12	-1.12	-5.19*	-4.99*
Spec 31:$comb_1$	-	-	-	-	-	-	-	-	-	-	1.38	1.38	-6.83*	1.38	-5.36*	-8.40*	-2.34*	-2.12*	-2.87*	-5.07*	-4.94*
Spec 32:$comb_2$	-	-	-	-	-	-	-	-	-	-	-	-0.34	-6.77*	1.60	-5.29*	-8.30*	-4.13*	-3.17*	-4.56*	-5.00*	-4.88*
Spec 33:$comb_3$	-	-	-	-	-	-	-	-	-	-	-	-	-6.77*	0.34	-5.29*	-8.30*	-2.92*	-5.89*	-4.59*	-5.00*	-4.88*
Spec 34:HAR	-	-	-	-	-	-	-	-	-	-	-	-	-	6.77*	4.27*	-1.70	5.12*	5.70*	5.42*	4.88*	5.04*
Spec 35:logHAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-5.29*	-8.30*	-5.12*	-5.65*	-4.67*	-5.00*	-4.88*
Spec 36:HEAVY	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-7.48*	-1.25	-1.63	-2.81*	3.52*	4.18*
Spec 37:$GA-MIDAS$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-1.15	-1.16	-1.11	7.86*	8.05*
Spec 38:$GA-RV_N$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-3.12*	-1.34	-6.16*	-5.34*
Spec 39:$GA-RV_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-1.37	-4.55*	-4.59*
Spec 40:$GA-RV_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-4.78*	-3.23*
Spec 41:$quant_1$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	4.77*
Spec 42:$quant_2$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Note: This table reports the out-of-sample t-statistic of the Unconditional Giacomini and White (2006) test under the MSE loss function for DAX. The period is from 02/01/2004 to 29/12/2017. A positive t-statistic indicates that average loss of model A (in row) is larger than of model B (in column) and hence model B is superior. *denotes significance at the 5% level

Table 3.17 – Superior Predictive Ability (SPA) test and Model Confidence Set (MCS) results for the DAX (4-year in-sample) - MSE, Robust Patton (b=1), Robust Patton (b=-0.5)

	MSE - Robust Patton (b=1) - Robust Patton (b=-0.5)					
	P-value (Hansen)	R-statistic	P-value (Hansen)	R-statistic	P-value (Hansen)	R-statistic
Spec 1:A_N	0.01*	✓	0.08	✓	≤ 0.00*	
Spec 2:GA_N	0.08	✓	0.20	✓	0.01*	✓
Spec 3:GJR_M	0.21	✓	0.48	✓	0.06	✓
Spec 4:TGA_N	0.24	✓	0.51	✓	0.05	✓
Spec 5:IGA_N	0.06	✓	0.16	✓	0.01*	
Spec 6:AGA_N	0.12	✓	0.33	✓	0.02*	✓
Spec 7:$NAGA_N$	0.11	✓	0.43	✓	0.01*	✓
Spec 8:EGA_N	0.16	✓	0.31	✓	0.01	✓
Spec 9:$APGA_N$	0.15	✓	0.21	✓	0.06	✓
Spec 10:$FIGA_N$	0.11	✓	0.21	✓	0.02*	
Spec 11:A_t	≤ 0.00*		0.09	✓	≤ 0.00*	
Spec 12:GA_t	0.05	✓	0.2	✓	0.02*	
Spec 13:GJR_t	0.21	✓	0.44	✓	0.08	✓
Spec 14:TGA_t	0.23	✓	0.49	✓	0.06	✓
Spec 15:IGA_t	0.06	✓	0.15	✓	0.02*	
Spec 16:AGA_t	0.10	✓	0.25	✓	0.03*	✓
Spec 17:$NAGA_t$	0.07	✓	0.29	✓	0.02*	✓
Spec 18:EGA_t	0.21	✓	0.34	✓	0.10	✓
Spec 19:$APGA_t$	0.20	✓	0.31	✓	0.07	✓
Spec 20:$FIGA_t$	0.08	✓	0.24	✓	0.01*	
Spec 21:A_{st}	≤ 0.00*		0.05	✓	≤ ≤ 0.00*	
Spec 22:GA_{st}	0.06	✓	0.20	✓	≤ 0.00*	
Spec 23:GJR_{st}	0.23	✓	0.46	✓	0.05	✓
Spec 24:TGA_{st}	0.19	✓	0.47	✓	0.04*	✓
Spec 25:IGA_{st}	0.08	✓	0.18	✓	0.02*	
Spec 26:AGA_{st}	0.12	✓	0.35	✓	0.03*	✓
Spec 27:$NAGA_{st}$	0.10	✓	0.35	✓	≤ 0.00*	✓
Spec 28:EGA_{st}	0.20	✓	0.31	✓	0.09	✓
Spec 29:$APGA_{st}$	0.16	✓	0.30	✓	0.04*	✓
Spec 30:$FIGA_{st}$	0.08	✓	0.22	✓	0.01*	
Spec 31:$comb_1$	0.40	✓	0.36	✓	0.18	✓
Spec 32:$comb_2$	1.00	✓	1.00	✓	1.00	✓
Spec 33:$comb_3$	0.65	✓	0.66	✓	0.74	✓
Spec 34:HAR	≤ 0.00*		0.07	✓	0.03*	
Spec 35:logHAR	1.00	✓	1.00	✓	1.00	✓
Spec 36:HEAVY	0.01*		0.04*	✓	0.08*	
Spec 37:$GA-MIDAS$	≤ 0.00*		0.02*	✓	0.02*	
Spec 38:$GA-RV_N$	0.07	✓	0.13	✓	0.02*	
Spec 39:$GA-RV_t$	0.08	✓	0.14	✓	0.01*	
Spec 40:$GA-RV_{st}$	0.06	✓	0.18	✓	0.01*	
Spec 41:$quant_1$	0.01*	✓	0.03*	✓	0.01*	
Spec 42:$quant_2$	0.02*	✓	0.05	✓	0.01*	

Note: This table reports the out-of-sample P-values of the Superior Predictive Ability test of Hansen (2005) under the MSE loss function, the robust loss of Patton (2011) for b=1 and for b=-0.5 for DAX. The period is from 02/01/2004 to 29/12/2017. Each row name indicates the benchmark which is compared to each of the alternative models. In the same table the models that are included in the Model Confidence set (MCS) of Hansen (2011) using the range statistic are indicated by the "tick" symbol. *significant at 5% level

3.5 Conclusion

This study focuses on the forecast evaluation of univariate volatility models. We employ non-parametric approaches, GARCH models assuming various distributions, RV models and simple equal-weighted forecast combinations. Then the GW test of equal predictive accuracy test as well as the MCS and the SPA tests are used with different loss functions to evaluate the out-of-sample forecasting accuracy of those specifications. The aim is to evaluate and rank the models and their specifications and examine the sensitivity of our results with respect to the selected loss function, size of the in-sample estimation period and stock index.

Our empirical results emphasize the lack of superiority of a single specification based on different measures, across all datasets. However, all the findings indicate that TGARCH specifications, logHAR model and the forecast combination based on the mean of all the models included in the optimal set (the 1st forecast combinations) are the best performing models. In most cases those specifications were equally performing at 5% significance level, , with minor exceptions (e.g., the TGARCH-N under the robust loss of Patton (b=-0.5) for the 4-years in-sample. What is more, for the TGARCH models for both samples and for any loss the null hypothesis could not be rejected among the three specifications, except from the 1-year in-sample the GW under the robust loss of Patton (b=1) reported that the TGARCH-St outperforms the TGARCH-t. Non robust loss functions produced a different ranking, in some cases, of the best models but, overall they reported a fairly accurate representation of the forecasting accuracy of the models among the rest of the specifications. Finally, changing the assumptions of the underlying error distribution did not seem to play a significant role on the certain model's predictive accuracy and/or comparison with the rest models. The model itself seems to be the one determining its rank. Different distributional assumptions of the same model seem to over/under - perform the same various models.

3.6 Appendix

Loss Functions:

$$MSE - LOG : L(\widehat{\sigma}_t, h) = \frac{1}{m}\sum_{t=1}^{m}(\log \widehat{\sigma}_t - \log h_t)^2 \tag{3.25}$$

$$MSE - SD : L(\widehat{\sigma}_t, h) = \frac{1}{m}\sum_{t=1}^{m}(\widehat{\sigma}_t - \sqrt{h_t})^2 \tag{3.26}$$

$$MSE - SD : L(\widehat{\sigma}_t, h) = \frac{1}{m}\sum_{t=1}^{m}(\frac{\widehat{\sigma}_t}{ht} - 1)^2 \tag{3.27}$$

$$RMSE : L(\widehat{\sigma}_t, h) = \sqrt{\frac{1}{m}\sum_{t=1}^{m}((\widehat{\sigma}_t - h_t)^2)} \tag{3.28}$$

$$RMSE - prop : L(\widehat{\sigma}_t, h) = \sqrt{\frac{1}{m}\sum_{t=1}^{m}(\frac{\widehat{\sigma}_t}{ht} - 1)^2} \tag{3.29}$$

$$MSE - prop : L(\widehat{\sigma}_t, h) = \frac{1}{m}\sum_{t=1}^{m}(\frac{\widehat{\sigma}_t}{h_t} - 1)^2 \tag{3.30}$$

$$Mean\ Absolute\ Error - MAE : L(\widehat{\sigma}_t, h) = \frac{1}{m}\sum_{t=1}^{m}(|\widehat{\sigma}_t - h_t|) \tag{3.31}$$

$$MAE - LOG : L(\widehat{\sigma}_t, h) = \frac{1}{m}\sum_{t=1}^{m}|\log \widehat{\sigma}_t - \log h_t| \tag{3.32}$$

$$MAE - SD : L(\widehat{\sigma}_t, h) = \frac{1}{m}\sum_{t=1}^{m}|\widehat{\sigma}_t - \sqrt{h_t}| \tag{3.33}$$

$$MAE - prop \quad MAPE : L(\widehat{\sigma}_t, h) = \frac{1}{m}\sum_{t=1}^{m}|\frac{\widehat{\sigma}_t}{ht} - 1| \ . \tag{3.34}$$

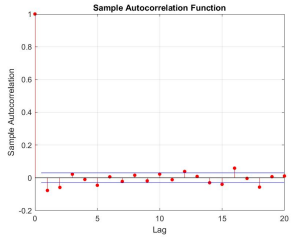


Figure 3.5 – Autocorrelation function for S&P 500 daily returns (2000-2017)

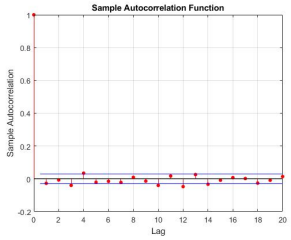


Figure 3.6 – Autocorrelation function for FTSE 100 daily returns (2000-2017)

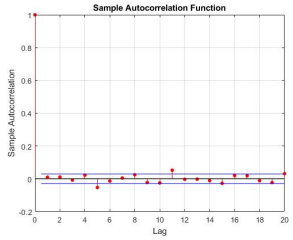


Figure 3.7 – Autocorrelation function for DAX daily returns (2000-2017)

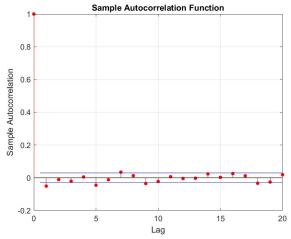


Figure 3.8 – Autocorrelation function for CAC 40 daily returns (2000-2017)

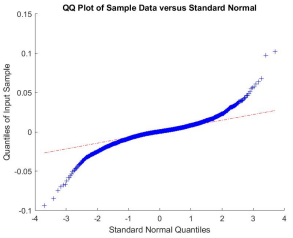


Figure 3.9 – Quantile-quantile plot for S&P 500 daily returns (2000-2017)

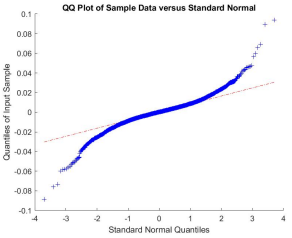


Figure 3.10 – Quantile-quantile plot for FTSE 100 daily returns (2000-2017)

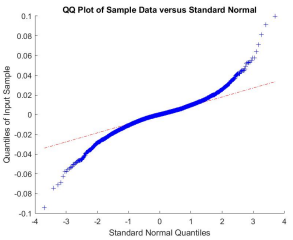


Figure 3.11 – Quantile-quantile plot for DAX daily returns (2000-2017)

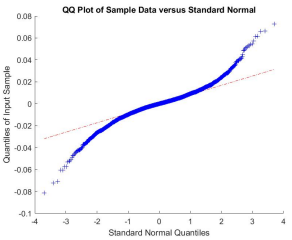


Figure 3.12 – Quantile-quantile plot for CAC 40 daily returns (2000-2017)

Table 3.18 – Pairwise comparisons of unconditional Giacomini and White test for S&P 500 (4-year in-sample) - all Loss Functions

	logMSE		sdMSE		propMSE		RMSE		propRMSE		MAE		logMAE		sdMAE		MAPE		GMLE	
	win	lose	win	lose	win	lose	win	lose	win	lose	win	lose	win	lose	win	lose	win	lose	win	lose
Spec 1:A_N	5	35	5	36	4	37	4	37	4	37	4	36	5	36	5	36	4	36	4	37
Spec 2:GA_N	12	29	17	22	9	30	9	31	9	31	15	24	11	29	11	23	15	24	9	31
Spec 3:GJR_M	24	12	26	7	18	12	21	15	21	15	26	7	22	11	27	6	26	7	19	18
Spec 4:TGA_N	32	7	34	4	31	7	32	8	32	8	33	4	32	6	35	5	33	4	28	9
Spec 5:IGA_N	8	31	7	31	6	33	6	34	6	34	6	35	8	31	6	34	6	35	6	34
Spec 6:AGA_N	19	20	22	8	18	12	16	23	16	23	22	10	19	22	21	15	22	10	15	25
Spec 7:$NAGA_N$	24	12	27	7	18	12	21	15	21	15	26	5	22	11	26	7	26	5	19	18
Spec 8:EGA_N	19	19	22	4	18	14	21	14	21	14	30	4	20	14	25	5	30	4	25	10
Spec 9:$APGA_N$	30	9	28	6	27	9	28	10	28	10	26	5	30	8	28	7	26	5	25	14
Spec 10:$FIGA_N$	15	23	10	23	12	26	13	26	13	26	9	26	15	24	9	23	9	26	13	26
Spec 11:A_t	4	36	4	37	2	38	2	38	2	38	2	38	3	37	3	37	2	38	2	39
Spec 12:GA_t	14	27	12	23	10	28	11	30	11	30	10	25	14	27	10	25	10	25	11	30
Spec 13:GJR_t	27	10	19	12	19	12	23	13	23	13	18	16	24	8	20	11	18	16	21	17
Spec 14:TGA_t	36	5	32	4	35	6	35	6	35	6	28	5	36	5	31	6	28	5	31	6
Spec 15:IGA_t	11	30	8	26	7	33	8	33	8	33	8	33	10	30	8	32	8	33	8	33
Spec 16:AGA_t	20	17	20	14	18	12	17	21	17	21	18	15	20	19	19	17	18	15	17	23
Spec 17:$NAGA_t$	27	11	19	12	18	11	23	13	23	13	18	16	23	11	19	16	18	16	21	17
Spec 18:EGA_t	22	13	22	5	19	9	29	9	29	9	28	6	23	11	25	6	28	6	31	6
Spec 19:$APGA_t$	32	7	21	11	28	7	31	8	31	8	18	15	30	7	24	10	18	15	27	11
Spec 20:$FIGA_t$	16	23	10	23	12	26	13	26	13	26	10	26	17	23	10	23	10	26	13	26
Spec 21:A_{st}	3	37	3	38	1	39	1	39	1	39	1	39	2	38	2	39	1	39	1	40
Spec 22:GA_{st}	13	28	12	23	9	31	9	31	9	31	9	26	12	28	9	26	9	26	9	31
Spec 23:GJR_{st}	24	12	19	11	18	13	21	15	21	15	18	16	22	12	19	16	18	16	19	18
Spec 24:TGA_{st}	35	6	31	4	32	7	34	7	34	7	26	5	34	6	28	7	26	5	30	7
Spec 25:IGA_{st}	8	31	8	26	6	34	6	34	6	34	7	34	8	31	7	33	7	34	6	34
Spec 26:AGA_{st}	20	17	20	14	18	12	17	21	17	21	18	15	20	19	19	17	18	15	17	23
Spec 27:$NAGA_{st}$	24	12	21	9	18	12	21	14	21	14	19	14	22	12	19	16	19	14	19	18
Spec 28:EGA_{st}	20	18	20	5	19	9	25	11	25	11	29	4	23	11	24	6	29	4	29	7
Spec 29:$APGA_{st}$	31	7	26	7	28	8	28	10	28	10	22	14	30	7	25	9	22	14	25	14
Spec 30:$FIGA_{st}$	15	24	10	23	11	28	12	29	12	29	9	27	15	24	9	24	9	27	12	29
Spec 31:$comb_1$	38	3	40	0	36	5	36	5	36	5	38	3	38	3	38	3	38	3	36	5
Spec 32:$comb_2$	40	0	38	2	38	1	38	0	38	0	39	0	40	0	40	0	39	0	38	2
Spec 33:$comb_3$	39	2	40	0	37	4	37	4	37	4	39	0	39	2	39	2	39	0	37	4
Spec 34:HAR	0	41	0	41	0	41	0	41	0	41	0	41	0	41	0	41	0	41	0	41
Spec 35:logHAR	40	0	38	2	38	1	38	0	38	0	39	0	40	0	40	0	39	0	38	2
Spec 36:HEAVY	37	4	9	19	38	0	39	0	39	0	29	4	37	4	37	4	29	4	40	1
Spec 37:$GA-MIDAS$	15	23	12	22	13	24	13	26	13	26	9	25	15	23	11	23	9	25	13	25
Spec 38:$GA-RV_N$	2	35	7	30	15	24	16	22	16	22	13	24	7	34	9	23	13	24	27	6
Spec 39:$GA-RV_t$	8	31	7	30	15	24	17	21	17	21	9	25	8	28	9	23	9	25	27	6
Spec 40:$GA-RV_{st}$	7	34	6	35	40	0	38	1	38	1	18	10	6	35	6	23	18	10	41	0
Spec 41:$quant_1$	1	40	1	40	1	38	1	38	1	38	1	38	1	40	1	40	1	38	3	38
Spec 42:$quant_2$	2	38	2	39	5	36	5	36	5	36	4	36	2	37	3	37	4	36	5	36

Note: This table reports a summary of the out-of-sample results for the Unconditional Giacomini and White (2006) performed under the logMSE, sdMSE, propMSE, RMSE, propRMSE, MAE, logMAE, sdMAE, MAPE and GMLE loss functions for S&P 500. The period is from 02/01/2004 to 29/12/2017. The "win" column reports the cases where a certain specification outperformed significantly the competing specifications. On the contrary, the "lose" column reports the cases where a certain specification was outperformed significantly by the competing specifications. Adding the two columns for each specification and for each loss function, respectively, shows how many of the 41 in total pairwise comparisons where statistically significant. For the hypothesis testing the 5% significance level is selected.

Table 3.19 – Pairwise comparisons of unconditional Giacomini and White test for S&P 500 (1-year in-sample) - all Loss Functions

	logMSE		sdMSE		propMSE		RMSE		propRMSE		MAE		logMAE		sdMAE		MAPE		GMLE	
	win	lose	win	lose	win	lose	win	lose	win	lose	win	lose	win	lose	win	lose	win	lose	win	lose
Spec 1:A_N	4	36	3	37	3	38	2	39	2	39	2	39	3	37	2	39	2	39	2	39
Spec 2:GA_N	17	19	17	20	15	21	13	23	13	23	16	22	16	23	16	21	16	22	13	26
Spec 3:GJR_M	29	4	33	5	23	6	25	6	25	6	30	3	32	4	34	3	30	3	22	9
Spec 4:TGA_N	37	2	39	0	34	4	35	4	35	4	38	2	37	2	38	2	38	2	31	7
Spec 5:IGA_N	13	26	12	26	10	30	9	30	9	30	7	31	13	27	10	28	7	31	8	31
Spec 6:AGA_N	24	8	22	5	24	5	23	7	23	7	22	5	24	9	25	6	22	5	19	10
Spec 7:$NAGA_N$	30	4	25	5	25	6	31	4	31	4	22	5	32	4	30	4	22	5	29	7
Spec 8:EGA_N	20	17	22	6	19	14	23	5	23	5	22	5	21	17	23	7	22	5	36	4
Spec 9:$APGA_N$	25	7	24	6	25	5	23	6	23	6	23	6	25	7	25	6	23	6	22	7
Spec 10:$FIGA_N$	20	17	14	21	18	20	19	22	19	22	12	24	19	18	18	21	12	24	16	23
Spec 11:A_t	3	37	1	38	1	39	1	40	1	40	1	40	2	38	1	40	1	40	1	40
Spec 12:GA_t	15	22	17	20	13	24	12	27	12	27	11	23	15	25	15	23	11	23	11	28
Spec 13:GJR_t	26	5	26	6	22	7	22	8	22	8	21	6	28	5	25	6	21	6	18	15
Spec 14:TGA_t	37	2	37	2	32	4	34	4	34	4	35	3	37	2	36	3	35	3	24	8
Spec 15:IGA_t	11	28	11	26	8	31	7	32	7	32	6	33	11	28	9	31	6	33	6	34
Spec 16:AGA_t	25	5	22	6	24	4	23	7	23	7	20	8	24	8	23	8	20	8	18	11
Spec 17:$NAGA_t$	29	4	22	8	25	6	25	7	25	7	19	9	31	4	23	7	19	9	22	10
Spec 18:EGA_t	18	17	22	12	20	16	21	11	21	11	19	12	18	18	21	19	19	12	35	5
Spec 19:$APGA_t$	25	8	22	10	25	6	23	7	23	7	20	9	25	9	23	9	20	9	22	8
Spec 20:$FIGA_t$	18	19	14	23	14	23	15	23	15	23	10	25	18	20	16	22	10	25	15	25
Spec 21:A_{st}	2	38	0	40	0	40	0	41	0	41	0	41	1	39	0	41	0	41	0	41
Spec 22:GA_{st}	14	24	16	21	12	27	11	29	11	29	10	24	14	26	13	25	10	24	10	29
Spec 23:GJR_{st}	24	9	23	7	20	16	20	18	20	18	20	8	24	9	23	7	20	8	16	22
Spec 24:TGA_{st}	33	4	37	2	27	6	24	6	24	6	35	3	32	4	34	4	35	3	18	15
Spec 25:IGA_{st}	10	29	10	28	7	33	6	33	6	33	5	34	10	29	8	33	5	34	5	35
Spec 26:AGA_{st}	25	5	22	6	24	4	23	7	23	7	20	8	24	8	23	8	20	8	18	11
Spec 27:$NAGA_{st}$	25	7	26	5	22	9	22	8	22	8	23	5	25	7	23	7	23	5	18	15
Spec 28:EGA_{st}	0	0	22	10	18	18	20	19	20	19	20	8	0	0	21	19	20	8	28	6
Spec 29:$APGA_{st}$	24	14	25	6	21	9	20	15	20	15	22	6	23	12	23	8	22	6	18	16
Spec 30:$FIGA_{st}$	16	21	14	23	13	25	13	25	13	25	10	25	16	23	14	23	10	25	13	26
Spec 31:$comb_1$	8	32	7	31	4	36	3	38	3	38	5	33	6	32	5	35	5	33	3	38
Spec 32:$comb_2$	10	28	10	20	9	30	9	30	9	30	15	20	10	28	10	28	15	20	8	31
Spec 33:$comb_3$	9	31	7	31	5	35	4	37	4	37	6	32	7	31	6	34	6	32	4	37
Spec 34:HAR	39	1	37	0	38	3	38	3	38	3	40	1	39	1	40	1	40	1	38	3
Spec 35:logHAR	40	0	39	0	39	2	40	1	40	1	41	0	40	0	41	0	41	0	39	2
Spec 36:HEAVY	14	17	10	20	40	0	41	0	41	0	26	2	19	16	23	2	26	2	40	1
Spec 37:$GA-MIDAS$	13	23	7	27	0	39	6	23	6	23	10	24	18	17	14	21	10	24	8	10
Spec 38:$GA-RV_N$	0	40	5	36	12	23	12	25	12	25	9	24	0	40	5	34	9	24	34	4
Spec 39:$GA-RV_t$	1	39	7	31	12	23	15	23	15	23	11	22	5	31	10	28	11	22	17	10
Spec 40:$GA-RV_{st}$	6	33	6	35	40	0	39	2	39	2	20	6	5	31	9	27	20	6	41	0
Spec 41:$quant_1$	6	33	2	37	4	35	5	36	5	36	3	38	5	33	3	37	3	38	5	34
Spec 42:$quant_2$	5	35	0	39	7	32	6	32	6	32	4	37	4	36	3	37	4	37	10	28

Note: This table reports a summary of the out-of-sample results for the Unconditional Giacomini and White (2006) performed under the logMSE, sdMSE, propMSE, RMSE, propRMSE, MAE, logMAE, sdMAE, MAPE and GMLE loss functions for S&P 500. The period is from 02/01/2001 to 29/12/2017. The "win" column reports the cases where a certain specification outperformed significantly the competing specifications. On the contrary, the "lose" column reports the cases where a certain specification was outperformed significantly by the competing specifications. Adding the two columns for each specification and for each loss function, respectively, shows how many of the 41 in total pairwise comparisons where statistically significant. For the hypothesis testing the 5% significance level is selected.

Table 3.20 – Superior Predictive Ability (SPA) test and Model Confidence Set (MCS) results for the S&P 500 (1-year in-sample) - MSE, Robust Patton (b=1), Robust Patton (b=-0.5)

	MSE - Robust Patton (b=1) - Robust Patton (b=-0.5)					
	P-value (White)	SQ-method	P-value (White)	SQ-method	P-value (White)	SQ-method
Spec 1:A_N	0.01*		0.05	✓	≤ 0.00*	
Spec 2:GA_N	0.07		0.21		≤ 0.00*	
Spec 3:GJR_M	0.21	✓	0.21	✓	0.20	✓
Spec 4:TGA_N	0.95	✓	0.79	✓	0.99	✓
Spec 5:IGA_N	0.03*		0.14		≤ 0.00*	
Spec 6:AGA_N	0.32	✓	0.41		0.26	✓
Spec 7:$NAGA_N$	0.30	✓	0.48	✓	0.15	✓
Spec 8:EGA_N	0.29	✓	0.33	✓	0.15	✓
Spec 9:$APGA_N$	0.14	✓	0.19	✓	0.10	✓
Spec 10:$FIGA_N$	≤ 0.00*		0.07		≤ 0.00*	
Spec 11:A_t	≤ 0.00*		0.06		≤ 0.00*	
Spec 12:GA_t	0.06		0.18		≤ 0.00*	
Spec 13:GJR_t	0.16	✓	0.16	✓	0.17	✓
Spec 14:TGA_t	0.83	✓	0.43	✓	0.96	✓
Spec 15:IGA_t	0.02*		0.14	✓	≤ 0.00*	
Spec 16:AGA_t	0.29		0.38		0.16	
Spec 17:$NAGA_t$	0.21	✓	0.35	✓	0.10	✓
Spec 18:EGA_t	0.03*	✓	0.27	✓	0.01*	✓
Spec 19:$APGA_t$	0.14	✓	0.16	✓	0.10	✓
Spec 20:$FIGA_t$	≤ 0.00*		0.10		≤ 0.00*	
Spec 21:A_{st}	≤ 0.00*		0.05		≤ 0.00*	
Spec 22:GA_{st}	0.07		0.19		≤ 0.00*	
Spec 23:GJR_{st}	0.15	✓	0.15	✓	0.13	✓
Spec 24:TGA_{st}	1.00	✓	1.00	✓	1.00	✓
Spec 25:IGA_{st}	0.03*		0.18		≤ 0.00*	
Spec 26:AGA_{st}	0.28	✓	0.37	✓	0.18	✓
Spec 27:$NAGA_{st}$	0.25	✓	0.60	✓	0.06	✓
Spec 28:EGA_{st}	0.03*	✓	0.21	✓	≤ 0.00*	✓
Spec 29:$APGA_{st}$	0.14	✓	0.22	✓	0.07	✓
Spec 30:$FIGA_{st}$	0.01*		0.10		≤ 0.00*	
Spec 31:$comb_1$	≤ 0.00*		0.49	✓	≤ 0.00*	
Spec 32:$comb_2$	0.03*	✓	0.12	✓	≤ 0.00*	
Spec 33:$comb_3$	≤ 0.00*		0.23	✓	≤ 0.00*	
Spec 34:HAR	0.38	✓	0.19	✓	0.72	✓
Spec 35:logHAR	0.81	✓	0.82	✓	0.81	✓
Spec 36:HEAVY	0.07		0.26	✓	0.01*	
Spec 37:$GA-MIDAS$	0.02*		0.12		≤ 0.00*	
Spec 38:$GA-RV_N$	0.02*		0.17		≤ 0.00*	
Spec 39:$GA-RV_t$	0.04*		0.19		≤ 0.00*	
Spec 40:$GA-RV_{st}$	0.05		0.19	✓	≤ 0.00*	
Spec 41:$quant_1$	0.05		0.09		0.01*	
Spec 42:$quant_2$	0.05		0.13	✓	0.02*	

Note: This table reports the out-of-sample White's Upper P-values (Original RC P-values) of the Superior Predictive Ability test of Hansen (2005) under the MSE loss function, the robust loss of Patton (2011) for b=1 and for b=-0.5 for S&P 500. The period is from 02/01/2001 to 29/12/2017. Each row name indicates the benchmark which is compared to each of the alternative models. In the same table the models that are included in the Model Confidence set (MCS) of Hansen (2011) using the SQ-method are indicated by the "tick" symbol. *significant at 5% level

Table 3.21 – Superior Predictive Ability (SPA) test and Model Confidence Set (MCS) results for the S&P 500 (4-year in-sample) - MSE, Robust Patton (b=1), Robust Patton (b=-0.5)

	MSE - Robust Patton (b=1) - Robust Patton (b=-0.5)					
	P-value (White)	SQ-method	P-value (White)	SQ-method	P-value (White)	SQ-method
Spec 1:A_N	0.02*		0.09		$\leq$ 0.00*	
Spec 2:GA_N	0.05		0.12		$\leq$ 0.00*	
Spec 3:GJR_M	0.28		0.26		0.43	✓
Spec 4:TGA_N	0.87	✓	0.86	✓	0.33	✓
Spec 5:IGA_N	0.04*		0.11		0.01*	
Spec 6:AGA_N	0.04*		0.10		0.01*	
Spec 7:$NAGA_N$	0.34	✓	0.58	✓	0.30	✓
Spec 8:EGA_N	0.27	✓	0.39	✓	0.05	✓
Spec 9:$APGA_N$	0.32	✓	0.32		0.31	✓
Spec 10:$FIGA_N$	0.04*		0.13		0.02*	
Spec 11:A_t	0.01*		0.07		$\leq$ 0.00*	
Spec 12:GA_t	0.05		0.12		$\leq$ 0.00*	
Spec 13:GJR_t	0.17		0.14		0.21	
Spec 14:TGA_t	0.45	✓	0.64	✓	0.45	✓
Spec 15:IGA_t	0.05		0.13		0.01*	
Spec 16:AGA_t	0.06		0.11		0.02*	
Spec 17:$NAGA_t$	0.14	✓	0.15	✓	0.15	
Spec 18:EGA_t	0.26	✓	0.85	✓	0.02*	✓
Spec 19:$APGA_t$	0.21		0.2		0.24	
Spec 20:$FIGA_t$	0.06		0.15		0.02*	
Spec 21:A_{st}	0.01*		0.07		$\leq$ 0.00*	
Spec 22:GA_{st}	0.04*		0.12		0.01*	
Spec 23:GJR_{st}	0.21		0.19		0.24	
Spec 24:TGA_{st}	0.94	✓	1.00	✓	0.79	✓
Spec 25:IGA_{st}	0.04*		0.10		0.01*	
Spec 26:AGA_{st}	0.05		0.14		0.02*	
Spec 27:$NAGA_{st}$	0.43	✓	0.63	✓	0.28	✓
Spec 28:EGA_{st}	0.39	✓	0.82	✓	0.04*	✓
Spec 29:$APGA_{st}$	0.22	✓	0.18	✓	0.30	✓
Spec 30:$FIGA_{st}$	0.07		0.15		0.01*	
Spec 31:$comb_1$	1.00	✓	0.93	✓	1.00	✓
Spec 32:$comb_2$	0.22		0.37		0.06	
Spec 33:$comb_3$	0.60	✓	0.57	✓	0.54	✓
Spec 34:HAR	0.01*		0.07		$\leq$ 0.00*	
Spec 35:logHAR	0.22	✓	0.37	✓	0.07	✓
Spec 36:HEAVY	0.04*		0.21		0.01*	
Spec 37:$GA-MIDAS$	0.05		0.14		0.02*	
Spec 38:$GA-RV_N$	0.05		0.12		0.01*	
Spec 39:$GA-RV_t$	0.07		0.11		0.02*	
Spec 40:$GA-RV_{st}$	0.06		0.15		0.01*	
Spec 41:$quant_1$	0.06		0.15		0.03*	
Spec 42:$quant_2$	0.06		0.16		0.03*	

Note: This table reports the out-of-sample White's Upper P-values (Original RC P-values) of the Superior Predictive Ability test of Hansen (2005) under the MSE loss function, the robust loss of Patton (2011) for b=1 and for b=-0.5 for S&P 500. The period is from 02/01/2004 to 29/12/2017. Each row name indicates the benchmark which is compared to each of the alternative models. In the same table the models that are included in the Model Confidence set (MCS) of Hansen (2011) using the SQ-method are indicated by the "tick" symbol. *significant at 5% level

Table 3.22 – Pairwise comparisons of unconditional Giacomini and White test for CAC 40 (4-year in-sample) - robust Patton (b=1)

	Spec 1	Spec 2	Spec 3	Spec 4	Spec 5	Spec 6	Spec 7	Spec 8	Spec 9	Spec 10	Spec 11	Spec12	Spec 13	Spec 14	Spec 15	Spec 16	Spec 17	Spec 18	Spec 19	Spec 20	Spec 21
Spec 1:A_N	-	2.62*	3.02*	3.64*	2.60*	2.76*	3.26*	3.32*	3.60*	1.89	0.52	2.44*	3.06*	3.64*	2.56*	2.78*	3.25*	3.25*	3.59*	1.90	0.82
Spec 2:GA_N	-	-	1.59	2.49*	0.08	1.90	2.07*	2.99*	2.38*	-2.81*	-3.22*	-1.56	1.59	2.37*	-0.12	1.84	1.98*	2.70*	2.24*	-2.72*	-3.26*
Spec 3:GJR_M	-	-	-	2.28*	-2.07*	-1.05	1.94	0.12	1.96*	-2.29*	-3.45*	-1.65	1.02	2.14*	-1.89	-1.05	1.91	-0.10	1.83	-2.17*	-3.40*
Spec 4:TGA_N	-	-	-	-	-2.91*	-2.06*	-0.97	-1.72	-1.47	-2.85*	-4.06*	-2.45*	-2.20*	0.82	-2.76*	-2.08*	-0.06	-1.75	-0.16	-2.73*	-4.01*
Spec 5:IGA_N	-	-	-	-	-	0.81	2.46*	3.08*	2.74*	-2.10*	-3.09*	-0.74	2.00*	2.71*	-0.42	0.80	2.28*	2.55*	2.54*	-1.95	-3.09*
Spec 6:AGA_N	-	-	-	-	-	-	1.64	2.37*	1.95	-3.98*	-3.45*	-3.82*	1.09	1.98*	-1.13	0.04	1.63	2.19*	1.86	-4.10*	-3.53*
Spec 7:$NAGA_N$	-	-	-	-	-	-	-	-1.05	0.62	-2.46*	-3.58*	-2.07*	-1.77	1.05	-2.34*	-1.65	1.32	-1.14	0.70	-2.38*	-3.52*
Spec 8:EGA_N	-	-	-	-	-	-	-	-	1.56	-3.67*	-4.10*	-3.01*	0.06	1.66	-3.16*	-2.45*	1.17	-1.89	1.49	-3.48*	-4.19*
Spec 9:$APGA_N$	-	-	-	-	-	-	-	-	-	-2.77*	-4.01*	-2.35*	-1.86	1.60	-2.61*	-1.97*	0.16	-1.61	0.65	-2.65*	-3.96*
Spec 10:$FIGA_N$	-	-	-	-	-	-	-	-	-	-	-2.27*	2.83*	2.23*	2.74*	2.28*	3.85*	2.33*	3.68*	2.65*	0.80	-2.25*
Spec 11:A_t	-	-	-	-	-	-	-	-	-	-	-	2.99*	3.47*	4.02*	3.06*	3.48*	3.50*	4.05*	3.95*	2.27*	2.11*
Spec 12:GA_t	-	-	-	-	-	-	-	-	-	-	-	-	1.63	2.34*	0.81	3.11*	1.98*	2.87*	2.22*	-3.03*	-3.02*
Spec 13:GJR_t	-	-	-	-	-	-	-	-	-	-	-	-	-	2.11*	-1.83	-1.10	1.89	-0.25	1.77	-2.11*	-3.42*
Spec 14:TGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-2.58*	-2.01*	-0.30	-1.70	-1.62	-2.62*	-3.96*
Spec 15:IGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.08	2.17*	2.72*	2.43*	-2.18*	-3.07*
Spec 16:AGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.63	2.28*	1.89	-3.86*	-3.57*
Spec 17:$NAGA_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-1.23	0.00	-2.26*	-3.42*
Spec 18:EGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.55	-3.48*	-4.15*
Spec 19:$APGA_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-2.53*	-3.89*
Spec 20:$FIGA_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-2.24*
Spec 21:A_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 22:GA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 23:GJR_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 24:TGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 25:IGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 26:AGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 27:$NAGA_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 28:EGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 29:$APGA_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 30:$FIGA_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 31:$comb_1$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 32:$comb_2$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 33:$comb_3$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 34:HAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 35:logHAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 36:HEAVY	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 37:$GA-MIDAS$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 38:$GA-RV_N$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 39:$GA-RV_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 40:$GA-RV_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 41:$quant_1$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 42:$quant_2$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Note: This table reports the out-of-sample t-statistic of the Unconditional Giacomini and White (2006) test under the robust loss of Patton (2011) for b=1 for CAC 40. The period is from 02/01/2004 to 29/12/2017. A positive t-statistic indicates that average loss of model A (in row) is larger than of model B (in column) and hence model B is superior.
*denotes significance at the 5% level

Table 3.22 cont. - Pairwise comparisons of unconditional Giacomini and White test for CAC 40 (4-year in-sample) - robust Patton (b=1)

	Spec 22	Spec 23	Spec 24	Spec 25	Spec 26	Spec 27	Spec 28	Spec 29	Spec 30	Spec 31	Spec 32	Spec 33	Spec 34	Spec 35	Spec 36	Spec 37	Spec 38	Spec 39	Spec 40	Spec 41	Spec 42
Spec 1:A_N	2.39*	3.08*	3.67*	2.56*	2.78*	3.31*	3.13*	3.62*	1.88	3.51*	3.40*	3.47*	-2.09*	2.70*	1.00	2.06*	1.41	1.35	1.41	-0.64	-0.65
Spec 2:GA_N	-1.57	1.69	2.44*	-0.17	1.84	2.08*	2.24*	2.30*	-2.66*	2.67*	2.56*	2.48*	-2.31*	0.56	-1.52	-3.38*	-1.00	-1.32	-0.68	-3.52*	-3.53*
Spec 3:GJR_M	-1.65	1.07	2.07*	-1.88	-1.05	2.06*	-0.39	1.72	-2.14*	0.84	0.45	0.69	-2.36*	-0.57	-1.85	-2.24*	-0.74	-1.12	-1.23	-3.06*	-3.07*
Spec 4:TGA_N	-2.42*	-2.18*	0.05	-2.74*	-2.08*	-0.01	-1.83	-0.68	-2.70*	-1.22	-1.32	-1.25	-2.42*	-1.70	-2.65*	-2.97*	-3.57*	-3.56*	-3.22*	-3.30*	-3.31*
Spec 5:IGA_N	-0.82	2.14*	2.78*	-0.46	0.80	2.40*	1.90	2.59*	-1.92	2.99*	2.78*	2.72*	-2.31*	0.54	-1.44	-2.07*	-3.32*	-3.67*	-2.98*	-3.18*	-3.19*
Spec 6:AGA_N	-3.98*	1.15	2.03*	-1.19	0.04	1.71	1.81	1.90	-4.01*	2.11*	1.92	1.95	-2.33*	0.24	-1.84	-5.03*	-2.34*	-3.54*	-2.31*	-3.92*	-3.93*
Spec 7:$NAGA_N$	-2.06*	-1.70	0.77	-2.33*	-1.65	1.61	-1.27	0.44	-2.36*	-0.59	-0.82	-0.65	-2.40*	-1.31	-2.26*	-2.52*	-1.32	-3.03*	-0.23	-3.11*	-3.12*
Spec 8:EGA_N	-2.98*	0.04	1.71	-3.15*	-2.45*	1.25	-2.27*	1.53	-3.43*	1.05	0.45	0.82	-2.37*	-0.80	-2.52*	-4.08*	-3.83*	-4.34*	-3.61*	-3.79*	-3.80*
Spec 9:$APGA_N$	-2.32*	-1.85	0.94	-2.60*	-1.97*	0.23	-1.70	-0.12	-2.63*	-1.06	-1.18	-1.10	-2.42*	-1.61	-2.62*	-2.89*	-3.18*	-5.42*	-3.34	-3.27*	-3.28*
Spec 10:$FIGA_N$	2.82*	2.33*	2.82*	2.30*	3.85*	2.41*	3.67*	2.73*	0.61	3.53*	3.60*	3.44*	-2.26*	1.88	-0.72	0.70	-0.44	-0.69	-1.06	-3.40*	-3.42*
Spec 11:A_t	2.93*	3.54*	4.09*	3.06*	3.48*	3.58*	3.91*	4.04*	2.24*	4.20*	4.14*	4.15*	-2.11*	3.11*	0.96	2.53*	2.76*	1.03	2.43*	-0.96	-0.97
Spec 12:GA_t	-1.40	1.72	2.39*	0.81	3.11*	2.07*	2.62*	2.28*	-2.96*	2.78*	2.76*	2.63*	-2.30*	0.82	-1.40	-2.92*	-0.67	-0.68	-1.24	-3.67*	-3.68*
Spec 13:GJR_t	-1.63	-0.23	2.02*	-1.82	-1.10	2.04*	-0.51	1.63	-2.09*	0.59	0.23	0.47	-2.36*	-0.64	-1.86	-2.22*	-1.12	-1.13	-0.78	-3.00*	-3.01*
Spec 14:TGA_t	-2.31*	-2.08*	-1.07	-2.57*	-2.01*	-0.26	-1.78	-1.52	-2.60*	-1.20	-1.29	-1.23	-2.42*	-1.67	-2.64*	-2.86*	-3.26*	-3.27*	-3.65*	-3.23*	-3.24*
Spec 15:IGA_t	-0.90	1.95	2.65*	-0.65	1.08	2.28*	2.12*	2.48*	-2.14*	3.02*	2.89*	2.77*	-2.31*	0.60	-1.46	-2.24*	-1.25	-1.43	-1.66	-3.30*	-3.32*
Spec 16:AGA_t	-3.29*	1.16	2.06*	-1.13	0.32	1.72	1.89	1.93	-3.79*	2.10*	1.90	1.93	-2.33*	0.23	-1.84	-5.23*	-4.01*	-3.89*	-1.12	-3.88*	-3.90*
Spec 17:$NAGA_t$	-1.97*	-1.79	0.06	-2.16*	-1.63	0.37	-1.33	-0.16	-2.24*	-0.80	-0.96	-0.84	-2.41*	-1.37	-2.26*	-2.40*	-1.63	-1.61	-1.45	-2.99*	-3.00*
Spec 18:EGA_t	-2.87*	0.25	1.75	-2.73*	-2.28*	1.31	-2.46*	1.59	-3.44*	1.27	0.72	1.05	-2.36*	-0.63	-2.48*	-4.06*	-3.76*	-4.12*	-4.35*	-3.83*	-3.85*
Spec 19:$APGA_t$	-2.21*	-1.76	0.29	-2.42*	-1.89	0.04	-1.64	-0.90	-2.51*	-1.03	-1.13	-1.07	-2.42*	-1.57	-2.61*	-2.76*	1.24	-1.26	-1.31	-3.19*	-3.20*
Spec 20:$FIGA_t$	3.14*	2.20*	2.69*	2.20*	3.86*	2.34*	3.48*	2.60*	-0.80	3.31*	3.35*	3.21*	-2.26*	1.71	-0.82	0.41	0.24	-0.12	0.67	-3.60*	-3.61*
Spec 21:A_{st}	2.96*	3.49*	4.05*	3.06*	3.57*	3.50*	4.02*	3.99*	2.22*	4.25*	4.21*	4.20*	-2.13*	3.09*	0.84	2.53*	1.96*	1.73	1.53	-1.16	-1.18
Spec 22:GA_{st}	-	1.71	2.37*	0.91	3.29*	2.05*	2.69*	2.26*	-3.10*	2.76*	2.73*	2.62*	-2.30*	0.89	-1.37	-2.77*	-0.54	-0.68	-0.37	-3.73*	-3.75*
Spec 23:GJR_{st}	-	-	2.02*	-1.93	-1.16	1.95	-0.52	1.64	-2.17*	0.64	0.27	0.51	-2.36*	-0.65	-1.89	-2.33*	-1.34	-0.35	-1.81	-3.06*	-3.08*
Spec 24:TGA_{st}	-	-	-	-2.63*	-2.06*	-0.02	-1.83	-1.40	-2.66*	-1.16	-1.25	-1.19	-2.42*	-1.65	-2.69*	-2.96*	-2.41*	-2.20*	-1.31	-3.28*	-3.29*
Spec 25:IGA_{st}	-	-	-	-	1.13	2.27*	2.14*	2.47*	-2.16*	3.01*	2.89*	2.77*	-2.31*	0.61	-1.46	-2.26*	-1.52	-0.43	-0.35	-3.32*	-3.33*
Spec 26:AGA_{st}	-	-	-	-	-	1.72	1.89	1.93	-3.79*	2.10*	1.90	1.93	-2.33*	0.23	-1.84	-5.23*	-4.88*	-3.95*	-1.23	-3.88*	-3.90*
Spec 27:$NAGA_{st}$	-	-	-	-	-	-	-1.41	-0.22	-2.32*	-0.87	-1.03	-0.91	-2.42*	-1.42	-2.32*	-2.50*	-1.20	-1.46	-1.47	-3.05*	-3.06*
Spec 28:EGA_{st}	-	-	-	-	-	-	-	1.69	-3.44*	1.57	1.10	1.37	-2.35*	-0.37	-2.39*	-4.00*	-2.62*	-3.59*	-2.90*	-3.89*	-3.91*
Spec 29:$APGA_{st}$	-	-	-	-	-	-	-	-	-2.58*	-0.98	-1.09	-1.02	-2.42*	-1.56	-2.68*	-2.86*	-1.71	-2.38*	-0.35	-3.26*	-3.27*
Spec 30:$FIGA_{st}$	-	-	-	-	-	-	-	-	-	3.27*	3.31*	3.18*	-2.26*	1.72	-0.81	0.47	1.27	-0.11	1.20	-3.64*	-3.66*
Spec 31:$comb_1$	-	-	-	-	-	-	-	-	-	-	-1.04	-1.23	-2.39*	-2.05*	-2.57*	-3.43*	-2.20*	-2.74*	-2.75*	-3.72*	-3.73*
Spec 32:$comb_2$	-	-	-	-	-	-	-	-	-	-	-	0.79	-2.38*	-1.52	-2.30*	-3.43*	-4.12*	-2.98*	-2.56*	-3.74*	-3.75*
Spec 33:$comb_3$	-	-	-	-	-	-	-	-	-	-	-	-	-2.39*	-2.10*	-2.50*	-3.29*	-3.21*	-2.20*	-3.33*	-3.70*	-3.71*
Spec 34:HAR	-	-	-	-	-	-	-	-	-	-	-	-	-	2.34*	2.21*	2.27*	2.28*	2.57*	3.12*	2.03*	2.03*
Spec 35:logHAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-1.73	-1.51	-2.71*	-2.12*	-2.30*	-3.38*	-3.40*
Spec 36:HEAVY	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.90	-2.81*	-2.10*	-2.12*	-2.31*	-2.33*
Spec 37:$GA-MIDAS$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-1.13	-1.22	-1.58	-3.28*	-3.29*
Spec 38:$GA-RV_N$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-4.09*	-0.10	-1.31	-1.70
Spec 39:$GA-RV_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.43	-1.22	-1.61
Spec 40:$GA-RV_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-2.77*	-3.31*
Spec 41:$quant_1$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-2.74*
Spec 42:$quant_2$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Note: This table reports the out-of-sample t-statistic of the Unconditional Giacomini and White (2006) test under the robust loss of Patton (2011) for b=1 for CAC 40. The period is from 02/01/2004 to 29/12/2017. A positive t-statistic indicates that average loss of model A (in row) is larger than of model B (in column) and hence model B is superior. *denotes significance at the 5% level

Table 3.23 – Pairwise comparisons of unconditional Giacomini and White test for CAC 40 (4-year in-sample) - robust Patton (b=-0.5)

	Spec 1	Spec 2	Spec 3	Spec 4	Spec 5	Spec 6	Spec 7	Spec 8	Spec 9	Spec 10	Spec 11	Spec12	Spec 13	Spec 14	Spec 15	Spec 16	Spec 17	Spec 18	Spec 19	Spec 20	Spec 21
Spec 1:A_N	-	7.57*	7.45*	8.31*	6.76*	8.52*	7.50*	8.61*	8.22*	7.08*	-2.24*	7.20*	7.48*	8.30*	6.68*	8.68*	7.41*	8.61*	8.19*	6.79*	-1.72
Spec 2:GA_N	-	-	3.11*	4.16*	-1.27	4.54*	3.47*	3.83*	3.80*	-3.05*	-8.34*	-1.87	2.93*	3.84*	-1.34	4.07*	3.18*	3.34*	3.46*	-2.91*	-8.26*
Spec 3:GJR_M	-	-	-	2.76*	-3.93*	-1.28	2.15*	0.07	1.99*	-3.51*	-7.93*	-3.11*	-0.37	2.35*	-3.67*	-1.32	1.96*	-0.24	1.61	-3.44*	-7.82*
Spec 4:TGA_N	-	-	-	-	-5.16*	-2.76*	-1.90	-2.56*	-2.99*	-4.24*	-8.62*	-4.08*	-2.98*	-1.16	-4.93*	-2.79*	-1.57	-2.56*	-3.13*	-4.16*	-8.48*
Spec 5:IGA_N	-	-	-	-	-	3.53*	4.44*	4.50*	4.73*	-1.54	-7.35*	0.19	3.71*	4.74*	-0.40	3.24*	3.99*	3.89*	4.30*	-1.38	-7.25*
Spec 6:AGA_N	-	-	-	-	-	-	2.04*	2.07*	2.39*	-5.36*	-9.43*	-4.93*	1.19	2.51*	-3.61*	-0.68	1.90	1.58	2.11*	-5.44*	-9.37*
Spec 7:$NAGA_N$	-	-	-	-	-	-	-	-1.31	0.78	-3.66*	-7.86*	-3.48*	-2.28*	1.32	-4.24*	-2.05*	0.54	-1.46	0.30	-3.63*	-7.74*
Spec 8:EGA_N	-	-	-	-	-	-	-	-	1.99*	-4.51*	-9.33*	-4.07*	-0.12	2.25*	-4.52*	-2.19*	1.27	-1.82	1.64	-4.46*	-9.23*
Spec 9:$APGA_N$	-	-	-	-	-	-	-	-	-	-3.99*	-8.53*	-3.77*	-2.13*	1.57	-4.54*	-2.42*	-0.53	-2.07*	-1.73	-3.91*	-8.39*
Spec 10:$FIGA_N$	-	-	-	-	-	-	-	-	-	-	-8.03*	2.61*	3.37*	4.01*	1.53	5.21*	3.42*	4.31*	3.74*	0.80	-8.00*
Spec 11:A_t	-	-	-	-	-	-	-	-	-	-	-	8.04*	7.95*	8.60*	7.30*	9.65*	7.74*	9.42*	8.48*	7.76*	1.48
Spec 12:GA_t	-	-	-	-	-	-	-	-	-	-	-	-	2.93*	3.78*	-0.35	4.40*	3.19*	3.71*	3.46*	-2.77*	-7.99*
Spec 13:GJR_t	-	-	-	-	-	-	-	-	-	-	-	-	-	2.57*	-3.47*	-1.24	2.14*	-0.20	1.74	-3.29*	-7.83*
Spec 14:TGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-4.53*	-2.56*	-1.16	-2.31*	-3.27*	-3.92*	-8.46*
Spec 15:IGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3.30*	3.81*	3.98*	4.13*	-1.40	-7.22*
Spec 16:AGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.92	1.72	2.15*	-5.18*	-9.59*
Spec 17:$NAGA_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-1.41	0.10	-3.38*	-7.62*
Spec 18:EGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.77	-4.28*	-9.33*
Spec 19:$APGA_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-3.64*	-8.34*
Spec 20:$FIGA_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-7.74*
Spec 21:A_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 22:GA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 23:GJR_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 24:TGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 25:IGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 26:AGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 27:$NAGA_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 28:EGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 29:$APGA_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 30:$FIGA_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 31:$comb_1$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 32:$comb_2$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 33:$comb_3$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 34:HAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 35:logHAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 36:HEAVY	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 37:$GA-MIDAS$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 38:$GA-RV_N$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 39:$GA-RV_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 40:$GA-RV_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 41:$quant_1$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 42:$quant_2$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Note: This table reports the out-of-sample t-statistic of the Unconditional Giacomini and White (2006) test under the robust loss of Patton (2011) for b=-0.5 for CAC 40. The period is from 02/01/2004 to 29/12/2017. A positive t-statistic indicates that average loss of model A (in row) is larger than of model B (in column) and hence model B is superior.
*denotes significance at the 5% level

Table 3.23 cont. - Pairwise comparisons of unconditional Giacomini and White test for CAC 40 (4-year in-sample) - robust Patton (b=-0.5)

	Spec 22	Spec 23	Spec 24	Spec 25	Spec 26	Spec 27	Spec 28	Spec 29	Spec 30	Spec 31	Spec 32	Spec 33	Spec 34	Spec 35	Spec 36	Spec 37	Spec 38	Spec 39	Spec 40	Spec 41	Spec 42
Spec 1:A_N	7.19*	7.55*	8.44*	6.71*	8.68*	7.49*	8.57*	8.36*	6.78*	9.06*	8.63*	9.03*	-8.53*	8.10*	0.07	7.06*	1.12	1.23	-0.91	-5.00*	-5.37*
Spec 2:GA_N	-1.95	2.99*	3.85*	-1.27	4.07*	3.26*	2.79*	3.43*	-2.85*	7.34*	6.67*	7.22*	-10.28*	4.32*	-5.89*	-3.72*	-1.41	-3.51*	-4.43*	-7.28*	-7.49*
Spec 3:GJR_M	-3.06*	-0.66	2.19*	-3.63*	-1.32	2.07*	-0.68	1.40	-3.38*	4.22*	2.46*	3.90*	-10.04*	1.08	-6.07*	-3.85*	-5.20*	-5.22*	-4.87*	-7.05*	-7.26*
Spec 4:TGA_N	-4.01*	-2.98*	-1.76	-4.87*	-2.79*	-1.52	-2.74*	-3.12*	-4.10*	1.76	0.00	1.56	-9.96*	-0.53	-6.83*	-4.85*	-3.60*	-4.22*	-4.99*	-7.10*	-7.30*
Spec 5:IGA_N	-0.01	3.74*	4.69*	-0.10	3.24*	4.03*	3.23*	4.22*	-1.38	8.43*	7.33*	8.24*	-10.00*	4.81*	-5.40*	-2.23*	-1.71	-1.12	-3.93*	-6.96*	-7.18*
Spec 6:AGA_N	-5.04*	1.21	2.52*	-3.59*	-0.68	1.97*	0.90	2.07*	-5.33*	5.50*	4.66*	5.36*	-10.58*	2.49*	-6.70*	-5.95*	-3.70*	-3.90*	-5.13*	-7.63*	-7.82*
Spec 7:$NAGA_N$	-3.43*	-2.24*	0.94	-4.18*	-2.05*	0.68	-1.73	0.00	-3.59*	2.61*	0.85	2.38*	-9.87*	0.06	-6.34*	-4.15*	-4.02*	-4.43*	-4.35*	-6.98*	-7.19*
Spec 8:EGA_N	-4.06*	-0.18	2.25*	-4.45*	-2.19*	1.35	-2.90*	1.57	-4.40*	5.41*	3.10*	5.02*	-10.39*	1.26	-7.14*	-5.37*	-4.61*	-4.70*	-4.71*	-7.46*	-7.65*
Spec 9:$APGA_N$	-3.72*	-2.16*	0.65	-4.49*	-2.42*	-0.47	-2.30*	-2.09*	-3.86*	2.30*	0.44	2.07*	-9.92*	-0.21	-6.77*	-4.65*	-5.21*	-5.12*	-5.22*	-7.06*	-7.27*
Spec 10:$FIGA_N$	2.52*	3.44*	4.05*	1.61	5.21*	3.48*	4.04*	3.75*	0.57	6.75*	6.37*	6.77*	-10.55*	5.42*	-5.48*	-1.05	-4.07*	-4.23*	-3.99*	-7.43*	-7.65*
Spec 11:A_t	8.07*	8.05*	8.76*	7.33*	9.65*	7.84*	9.47*	8.67*	7.76*	9.60*	9.30*	9.58*	-8.65*	8.71*	0.53	8.09*	3.00*	2.89*	0.01	-5.02*	-5.41*
Spec 12:GA_t	-1.63	3.00*	3.79*	-0.25	4.40*	3.26*	3.26*	3.44*	-2.75*	7.22*	6.93*	7.21*	-10.39*	4.90*	-5.90*	-3.29*	-0.56	-1.98*	-5.34*	-7.39*	-7.60*
Spec 13:GJR_t	-2.90*	-0.65	2.41*	-3.42*	-1.24	2.28*	-0.63	1.52	-3.24*	4.12*	2.38*	3.80*	-9.99*	1.07	-6.03*	-3.75*	-3.41*	-5.04*	-4.42*	-7.00*	-7.21*
Spec 14:TGA_t	-3.73*	-2.58*	-1.50	-4.47*	-2.56*	-1.10	-2.50*	-3.19*	-3.87*	1.76	0.14	1.58	-9.90*	-0.40	-6.77*	-4.63*	-2.73*	-2.77*	-2.92*	-7.04*	-7.24*
Spec 15:IGA_t	0.10	3.50*	4.49*	2.40*	3.30*	3.85*	3.34*	4.07*	-1.40	8.27*	7.38*	8.15*	-10.07*	5.01*	-5.47*	-2.27*	-1.20	-1.98*	-2.97*	-7.03*	-7.25*
Spec 16:AGA_t	-4.54*	1.26	2.57*	-3.27*	0.67	2.00*	1.04	2.13*	-5.08*	5.37*	4.48*	5.23*	-10.58*	2.46*	-6.72*	-5.94*	-5.92*	-5.33*	-5.36*	-7.62*	-7.82*
Spec 17:$NAGA_t$	-3.15*	-2.10*	0.75	-3.76*	-1.92	0.44	-1.66	-0.17	-3.34*	2.10*	0.63	1.93	-9.76*	-0.01	-6.25*	-3.90*	-5.78*	-5.33*	-4.92*	-6.88*	-7.09*
Spec 18:EGA_t	-3.75*	0.16	2.32*	-3.92*	-1.72	1.49	-3.72*	1.73	-4.23*	5.19*	3.28*	4.90*	-10.46*	1.51	-7.21*	-5.20*	-6.49*	-5.99*	-5.73*	-7.52*	-7.71*
Spec 19:$APGA_t$	-3.42*	-1.77	2.01*	-4.08*	-2.15*	-0.04	-2.02*	-1.44	-3.61*	2.33*	0.61	2.13*	-9.86*	-0.05	-6.70*	-4.39*	-5.70*	-5.42*	-5.94*	-7.00*	-7.21*
Spec 20:$FIGA_t$	2.73*	3.36*	3.94*	1.49	5.18*	3.44*	4.00*	3.65*	-1.03	6.80*	6.54*	6.82*	-10.60*	5.44*	-5.61*	-1.34	-6.02*	-6.01*	-6.70*	-7.50*	-7.71*
Spec 21:A_{st}	8.02*	7.94*	8.62*	7.25*	9.59*	7.71*	9.39*	8.52*	7.74*	9.49*	9.22*	9.48*	-8.74*	8.65*	0.48	8.07*	2.21*	2.35*	1.29	-5.11*	-5.49*
Spec 22:GA_{st}	-	2.96*	3.74*	-0.01	4.54*	3.22*	3.33*	3.41*	-2.74*	7.02*	6.77*	7.02*	-10.44*	4.92*	-5.93*	-3.23*	-0.95	-2.98*	-3.01*	-7.44*	-7.65*
Spec 23:GJR_{st}	-	-	2.45*	-3.45*	-1.26	2.24*	-0.62	1.57	-3.30*	4.28*	2.52*	3.94*	-10.03*	1.12	-6.08*	-3.82*	-3.78*	-4.12*	-4.41*	-7.04*	-7.25*
Spec 24:TGA_{st}	-	-	-	-4.43*	-2.57*	-0.71	-2.54*	-3.11*	-3.90*	2.02*	0.30	1.81	-9.94*	-0.30	-6.83*	-4.70*	-5.15*	-5.23*	-5.35*	-7.07*	-7.28*
Spec 25:IGA_{st}	-	-	-	-	3.27*	3.80*	3.28*	4.01*	-1.49	8.23*	7.36*	8.12*	-10.09*	4.99*	-5.49*	-2.35*	-1.72	-2.91*	-2.97*	-7.05*	-7.27*
Spec 26:AGA_{st}	-	-	-	-	-	2.00*	1.04	2.13*	-5.08*	5.37*	4.48*	5.23*	-10.58*	2.46*	-6.72*	-5.94*	-6.82*	-6.63*	-6.66*	-7.62*	-7.82*
Spec 27:$NAGA_{st}$	-	-	-	-	-	-	-1.74	-0.24	-3.40*	2.12*	0.61	1.94	-9.80*	-0.03	-6.32*	-3.97*	-2.12*	-3.69*	-3.71*	-6.91*	-7.12*
Spec 28:EGA_{st}	-	-	-	-	-	-	-	2.01*	-3.98*	5.29*	3.70*	5.07*	-10.52*	1.90	-7.21*	-4.94*	-4.54*	-3.99*	-3.72*	-7.56*	-7.76*
Spec 29:$APGA_{st}$	-	-	-	-	-	-	-	-	-3.62*	2.61*	0.79	2.39*	-9.90*	0.06	-6.79*	-4.45*	-2.19*	-2.20*	-2.22*	-7.04*	-7.25*
Spec 30:$FIGA_{st}$	-	-	-	-	-	-	-	-	-	6.68*	6.43*	6.71*	-10.63*	5.41*	-5.63*	-1.28	-3.94*	-4.13*	-4.24*	-7.54*	-7.74*
Spec 31:$comb_1$	-	-	-	-	-	-	-	-	-	-	-3.86*	-0.81	-10.54*	-3.08*	-7.92*	-7.22*	-6.24*	-6.78*	-6.82*	-7.71*	-7.89*
Spec 32:$comb_2$	-	-	-	-	-	-	-	-	-	-	-	4.02*	-10.59*	-1.16	-7.57*	-6.75*	-4.77*	-4.13*	-5.22*	-7.72*	-7.90*
Spec 33:$comb_3$	-	-	-	-	-	-	-	-	-	-	-	-	-10.57*	-3.30*	-7.95*	-7.20*	-7.12*	-6.84*	-6.87*	-7.73*	-7.91*
Spec 34:HAR	-	-	-	-	-	-	-	-	-	-	-	-	-	10.68*	9.42*	10.30*	9.31*	9.76*	9.82*	9.06*	8.17*
Spec 35:logHAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-7.69*	-5.50*	-6.77*	-6.00*	-6.45	-7.80*	-7.99*
Spec 36:HEAVY	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	5.38*	-1.54	-1.62	-5.72*	-5.80*	-6.21*
Spec 37:$GA-MIDAS$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-1.03	-6.71*	-6.80*	-7.26*	-7.48*
Spec 38:$GA-RV_N$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-0.31	-7.01*	-7.10*	-7.11*
Spec 39:$GA-RV_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-7.23*	-7.67*	-7.48*
Spec 40:$GA-RV_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-5.22*	-6.93*
Spec 41:$quant_1$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-7.33*
Spec 42:$quant_2$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Note: This table reports the out-of-sample t-statistic of the Unconditional Giacomini and White (2006) test under the robust loss of Patton (2011) for b=-0.5 for CAC 40. The period is from 02/01/2004 to 29/12/2017. A positive t-statistic indicates that average loss of model A (in row) is larger than of model B (in column) and hence model B is superior. *denotes significance at the 5% level

Table 3.24 – Pairwise comparisons of unconditional Giacomini and White test for CAC 40 (4-year in-sample) - all Loss Functions

	logMSE		sdMSE		propMSE		RMSE		propRMSE		MAE		logMAE		sdMAE		MAPE		GMLE	
	win	lose	win	lose	win	lose	win	lose	win	lose	win	lose	win	lose	win	lose	win	lose	win	lose
Spec 1:A_N	5	33	5	33	4	34	4	34	4	34	4	33	5	33	5	33	4	33	2	35
Spec 2:GA_N	10	25	14	22	9	23	9	27	9	27	11	22	10	26	10	23	11	22	9	28
Spec 3:GJR_M	18	15	18	11	16	19	17	18	17	18	18	13	20	13	18	13	18	13	17	19
Spec 4:TGA_N	32	4	34	4	25	5	28	7	28	7	31	4	31	4	33	4	31	4	27	11
Spec 5:IGA_N	8	29	7	25	6	30	7	30	7	30	6	29	8	29	7	29	6	29	7	30
Spec 6:AGA_N	17	20	17	11	16	10	16	22	16	22	17	8	15	21	17	13	17	8	16	22
Spec 7:$NAGA_N$	22	12	21	7	21	5	21	15	21	15	19	9	20	13	19	13	19	9	21	16
Spec 8:EGA_N	24	8	22	7	21	6	26	10	26	10	25	4	26	7	26	6	25	4	28	8
Spec 9:$APGA_N$	26	8	28	5	22	6	24	13	24	13	24	5	26	8	27	7	24	5	24	13
Spec 10:$FIGA_N$	14	23	7	23	9	23	13	24	13	24	10	24	14	24	11	23	10	24	13	24
Spec 11:A_t	3	35	3	35	2	36	2	36	2	36	2	35	3	35	3	35	2	35	0	37
Spec 12:GA_t	10	26	11	23	9	24	9	27	9	27	11	23	10	26	10	25	11	23	9	27
Spec 13:GJR_t	17	17	17	16	15	18	17	18	17	18	17	16	20	14	17	16	17	16	20	18
Spec 14:TGA_t	32	5	31	5	22	6	28	7	28	7	28	6	31	5	29	6	28	6	28	9
Spec 15:IGA_t	7	31	7	27	6	31	6	32	6	32	6	30	7	31	7	30	6	30	5	32
Spec 16:AGA_t	18	17	17	12	17	16	17	18	17	18	17	10	17	19	17	13	17	10	17	19
Spec 17:$NAGA_t$	21	14	18	10	20	6	21	15	21	15	17	15	20	14	17	15	17	15	22	15
Spec 18:EGA_t	26	7	20	7	21	7	28	7	28	7	24	4	26	7	26	6	24	4	32	6
Spec 19:$APGA_t$	26	8	23	8	19	10	26	11	26	11	22	9	26	8	26	9	22	9	25	12
Spec 20:$FIGA_t$	14	23	7	25	9	24	13	24	13	24	11	24	15	22	12	23	11	24	12	24
Spec 21:A_{st}	4	34	4	34	3	35	3	35	3	35	3	34	4	34	4	34	3	34	1	36
Spec 22:GA_{st}	11	25	13	22	10	23	11	26	11	26	14	22	12	25	11	23	14	22	11	26
Spec 23:GJR_{st}	21	13	18	12	16	15	21	15	21	15	18	13	21	13	18	13	18	13	21	15
Spec 24:TGA_{st}	33	4	31	5	26	5	32	5	32	5	29	5	33	4	33	4	29	5	31	7
Spec 25:IGA_{st}	8	29	8	25	7	30	7	30	7	30	7	28	8	29	8	29	7	28	6	30
Spec 26:AGA_{st}	18	17	17	12	17	16	17	18	17	18	17	10	17	19	17	13	17	10	17	19
Spec 27:$NAGA_{st}$	23	12	20	8	23	5	24	13	24	13	18	12	21	13	18	13	18	12	24	12
Spec 28:EGA_{st}	27	7	17	10	23	5	32	5	32	5	23	4	26	5	26	6	23	4	33	5
Spec 29:$APGA_{st}$	28	7	23	8	22	6	27	7	27	7	24	7	28	7	27	7	24	7	29	8
Spec 30:$FIGA_{st}$	16	22	9	23	10	23	15	23	15	23	13	22	16	19	16	22	13	22	15	23
Spec 31:$comb_1$	36	2	36	1	35	3	35	3	35	3	36	1	36	2	36	2	36	1	35	3
Spec 32:$comb_2$	35	3	35	2	34	4	34	4	34	4	35	3	35	3	35	3	35	3	34	4
Spec 33:$comb_3$	37	1	38	0	36	2	36	2	36	2	37	0	37	1	37	1	37	0	36	2
Spec 34:HAR	0	38	0	38	0	38	0	38	0	38	0	38	0	38	0	38	0	38	0	34
Spec 35:logHAR	38	0	35	1	37	1	37	0	37	0	36	0	38	0	38	0	36	0	37	1
Spec 36:HEAVY	6	32	6	32	38	0	37	0	37	0	6	29	6	32	6	32	6	29	38	0
Spec 37:$GA-MIDAS$	10	25	7	25	9	21	9	26	9	26	9	27	10	25	10	25	9	27	10	25
Spec 38:$GA-RV_N$	1	31	8	30	18	21	15	21	17	21	14	23	8	33	10	22	14	25	28	7
Spec 39:$GA-RV_t$	9	32	8	31	10	23	19	22	16	20	10	23	9	26	12	23	8	26	28	7
Spec 40:$GA-RV_{st}$	8	32	7	34	39	2	36	1	34	0	16	11	8	35	3	21	18	11	39	2
Spec 41:$quant_1$	1	37	1	36	1	37	1	37	1	37	1	37	1	37	1	37	1	37	3	34
Spec 42:$quant_2$	2	36	1	36	5	33	5	33	5	33	2	33	2	36	2	36	2	33	5	31

Note: This table reports a summary of out-of-sample the results for the Unconditional Giacomini and White (2006) performed under the logMSE, sdMSE, propMSE, RMSE, propRMSE, MAE, logMAE, sdMAE, MAPE and GMLE loss functions for CAC 40. The period is from 02/01/2004 to 29/12/2017. The "win" column reports the cases where a certain specification outperformed significantly the competing specifications. On the contrary, the "lose" column reports the cases where a certain specification was outperformed significantly by the competing specifications. Adding the two columns for each specification and for each loss function, respectively, shows how many of the 41 in total pairwise comparisons where statistically significant. For the hypothesis testing the 5% significance level is selected.

Table 3.25 – Pairwise comparisons of unconditional Giacomini and White test for FTSE 100 (4-year in-sample) - robust Patton (b=1)

	Spec 1	Spec 2	Spec 3	Spec 4	Spec 5	Spec 6	Spec 7	Spec 8	Spec 9	Spec 10	Spec 11	Spec12	Spec 13	Spec 14	Spec 15	Spec 16	Spec 17	Spec 18	Spec 19	Spec 20	Spec 21
Spec 1:A_N	-	1.91	1.93	2.08*	1.90	2.03*	1.98*	2.18*	2.02*	1.94	1.32	1.91	1.93	2.07*	1.90	2.04*	1.96*	2.20*	2.02*	1.95	1.30
Spec 2:GA_N	-	-	0.34	2.96*	-4.63*	3.05*	0.76	3.69*	1.84	0.94	-2.15*	-0.05	0.28	2.40*	-4.17*	2.89*	0.54	3.63*	1.65	1.07	-2.17*
Spec 3:GJR_M	-	-	-	1.04	-0.41	0.42	0.49	1.29	0.77	-0.05	-2.08*	-0.33	-0.24	1.13	-0.40	0.49	0.28	1.35	0.88	0.00	-2.09*
Spec 4:TGA_N	-	-	-	-	-3.10*	-1.00	-1.12	1.73	-1.87	-2.04*	-2.42*	-2.89*	-0.97	-0.49	-3.09*	-0.70	-1.14	1.83	-1.26	-1.81	-2.44*
Spec 5:IGA_N	-	-	-	-	-	3.21*	0.84	3.77*	1.97*	1.15	-2.14*	1.08	0.34	2.53*	0.26	3.03*	0.61	3.71*	1.77	1.25	-2.15*
Spec 6:AGA_N	-	-	-	-	-	-	-0.24	3.55*	0.08	-2.92*	-2.37*	-3.38*	-0.41	0.64	-3.28*	1.85	-0.34	3.52*	0.10	-3.05*	-2.39*
Spec 7:$NAGA_N$	-	-	-	-	-	-	-	1.45	0.58	-0.36	-2.20*	-0.73	-0.46	1.21	-0.83	0.34	-1.06	1.53	0.67	-0.27	-2.21*
Spec 8:EGA_N	-	-	-	-	-	-	-	-	-1.79	-3.91*	-2.62*	-3.77*	-1.22	-1.41	-3.80*	-3.39*	-1.42	2.18*	-1.55	-3.96*	-2.64*
Spec 9:$APGA_N$	-	-	-	-	-	-	-	-	-	-1.02	-2.30*	-1.77	-0.71	3.16*	-1.96*	0.09	-0.72	1.86	0.23	-0.85	-2.32*
Spec 10:$FIGA_N$	-	-	-	-	-	-	-	-	-	-	-2.21*	-1.04	0.03	1.56	-1.16	3.19*	0.20	3.89*	0.95	0.86	-2.22*
Spec 11:A_t	-	-	-	-	-	-	-	-	-	-	-	2.15*	2.06*	2.39*	2.13*	2.40*	2.15*	2.65*	2.30*	2.23*	0.68
Spec 12:GA_t	-	-	-	-	-	-	-	-	-	-	-	-	0.28	2.32*	-1.21	3.17*	0.52	3.70*	1.59	1.20	-2.17*
Spec 13:GJR_t	-	-	-	-	-	-	-	-	-	-	-	-	-	1.03	-0.34	0.48	0.29	1.27	0.81	0.02	-2.07*
Spec 14:TGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-2.52*	-0.42	-1.24	1.52	-1.98*	-1.36	-2.40*
Spec 15:IGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3.08*	0.60	3.73*	1.76	1.27	-2.15*
Spec 16:AGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-0.42	3.42*	-0.04	-3.70*	-2.42*
Spec 17:$NAGA_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.50	0.82	-0.13	-2.16*
Spec 18:EGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-1.63	-3.94*	-2.68*
Spec 19:$APGA_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-0.80	-2.31*
Spec 20:$FIGA_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-2.25*
Spec 21:A_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 22:GA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 23:GJR_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 24:TGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 25:IGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 26:AGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 27:$NAGA_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 28:EGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 29:$APGA_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 30:$FIGA_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 31:$comb_1$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 32:$comb_2$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 33:$comb_3$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 34:HAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 35:logHAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 36:HEAVY	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 37:$GA-MIDAS$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 38:$GA-RV_N$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 39:$GA-RV_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 40:$GA-RV_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 41:$quant_1$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 42:$quant_2$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Note:This table reports the out-of-sample t-statistic of the Unconditional Giacomini and White (2006) test under the robust loss of Patton (2011) for b=1 for FTSE 100. The period is from 02/01/2004 to 29/12/2017. A positive t-statistic indicates that average loss of model A (in row) is larger than of model B (in column) and hence model B is superior.
*denotes significance at the 5% level

Table 3.26 cont. - Pairwise comparisons of unconditional Giacomini and White test for FTSE 100 (4-year in-sample) - robust Patton (b=1)

	Spec 22	Spec 23	Spec 24	Spec 25	Spec 26	Spec 27	Spec 28	Spec 29	Spec 30	Spec 31	Spec 32	Spec 33	Spec 34	Spec 35	Spec 36	Spec 37	Spec 38	Spec 39	Spec 40	Spec 41	Spec 42
Spec 1:A_N	1.91	1.92	2.06*	1.90	2.06*	1.95	2.22*	2.05*	1.96*	2.17*	2.11*	2.17*	-0.69	2.05*	1.45	-0.94	1.45	0.90	1.12	1.59	1.59
Spec 2:GA_N	0.69	0.21	2.02*	-3.77*	3.08*	0.44	3.62*	1.91	1.13	3.85*	4.57*	3.91*	-2.38*	1.01	-1.51	-6.17*	-0.95	-1.12	-1.23	-1.00	-1.01
Spec 3:GJR_M	-0.29	-0.35	1.25	-0.42	0.55	0.21	1.44	1.22	0.00	1.30	1.16	1.32	-2.36*	0.40	-1.18	-5.38*	-4.43*	-3.01*	-0.71	-0.82	-0.82
Spec 4:TGA_N	-2.81*	-0.92	-0.40	-3.10*	-0.58	-1.02	2.02*	-0.51	-1.82	1.60	0.81	1.70	-2.53*	-0.14	-2.16*	-6.52*	-5.40*	-3.91*	-3.13*	-1.69	-1.69
Spec 5:IGA_N	1.72	0.27	2.14*	-0.78	3.22*	0.50	3.70*	2.02*	1.31	3.96*	4.73*	4.02*	-2.37*	1.07	-1.47	-6.14*	-1.45	-0.98	-0.73	-0.96	-0.97
Spec 6:AGA_N	-3.33*	-0.42	0.48	-3.17*	2.41*	-0.34	3.46*	0.40	-2.97*	2.88*	2.08*	2.92*	-2.50*	0.31	-2.24*	-6.70*	-5.61*	-5.71*	-5.73*	-1.70	-1.70
Spec 7:$NAGA_N$	-0.68	-0.48	1.34	-0.85	0.40	-0.69	1.68	1.18	-0.27	1.48	1.17	1.52	-2.41*	0.30	-1.43	-5.73*	-0.86	-2.45*	-1.21	-1.03	-1.04
Spec 8:EGA_N	-3.75*	-1.17	-1.22	-3.75*	-3.24*	-1.32	2.34*	-1.23	-3.93*	-0.79	-1.38	-0.72	-2.63*	-1.01	-2.93*	-7.08*	-5.83*	-4.76*	-4.77*	-2.41*	-2.41*
Spec 9:$APGA_N$	-1.68	-0.68	2.00*	-1.99*	0.18	-0.66	1.99*	1.55	-0.86	1.78	1.59	1.85	-2.47*	0.16	-1.81	-6.22*	-3.71*	-3.42*	-0.79	-1.36	-1.37
Spec 10:$FIGA_N$	-0.92	-0.01	1.30	-1.15	3.46*	0.15	3.86*	1.21	1.17	3.44*	3.07*	3.45*	-2.42*	0.90	-1.86	-6.49*	-0.45	-0.47	-0.11	-1.31	-1.31
Spec 11:A_t	2.16*	2.04*	2.37*	2.13*	2.42*	2.12*	2.69*	2.35*	2.24*	2.59*	2.49*	2.59*	-1.29	2.40*	1.35	-2.37*	2.31*	1.49	2.10*	1.60	1.60
Spec 12:GA_t	2.38*	0.21	1.94	-1.11	3.35*	0.43	3.67*	1.84	1.25	3.84*	4.52*	3.89*	-2.38*	1.04	-1.54	-6.21*	-0.69	-0.71	-0.88	-1.02	-1.03
Spec 13:GJR_t	-0.24	-0.46	1.13	-0.35	0.53	0.23	1.37	1.11	0.02	1.22	1.07	1.23	-2.35*	0.40	-1.13	-5.28*	-1.60	-1.62	-1.05	-0.78	-0.79
Spec 14:TGA_t	-2.24*	-0.98	-0.29	-2.55*	-0.32	-1.09	1.71	-0.52	-1.38	1.34	0.80	1.40	-2.51*	-0.06	-2.00*	-6.36*	-4.31*	-3.87*	-2.67*	-1.54	-1.55
Spec 15:IGA_t	1.89	0.26	2.12*	-0.55	3.27*	0.50	3.72*	2.00*	1.33	3.97*	4.73*	4.03*	-2.37*	1.07	-1.48	-6.15*	-0.44	-1.05	-0.13	-0.97	-0.97
Spec 16:AGA_t	-3.13*	-0.48	0.31	-2.99*	1.71	-0.41	3.38*	0.23	-3.53*	2.50*	1.56	2.53*	-2.51*	0.20	-2.37*	-6.79*	-3.72*	-3.62*	-1.93	-1.82	-1.82
Spec 17:$NAGA_t$	-0.48	-0.34	1.38	-0.62	0.48	-0.31	1.63	1.26	-0.13	1.45	1.17	1.47	-2.39*	0.36	-1.32	-5.59*	-1.13	-1.41	-1.01	-0.94	-0.95
Spec 18:EGA_t	-3.68*	-1.22	-1.33	-3.69*	-3.30*	-1.39	2.32*	-1.33	-3.93*	-1.26	-1.53	-1.18	-2.64*	-1.18	-3.02*	-7.13*	-2.51*	-2.41*	-3.99*	-2.50*	-2.51*
Spec 19:$APGA_t$	-1.52	-0.77	2.75*	-1.79	0.13	-0.74	1.78	3.03*	-0.80	1.53	1.27	1.58	-2.47*	0.14	-1.77	-6.16*	-1.36	-1.61	-1.25	-1.33	-1.34
Spec 20:$FIGA_t$	-1.09	-0.06	1.13	-1.25	3.93*	0.09	3.88*	1.04	0.13	3.28*	2.80*	3.29*	-2.44*	0.88	-1.97*	-6.58*	0.81	-1.17	1.22	-1.41	-1.41
Spec 21:A_{st}	2.18*	2.04*	2.38*	2.15*	2.44*	2.13*	2.72*	2.37*	2.25*	2.61*	2.51*	2.61*	-1.31	2.42*	1.34	-2.45*	1.91*	1.92	1.87	1.60	1.60
Spec 22:GA_{st}	-	0.18	1.88	-1.69	3.33*	0.39	3.66*	1.77	1.15	3.82*	4.45*	3.87*	-2.38*	1.01	-1.56	-6.23*	-1.54	-1.63	-0.99	-1.05	-1.05
Spec 23:GJR_{st}	-	-	1.05	-0.27	0.53	0.29	1.31	1.03	0.06	1.16	1.02	1.18	-2.33*	0.40	-1.08	-5.18*	-0.25	-0.31	-0.69	-0.74	-0.74
Spec 24:TGA_{st}	-	-	-	-2.15*	-0.22	-1.20	1.51	-0.85	-1.14	1.15	0.69	1.20	-2.50*	-0.03	-1.89	-6.23*	-3.62*	-3.40*	-1.42	-1.45	-1.46
Spec 25:IGA_{st}	-	-	-	-	3.19*	0.51	3.68*	2.04*	1.32	3.94*	4.71*	4.00*	-2.37*	1.07	-1.46	-6.14*	-1.78	-1.82	-1.37	-0.95	-0.96
Spec 26:AGA_{st}	-	-	-	-	-	-0.47	3.29*	0.15	-3.89*	2.40*	1.43	2.43*	-2.52*	0.12	-2.39*	-6.79*	-5.12*	-5.35*	-1.56	-1.84	-1.85
Spec 27:$NAGA_{st}$	-	-	-	-	-	-	1.51	1.11	-0.08	1.32	1.07	1.34	-2.38*	0.36	-1.26	-5.48*	-1.05	-1.08	-1.22	-0.89	-0.89
Spec 28:EGA_{st}	-	-	-	-	-	-	-	-1.49	-3.88*	-1.84	-1.75	-1.74	-2.66*	-1.37	-3.09*	-7.16*	-3.42*	-3.52*	-3.82*	-2.58*	-2.58*
Spec 29:$APGA_{st}$	-	-	-	-	-	-	-	-	-1.06	1.16	0.76	1.20	-2.50*	-0.00	-1.85	-6.20*	-1.62	-2.12*	-0.89	-1.42	-1.42
Spec 30:$FIGA_{st}$	-	-	-	-	-	-	-	-	-	3.34*	2.84*	3.34*	-2.44*	0.86	-1.95	-6.56*	1.35	-1.33	1.14	-1.39	-1.40
Spec 31:$comb_1$	-	-	-	-	-	-	-	-	-	-	-1.20	0.55	-2.60*	-0.74	-2.64*	-6.87*	-2.25*	-2.85*	-3.85*	-2.14*	-2.15*
Spec 32:$comb_2$	-	-	-	-	-	-	-	-	-	-	-	1.30	-2.56*	-0.34	-2.32*	-6.65*	-2.48*	-2.24*	-3.12*	-1.84	-1.85
Spec 33:$comb_3$	-	-	-	-	-	-	-	-	-	-	-	-	-2.60*	-0.75	-2.63*	-6.87*	-4.24*	-4.89*	-3.78*	-2.14*	-2.15*
Spec 34:HAR	-	-	-	-	-	-	-	-	-	-	-	-	-	2.56*	2.10*	0.06	5.14*	5.76*	5.51*	2.21*	2.21*
Spec 35:logHAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-3.72*	-7.45*	-2.45*	-2.71*	-3.31*	-3.04*	-3.05*
Spec 36:HEAVY	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-6.67*	-1.99*	-2.12*	-2.38*	2.34*	2.31*
Spec 37:$GA-MIDAS$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-1.79	-1.78	-1.65	6.80*	6.80*
Spec 38:$GA-RV_N$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-2.91*	-0.07	-1.32	-0.72
Spec 39:$GA-RV_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.47	-1.63	-1.01
Spec 40:$GA-RV_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-2.24*	-2.82*
Spec 41:$quant_1$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-3.86*
Spec 42:$quant_2$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Note: This table reports the out-of-sample t-statistic of the Unconditional Giacomini and White (2006) test under the robust loss of Patton (2011) for b=1 for FTSE 100. The period is from 02/01/2004 to 29/12/2017. A positive t-statistic indicates that average loss of model A (in row) is larger than of model B (in column) and hence model B is superior. *denotes significance at the 5% level

Table 3.26 – Pairwise comparisons of unconditional Giacomini and White test for FTSE 100 (4-year in-sample) - robust Patton (b=-0.5)

	Spec 1	Spec 2	Spec 3	Spec 4	Spec 5	Spec 6	Spec 7	Spec 8	Spec 9	Spec 10	Spec 11	Spec12	Spec 13	Spec 14	Spec 15	Spec 16	Spec 17	Spec 18	Spec 19	Spec 20	Spec 21
Spec 1:A_N	-	4.57*	5.34*	5.54*	4.36*	5.15*	5.07*	5.74*	5.43*	4.55*	0.03	4.51*	5.34*	5.57*	4.35*	5.21*	5.00*	5.76*	5.49*	4.48*	0.13
Spec 2:GA_N	-	-	2.33*	3.61*	-2.15*	2.88*	1.73	3.98*	3.31*	0.31	-5.48*	-0.58	2.28*	3.50*	-2.09*	3.13*	1.57	3.76*	3.34*	0.23	-5.42*
Spec 3:GJR_M	-	-	-	0.30	-2.57*	-0.96	-1.60	0.65	-0.06	-1.94	-5.46*	-2.25*	0.91	0.43	-2.56*	-0.79	-1.73	0.69	0.17	-1.84	-5.40*
Spec 4:TGA_N	-	-	-	-	-3.64*	-2.54*	-1.40	1.39	-1.81	-3.32*	-6.30*	-3.53*	-0.18	0.41	-3.64*	-2.10*	-1.33	1.42	-0.61	-3.14*	-6.24*
Spec 5:IGA_N	-	-	-	-	-	2.90*	1.94	3.98*	3.40*	1.22	-5.21*	1.85	2.52*	3.57*	0.00	3.16*	1.78	3.79*	3.46*	1.12	-5.16*
Spec 6:AGA_N	-	-	-	-	-	-	0.40	3.96*	1.76	-2.75*	-6.38*	-2.96*	0.97	2.21*	-2.91*	1.76	0.32	3.65*	1.82	-2.68*	-6.33*
Spec 7:$NAGA_N$	-	-	-	-	-	-	-	1.46	1.11	-1.46	-5.23*	-1.68	1.60	1.75	-1.93	-0.24	-0.53	1.50	1.48	-1.38	-5.17*
Spec 8:EGA_N	-	-	-	-	-	-	-	-	-1.57	-4.07*	-6.83*	-3.99*	-0.54	-0.91	-3.99*	-3.62*	-1.41	0.76	-1.14	-3.97*	-6.78*
Spec 9:$APGA_N$	-	-	-	-	-	-	-	-	-	-2.92*	-6.06*	-3.21*	0.16	3.22*	-3.39*	-1.41	-1.07	1.57	1.65	-2.73*	-5.99*
Spec 10:$FIGA_N$	-	-	-	-	-	-	-	-	-	-	-5.49*	-0.49	1.91	3.14*	-1.22	3.18*	1.33	3.85*	2.94*	-0.13	-5.42*
Spec 11:A_t	-	-	-	-	-	-	-	-	-	-	-	5.48*	5.41*	6.20*	5.21*	6.52*	5.10*	6.85*	6.07*	5.53*	1.50
Spec 12:GA_t	-	-	-	-	-	-	-	-	-	-	-	-	2.20*	3.39*	-1.87	3.27*	1.53	3.77*	3.23*	0.42	-5.42*
Spec 13:GJR_t	-	-	-	-	-	-	-	-	-	-	-	-	-	0.28	-2.51*	-0.81	-1.77	0.58	0.04	-1.82	-5.35*
Spec 14:TGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-3.57*	-1.85	-1.65	0.99	-1.87	-2.95*	-6.14*
Spec 15:IGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3.18*	1.78	3.80*	3.45*	1.13	-5.16*
Spec 16:AGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.18	3.36*	1.50	-3.17*	-6.48*
Spec 17:$NAGA_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.45	1.41	-1.26	-5.04*
Spec 18:EGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-1.18	-3.77*	-6.79*
Spec 19:$APGA_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-2.75*	-6.01*
Spec 20:$FIGA_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-5.47*
Spec 21:A_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 22:GA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 23:GJR_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 24:TGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 25:IGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 26:AGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 27:$NAGA_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 28:EGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 29:$APGA_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 30:$FIGA_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 31:$comb_1$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 32:$comb_2$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 33:$comb_3$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 34:HAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 35:logHAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 36:HEAVY	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 37:$GA-MIDAS$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 38:$GA-RV_N$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 39:$GA-RV_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 40:$GA-RV_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 41:$quant_1$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 42:$quant_2$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Note: This table reports the out-of-sample t-statistic of the Unconditional Giacomini and White (2006) test under the robust loss of Patton (2011) for b=-0.5 for FTSE 100. The period is from 02/01/2004 to 29/12/2017. A positive t-statistic indicates that average loss of model A (in row) is larger than of model B (in column) and hence model B is superior. *denotes significance at the 5% level

Table 3.26 cont. - Pairwise comparisons of unconditional Giacomini and White test for FTSE 100 (4-year in-sample) - robust Patton (b=-0.5)

	Spec 22	Spec 23	Spec 24	Spec 25	Spec 26	Spec 27	Spec 28	Spec 29	Spec 30	Spec 31	Spec 32	Spec 33	Spec 34	Spec 35	Spec 36	Spec 37	Spec 38	Spec 39	Spec 40	Spec 41	Spec 42
Spec 1:A_N	4.56*	5.34*	5.62*	4.38*	5.31*	5.00*	5.88*	5.62*	4.55*	6.04*	5.80*	6.01*	-6.35*	4.51*	-4.60*	-8.12*	1.31	1.34	-1.12	-2.74*	-2.87*
Spec 2:GA_N	0.31	2.22*	3.42*	-1.73	3.46*	1.57	3.88*	3.54*	0.46	5.57*	5.10*	5.39*	-7.95*	1.01	-6.26*	-9.94*	-3.27*	-3.56*	-3.77*	-4.61*	-4.67*
Spec 3:GJR_M	-2.19*	1.03	0.71	-2.55*	-0.67	-1.39	0.93	0.68	-1.83	1.19	0.78	1.19	-7.60*	-0.55	-6.01*	-9.39*	-2.25*	-2.45*	-2.94*	-4.49*	-4.55*
Spec 4:TGA_N	-3.43*	-0.05	0.73	-3.59*	-1.84	-1.01	1.94	0.64	-3.10*	2.81*	1.81	3.09*	-8.19*	-0.96	-6.55*	-10.19*	-5.23*	-5.54*	-5.23*	-4.96*	-5.02*
Spec 5:IGA_N	2.45*	2.45*	3.53*	1.11	3.45*	1.78	3.92*	3.65*	1.40	5.44*	4.96*	5.25*	-7.86*	1.24	-6.17*	-9.84*	-1.23	-1.34	-1.99*	-4.53*	-4.59*
Spec 6:AGA_N	-2.80*	0.99	2.09*	-2.81*	2.95*	0.41	3.69*	2.10*	-2.55*	4.97*	4.59*	4.94*	-8.33*	-0.06	-6.65*	-10.41*	-4.55*	-4.65*	-4.23*	-5.00*	-5.05*
Spec 7:$NAGA_N$	-1.61	1.62	2.31*	-1.91	-0.09	0.42	1.78	2.11*	-1.35	2.18*	1.69	2.21*	-7.59*	-0.24	-6.00*	-9.40*	-5.24*	-5.89*	-5.23*	-4.44*	-4.51*
Spec 8:EGA_N	-3.90*	-0.42	-0.51	-3.93*	-3.34*	-1.17	2.00*	-0.55	-3.91*	1.27	-0.06	1.35	-8.47*	-1.56	-6.82*	-10.55*	-6.23*	-6.24*	-6.78*	-5.22*	-5.26*
Spec 9:$APGA_N$	-3.10*	0.27	2.59*	-3.34*	-1.17	-0.75	1.97*	2.69*	-2.70*	2.95*	2.32*	3.11*	-8.07*	-0.72	-6.44*	-10.03*	-6.24*	-6.67*	-6.02*	-4.85*	-4.91*
Spec 10:$FIGA_N$	-0.24	1.88	3.03*	-1.05	3.46*	1.34	3.94*	3.13*	0.49	5.21*	4.62*	5.06*	-8.14*	1.07	-6.44*	-10.21*	-2.34*	-3.68*	-3.97*	-4.77*	-4.83*
Spec 11:A_t	5.54*	5.35*	6.15*	5.24*	6.66*	5.04*	6.95*	6.16*	5.61*	7.05*	6.82*	7.01*	-6.92*	6.07*	-5.08*	-8.92*	6.02*	6.23*	1.85	-3.09*	-3.22*
Spec 12:GA_t	2.83*	2.15*	3.30*	-1.45	3.58*	1.53	3.87*	3.41*	0.69	5.43*	5.00*	5.26*	-7.99*	1.08	-6.30*	-10.00*	-1.35	-3.79*	-3.64*	-4.65*	-4.71*
Spec 13:GJR_t	-2.15*	1.10	0.52	-2.50*	-0.70	-1.52	0.81	0.49	-1.81	1.02	0.63	1.02	-7.56*	-0.57	-5.98*	-9.33*	-4.73*	-4.77*	-5.15*	-4.46*	-4.53*
Spec 14:TGA_t	-3.30*	-0.13	1.09	-3.53*	-1.64	-1.25	1.47	0.84	-2.93*	2.11*	1.18	2.24*	-8.10*	-0.93	-6.47*	-10.06*	-6.89*	-6.23*	-6.78*	-4.89*	-4.95*
Spec 15:IGA_t	2.48*	2.44*	3.52*	1.46	3.47*	1.77	3.93*	3.64*	1.41	5.44*	4.98*	5.25*	-7.87*	1.24	-6.18*	-9.85*	-0.06	-3.06*	-3.13*	-4.53*	-4.60*
Spec 16:AGA_t	-3.11*	0.84	1.78	-3.07*	2.96*	0.27	3.43*	1.79	-3.02*	4.46*	3.95*	4.40*	-8.39*	-0.20	-6.71*	-10.50*	-5.78*	-4.63*	-4.69*	-5.06*	-5.12*
Spec 17:$NAGA_t$	-1.46	1.82	2.15*	-1.75	-0.04	1.23	1.71	1.99*	-1.24	2.04*	1.60	2.06*	-7.53*	-0.21	-5.94*	-9.32*	-3.78*	-3.57*	-2.98*	-4.40*	-4.46*
Spec 18:EGA_t	-3.69*	-0.46	-0.60	-3.74*	-3.12*	-1.21	2.62*	-0.63	-3.71*	0.91	-0.22	0.96	-8.47*	-1.60	-6.83*	-10.56*	-2.89*	-3.12*	-3.17*	-5.22*	-5.27*
Spec 19:$APGA_t$	-3.14*	0.09	2.76*	-3.42*	-1.30	-1.03	1.59	3.62*	-2.73*	2.33*	1.59	2.42*	-8.04*	-0.79	-6.41*	-9.98*	-2.88*	-3.82*	-2.54*	-4.83*	-4.89*
Spec 20:$FIGA_t$	-0.18	1.79	2.85*	-0.96	3.44*	1.28	3.85*	2.93*	1.23	4.95*	4.40*	4.81*	-8.22*	1.15	-6.51*	-10.31*	-5.68*	-4.96*	-3.26*	-4.83*	-4.89*
Spec 21:A_{st}	5.49*	5.29*	6.09*	5.19*	6.61*	4.98*	6.89*	6.10*	5.54*	6.99*	6.76*	6.95*	-6.94*	6.04*	-5.10*	-8.95*	2.78*	3.12*	0.76	-3.11*	-3.24*
Spec 22:GA_{st}	-	2.09*	3.22*	-2.00*	3.44*	1.47	3.80*	3.33*	0.44	5.34*	4.88*	5.16*	-8.00*	1.03	-6.31*	-10.01*	-1.54	-3.49*	-3.53*	-4.66*	-4.72*
Spec 23:GJR_{st}	-	-	0.32	-2.43*	-0.73	-1.66	0.67	0.29	-1.78	0.84	0.48	0.84	-7.51*	-0.60	-5.94*	-9.26*	-4.56*	-4.23*	-4.76*	-4.43*	-4.50*
Spec 24:TGA_{st}	-	-	-	-3.49*	-1.61	-1.64	1.04	-0.24	-2.84*	1.46	0.62	1.51	-8.03*	-0.95	-6.41*	-9.96*	-2.28*	-2.74*	-3.12*	-4.84*	-4.90*
Spec 25:IGA_{st}	-	-	-	-	3.37*	1.75	3.88*	3.62*	1.22	5.39*	4.92*	5.20*	-7.86*	1.20	-6.17*	-9.83*	-1.43	-2.76*	-2.97*	-4.53*	-4.59*
Spec 26:AGA_{st}	-	-	-	-	-	0.15	3.25*	1.61	-3.32*	4.30*	3.69*	4.22*	-8.39*	-0.34	-6.72*	-10.49*	-3.87*	-4.88*	-4.21*	-5.07*	-5.13*
Spec 27:$NAGA_{st}$	-	-	-	-	-	-	1.46	1.53	-1.26	1.71	1.31	1.73	-7.48*	-0.26	-5.90*	-9.25*	-2.29*	-2.58*	-3.05*	-4.36*	-4.43*
Spec 28:EGA_{st}	-	-	-	-	-	-	-	-1.05	-3.81*	0.20	-0.79	0.20	-8.46*	-1.75	-6.82*	-10.53*	-4.05*	-4.30*	-4.56*	-5.23*	-5.27*
Spec 29:$APGA_{st}$	-	-	-	-	-	-	-	-	-2.92*	1.54	0.69	1.58	-8.03*	-0.95	-6.41*	-9.97*	-3.06*	-3.18*	-3.54*	-4.85*	-4.90*
Spec 30:$FIGA_{st}$	-	-	-	-	-	-	-	-	-	4.98*	4.38*	4.82*	-8.19*	1.07	-6.49*	-10.27*	-3.30*	-3.34*	-4.08*	-4.81*	-4.87*
Spec 31:$comb_1$	-	-	-	-	-	-	-	-	-	-	-1.84	-0.09	-8.37*	-1.73	-6.73*	-10.40*	-4.56*	-4.89*	-5.10*	-5.16*	-5.21*
Spec 32:$comb_2$	-	-	-	-	-	-	-	-	-	-	-	1.93	-8.33*	-1.38	-6.69*	-10.35*	-4.45*	-4.72*	-4.89*	-5.10*	-5.15*
Spec 33:$comb_3$	-	-	-	-	-	-	-	-	-	-	-	-	-8.37*	-1.72	-6.73*	-10.40*	-4.07*	-4.14*	-4.57*	-5.16*	-5.20*
Spec 34:HAR	-	-	-	-	-	-	-	-	-	-	-	-	-	9.31*	5.58*	-2.70*	8.65*	8.98*	9.22*	9.86*	9.60*
Spec 35:logHAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-7.59*	-11.79*	-4.77*	-4.88*	-4.91*	-5.85*	-5.88*
Spec 36:HEAVY	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-8.71*	-1.56	-1.67	-2.12*	12.73*	13.49*
Spec 37:$GA-MIDAS$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-1.78	-2.37*	-2.98*	13.56*	13.16*
Spec 38:$GA-RV_N$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-1.95	-6.87*	-6.90*	-7.12*
Spec 39:$GA-RV_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-5.12*	-5.33*	-5.75*
Spec 40:$GA-RV_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-7.56*	-8.13*
Spec 41:$quant_1$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-6.57*
Spec 42:$quant_2$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Note: This table reports the out-of-sample t-statistic of the Unconditional Giacomini and White (2006) test under the robust loss of Patton (2011) for b=-0.5 for FTSE 100. The period is from 02/01/2004 to 29/12/2017. A positive t-statistic indicates that average loss of model A (in row) is larger than of model B (in column) and hence model B is superior. *denotes significance at the 5% level

Table 3.27 – Pairwise comparisons of unconditional Giacomini and White test for FTSE 100 (4-year in-sample) - all Loss Functions

	logMSE		sdMSE		propMSE		RMSE		propRMSE		MAE		logMAE		sdMAE		MAPE		GMLE	
	win	lose	win	lose	win	lose	win	lose	win	lose	win	lose	win	lose	win	lose	win	lose	win	lose
Spec 1:A_N	7	31	5	31	6	32	5	32	5	32	4	31	7	31	7	31	4	31	4	34
Spec 2:GA_N	11	25	11	22	12	25	11	25	11	25	11	22	11	26	11	24	11	22	11	25
Spec 3:GJR_M	20	11	19	9	20	13	20	14	20	14	21	13	20	11	22	13	21	13	20	18
Spec 4:TGA_N	29	5	27	3	31	5	31	5	31	5	31	7	29	4	30	4	31	7	28	7
Spec 5:IGA_N	10	28	9	28	10	28	10	28	10	28	10	28	10	28	10	28	10	28	10	28
Spec 6:AGA_N	17	20	17	14	18	19	18	19	18	19	19	8	17	19	17	13	19	8	17	20
Spec 7:$NAGA_N$	21	11	18	13	20	12	20	13	20	13	19	14	21	11	19	13	19	14	21	14
Spec 8:EGA_N	25	10	28	3	27	7	27	7	27	7	32	3	25	10	29	4	32	3	28	5
Spec 9:$APGA_N$	29	5	23	9	26	7	26	8	26	8	24	8	29	4	27	5	24	8	26	12
Spec 10:$FIGA_N$	16	22	11	22	16	22	16	22	16	22	11	21	15	22	15	22	11	21	16	22
Spec 11:A_t	5	32	5	31	4	33	3	35	3	35	3	32	5	32	4	32	3	32	3	35
Spec 12:GA_t	11	26	11	23	11	27	11	26	11	26	11	23	11	26	11	26	11	23	11	26
Spec 13:GJR_t	20	11	18	10	20	14	20	14	20	14	19	14	20	11	18	14	19	14	21	16
Spec 14:TGA_t	30	5	27	3	31	5	31	5	31	5	24	8	29	5	27	5	24	8	30	5
Spec 15:IGA_t	8	30	8	30	8	30	8	30	8	30	8	30	8	30	8	30	8	30	8	30
Spec 16:AGA_t	17	20	17	14	17	21	17	21	17	21	19	9	17	20	17	15	19	9	17	20
Spec 17:$NAGA_t$	20	13	17	17	20	14	20	14	20	14	17	20	20	13	17	17	17	20	22	14
Spec 18:EGA_t	20	11	24	4	24	10	26	10	26	10	32	5	20	11	26	5	32	5	28	5
Spec 19:$APGA_t$	29	7	23	9	26	8	26	8	26	8	23	12	29	6	26	10	23	12	27	11
Spec 20:$FIGA_t$	14	24	11	22	14	24	14	24	14	24	11	22	14	24	12	24	11	22	14	24
Spec 21:A_{st}	5	32	5	31	3	34	2	36	2	36	3	31	5	32	4	32	3	31	2	36
Spec 22:GA_{st}	12	25	12	22	12	25	12	25	12	25	12	22	13	25	12	24	12	22	12	25
Spec 23:GJR_{st}	22	10	18	7	21	13	22	13	22	13	19	14	22	10	18	14	19	14	22	14
Spec 24:TGA_{st}	33	4	28	3	34	4	34	4	34	4	24	8	31	4	27	4	24	8	31	4
Spec 25:IGA_{st}	9	29	9	28	9	29	9	29	9	29	9	29	9	29	9	29	9	29	9	29
Spec 26:AGA_{st}	19	19	19	13	18	19	18	19	18	19	21	8	18	19	18	13	21	8	19	19
Spec 27:$NAGA_{st}$	21	11	17	14	21	12	21	13	21	13	16	20	21	11	17	15	16	20	25	13
Spec 28:EGA_{st}	20	11	29	2	27	7	27	7	27	7	33	3	20	11	26	5	33	3	32	4
Spec 29:$APGA_{st}$	30	4	28	3	28	5	29	5	29	5	24	8	30	4	27	5	24	8	28	7
Spec 30:$FIGA_{st}$	15	23	11	22	15	23	15	23	15	23	11	22	15	22	15	22	11	22	15	23
Spec 31:$comb_1$	36	2	36	0	36	2	36	2	36	2	36	2	36	2	36	2	36	2	36	2
Spec 32:$comb_2$	35	3	34	2	35	3	35	3	35	3	33	3	35	3	35	3	33	3	35	3
Spec 33:$comb_3$	37	1	36	0	37	1	37	1	37	1	37	1	37	1	37	1	37	1	37	1
Spec 34:HAR	1	37	1	37	1	37	1	37	1	37	1	37	1	37	1	37	1	37	1	37
Spec 35:logHAR	38	0	23	0	38	0	38	0	38	0	38	0	38	0	38	0	38	0	38	0
Spec 36:HEAVY	2	36	2	36	2	36	4	32	4	32	2	36	2	36	2	36	2	36	6	31
Spec 37:$GA-MIDAS$	0	38	0	38	0	38	0	38	0	38	0	38	0	38	0	38	0	38	0	38
Spec 38:$GA-RV_N$	2	34	7	28	16	23	15	22	16	20	13	26	7	34	10	23	12	25	28	5
Spec 39:$GA-RV_t$	7	33	9	28	13	26	16	21	17	22	9	25	7	28	11	21	9	25	26	7
Spec 40:$GA-RV_{st}$	8	33	5	34	40	1	38	0	38	2	18	10	5	35	6	23	17	11	40	1
Spec 41:$quant_1$	3	35	3	35	3	33	4	33	4	33	3	32	3	35	3	35	3	32	5	33
Spec 42:$quant_2$	4	34	4	34	7	31	7	31	7	31	4	31	4	34	4	32	4	31	6	31

Note: This table reports a summary of the out-of-sample results for the Unconditional Giacomini and White (2006) performed under the logMSE, sdMSE, propMSE, RMSE, propRMSE, MAE, logMAE, sdMAE, MAPE and GMLE loss functions for FTSE 100. The period is from 02/01/2004 to 29/12/2017. The "win" column reports the cases where a certain specification outperformed significantly the competing specifications. On the contrary, the "lose" column reports the cases where a certain specification was outperformed significantly by the competing specifications. Adding the two columns for each specification and for each loss function, respectively, shows how many of the 41 in total pairwise comparisons where statistically significant. For the hypothesis testing the 5% significance level is selected.

Table 3.28 – Pairwise comparisons of unconditional Giacomini and White test for DAX - (4-year in-sample) robust Patton (b=1)

	Spec 1	Spec 2	Spec 3	Spec 4	Spec 5	Spec 6	Spec 7	Spec 8	Spec 9	Spec 10	Spec 11	Spec12	Spec 13	Spec 14	Spec 15	Spec 16	Spec 17	Spec 18	Spec 19	Spec 20	Spec 21
Spec 1:A_N	-	0.82	0.51	0.52	0.57	0.91	0.73	0.14	0.45	0.00	0.81	0.36	0.44	0.56	0.26	0.81	0.68	0.14	0.53	-0.10	1.10
Spec 2:GA_N	-	-	0.09	0.12	-1.37	0.73	0.44	-0.46	0.01	-1.34	-0.63	-1.81	-0.04	0.18	-3.52*	0.49	0.35	-0.44	0.13	-1.50	-0.36
Spec 3:GJR_M	-	-	-	0.24	-0.30	0.35	1.05	-2.28*	-0.46	-1.31	-0.33	-0.57	-1.51	0.60	-0.65	0.22	0.75	-2.30*	0.26	-1.78	-0.21
Spec 4:TGA_N	-	-	-	-	-0.32	0.27	1.04	-2.44*	-2.47*	-1.28	-0.34	-0.59	-1.02	1.31	-0.66	0.15	0.72	-2.58*	0.27	-1.70	-0.23
Spec 5:IGA_N	-	-	-	-	-	0.94	0.63	-0.21	0.22	-0.75	-0.17	-0.36	0.18	0.38	-0.97	0.76	0.55	-0.20	0.33	-0.90	0.05
Spec 6:AGA_N	-	-	-	-	-	-	0.17	-1.15	-0.44	-2.66*	-0.84	-2.58*	-0.61	-0.19	-2.20*	-2.09*	0.05	-1.11	-0.25	-3.27*	-0.62
Spec 7:$NAGA_N$	-	-	-	-	-	-	-	-1.83	-1.34	-1.69	-0.59	-1.00	-1.40	-0.97	-1.06	-0.32	-1.39	-1.85	-1.10	-2.08*	-0.46
Spec 8:EGA_N	-	-	-	-	-	-	-	-	2.15*	-0.34	0.11	0.14	1.99*	2.36*	0.01	1.08	1.62	0.16	2.23*	-0.64	0.20
Spec 9:$APGA_N$	-	-	-	-	-	-	-	-	-	-1.10	-0.26	-0.45	-0.29	2.28*	-0.53	0.33	1.03	-2.30*	2.25*	-1.50	-0.15
Spec 10:$FIGA_N$	-	-	-	-	-	-	-	-	-	-	0.42	0.89	1.16	1.39	0.54	2.52*	1.50	0.35	1.29	-0.95	0.52
Spec 11:A_t	-	-	-	-	-	-	-	-	-	-	-	-0.05	0.23	0.39	-0.24	0.70	0.52	-0.10	0.35	-0.55	1.19
Spec 12:GA_t	-	-	-	-	-	-	-	-	-	-	-	-	0.43	0.67	-1.00	2.12*	0.87	-0.12	0.60	-1.24	0.23
Spec 13:GJR_t	-	-	-	-	-	-	-	-	-	-	-	-	-	1.35	-0.53	0.48	1.08	-1.98*	0.95	-1.66	-0.12
Spec 14:TGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-0.74	0.05	0.60	-2.46*	-1.32	-1.84	-0.28
Spec 15:IGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.86	0.94	-0.00	0.67	-0.81	0.42
Spec 16:AGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.19	-1.04	-0.13	-3.26*	-0.50
Spec 17:$NAGA_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-1.64	-0.75	-1.86	-0.40
Spec 18:EGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2.35*	-0.65	0.19
Spec 19:$APGA_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-1.70	-0.24
Spec 20:$FIGA_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.63
Spec 21:A_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 22:GA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 23:GJR_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 24:TGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 25:IGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 26:AGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 27:$NAGA_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 28:EGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 29:$APGA_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 30:$FIGA_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 31:$comb_1$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 32:$comb_2$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 33:$comb_3$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 34:HAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 35:logHAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 36:HEAVY	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 37:$GA-MIDAS$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 38:$GA-RV_N$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 39:$GA-RV_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 40:$GA-RV_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 41:$quant_1$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 42:$quant_2$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Note: This table reports the out-of-sample t-statistic of the Unconditional Giacomini and White (2006) test under the robust loss of Patton (2011) for b=1 for DAX. The period is from 02/01/2004 to 29/12/2017. A positive t-statistic indicates that average loss of model A (in row) is larger than of model B (in column) and hence model B is superior.
*denotes significance at the 5% level

Table 3.29 cont. - Pairwise comparisons of unconditional Giacomini and White test for DAX (4-year in-sample) - robust Patton (b=1)

	Spec 22	Spec 23	Spec 24	Spec 25	Spec 26	Spec 27	Spec 28	Spec 29	Spec 30	Spec 31	Spec 32	Spec 33	Spec 34	Spec 35	Spec 36	Spec 37	Spec 38	Spec 39	Spec 40	Spec 41	Spec 42
Spec 1:A_N	0.36	0.47	0.60	0.30	0.90	0.68	0.16	0.50	-0.11	3.26*	3.28*	3.28*	-2.67*	3.28*	-1.28	-5.22*	1.49	1.56	1.70	-1.05	-1.03
Spec 2:GA_N	-2.37*	0.00	0.24	-3.94*	0.73	0.36	-0.42	0.09	-1.73	2.84*	2.76*	2.76*	-2.92*	2.76*	-2.47*	-6.37*	-0.98	-1.42	-1.48	-2.17*	-2.13*
Spec 3:GJR_M	-0.55	-1.23	0.94	-0.59	0.32	0.71	-2.23*	0.06	-1.63	1.87	1.85	1.85	-3.00*	1.85	-4.73*	-7.29*	-3.22*	-3.32*	-1.90	-4.39*	-4.30*
Spec 4:TGA_N	-0.57	-0.78	1.87	-0.60	0.25	0.68	-2.53*	-0.28	-1.57	1.86	1.84	1.84	-3.01*	1.84	-4.71*	-7.25*	-4.70*	-4.56*	-4.62*	-4.40*	-4.32*
Spec 5:IGA_N	-0.45	0.23	0.44	-0.97	0.95	0.56	-0.17	0.29	-1.03	3.23*	3.14*	3.14*	-2.85*	3.14*	-2.05*	-6.06*	-0.90	-1.14	-1.18	-1.77	-1.73
Spec 6:AGA_N	-2.23*	-0.54	-0.09	-1.94	-0.66	0.07	-1.06	-0.29	-3.50*	2.22*	2.17*	2.17*	-3.02*	2.17*	-3.69*	-7.01*	-3.77*	-3.78*	-2.49*	-3.36*	-3.30*
Spec 7:$NAGA_N$	-0.96	-1.33	-0.75	-0.99	-0.19	-0.61	-1.81	-1.12	-2.00*	1.90	1.87	1.87	-3.04*	1.87	-3.99*	-7.15*	-1.91	-2.59*	-1.95	-3.65*	-3.60*
Spec 8:EGA_N	0.12	2.02*	2.36*	0.04	1.10	1.53	0.49	2.02*	-0.58	2.02*	1.99*	1.99*	-2.89*	1.99*	-5.35*	-7.16*	-4.75*	-4.43*	-4.02*	-5.22*	-5.09*
Spec 9:$APGA_N$	-0.44	-0.05	2.43*	-0.48	0.41	0.96	-2.23*	1.24	-1.38	1.91	1.88	1.88	-2.99*	1.88	-4.70*	-7.22*	-5.18*	-5.44*	1.20	-4.41*	-4.33*
Spec 10:$FIGA_N$	0.82	1.24	1.47	0.58	2.66*	1.47	0.38	1.18	-1.06	2.63*	2.56*	2.56*	-2.81*	2.56*	-2.61*	-6.54*	-0.10	-0.60	-0.42	-2.22*	-2.17*
Spec 11:A_t	-0.06	0.27	0.44	-0.18	0.84	0.53	-0.08	0.32	-0.59	3.11*	3.07*	3.07*	-2.81*	3.07*	-1.84	-5.86*	3.01*	1.79	2.093*	-1.57	-1.54
Spec 12:GA_t	-0.14	0.49	0.74	-0.53	2.72*	0.86	-0.09	0.54	-1.50	2.84*	2.75*	2.75*	-2.87*	2.75*	-2.55*	-6.48*	-0.53	-0.76	-0.99	-2.21*	-2.16*
Spec 13:GJR_t	-0.42	1.74	1.51	-0.47	0.57	1.00	-1.86	0.62	-1.50	1.95	1.92	1.92	-2.98*	1.92	-4.82*	-7.27*	-1.26	-1.74	-1.77	-4.47*	-4.38*
Spec 14:TGA_t	-0.65	-1.14	1.89	-0.68	0.17	0.57	-2.40*	-0.93	-1.70	1.87	1.84	1.84	-3.01*	1.84	-4.62*	-7.25*	-5.75*	-5.90*	-4.98*	-4.30*	-4.23*
Spec 15:IGA_t	1.58	0.58	0.81	0.91	2.29*	0.94	0.02	0.61	-0.99	3.07*	2.97*	2.97*	-2.83*	2.97*	-2.24*	-6.29*	-0.04	-0.41	-0.59	-1.92	-1.88
Spec 16:AGA_t	-1.84	-0.41	0.04	-1.63	1.34	0.20	-0.99	-0.17	-3.43*	2.23*	2.18*	2.18*	-3.01*	2.18*	-3.73*	-7.02*	-3.45*	-3.99*	-1.94	-3.40*	-3.34*
Spec 17:$NAGA_t$	-0.84	-1.00	-0.38	-0.88	-0.07	0.31	-1.60	-0.82	-1.80	1.93	1.90	1.90	-3.03*	1.90	-3.84*	-7.10*	-1.21	-1.56	-1.70	-3.50*	-3.44*
Spec 18:EGA_t	0.11	2.01*	2.45*	0.03	1.07	1.54	1.70	2.14*	-0.59	2.01*	1.98*	1.98*	-2.89*	1.98*	-5.27*	-7.15*	-5.16*	-5.34*	-5.12*	-5.15*	-5.02*
Spec 19:$APGA_t$	-0.58	-0.74	2.36*	-0.62	0.23	0.70	-2.29*	-0.58	-1.58	1.87	1.85	1.85	-3.01*	1.85	-4.59*	-7.23*	-0.98	-1.65	-1.02	-4.27*	-4.20*
Spec 20:$FIGA_t$	1.12	1.74	1.90	0.83	3.24*	1.80	0.68	1.55	-0.09	2.64*	2.57*	2.57*	-2.79*	2.57*	-2.76*	-6.60*	1.25	-1.17	1.10	-2.35*	-2.28*
Spec 21:A_{st}	-0.25	0.15	0.32	-0.37	0.61	0.41	-0.17	0.21	-0.68	3.07*	3.05*	3.05*	-2.82*	3.05*	-1.84	-5.81*	1.86*	1.70	1.52	-1.58	-1.55
Spec 22:GA_{st}	-	0.47	0.72	-0.72	2.33*	0.84	-0.08	0.52	-1.35	2.91*	2.82*	2.82*	-2.86*	2.82*	-2.44*	-6.41*	-1.54	-1.91	-1.33	-2.11*	-2.06*
Spec 23:GJR_{st}	-	-	1.36	-0.52	0.50	0.93	-1.90	0.45	-1.58	1.94	1.91	1.91	-2.98*	1.91	-4.74*	-7.27*	-1.45	-1.90	-1.77	-4.39*	-4.31*
Spec 24:TGA_{st}	-	-	-	-0.75	0.06	0.38	-2.40*	-1.79	-1.77	1.84	1.82	1.82	-3.02*	1.82	-4.54*	-7.25*	-4.21*	-4.43*	-4.76	-4.22*	-4.14*
Spec 25:IGA_{st}	-	-	-	-	2.00*	0.88	-0.00	0.56	-1.00	3.10*	3.00*	3.00*	-2.83*	3.00*	-2.21*	-6.26*	-0.09	-0.17	-0.22	-1.89	-1.85
Spec 26:AGA_{st}	-	-	-	-	-	0.09	-1.02	-0.26	-3.52*	2.25*	2.19*	2.19*	-3.02*	2.19*	-3.62*	-6.99*	-3.20*	-3.58*	-1.80	-3.29*	-3.24*
Spec 27:$NAGA_{st}$	-	-	-	-	-	-	-1.50	-0.78	-1.75	1.94	1.91	1.91	-3.03*	1.91	-3.70*	-7.05*	-0.31	-0.33	-0.21	-3.36*	-3.30*
Spec 28:EGA_{st}	-	-	-	-	-	-	-	2.10*	-0.62	1.99*	1.96*	1.96*	-2.90*	1.96*	-5.24*	-7.16*	-2.98*	-2.79*	-3.12*	-5.10*	-4.98*
Spec 29:$APGA_{st}$	-	-	-	-	-	-	-	-	-1.44	1.88	1.85	1.85	-3.00*	1.85	-4.46*	-7.20*	-1.55	-2.10*	-1.55	-4.14*	-4.06*
Spec 30:$FIGA_{st}$	-	-	-	-	-	-	-	-	-	2.75*	2.67*	2.67*	-2.78*	2.67*	-2.53*	-6.51*	0.27	-0.32	0.29	-2.13*	-2.07*
Spec 31:$comb_1$	-	-	-	-	-	-	-	-	-	-	1.48	1.48	-3.36*	1.48	-3.02*	-6.33*	-4.72*	-4.99*	-5.01*	-2.83*	-2.80*
Spec 32:$comb_2$	-	-	-	-	-	-	-	-	-	-	-	-0.55	-3.36*	-	-2.95*	-6.21*	-2.99*	-3.14*	-3.30*	-2.76*	-2.74*
Spec 33:$comb_3$	-	-	-	-	-	-	-	-	-	-	-	-	-3.36*	0.55	-2.95*	-6.21*	-5.20*	-5.19*	-5.32*	-2.76*	-2.74*
Spec 34:HAR	-	-	-	-	-	-	-	-	-	-	-	-	-	3.36*	2.38*	-0.00	2.38*	2.71*	2.89*	2.47*	2.48*
Spec 35:logHAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-2.95*	-6.21*	-2.28*	-2.71*	-2.90*	-2.76*	-2.74*
Spec 36:HEAVY	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-6.08*	-4.71*	-4.77*	-4.94*	1.58	1.75
Spec 37:$GA-MIDAS$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-0.99	-1.22	-1.03	6.15*	6.18*
Spec 38:$GA-RV_N$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-4.09*	-0.07	-1.22	-1.23
Spec 39:$GA-RV_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.41	-1.74	-1.31
Spec 40:$GA-RV_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-6.32*	-6.80*
Spec 41:$quant_1$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	7.23*
Spec 42:$quant_2$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Note: This table reports the out-of-sample t-statistic of the Unconditional Giacomini and White (2006) test under the robust loss of Patton (2011) for b=1 for DAX. The period is from 02/01/2004 to 29/12/2017. A positive t-statistic indicates that average loss of model A (in row) is larger than of model B (in column) and hence model B is superior. *denotes significance at the 5% level

Table 3.29 – Pairwise comparisons of unconditional Giacomini and White test for DAX (4-year in-sample) - robust Patton (b=-0.5)

	Spec 1	Spec 2	Spec 3	Spec 4	Spec 5	Spec 6	Spec 7	Spec 8	Spec 9	Spec 10	Spec 11	Spec12	Spec 13	Spec 14	Spec 15	Spec 16	Spec 17	Spec 18	Spec 19	Spec 20	Spec 21
Spec 1:A_N	-	6.66*	4.94*	4.65*	6.03*	7.35*	5.13*	3.59*	4.42*	4.91*	-1.53	5.57*	4.89*	4.73*	5.41*	7.08*	4.87*	3.47*	4.54*	4.28*	-0.86
Spec 2:GA_N	-	-	1.00	0.86	-1.60	3.20*	1.45	-0.35	0.60	-2.06*	-7.49*	-3.00*	0.77	0.93	-4.44*	2.66*	1.18	-0.33	0.76	-2.83*	-7.25*
Spec 3:GJR_M	-	-	-	-0.27	-1.21	0.02	0.85	-2.75*	-0.93	-2.62*	-5.38*	-1.69	-1.75	-0.14	-1.78	-0.15	0.33	-2.57*	-0.53	-3.10*	-5.16*
Spec 4:TGA_N	-	-	-	-	-1.07	0.12	1.24	-3.03*	-2.78*	-2.27*	-5.12*	-1.55	-0.78	0.37	-1.63	-0.04	0.65	-3.02*	-1.23	-2.83*	-4.91*
Spec 5:IGA_N	-	-	-	-	-	3.07*	1.72	-0.14	0.81	-1.31	-6.64*	-1.32	1.00	1.15	-2.43*	2.64*	1.44	-0.12	0.97	-2.01*	-6.43*
Spec 6:AGA_N	-	-	-	-	-	-	0.46	-1.47	-0.44	-5.33*	-8.35*	-4.90*	-0.46	-0.08	-5.00*	-2.13*	0.16	-1.39	-0.26	-5.89*	-8.03*
Spec 7:$NAGA_N$	-	-	-	-	-	-	-	-2.35*	-1.74	-2.93*	-5.52*	-2.16*	-1.62	-1.36	-2.32*	-0.65	-3.14*	-2.30*	-1.63	-3.38*	-5.31*
Spec 8:EGA_N	-	-	-	-	-	-	-	-	2.55*	-0.51	-4.09*	-0.20	2.43*	2.75*	-0.27	1.36	2.09*	0.09	2.54*	-1.06	-3.90*
Spec 9:$APGA_N$	-	-	-	-	-	-	-	-	-	-1.89	-4.90*	-1.28	0.18	2.37*	-1.36	0.28	1.25	-2.63*	1.88	-2.48*	-4.69*
Spec 10:$FIGA_N$	-	-	-	-	-	-	-	-	-	-	-5.66*	0.57	2.47*	2.38*	0.35	5.09*	2.54*	0.50	2.11*	-2.46*	-5.36*
Spec 11:A_t	-	-	-	-	-	-	-	-	-	-	-	6.61*	5.42*	5.21*	6.26*	8.14*	5.26*	3.95*	5.00*	5.11*	2.96*
Spec 12:GA_t	-	-	-	-	-	-	-	-	-	-	-	-	1.56	1.65	-0.74	4.91*	1.87	0.20	1.45	-1.83	-6.37*
Spec 13:GJR_t	-	-	-	-	-	-	-	-	-	-	-	-	-	0.88	-1.65	0.26	1.06	-2.25*	0.30	-3.13*	-5.19*
Spec 14:TGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-1.74	-0.09	0.66	-2.73*	-2.25*	-2.97*	-4.99*
Spec 15:IGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	4.82*	2.02*	0.27	1.54	-1.45	-6.04*
Spec 16:AGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.34	-1.29	-0.09	-6.09*	-7.82*
Spec 17:$NAGA_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-2.05*	-1.04	-3.01*	-5.06*
Spec 18:EGA_t	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2.58*	-1.03	-3.77*
Spec 19:$APGA_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-2.69*	-4.80*
Spec 20:$FIGA_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-4.83*
Spec 21:A_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 22:GA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 23:GJR_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 24:TGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 25:IGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 26:AGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 27:$NAGA_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 28:EGA_{st}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 29:$APGA_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 30:$FIGA_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 31:$comb_1$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 32:$comb_2$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 33:$comb_3$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 34:HAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 35:logHAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 36:HEAVY	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 37:$GA-MIDAS$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 38:$GA-RV_N$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 39:$GA-RV_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 40:$GA-RV_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 41:$quant_1$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Spec 42:$quant_2$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Note: This table reports the out-of-sample t-statistic of the Unconditional Giacomini and White (2006) test under the robust loss of Patton (2011) for b=-0.5 for DAX. The period is from 02/01/2004 to 29/12/2017. A positive t-statistic indicates that average loss of model A (in row) is larger than of model B (in column) and hence model B is superior.
*denotes significance at the 5% level

Table 3.29 cont. - Pairwise comparisons of unconditional Giacomini and White test for DAX (4-year in-sample) - robust Patton (b=-0.5)

	Spec 22	Spec 23	Spec 24	Spec 25	Spec 26	Spec 27	Spec 28	Spec 29	Spec 30	Spec 31	Spec 32	Spec 33	Spec 34	Spec 35	Spec 36	Spec 37	Spec 38	Spec 39	Spec 40	Spec 41	Spec 42
Spec 1:A_N	5.70*	4.96*	4.79*	5.55*	7.32*	4.86*	3.49*	4.29*	4.44*	8.15*	8.07*	8.07*	-7.57*	8.07*	5.98*	8.48*	1.01	0.95	-1.12	-4.90	-4.68*
Spec 2:GA_N	-3.38*	0.87	1.04	-4.54*	3.37*	1.20	-0.25	0.57	-3.14*	5.27*	5.16*	5.16*	-9.25*	5.16*	-7.72*	-10.51*	-3.56*	-3.48*	-3.49*	-6.96*	-6.69*
Spec 3:GJR_M	-1.61	-1.48	0.21	-1.67	0.03	0.41	-2.42*	-0.82	-2.88*	3.62*	3.53*	3.53*	-10.08*	3.53*	-8.52*	-11.55*	-5.50*	-5.69*	-6.12*	-7.81*	-7.49*
Spec 4:TGA_N	-1.47	-0.43	1.50	-1.51	0.13	0.70	-2.79*	-1.90	-2.63*	3.69*	3.61*	3.61*	-10.16*	3.61*	-8.61*	-11.62*	-5.17*	-4.76*	-4.80*	-7.92*	-7.59*
Spec 5:IGA_N	-1.29	1.10	1.26	-2.34*	3.23*	1.47	-0.05	0.78	-2.20*	5.93*	5.80*	5.80*	-8.96*	5.80*	-7.45*	-10.21*	-1.12	-1.90	-3.19*	-6.67*	-6.41*
Spec 6:AGA_N	-4.76*	-0.33	0.06	-4.70*	0.12	0.21	-1.28	-0.41	-6.19*	4.42*	4.30*	4.30*	-9.81*	4.30*	-8.27*	-11.14*	-2.45*	-2.74*	-3.59*	-7.57*	-7.27*
Spec 7:$NAGA_N$	-2.08*	-1.48	-1.05	-2.20*	-0.45	-1.86	-2.16*	-1.77	-3.28*	4.28*	4.16*	4.16*	-9.65*	4.16*	-8.13*	-11.02*	-4.13*	-4.18*	-4.13*	-7.42*	-7.14*
Spec 8:EGA_N	-0.13	2.45*	2.82*	-0.17	1.45	2.05*	0.69	2.27*	-0.92	3.66*	3.60*	3.60*	-10.80*	3.60*	-9.20*	-12.44*	-6.19*	-6.34*	-5.74*	-8.49*	-8.08*
Spec 9:$APGA_N$	-1.20	0.43	2.72*	-1.24	0.44	1.24	-2.36*	-0.00	-2.29*	3.78*	3.70*	3.70*	-10.20*	3.70*	-8.64*	-11.66*	-4.12*	-4.67*	-4.33*	-7.96*	-7.63*
Spec 10:$FIGA_N$	0.73	2.60*	2.50*	0.61	5.41*	2.52*	0.59	1.77	-2.09*	5.47*	5.34*	5.34*	-9.48*	5.34*	-7.91*	-10.84*	-2.76*	-2.90*	-4.34*	-7.14*	-6.84*
Spec 11:A_t	6.69*	5.47*	5.25*	6.37*	8.31*	5.22*	3.96*	4.74*	5.27*	8.12*	8.03*	8.03*	-7.78*	8.03*	-6.15*	-9.10*	5.36*	6.33*	1.20	-5.03*	-4.79*
Spec 12:GA_t	1.35	1.65	1.75	0.13	5.17*	1.86	0.27	1.22	-1.95	5.26*	5.16*	5.16*	-9.42*	5.16*	-7.87*	-10.73*	-1.22	-4.75*	-4.40*	-7.10*	-6.80*
Spec 13:GJR_t	-1.46	1.99*	1.26	-1.51	0.46	1.07	-2.04*	-0.15	-2.87*	3.86*	3.77*	3.77*	-10.17*	3.77*	-8.59*	-11.65*	-2.85*	-2.87*	-2.90*	-7.89*	-7.55*
Spec 14:TGA_t	-1.56	-0.54	2.54*	-1.62	0.09	0.72	-2.51*	-2.07*	-2.78*	3.84*	3.76*	3.76*	-10.06*	3.76*	-8.52*	-11.50*	-3.35*	-3.39*	-4.12*	-7.82*	-7.50*
Spec 15:IGA_t	1.81	1.75	1.84	2.60*	5.33*	2.02*	0.34	1.30	-1.55	5.72*	5.60*	5.60*	-9.19*	5.60*	-7.65*	-10.47*	-1.20	-3.55*	-2.43*	-6.86*	-6.58*
Spec 16:AGA_t	-4.63*	-0.12	0.24	-4.41*	2.35*	0.39	-1.18	-0.26	-6.37*	4.40*	4.29*	4.29*	-9.86*	4.29*	-8.32*	-11.21*	-4.75*	-4.61*	-5.09*	-7.62*	-7.31*
Spec 17:$NAGA_t$	-1.80	-0.88	-0.27	-1.90	-0.15	0.86	-1.92	-1.29	-2.91*	4.39*	4.27*	4.27*	-9.61*	4.27*	-8.08*	-10.98*	-2.51*	-3.87*	-3.21*	-7.36*	-7.07*
Spec 18:EGA_t	-0.14	2.28*	2.79*	-0.18	1.38	2.02*	2.26*	2.36*	-0.89	3.58*	3.52*	3.52*	-10.82*	3.52*	-9.23*	-12.45*	-4.12*	-4.75*	-4.90*	-8.55*	-8.14*
Spec 19:$APGA_t$	-1.37	-0.01	3.00*	-1.42	0.26	1.05	-2.35*	-1.66	-2.51*	3.84*	3.76*	3.76*	-10.08*	3.76*	-8.53*	-11.52*	-5.32*	-5.13*	-6.94*	-7.84*	-7.51*
Spec 20:$FIGA_t$	1.93	3.21*	3.05*	1.65	6.01*	2.95*	1.10	2.32*	0.81	5.42*	5.30*	5.30*	-9.63*	5.30*	-8.03*	-11.00*	-6.39*	-6.75*	-6.01*	-7.25*	-6.93*
Spec 21:A_{st}	6.45*	5.24*	5.04*	6.14*	7.99*	5.03*	3.78*	4.54*	4.99*	7.95*	7.87*	7.87*	-7.79*	7.87*	-6.16*	-9.10*	5.30*	4.44*	0.99	-5.06*	-4.82*
Spec 22:GA_{st}	-	1.56	1.66	-0.40	5.03*	1.80	0.21	1.14	-2.10*	5.32*	5.22*	5.22*	-9.36*	5.22*	-7.81*	-10.65*	-1.73	-3.00*	-3.54*	-7.04*	-6.75*
Spec 23:GJR_{st}	-	-	0.95	-1.61	0.33	0.92	-2.09*	-0.37	-2.96*	3.87*	3.77*	3.77*	-10.10*	3.77*	-8.53*	-11.58*	-3.42*	-3.78*	-4.12*	-7.83*	-7.50*
Spec 24:TGA_{st}	-	-	-	-1.72	-0.06	0.39	-2.60*	-2.68*	-2.87*	3.82*	3.74*	3.74*	-10.00*	3.74*	-8.46*	-11.43*	-5.13*	-5.82*	-6.33*	-7.77*	-7.45*
Spec 25:IGA_{st}	-	-	-	-	5.01*	1.91	0.25	1.18	-1.79	5.74*	5.62*	5.62*	-9.15*	5.62*	-7.62*	-10.43*	-1.51	-2.48*	-3.69*	-6.83*	-6.55*
Spec 26:AGA_{st}	-	-	-	-	-	0.21	-1.27	-0.41	-6.39*	4.44*	4.33*	4.33*	-9.78*	4.33*	-8.25*	-11.12*	-5.99*	-4.63*	-5.10*	-7.55*	-7.25*
Spec 27:$NAGA_{st}$	-	-	-	-	-	-	-1.89	-1.30	-2.86*	4.44*	4.31*	4.31*	-9.54*	4.31*	-8.02*	-10.90*	-4.88*	-4.32*	-4.90*	-7.29*	-7.01*
Spec 28:EGA_{st}	-	-	-	-	-	-	-	2.13*	-0.96	3.50*	3.45*	3.45*	-10.82*	3.45*	-9.23*	-12.44*	-2.89*	-3.15*	-3.73*	-8.55*	-8.14*
Spec 29:$APGA_{st}$	-	-	-	-	-	-	-	-	-2.16*	3.83*	3.75*	3.75*	-10.07*	3.75*	-8.52*	-11.52*	-4.02*	-4.35*	-4.69*	-7.83*	-7.50*
Spec 30:$FIGA_{st}$	-	-	-	-	-	-	-	-	-	5.60*	5.48*	5.48*	-9.49*	5.48*	-7.91*	-10.84*	-3.30*	-3.76*	-4.78*	-7.11*	-6.81*
Spec 31:$comb_1$	-	-	-	-	-	-	-	-	-	-	1.26	1.26	-9.12*	1.26	-7.75*	-10.26*	-2.56*	-2.57*	-3.03*	-7.11*	-6.89*
Spec 32:$comb_2$	-	-	-	-	-	-	-	-	-	-	-	-0.12	-9.07*	-	-7.70*	-10.20*	-6.29*	-5.91*	-6.72*	-7.06*	-6.84*
Spec 33:$comb_3$	-	-	-	-	-	-	-	-	-	-	-	-	-9.07*	0.12	-7.70*	-10.20*	-6.23*	-6.34*	-6.82*	-7.06*	-6.84*
Spec 34:HAR	-	-	-	-	-	-	-	-	-	-	-	-	-	9.07*	5.17*	-3.02*	8.42*	8.11*	9.78*	8.93*	9.35*
Spec 35:logHAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-7.70*	-10.20*	-5.23*	-5.91*	-6.08*	-7.06*	-6.84*
Spec 36:HEAVY	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-7.59*	-1.25	-1.37	-2.56*	8.06*	8.62*
Spec 37:$GA-MIDAS$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-1.86	-4.73*	-4.90*	10.50*	10.73*
Spec 38:$GA-RV_N$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-1.70	-6.71*	-7.03*	-7.79*
Spec 39:$GA-RV_t$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-5.12*	-6.89*	-8.16*
Spec 40:$GA-RV_{st}$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-8.90*	-9.67*
spec 41:$quant_1$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2.14*
Spec 42:$quant_2$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Note: This table reports the out-of-sample t-statistic of the Unconditional Giacomini and White (2006) test under the robust loss of Patton (2011) for b=-0.5 for DAX. The period is from 02/01/2004 to 29/12/2017. A positive t-statistic indicates that average loss of model A (in row) is larger than of model B (in column) and hence model B is superior. *denotes significance at the 5% level

Table 3.30 – Pairwise comparisons of unconditional Giacomini and White test for DAX (4-year in-sample) - all Loss Functions

	logMSE		sdMSE		propMSE		RMSE		propRMSE		MAE		logMAE		sdMAE		MAPE		GMLE	
	win	lose	win	lose	win	lose	win	lose	win	lose	win	lose	win	lose	win	lose	win	lose	win	lose
Spec 1:A_N	7	31	7	31	5	32	4	33	4	33	6	31	6	31	6	31	6	31	3	34
Spec 2:GA_N	15	22	16	19	13	22	15	22	15	22	16	22	15	22	16	22	16	22	14	23
Spec 3:GJR_M	17	11	22	4	17	16	17	15	17	15	21	4	17	9	22	5	21	4	19	17
Spec 4:TGA_N	31	5	24	4	28	5	29	5	29	5	30	4	32	4	34	4	30	4	27	7
Spec 5:IGA_N	12	25	12	20	8	25	9	27	9	27	9	25	10	25	10	25	9	25	8	29
Spec 6:AGA_N	17	11	19	4	17	17	17	18	17	18	18	4	17	11	20	5	18	4	17	20
Spec 7:$NAGA_N$	18	11	22	4	21	13	22	13	22	13	19	5	18	10	19	8	19	5	22	14
Spec 8:EGA_N	17	18	13	17	22	12	25	10	25	10	17	8	17	13	17	13	17	8	29	7
Spec 9:$APGA_N$	25	8	20	6	25	9	26	10	26	10	25	5	26	7	26	5	25	5	26	12
Spec 10:$FIGA_N$	8	28	10	25	8	27	9	26	9	26	10	25	8	26	9	25	10	25	10	26
Spec 11:A_t	4	33	5	33	3	34	2	35	2	35	4	33	4	34	4	33	4	33	0	37
Spec 12:GA_t	14	24	13	21	13	23	14	24	14	24	14	23	14	24	14	24	14	23	10	25
Spec 13:GJR_t	18	11	18	8	17	17	17	17	17	17	17	9	17	11	17	10	17	9	19	18
Spec 14:TGA_t	31	5	22	5	29	5	29	5	29	5	26	6	32	5	27	6	26	6	27	9
Spec 15:IGA_t	8	27	8	27	8	27	8	30	8	30	8	27	8	27	8	27	8	27	7	31
Spec 16:AGA_t	17	11	17	6	17	17	17	18	17	18	17	6	17	11	18	8	17	6	16	22
Spec 17:$NAGA_t$	18	11	17	6	23	12	21	14	21	14	17	12	17	15	17	17	17	12	22	15
Spec 18:EGA_t	18	11	13	13	27	6	29	6	29	6	17	9	18	11	17	11	17	9	32	6
Spec 19:$APGA_t$	27	8	20	7	26	8	26	10	26	10	19	8	25	7	21	8	19	8	24	13
Spec 20:$FIGA_t$	8	28	8	28	9	26	11	26	11	26	8	26	8	27	8	27	8	26	10	26
Spec 21:A_{st}	5	32	6	32	4	33	3	34	3	34	5	32	5	32	5	32	5	32	1	35
Spec 22:GA_{st}	15	22	14	20	15	22	15	22	15	22	14	23	15	22	15	23	14	23	13	24
Spec 23:GJR_{st}	20	9	19	5	18	15	20	14	20	14	17	8	20	7	20	8	17	8	20	17
Spec 24:TGA_{st}	34	4	26	4	33	4	33	4	33	4	27	5	33	4	29	5	27	5	29	7
Spec 25:IGA_{st}	11	25	11	26	9	25	9	27	9	27	9	25	9	25	9	26	9	25	8	29
Spec 26:AGA_{st}	20	8	19	4	19	14	19	16	19	16	18	4	20	7	21	5	18	4	17	20
Spec 27:$NAGA_{st}$	20	9	17	5	25	10	24	12	24	12	17	12	18	11	17	16	17	12	23	13
Spec 28:EGA_{st}	19	8	15	12	30	4	31	4	31	4	18	8	21	7	21	8	18	8	33	5
Spec 29:$APGA_{st}$	31	5	20	7	29	5	29	6	29	6	19	8	26	7	21	8	19	8	27	9
Spec 30:$FIGA_{st}$	10	26	8	28	11	23	13	25	13	25	8	26	10	25	9	25	8	26	12	23
Spec 31:$comb_1$	35	3	35	3	35	3	35	3	35	3	35	3	35	3	35	3	35	3	35	3
Spec 32:$comb_2$	36	0	36	0	36	0	36	0	36	0	36	0	36	0	36	0	36	0	36	0
Spec 33:$comb_3$	36	0	36	0	36	0	36	0	36	0	36	0	36	0	36	0	36	0	36	0
Spec 34:HAR	2	36	1	37	1	37	1	33	1	33	1	37	1	36	1	37	1	37	5	33
Spec 35:logHAR	36	0	36	0	36	0	36	0	36	0	36	0	36	0	36	0	36	0	36	0
Spec 36:HEAVY	1	37	2	36	3	31	7	31	7	31	3	35	1	36	2	35	3	35	34	4
Spec 37:$GA-MIDAS$	0	38	0	38	0	38	0	38	0	38	0	38	0	38	0	38	0	38	0	34
Spec 38:$GA-RV_N$	2	35	7	31	17	24	16	22	16	21	14	24	7	32	9	23	13	24	27	7
Spec 39:$GA-RV_t$	7	30	5	31	15	24	17	22	18	21	10	24	10	28	8	23	10	26	28	6
Spec 40:$GA-RV_{st}$	8	34	4	37	40	1	38	2	36	1	16	12	7	34	6	22	18	9	39	2
Spec 41:$quant_1$	3	35	3	35	2	36	1	36	1	36	2	36	3	35	2	35	2	36	1	35
Spec 42:$quant_2$	4	32	4	34	6	31	6	32	6	32	4	31	5	31	4	31	4	31	6	32

Note: This table reports a summary of the out-of-sample results for the Unconditional Giacomini and White (2006) performed under the logMSE, sdMSE, propMSE, RMSE, propRMSE, MAE, logMAE, sdMAE, MAPE and GMLE loss functions for DAX. The period is from 02/01/2004 to 29/12/2017. The "win" column reports the cases where a certain specification outperformed significantly the competing specifications. On the contrary, the "lose" column reports the cases where a certain specification was outperformed significantly by the competing specifications. Adding the two columns for each specification and for each loss function, respectively, shows how many of the 41 in total pairwise comparisons where statistically significant. For the hypothesis testing the 5% significance level is selected.

Table 3.31 – Pairwise comparisons of unconditional Giacomini and White test for S&P 500 (4-year and 1-year in-sample), CAC 40 (4-year in-sample), FTSE 100 (4-year in-sample) and DAX (4-year in-sample) - MSE Loss Function

	S & P 500 (4-year in-sample)		S & P 500 (1-year in-sample)		CAC 40 (4-year in-sample)		FTSE 100 (4-year in-sample)		DAX (4-year in-sample)	
	win	loss	win	loss	win	loss	win	loss	win	loss
Spec 1:A_N	5	30	4	36	5	32	4	31	7	25
Spec 2:GA_N	9	20	8	17	13	22	13	15	16	4
Spec 3:GJR_M	20	4	24	3	19	9	10	0	15	4
Spec 4:TGA_N	28	0	35	0	29	0	20	3	16	4
Spec 5:IGA_N	6	27	5	17	10	22	10	17	10	4
Spec 6:AGA_N	22	4	13	4	19	10	19	8	18	4
Spec 7:$NAGA_N$	24	1	19	4	21	0	10	0	17	4
Spec 8:EGA_N	20	1	21	2	23	5	22	2	7	14
Spec 9:$APGA_N$	22	4	19	7	26	1	19	6	13	8
Spec 10:$FIGA_N$	5	21	5	21	9	27	10	15	10	9
Spec 11:A_t	3	36	0	37	5	32	5	31	7	27
Spec 12:GA_t	7	22	9	16	13	22	11	16	10	8
Spec 13:GJR_t	19	7	19	5	19	8	10	0	14	4
Spec 14:TGA_t	29	0	33	0	30	0	21	0	16	5
Spec 15:IGA_t	7	22	5	17	11	22	10	18	10	8
Spec 16:AGA_t	20	9	18	7	19	11	19	7	17	5
Spec 17:$NAGA_t$	15	4	15	5	22	0	10	0	13	5
Spec 18:EGA_t	23	0	20	5	22	5	23	1	7	14
Spec 19:$APGA_t$	19	5	15	11	22	1	19	3	15	5
Spec 20:$FIGA_t$	5	23	5	21	9	25	10	15	7	19
Spec 21:A_{st}	3	36	0	37	5	32	5	31	7	26
Spec 22:GA_{st}	7	22	8	16	13	22	14	15	10	8
Spec 23:GJR_{st}	20	5	20	5	19	8	10	0	15	4
Spec 24:TGA_{st}	30	0	35	0	29	0	21	0	19	4
Spec 25:IGA_{st}	7	22	5	17	11	22	10	17	10	8
Spec 26:AGA_{st}	20	9	29	0	19	11	21	6	17	4
Spec 27:$NAGA_{st}$	18	0	26	3	22	0	10	0	12	4
Spec 28:EGA_{st}	22	0	22	3	19	9	26	0	7	13
Spec 29:$APGA_{st}$	22	2	5	21	22	2	21	0	13	5
Spec 30:$FIGA_{st}$	5	23	10	7	9	25	10	15	8	17
Spec 31:$comb_1$	29	0	9	3	30	0	24	0	37	0
Spec 32:$comb_2$	20	2	9	8	23	2	22	0	37	0
Spec 33:$comb_3$	25	0	5	0	29	0	24	0	37	0
Spec 34:HAR	0	41	23	0	2	38	2	37	2	37
Spec 35:logHAR	20	2	5	16	19	2	10	0	37	0
Spec 36:HEAVY	7	16	5	19	5	32	4	35	4	36
Spec 37:$GA-MIDAS$	6	22	5	19	9	28	2	37	2	37
Spec 38:$GA-RV_N$	6	20	5	20	0	40	0	40	0	40
Spec 39:$GA-RV_t$	5	22	5	20	1	39	1	39	1	39
Spec 40:$GA-RV_{st}$	3	23	1	37	4	36	6	31	5	35
Spec 41:$quant_1$	2	39	0	38	3	37	5	32	6	34
Spec 42:$quant_2$	1	40	2	39	1	38	6	34	7	35

Note: This table reports a summary of the out-of-sample results for the Unconditional Giacomini and White (2006) performed under the MSE for S&P 500 (for the period is from 02/01/2004 to 29/12/2017), S&P 500 (for the period is from 02/01/2001 to 29/12/2017), CAC 40 (for the period is from 02/01/2004 to 29/12/2017), FTSE 100 (for the period is from 02/01/2004 to 29/12/2017) and DAX (for the period is from 02/01/2004 to 29/12/2017). The "win" column reports the cases where a certain specification outperformed significantly the competing specifications. On the contrary, the "lose" column reports the cases where a certain specification was outperformed significantly by the competing specifications. Adding the two columns for each specification and for each loss function, respectively, shows how many of the 41 in total pairwise comparisons where statistically significant. For the hypothesis testing the 5% significance level is selected.

Chapter 4

Directional Volatility Forecasting in the US Market

4.1 Introduction

Forecasting volatility is important for academics, practitioners, investors and policy makers. Volatility plays a central role in many areas of finance, such as derivatives pricing, risk measures management (Value-at-Risk and Expected Shortfall), hedging, and investment strategies. Traditionally, stock market volatility is predicted using time series models which mainly belong to the family of Generalized Autoregressive Conditionally Heteroskedasticity (GARCH) models and their extensions (see Eugenie (2003); Poon and Granger (2003); Hansen and Lunde (2005); Awartani and Corradi (2005); Awartani and Corradi (2005); Liu and Hung (2010); Chortareas et al. (2011); Mcmillan and Speight (2012); Dunis et al. (2013); Dendramis et al. (2015); Kambouroudis et al. (2016); Jiang et al. (2017); Gulay and Emec (2018); Catania and Proietti (2020); Conrad and Kleen (2020) among others)

In a review of several studies dealing with the forecasting performance of volatility models, Poon and Granger (2003) document the superior performance of GARCH models against historical volatility models. Using a large set of volatility models, Hansen and Lunde (2005) show that the GARCH(1,1) model is never outperformed by more sophisticated models when using exchange rate data. However, in the case of stock returns (IBM stock returns), GARCH(1,1) is much less competitive. For a smaller number of forecasting competitors, Awartani and Corradi (2005) and Liu and Hung (2010) show that the asymmetric GARCH outperforms the GARCH(1,1) for the S&P 500 index, while Mcmillan and Speight (2012) uncover a very poor performance of the GARCH(1,1) model against all other types of GARCH models for three Euro exchange rates. The relevance of asymmetric effects when forecasting stock index return volatility is also documented in Kambouroudis et al. (2016) for US and European indices.

The use of intraday data also displays mixed results, as shown in Chortareas et al. (2011). Specifically, the authors find the Fractionally Integrated GARCH (FIGARCH) model along with the intraday GARCH to be superior in almost all exchange rate series they consider. Although, the widely preferred Autoregressive Fractionally Integrated Moving Average (ARFIMA) model performs better than the traditional daily volatility models, it cannot beat the FIGARCH model and the intraday GARCH model. Likewise, Eugenie (2003) shows that ARFIMA and stochastic volatility models perform well for the case of the S&P 500 stock index returns, while Jiang et al. (2017) document similar performance of ARFIMA models against GARCH models for Chinese commodity data. Using intraday data for three different euro exchange rates, Mcmillan and Speight (2012) show that the GARCH(1,1) model is outperformed by all other GARCH specifications, and Dunis et al. (2013) record a good performance of an asymmetric GARCH model against other models such as ARFIMA for one-step ahead forecasts, while for longer forecast horizons the opposite is true.

In line with the aforementioned literature, this chapter aims to evaluate the forecasting performance of several GARCH models for the future volatility direction. We aim to establish whether more sophisticated models perform better than more parsimonious ones and evaluate a variety of forecast combination methods. The selection of a forecasting model for future volatility is certainly at the core of the empirical analysis on time series models, and most of the aforementioned studies use standard measures of accuracy like Mean Absolute Error (MAE) or Mean Squared Error (MSE) to evaluate forecasting models[1]. This chapter focuses on directional forecast accuracy and uses different forecast accuracy measures for the evaluation of diverse GARCH models, specifically the directional accuracy (DA) and directional forecast value (DV) measures by Blaskowitz and Herwartz (2009) and Blaskowitz and Herwartz (2011)). Only few studies in the field of volatility forecasting have used DA as an evaluation measure however, no single study focuses on both DA and DV measures. This is our first contribution. The DA measure is more robust to outlying forecast than other standard loss functions (Blaskowitz and Herwartz (2011)) and DV grasps the ability to predict the sign and the magnitude of the realized changes (Blaskowitz and Herwartz (2011)). For completeness, our analysis also illustrates the results for MSE.

The relevance of directional predictability is largely documented in the forecasting literature, and dates back to Merton (1991) and Leitch (1995) (for a comprehensive survey on directional forecasts, see Blaskowitz and Herwartz (2011); Blaskowitz and Herwartz (2014)). Leitch and Tanner (1991) and Satchell and Timmermann (1995) were the first to question the ability of statistically accurate forecasts to lead to high profits. Moreover, Leitch and Tanner (1991) and Engle and Susmel (1993a) were the first to discuss that a measure based on the direction of change might be more suitable when

[1]For a comprehensive survey on accuracy measures, see Koutsandreas et al. (2021).

introduced in trading strategies. Moosa and Vaz (2015) have shown that directional accuracy is more strongly correlated with profitability than the magnitude of the error. Furthermore, some technical trading rules were suggested, which measure a forecast's performance as profitable, when large profits were made for a small number of correct (sign) predictions and when small profits were achieved for a large number of wrong (sign) (Cumby and Modest (1987)). This early intuition is connected to a measure that combines the direction of change with the magnitude or error for profitability such as the DV. More recently, Vrontos et al. (2021) focus on forecasting the direction of U.S. implied volatility (VIX index) employing both standard econometric models and several machine learning techniques. They find that the implementation of machine learning techniques can be more effective compared to mainstream econometric models and model selection techniques in both a statistical and an economic evaluation setting.

Second, in this chapter, the test by Eberhardt and Fligner (1977) based on the equality of two proportions is used to assess the equivalence of the directional forecasting performance of the models. This is the second contribution of the chapter since to the best of our knowledge it is the first time that this type of test is used for evaluation of directional forecasting purposes. Third, the chapter not only looks at the performance of single and multivariate GARCH models, but it also evaluates diverse methods of combination of forecasts based on a simple mean and median, on a weighted average forecasting scheme for directional change (see Greer (2005)), on DA (see Costantini et al. (2016b)), and on a hierarchical algorithm encompassing-MSE procedure (see Costantini and Pappalardo (2010)). Some previous studies have also investigated the potential of the combination of the forecasts for future volatility, but standard accuracy measures have mainly been used (see, for example, Becker and Clements (2008); Kambouroudis et al. (2016); Ma et al. (2018); Fameliti and Skintzi (2020)).

The analysis is carried out using daily data for the S&P 500 index over the period 3/1/2000-29/12/2017. A variety of univariate GARCH models are considered along with two multivariate GARCH models [2]. For the estimation of the two multivariate GARCH models, we use both S&P 500 and RUSSELL 2000 data. For the out-of-sample analysis, the realized variance (RV) is used as a proxy for the true unobservable volatility (see also Meligkotsidou et al. (2019); Karanasos et al. (2021); Yfanti and Karanasos (2021)). A rolling-window of 750 observations (3 years) observations is applied to generate one-step ahead forecasts for two different holdout periods of 500 (2 years) and 250 (1 year). We allow for a holdout period to estimate the combining weights in order to have a more realistic out-of-sample performance evaluation. There is a trade off involved in selecting the optimal size of the holdout period. To minimize the bias associated with the chosen holdout period, but to also explore the effect of the size of the window, two periods are used (Rapach et al. (2010); Fameliti

[2]For a comparison of univariate and multivariate GARCH models, see, for example, Santos et al. (2013).

and Skintzi (2020)).

Further, bear and bull markets are also considered along with recession and expansion periods. The relationship between stock market volatility and business cycles and market conditions is well documented. A number of authors have found that volatility is higher during bear markets and recessions than in bull markets and expansions (Guidolin and Timmermann (2005); Perez-Quiros and Timmermann (2007)). However, the question remains as how business cycles and market regimes affect the stock market volatility predictability. Conventionally, regime switching specifications of GARCH models have been used to define different states of the economy. Empirical studies using the conventional MSE criterion have explored the stock volatility in bear and bull markets and in expansions and recessions (Li and Zakamulin (2020); Liu and Pan (2020)). To the best of our knowledge, this is the first study that examines business cycles and market conditions on stock returns' volatility predictability in terms of directional accuracy. The main difference between bear and bull markets in contrast with periods of expansions and recessions is that bear and bull periods are documented in almost real time with no major time delays.

Lastly, this chapter assesses the volatility forecasts with respect to their economic value via constructing volatility investment strategies. Specifically, two strategies are considered; the first one goes long on the S&P 500 when volatility is expected to decrease and vice versa, while the second one is more passive and goes long on S&P when volatility is expected to decrease, or does nothing when volatility is expected to increase. In order to evaluate the volatility strategies, we calculate the average return, the cumulative return, the Sharpe Ratio (SR) of Sharpe (1964) (see also Guidolin and Pedio (2021)), the Sortino ratio (Sortino and Van Der Meer (1991)) and the Omega ratio (Keating and Shadwick (2002)) of the portfolio over the out-of-sample period.

Our empirical analysis conveys the following main results. First, among the single GARCH models, the EGARCH model outperforms the other GARCH models in almost all cases, both in terms of DA and DV. Second, among the methods of combination, the procedures by Greer (2005) and Costantini and Pappalardo (2010) generate the best predictions for DA and DV in several cases. Third, the larger holdout period generally provides more significant results in terms of DA for both the single models and the combinations. Fourth, GARCH models and combinations of forecasts generally perform better during bear markets and recession periods compared with the bull and expansion ones. Fifth, directional forecasts are equivalent among the best performers, that is the EGARCH model and the combination schemes by Greer (2005) and Costantini and Pappalardo (2010). The results by the test of Eberhardt and Fligner (1977) give clear-cut evidence in this sense. Therefore, when there is uncertainty about what model to chose, the methods of combination may be preferable considering

directional forecast equivalence among models and forecast combinations. Lastly, our real-time investment strategies point to superiority of the hierarchical encompassing-HMSE combination method.

The rest of the chapter is organized as follows. Section 2 describes data and methodology. Section 3 presents and discusses the empirical results. Section 4 concludes.

4.2 Data and methodology

4.2.1 Data

Daily returns data for the S&P 500 over the period 3/1/2000-29/12/2017 are used for the evaluation of the forecasts. For the estimation of the two multivariate GARCH models (see section 4.2.2), the Russell 2000 index along with S&P 500 is used. Data are taken from the Oxford-Man Institute's Realized Library (library version 0.2). The sample period for our analysis includes some important financial events: the dot-com bubble of 2000, which caused the S&P 500 index to an all-time intraday high of 1,552.87 on March 24, 2000, the record high closing price on October 9, 2007, a new intraday record high on October 11, 2007, and last but not least, the financial crisis of 2008-2010, which was conventionally initiated by the September 2008 Lehman Brothers bankruptcy.

S&P 500 return series has a zero mean and a standard deviation of 0.01, as seen in table 3.2. While the negative skewness (-0.18) and excess kurtosis (11.52) show that the distribution of the returns has fatter tails than the normal distribution and it is asymmetric. For the out-of-sample evaluation, realized variance (RV) is used as proxy for the true unobservable variance. More specifically, the 5-min RV is used taken from Oxford-Man Institute's Realized Library (library version 0.2) in line with previous studies (see, for example, Liu et al. (2015)).[3]

4.2.2 Methodology

Our forecasting exercise relies on the use of univariate and multivariate volatility models. The standard GARCH is considered (see Bollerslev (1986)) and other GARCH specifications, namely the GJR-GARCH Glosten et al. (1993), the exponential GARCH (EGARCH) by Nelson (1991), the asymmetric GARCH (AGARCH) by Engle (1990a), the non-linear Asymmetric GARCH (NAGARCH) by Engle and Susmel (1993b), the asymmetric power GARCH (APGARCH) by Ding et al. (1993), the Integrated GARCH (IGARCH) by Nelson (1990) and the fractional integrated GARCH (FIGARCH)

[3]The data are cleaned by Oxford-Man Institute; see Barndorff-Nielsen et al. (2009). For the realized variance, few missing observations were filled by averaging the prior and the post observation. Link to the data: https://realized.oxford-man.ox.ac.uk/.

model by Baillie et al. (1996). Two multivariate models largely used in empirical analysis are also considered Caporin and McAleer (2012a), namely the BEKK model Engle and Kroner (1995) and the DCC model Engle (2002a).[4] The models are illustrated in Table 4.1.

Table 4.1 – Univariate and multivariate volatility models

Model (Author)	Specification
GARCH	$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2$
Bollerslev (1986)	$(\omega, \alpha_i, \text{ and } \beta_i > 0, \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 < 1)$
IGARCH	$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2$
Nelson (1990)	$(\sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 = 1)$
EGARCH	$\log(\sigma_t^2) = \omega + \sum_{i=1}^{p}((\alpha_i(\lvert z_{t-i}\rvert - E\lvert z_{t-i}\rvert) + \gamma_i z_{t-i}) + \sum_{j=1}^{q} \beta_j \log(\sigma_{t-j}^2)$
Nelson (1991)	$(z_t = \varepsilon_t \sigma_t^{-1})$
GJR-GARCH	$\sigma_t^2 = \omega + \sum_{i}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{k=1}^{o} \gamma_k \varepsilon_{t-k}^2 I_{[\varepsilon_{t-k}<0]} + \sum_{i}^{q} \beta_i \sigma_{t-i}^2$
Glosten et al. (1993)	$(I_{t-1} = 1 \text{ if } \varepsilon_{t-1} < 0 \text{ and } I_{t-1} = 0 \text{ otherwise})$
NAGARCH	$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i(\varepsilon_{t-i}^2 + \gamma_i \sigma_{t-i})^2 + \sum_{j=1}^{q} \beta \sigma_{t-j}^2$
Engle and Susmel (1993b)	
APGARCH	$\sigma_t^{\delta} = \omega + \sum_{i=1}^{p} \alpha_i(\lvert\varepsilon_{t-1}\rvert + \gamma_i \varepsilon_{t-i})^{\delta} + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^{\delta}$
Ding et al. (1993)	
FIGARCH	$\sigma_t^2 = \omega + [1 - \beta(L) - \alpha(L)(1-L)^d]\varepsilon_t + \beta \sigma_{t-1}^2$
Baillie et al. (1996)	$(0 < d < 1, \alpha(L) = 1 + \alpha_1 L + \alpha_2 L^2 + ... + \alpha_q L^q, \beta(L) = 1 + \beta_1 L + \beta_2 L^2 + ... + \beta_p L^p)$
BEKK	$\Omega_t = CC' + \sum_{i=1}^{p} A_i(\varepsilon_{t-i}\varepsilon_{t-i}^{'})A_i^{'} + \sum_{j=1}^{q} B_j \Omega_{t-j} B_j^{'}$
Engle and Kroner (1995)	
DCC	$Q_t = Q(1 - \sum_{i}^{p} \alpha_i - \sum_{j=1}^{q} \beta_j) + \sum_{j=1}^{q} \beta_j Q_{t-j} + \sum_{j=1}^{p} \alpha_j \hat{\varepsilon}_{t-j} \hat{\varepsilon}_{t-j}^{'}$
Engle (2002a)	$R_t = (I_n \circ Q_t)^{1/2} Q_t (I_n \circ Q_t)^{1/2}$

Notes: This table reports the univariate and multivariate models used in the empirical analysis. In the BEKK model, the matrices $\{\sum_{i=1}^{p} A_i\}$ and $\{\sum_{j=1}^{q} B_j\}$ are non-negative and symmetric. In the DCC model, Q_t defines the dynamic process of the variance-covariance matrix of ε_t ("de-garched" returns), R_t is the correlation matrix, and Q is a $N \times N$ symmetric and positive-definite matrix parameter.

We also compare the performance of several volatility models and combinations of forecasts for future volatility (see also Becker and Clements (2008); Kambouroudis et al. (2016); Ma et al. (2018); Fameliti and Skintzi (2020)). Specifically, eight different methods of combination are considered based on the following equation

$$\hat{f}_{c,t+h|t}^{h} = w_{c,ot}^{h} + \sum_{i=1}^{k} w_{c,it}^{h} \hat{f}_{i,t+h|t} \tag{4.1}$$

where $\hat{f}_{c,t+h|t}^{h}$ is a given combination forecast whose weights, $\{w_{c,it}^{h}\}_{i=0}^{k}$, are computed using the individual out-of-sample forecasts, c indicates the combination method, k is the number of models, and h is the forecast horizon (h=1 in our case). The first method of combination is the simple mean forecasting combination that averages the forecasts of all volatility models. This naive forecast combination has been documented as the hardest one to beat in the literature. The second method is the

[4]The list of volatility models considered in our analysis does not intend to be exhaustive, given the large number of models developed in the literature. However, the models we use are largely applied in empirical studies (see, for example, Hansen and Lunde (2005)).

median forecast combination that calculates the median of all forecasts. Next, the hierarchical procedure based on encompassing tests and the MSE (H − MSE) proposed by Costantini and Pappalardo (2010) is considered. This combination method that works as follows has been shown to achieve gains in accuracy (Costantini and Pappalardo (2010),Morales-Arias and Moura (2013)). The models are first ranked on the basis of the MSE. Then, the best model (i.e. the one with the lowest MSE) is tested sequentially against all the other models using encompassing tests (i.e the ones with the highest MSE) [5]. If the best model encompasses the alternative model at some significance level, the alternative model is discarded from the list of forecast combinations. This step is then repeated with the second best model (the one with the second lowest MSE), and so on. The procedure terminates when no encompassed model remains in the list. In our analysis, the encompassing test by Clark and McCracken (2001) (Chen et al. (2010); Welch and Goyal (2008); Stock and Watson (2002)) :

$$Enc - new = P\frac{P^{-1}\sum_{t=R}^{T}(\widehat{e}_{1,t+1}^2 - \widehat{e}_{1,t+1}\widehat{e}_{2,t+1})}{P^{-1}\sum_{t=R}^{T}\widehat{e}_{1,t+1}^2} \tag{4.2}$$

where P denotes the number of 1-step ahead forecasts, R is the number of in-sample observation, $R+P=T+1$ is the total number of observations, and $\widehat{e}_{1,t+1}$ and $\widehat{e}_{2,t+1}$ are the 1-step ahead forecast errors of model, 1 and 2, respectively. For the combination of the forecasts, the average is used to combine the forecasts from the models that are not encompassed. This implies that $w_{c,ot}^h = 0$ and $w_{c,it}^h = \frac{1}{k}$ in equation (4.1). The heteroskedasticity-adjusted hierarchical procedure (H − HMSE) which uses the Heteroskedastic-MSE (HMSE) loss function is also considered

$$HMSE = P^{-1}\sum_{t=1}^{P}(\hat{\sigma}_{t+1}^2\hat{h}_{t+1}^{-1} - 1)^2 \tag{4.3}$$

where $\hat{\sigma}_t$ is the 1-step ahead forecasted variance and $\hat{h_t}$ is the realized variance (Marcucci (2005)).

The next class of forecast combination methods uses the DA measure and is based on the hit/success rates (HR) and exponential hit/success rates (EHR), respectively (see Costantini et al. (2016b)). The weights for the hit/success rates (HR) combination are given by

$$w_{HR,it}^h = \frac{\sum_{j=T_1+h-1}^{t} DA_{jh}^i}{\sum_{l=1}^{k}\sum_{j=T_1+h-1}^{t} DA_{jh}^l} \tag{4.4}$$

whereas those for the EHR method are (see Bacchini et al. (2010)):

$$w_{EHR,it}^h = \frac{exp\sum_{j=T_1+h-1}^{t} DA_{jh}^i}{\sum_{l=1}^{k} exp\sum_{j=T_1+h-1}^{t} DA_{jh}^l} \tag{4.5}$$

[5] Consider 3 models, M1, M2 and M3, where $MSE1 < MSE2 < MSE3$. In the 1st step, M1 (the best model in terms of forecast accuracy) is compared with M2 and M3 using the encompassing test. In a successive step, M2 (the second best model) is tested against M3.

where DA denotes the directional accuracy by Blaskowitz and Herwartz (2011) (see the discussion below).

The final class of combination forecasts is based on the approach by Greer (2005). Two variations of this forecasting scheme are employed. The first one uses the following weights:

$$w^h_{greer,it} = \begin{cases} \frac{\sum_{j=T_1+h-1}^{t} DA^i_{jh}}{\sum_{l=1}^{k}\sum_{j=T_1+h-1}^{t} DA^l_{jh}} & \text{if } DA \geq 0.5 \\ 0 & \text{if } DA < 0.5 \end{cases} \tag{4.6}$$

Finally, the second method of Greer combinations is a simple mean of all the models where the forecaster documents more success than misses.

As mentioned above, for the out-of-sample forecast evaluation, this chapter focuses on the directional accuracy (DA) measure proposed by Blaskowitz and Herwartz (2011) and Blaskowitz and Herwartz (2014).[6] Let X^h_t and Y_{t+h} be the h-step ahead forecast available at time t and the realized value, respectively. Consider h=1 (as in our forecasting exercise). Using the indicator function $I(\bullet)$, the realized and predicted directions are given by $\tilde{Y}_t = I(Y_{t+1} - Y_t > 0)$ and $\tilde{X}_t = I(X^1_t - Y_t > 0)$, respectively (see Blaskowitz and Herwartz (2011)). In our forecasting exercise, $Y_{t+1} - Y_t$ is the first difference of the daily 5-min RV, and X^1_t is the one-step ahead future variance derived from the single models or the forecast combinations.

The (In-)correct directional forecast can be defined by the binary variable $\tilde{Z}_t = I(\tilde{X}_t = \tilde{Y}_t)$. A loss function for directional forecasting accuracy is given by (see Blaskowitz and Herwartz (2011)):

$$L^{DA}(X^1_t, Y_{t+1}, Y_t) = \begin{cases} a & \text{if } \tilde{Z}_t = 1 \\ b & \text{if } \tilde{Z}_t = 0 \end{cases} \tag{4.7}$$

where $(a, b) \neq (0, 0)$. A correct direction forecast will take the value of a, while an incorrectly predicted direction will take the value of b. Usually $(a, b) = (1, 0)$ or $(a, b) = (1, -1)$. In this chapter, we use $(a, b) = (1, 0)$ as in Swanson and White (1995), Swanson and White (1997a), Swanson and White (1997b), Gradojevic and Yang (2006), Diebold (2007) and Costantini et al. (2016a). In order to assess the ability of the models and the combinations of forecasts to predict the direction of future volatility the test proposed by Pesaran and Timmermann (1992) is employed.

Further more, we also test for the equivalence of the proportion of successful direction predictions using the Z_1 statistics by Eberhardt and Fligner (1977), which tests the difference in two proportions,

[6] For the directional accuracy, see also Chen et al. (2016) and Costantini et al. (2016a).

and it is normally distributed in large samples.[7] Under the null hypothesis, the proportion of successes of the two models is equal, i.e. the forecasting models have equal directional forecast accuracy. The alternative hypothesis states that the number of successes are different.

Since the L^{DA} ignores the size of the realized directional movement, this chapter also considers the directional forecast value (DV), which can be used to measure the benefit/losses to the forecast user:

$$L^{DV}(X_t^1, Y_{t+1}, Y_t) = \begin{cases} H_t^{UU} = H^{UU}(Y_{t+1}, Y_t) & \text{if correct upward prediction} \\ H_t^{DD} = H^{DD}(Y_{t+1}, Y_t) & \text{if correct downward prediction} \\ H_t^{UD} = H^{UD}(Y_{t+1}, Y_t) & \text{if incorrect upward prediction} \\ H_t^{DD} = H^{DU}(Y_{t+1}, Y_t) & \text{if incorrect downward prediction} \end{cases} \tag{4.8}$$

In equation (4.8), H_t^{UU} and H_t^{DD} indicate the benefits (gains) to the forecast user when an upward (downward) forecast direction is expected, and an upward (downward) movement occurs. Likewise, H_t^{UD} and H_t^{DU}) indicate the costs (losses) for the forecast when an incorrect directional prediction is made (see Blaskowitz and Herwartz (2011)). Our models and forecasts combinations are assessed by calculating the mean DV over the out-of-sample period.

Finally, the volatility forecasts are assessed with respect to their economic value via constructing volatility investment strategies. Specifically, two strategies are considered; the first one goes long on the S&P 500 index when volatility is expected to decrease and vice versa, while the second one is more passive and goes long on S&P when volatility is expected to decrease, but does not do when volatility is expected to increase. In order to evaluate the volatility strategies, we calculate the average daily return, the cumulative return, the Sharpe Ratio (SR) of Sharpe (1964), the Sortino ratio by Sortino and Van Der Meer (1991) and the Omega ratio by Keating and Shadwick (2002) of the portfolio over the out-of-sample period. The Sharpe Ratio (SR) is very popular and sufficient measure for investments with returns that follow the normal distribution. It represents the risk premium return earned per unit of total risk and is calculated as follows.

$$Sharpe\ Ratio = \frac{E(R_p) - r_f}{\sigma} \tag{4.9}$$

where $E(R_p)$ is the portfolio returns and σ is the portfolio returns' volatility, r_f is the risk-free rate. However, Sharpe ratio cannot distinguish between upside and downside risk. Both the Sortino and the Omega ratio belong to the performance evaluation measures that are based on Lower Partial Moments (LPMs). LPMs only consider deviations that fall below an ex-ante defined threshold (in our case, the average of the risk-free rate). These measures are also known as downside deviation measures,

[7] For details, see Eberhardt and Fligner (1977).

as downside risk is defined as the risk below a threshold. The motivation for this adjustment is that upside risk is not a threat to the portfolio, thus it should be disregarded. More in detail, the Sortino ratio only considers downside risk instead of the standard deviation and is calculated as follows.

$$Sortino = \frac{E(R_p) - r_f}{\sqrt{E[(r_f - R_p)_+^2]}} \tag{4.10}$$

where $E(R_p)$ is the portfolio returns and σ is the portfolio returns volatility, r_f is the risk-free rate, $\sqrt{E[(r_f - R_p)_+^2]}$ is the excess returns of the portfolio over the threshold. This approach may be more suitable to asymmetric return distributions. However, it may not give the full perception of risk since it only recognizes downside volatility risk.

The Omega measure considers the whole distribution and does not rely on distributional assumptions. The Omega ratio considers returns above and below the risk-free rate and is calculated as a ratio of the cumulative probability of an investment's outcome above the risk-free rate, over the cumulative probability of an investment's outcome below the risk-free rate. It can be seen as a probability-weighted ratio of gains to losses relative to the risk-free rate and is given by the following formula.

$$Omega = \frac{E[(R_p - r_f)^+]}{E[(r_f - R_p)_+]} \tag{4.11}$$

where $E(R_p)$ is the portfolio returns and r_f is the risk-free rate.

4.3 Empirical results

In this section, the empirical findings are presented. Specifically, we present the performance of a range of volatility models (see Table 4.1) and of the eight forecast combination methods described in Section 4.2.2. With respect to the estimation of the univariate GARCH(p,q) models, we opt for (p,q)=(1,1) (see, for example, Engle and Ng (1993); Caporin and McAleer (2012a)) and a normal distribution assumption for the error term (see also Engle and Patton (2001); Pesaran and Pick (2009)). For the multivariate GARCH models, the symmetric DCC model and the diagonal specification of the BEKK models are used (Caporin and McAleer (2012a)). One-step ahead forecasts are evaluated (see also Rapach and Strauss (2008); Fuertes et al. (2009b); Herwartz (2017)) and a rolling window estimation scheme is used to generate the forecasts with a window size of 3 years for two different holdout periods of 500 (2 years) and 250 (1 year) observations, respectively. Such a forecast design is quite typical in the literature (Fameliti and Skintzi (2020)[8]). Precisely, for the methods of combinations that require estimation of the combination weights the first 250 or 500 observations, respectively, of the out-of-sample period are used as the initial holdout period.

[8] Fameliti and Skintzi (2020) consider an in sample period of 3.5 years with a holdout window of 1.5 years.

The forecast evaluation is conducted using the DA and DV measures, while the standard MSE measure is computed as a preliminary step. The results for the MSE are reported in Appendix (see Tables 4.10 - 4.11), and show that the H − MSE combination is rather competitive against all other models and combinations.

Table 4.2 reports the results for the DA and DV measures, and the significance for the DA, using 500 observations from the out-of-sample period as the initial holdout period. Also, the p-values of the test of equal proportions by Eberhardt and Fligner (1977) of DA for the pairwise comparisons among the best single model and all the other models and forecast combinations are reported. The EGARCH model is associated with the highest directional accuracy followed by the NAGARCH one. These models attain the highest values of DA, which are statistically significant. However, Jung (2016) found that GARCH is better than GJR in terms of DA. Among the forecast combination, the best performing ones are the two methods of Greer (2005), while the two methods based on the hierarchical procedure perform slightly worse. More specifically, the Greer-mean combination attains a DA measure of 0.524, while the respective value for the Greer combination is 0.522. Both hierarchical combination methods are associated with DA values of 0.513. More importantly, the test by Eberhardt and Fligner (1977) indicates that the best univariate volatility model, EGARCH, along with the NAGARCH and the four forecast combinations achieve equivalent directional performance. As such, the use of combination of forecasts may be recommended given the uncertainty about which model to be used (see, for example, Armstrong (2001)). However, in the literature, prior studies have found evidence that single models outperform simple forecast combinations (Zhang et al. (2021); Zhang et al. (2020)). As for the DV measure, the EGARCH model ranks first (0.036) closely followed by the H-MSE, HR and Greer combination of forecasts which achieve a value of 0.035. This finding is important as the DV measure is connected with profitability (Blaskowitz and Herwartz (2009)).

The impact of bull and bear markets, as identified using a rule of thumb on stock index prices, on directional volatility forecasts is now examined. Table 4.3 illustrates our findings. Quite interestingly, the directional forecasting ability of the majority of single models and all forecast combinations is more pronounced in bear than bull markets. Specifically, for the bear period, more single models (5 out of 9) achieve statistical significance for DA compared to the full sample, reported in Table 4.2. Among single models, the EGARCH model remains the best performer followed by the NAGARCH that ranks second. What is more, forecast combinations perform quite well during bear markets. In more details, all forecast combination methods are associated with statistically significant directional accuracy with the H-MSE achieving the best performance, which is equal to the single best model one (EGARCH). The Eberhardt and Fligner (1977) test results suggest that all forecast combinations are equivalent to the EGARCH model, and as such they would safely be used in the context of directional forecasting.

Table 4.2 – Directional accuracy (DA) and directional forecast value (DV) of volatility. The holdout period is 500 observations.

Single models	DA	Test of equal DA	DV
EGARCH	**0.537*****	1.00	**0.036**
NAGARCH	0.525***	0.30	0.034
DCC	0.509	0.02	0.028
APGARCH	0.508	0.02	0.034
GJR-GARCH	0.503	0.01	0.034
FIGARCH	0.489	0.00	0.029
GARCH	0.442	0.00	0.028
IGARCH	0.440	0.00	0.028
BEKK	0.438	0.00	0.028
Combinations of forecasts			
Mean	0.484	0.00	0.033
Median	0.479	0.00	0.032
H-MSE	0.513*	0.05	0.035
H-HMSE	0.513*	0.05	0.034
HR	0.489	0.00	0.035
EHR	0.481	0.00	0.032
Greer	0.522***	0.22	0.035
Greer-mean	0.524***	0.26	0.030

Notes: Out-of-sample period: $02/01/2003 - 29/12/2017$. Holdout=500 obs. The MSE is multiplied by 10^6 (see also Blaskowitz and Herwartz (2009)). GARCH models are described in Table 4.1. $H - MSE$ and $H - HMSE$ denote the hierarchical procedure by Costantini and Pappalardo (2010) based on the MSE and HMSE loss functions, respectively. HR and EHR indicate hit/success rate and exponential of hit/success rates methods of forecast combination (Equations (4.4) and (4.5), respectively). Greer is the method of combination based on equation (4.6). The column 'DA' indicates the proportion of corrected predictions. *,** and *** indicate significance for the DA using the test of Pesaran and Timmermann (1992) at 10%, 5% and 1% levels, respectively. The highest DA value is reported in bold. The column 'Test of equal DA' reports the p-value of the test of equal proportions by Eberhardt and Fligner (1977) of DA for the pairwise comparisons among the best single model and all the other models and forecast combinations. The null hypothesis states that two models have equal DA, while the alternative hypothesis they do not. The column 'DV' reports the mean of directional forecast value. The highest DV value is reported in bold. The results for DV have been multiplied by 10^3.

Turning to the DV measure, the best performance is gained by the APGARCH model followed by the GJR-GARCH, the EGARCH and the NAGARCH models. This superior performance probably stems from the asymmetric nature of these models, as extra volatility terms are introduced when returns are negative (bear markets). However, we have to note that four of the eight forecast combinations; namely the mean, H-MSE, HR and Greer combinations perform similarly to the best single model attaining values of DV ranging from 0.152 to 0.156. Our findings concerning the bull period resemble more those for the full sample reported in Table 4.2. More precisely, the EGARCH model is still the best one among the single models and combinations of forecasts for DA, while NAGARCH ranks second. However, the Greer forecast combinations are rather competitive achieving statistically significant DA values. More importantly, these combinations are statistically equivalent to the EGARCH model as judged by the Eberhardt and Fligner (1977) test. Turning to the DV results, it is noted that the EGARCH model ranks first with a DV value equal to 0.022, closely followed by NAGARCH, DCC and FIGARCH (0.021). Similar DV values are attained by the hierarchical combinations and the Greer ones. This finding is important as forecast combinations alleviate uncertainty over the single best GARCH model.

The stock return forecasting literature has identified superiority of predictive performance during recessions as opposed to expansions (Rapach et al. (2010)). We now check whether similar findings pertain with respect to volatility forecasts. Table 4.4 reports the related results for recession and expansion periods as dated by NBER. Focusing on both panels, it is observed that our results point to the same conclusion. Directional forecast accuracy is more prevalent during recessions than expansions. The well established fact in the literature that economic agents tend to respond more strongly to negative conditions compared to positive ones seems to prevail in volatility of the markets as well. More specifically, the H-MSE forecast combination method ranks first when considering recessions, while DA remains significant for other methods of combination and single models. Five out of nine single models (EGARCH, GJR-GARCH, NAGARCH, APGARCH and DCC) are associated with significant DA values, along with all forecast combinations, excluding the simple mean, median and the EHR method. The results by the test of Eberhardt and Fligner (1977) point to a substantial equivalence of directional performance among models and combinations. Once more, these findings point to the use of a combination method, that is $H-MSE$. As for DV, the highest score is gained by the APGARCH model, while the $H-MSE$ procedure takes a slightly smaller value of DV, which indicates that this procedure is highly competitive for the directional forecast of future volatility.

Our findings for the case of expansionary periods mimic those for the bull periods. This could be expected as both expansionary periods and bull markets are prolonged periods characterized with low

Table 4.3 – Directional accuracy (DA) and directional forecast value (DV) of volatility in bear and bull markets. The holdout period is 500 observations.

	Bear			Bull		
Single models	DA	Test of equal DA	DV	DA	Test of equal DA	DV
EGARCH	**0.579*****	1.00	0.153	**0.535*****	1.00	**0.022**
NAGARCH	0.576***	0.96	0.152	0.518**	0.19	0.021
GJR-GARCH	0.564***	0.71	0.156	0.491	0.07	0.019
APGARCH	0.558***	0.60	**0.160**	0.499	0.01	0.020
DCC	0.532*	0.22	0.089	0.511	0.00	0.021
FIGARCH	0.483	0.01	0.099	0.493	0.00	0.021
GARCH	0.474	0.01	0.136	0.432	0.00	0.015
BEKK	0.474	0.01	0.137	0.428	0.00	0.015
IGARCH	0.471	0.00	0.136	0.428	0.00	0.016
Combinations of forecasts						
Mean	0.544***	0.37	0.152	0.476	0.00	0.018
Median	0.538*	0.29	0.141	0.471	0.00	0.017
H-MSE	**0.579*****	1.00	0.156	0.505	0.02	0.021
H-HMSE	0.558***	0.60	0.149	0.508	0.04	0.021
HR	0.555***	0.54	0.154	0.481	0.00	0.019
EHR	0.538**	0.29	0.149	0.474	0.00	0.018
Greer	0.552**	0.49	0.154	0.519**	0.22	0.021
Greer-mean	0.555***	0.54	0.153	0.520**	0.25	0.021

Notes: The out-of-sample period is 02/01/2003 − 29/12/2017. Holdout=500 obs. The MSE $\times$ 10^6 (see also Blaskowitz and Herwartz (2009)). GARCH models are described in Table 4.1. H − MSE and H − HMSE loss functions denote the hierarchical procedure by Costantini and Pappalardo (2010) based on MSE and HMSE, respectively. HR and EHR indicate hit/success rate and exponential of hit/success rates methods of combination (equations (4.4) and (4.5), respectively). Greer is the method of forecast combination based on equation (4.6). The column 'DA' indicates the proportion of corrected predictions. *,** and *** indicate significance for the DA using the test of Pesaran and Timmermann (1992) at 10%, 5% and 1% levels, respectively. The highest DA value is reported in bold. The column 'Test of equal DA' reports the p-value of the test of equal proportions by Eberhardt and Fligner (1977) of DA for the pairwise comparisons among the best single model and all the other models and forecast combinations. The null hypothesis states that two models have equal DA, while the alternative hypothesis they do not. The column 'DV' reports the mean of directional forecast value. The highest DV value is reported in bold. The results for DV have been multiplied by 10^3. The results are presented for bear and bull markets.

volatility and positive returns. Fewer models and forecast combinations achieve statistically significant directional forecasts. Specifically, the EGARCH ranks first followed by the NAGARCH model succeeding in forecasting accurately in 53.9% and 52.1% of the respective time periods. More importantly, the Greer combinations achieve an equivalent performance and succeed in 52.2% and 52.1% of the cases. The directional value results corroborate this finding as EGARCH and the Greer combinations have the highest DV of 0.021. Once more, this set of results confirm the potential of the combination of the forecasts for future volatility in terms of DA and DV measures.

Table 4.5 reports the economic evaluation of the two volatility strategies (Panels A and B). The majority of models and combinations produce negative returns for both strategies. Our findings point towards superiority of the H-HMSE combination method where the annualized return is 27.78 per

Table 4.4 – Recession and expansion periods directional accuracy (DA) and direction forecast value (DV) of volatility. The holdout period is 500 observations.

	Recession			Expansion		
Single models	DA	Test of equal DA	DV	DA	Test of equal DA	DV
EGARCH	0.554**	0.87	0.145	**0.539*****	1.00	**0.021**
GJR-GARCH	0.552**	0.83	0.148	0.492	0.00	0.019
NAGARCH	0.549**	0.76	0.144	0.521***	0.17	0.020
APGARCH	0.544**	0.66	**0.151**	0.501	0.00	0.019
DCC	0.544**	0.66	0.090	0.509	0.02	0.020
FIGARCH	0.496	0.08	0.099	0.490	0.00	0.020
GARCH	0.462	0.01	0.129	0.434	0.00	0.015
IGARCH	0.462	0.01	0.128	0.430	0.00	0.015
BEKK	0.462	0.01	0.128	0.430	0.00	0.015
Combinations of forecasts						
Mean	0.528	0.38	0.143	0.478	0.00	0.018
Median	0.509	0.16	0.133	0.475	0.00	0.017
H-MSE	**0.560*****	1.00	0.148	0.507	0.01	0.020
H-HMSE	0.541*	0.60	0.142	0.510	0.03	0.020
HR	0.541*	0.60	0.146	0.483	0.00	0.018
EHR	0.523	0.31	0.140	0.476	0.00	0.018
Greer	0.538*	0.54	0.146	0.521***	0.17	**0.021**
Greer-mean	0.541*	0.60	0.145	0.522***	0.19	**0.021**

Notes: Out-of-sample period: 02/01/2003 − 29/12/2017. Holdout=500 obs. The MSE $\times 10^6$ (see also Blaskowitz and Herwartz (2009)). GARCH models are described in Table 4.1. H − MSE and H − HMSE denote the hierarchical procedure by Costantini and Pappalardo (2010) based on MSE and HMSE, respectively. HR and EHR indicate hit/success rate and exponential of hit/success rates methods of combination (equations (4.4) and (4.5), respectively). Greer is the method of combination based on equation (4.6). The column 'DA' indicates the proportion of corrected predictions. *,** and *** indicate the significance of DA by the test of Pesaran and Timmermann (1992) at 10%, 5% and 1%, respectively. The highest DA value is reported in bold. The column 'Test of equal DA' reports the p-value of the test of equal proportions by Eberhardt and Fligner (1977) of DA for the pairwise comparisons among the best single model and all the other models and forecast combinations. The null hypothesis states that two models have equal DA, while the alternative one they do not. The column 'DV' reports the mean of directional forecast value. The highest DV value is reported in bold. The results for DV have been multiplied by 10^3. The results are presented for recession and expansion periods as dated by NBER.

cent under the first strategy and 16.46 per cent under the second. In both cases, the associated annualized SR is around 1.50. It should also be noted that FIGARCH and GJR are also associated with positive returns greater than 5 per cent. Now focusing on the combinations, the simple mean and the Greer combination also produce positive returns greater than 5 per cent on the basis of both volatility strategies. Finally, both the Omega and the Sortino ratios point to superiority of the H-HMSE combination method, followed by the mean and Greer combination methods.

The sensitivity of our findings to the length of the holdout window is now being explored. Specifically, a holdout period of 250 observations is considered, roughly corresponding to one year. Our results for the full out of sample period are reported in Table 4.6. Compared with the longer holdout window, our results are qualitatively similar although the performance of forecast combinations tend to deteriorate. This finding indicates that a longer holdout windows guarantees a more accurate evaluation and a better estimation of the time-varying weights required for the forecast combinations. However, even in this case, the Greer-mean method attains a significant DA and is still competitive and equivalent to the best single model (EGARCH). More importantly, in terms of DV, the best single model (along with NAGARCH and APGARCH) and the Greer combinations have the highest directional values of 0.034 and 0.033, respectively.

Table 4.5 – Evaluation of strategies. The holdout period is 500 observations.

Panel A: Strategy 1				
Single models	Return	SR	Omega	Sortino
APAGARCH	-2.52	-0.14	1.01	0.00
EGARCH	-13.24	-0.75	0.92	-0.02
FIGARCH	4.97	0.28	1.06	0.02
GARCH	-6.32	-0.36	0.98	-0.01
GIR-GARCH	5.14	0.29	1.05	0.01
IGARCH	-10.32	-0.59	0.95	-0.01
NAGARCH	-9.45	-0.54	0.96	-0.01
DCC	-8.83	-0.50	0.95	-0.01
BEKK	-10.37	-0.59	0.95	-0.01
Methods of combination of forecasts				
Mean	5.14	0.29	1.05	0.01
Median	-135.97	-8.80	0.16	-0.31
H-MSE	-10.08	-0.57	0.95	-0.02
H-HMSE	27.78	1.58	1.51	0.11
HR	-4.88	-0.28	0.99	0.00
EHR	-10.07	-0.57	0.94	-0.02
Greer (2005)	5.14	0.29	1.05	0.01
Panel B: Strategy 2				
Single models	Return	SR	Omega	Sortino
APAGARCH	1.31	0.09	1.00	0.00
EGARCH	-4.05	-0.30	0.90	-0.03
FIGARCH	5.05	0.36	1.07	0.02
GARCH	-0.59	-0.04	0.97	-0.01
GIR-GARCH	5.14	0.29	1.04	0.01
IGARCH	-2.59	-0.19	0.93	-0.02
NAGARCH	-2.16	-0.15	0.94	-0.02
DCC	-1.85	-0.13	0.95	-0.01
BEKK	-2.62	-0.19	0.93	-0.02
Methods of combination of forecasts				
Mean	5.14	0.29	1.04	0.01
Median	-65.42	-5.31	0.16	-0.31
H-MSE	-2.47	-0.17	0.94	-0.02
H-HMSE	16.46	1.48	1.48	0.10
HR	0.13	0.01	0.98	-0.01
EHR	-2.47	-0.18	0.93	-0.02
Greer (2005)	5.14	0.29	1.04	0.01

Notes:Out-of-sample period: 02/01/2003 − 29/12/2017. Holdout=500 obs. The MSE $\times 10^6$ (see also Blaskowitz and Herwartz (2009)). GARCH models are described in Table 4.1. H − MSE and H − HMSE denote the hierarchical procedure by Costantini and Pappalardo (2010) based on MSE and HMSE, respectively. HR and EHR indicate hit/success rate and exponential of hit/success rates methods of combination (equations ((4.4)) and ((4.5)), respectively). Greer is the method of combination based on equation ((4.6)). The column Return is the annualized return, column SR is the annualized Sharpe Ratio as in equation (4.9), the column Omega is the Omega ratio as in equation (4.11), where risk-free rate is the 1-Month Treasury Bill Rate and the column Sortino is the Sortino ratio as in equation (5.16), where risk-free rate is the 1-month Treasury Bill Rate. Annualized returns are reported in percentages.

Table 4.6 – Directional accuracy (DA) and directional forecast value (DV) of volatility. The holdout period is 250 observations

Single models	DA	Test of equal DA	DV
EGARCH	**0.536*****	1.00	**0.034**
NAGARCH	0.524***	0.32	0.033
APGARCH	0.509	0.02	0.033
DCC	0.507	0.01	0.027
GJR-GARCH	0.503	0.01	0.032
FIGARCH	0.488	0.00	0.028
GARCH	0.444	0.00	0.027
IGARCH	0.442	0.00	0.027
BEKK	0.441	0.00	0.027
Combinations of forecasts			
Mean	0.507	0.01	0.031
Median	0.489	0.00	0.029
H-MSE	0.482	0.00	0.032
H-HMSE	0.509	0.02	0.032
HR	0.509	0.02	0.032
EHR	0.485	0.00	0.031
Greer	0.477	0.00	0.033
Greer-mean	0.514**	0.07	0.033

Notes: Out-of-sample period: 02/01/2003 − 29/12/2017. Holdout=250 obs. The MSE $\times$ 10^6 (see also Blaskowitz and Herwartz (2009)). GARCH models are described in Table 4.1. H − MSE and H − HMSE denote the hierarchical procedure by Costantini and Pappalardo (2010) based on MSE and HMSE, respectively. HR and EHR indicate the hit/success rate and exponential of hit/success rates methods of combination (equations (4.4) and (4.5), respectively). Greer is the method of combination based on equation (4.6). The column 'DA' indicates the proportion of corrected predictions. *,** and *** indicate the significance of DA by the test of Pesaran and Timmermann (1992) at 10%, 5% and 1%, respectively. The highest DA value is reported in bold. The column 'Test of equal DA' reports the p-value of the test of equal proportions by Eberhardt and Fligner (1977) of DA for the pairwise comparisons among the best single model and all the other models and forecast combinations. The null hypothesis states that two models have equal DA, while the alternative one they do not. The column 'DV' reports the mean of directional forecast value. The highest DV value is reported in bold. The results for DV have been multiplied by 10^3.

Table 4.7 illustrates the performance of the GARCH models and combinations of forecasts for the short holdout period in the case of bear and bull periods. The findings for the bear period reveal similarities with the results reported in Table 4.3. Directional predictability is enhanced with five out of nine single models and all forecast combinations exhibiting statistically significant directional accuracy. Furthermore, the test of Eberhardt and Fligner (1977) shows that all these models and combinations are equivalent to the single best model. As for the bull periods, the combinations of forecasts are outperformed by the best model, that is EGARCH (with the NAGARCH being the second best). However, the mean DV results show that the Greer combinations achieve a value of 0.020 very close to 0.021 gained by EGARCH.

Table 4.7 – Bear and bull markets directional accuracy (DA) and directional forecast value (DV) of volatility. The holdout period is 250 observations.

	BEAR			BULL		
Single models	DA	Test of equal DA	DV	DA	Test of equal DA	DV
EGARCH	**0.578*****	1.00	0.153	**0.534*****	1.00	**0.021**
NAGARCH	0.576***	0.96	0.152	0.518**	0.20	0.020
GJR-GARCH	0.564***	0.71	0.156	0.493	0.00	0.019
APGARCH	0.558***	0.60	**0.160**	0.501	0.01	0.019
DCC	0.532*	0.22	0.089	0.509	0.05	0.020
FIGARCH	0.483	0.01	0.099	0.491	0.00	0.020
GARCH	0.474	0.01	0.136	0.435	0.00	0.015
BEKK	0.474	0.01	0.137	0.432	0.00	0.015
IGARCH	0.471	0.00	0.136	0.431	0.00	0.015
Combinations of forecasts						
Mean	0.544**	0.37	0.152	0.478	0.00	0.018
Median	0.538*	0.29	0.141	0.470	0.00	0.017
H-MSE	0.570**	0.83	0.150	0.500	0.01	0.019
H-HMSE	0.564***	0.71	0.151	0.503	0.01	0.019
HR	0.555***	0.54	0.156	0.482	0.00	0.018
EHR	0.538**	0.29	0.149	0.476	0.00	0.018
Greer	0.552**	0.49	0.149	0.505	0.02	0.020
Greer-mean	0.558***	0.60	0.150	0.510	0.06	0.020

Notes: Out-of-sample period: 02/01/2003 − 29/12/2017. Holdout=250 obs. The MSE $\times 10^6$ (see also Blaskowitz and Herwartz (2009)). GARCH models are described in Table 4.1. H − MSE and H − HMSE denote the hierarchical procedure by Costantini and Pappalardo (2010) based on MSE and HMSE, respectively. HR and EHR indicate hit/success rate and exponential of hit/success rates methods of combination (equations (4.4) and (4.5), respectively). Greer is the method of combination based on equation (4.6). The column 'DA' indicates the proportion of corrected predictions. *,** and *** indicate the significance of DA by the test of Pesaran and Timmermann (1992) at 10%, 5% and 1%, respectively. The highest DA value is reported in bold. The column 'Test of equal DA' reports the p-value of the test of equal proportions by Eberhardt and Fligner (1977) of DA for the pairwise comparisons among the best single model and all the other models and forecast combinations. The null hypothesis states that two models have equal DA, while the alternative one they do not. The column 'DV' reports the mean of directional forecast value. The highest DV value is reported in bold. The results for DV have been multiplied by 10^3. The results are presented for bear and bull markets.

The results for the recessions and expansions, reported in Table 4.8, are similar to the benchmark case (Table 4.4) with superior predictive performance concentrated in the recessions rather than the expansions. In this case, the H-MSE combination achieves the highest DA (0.557) followed by EGARCH. The best DV is attained by the APGARCH, possibly due to its asymmetric structure. In expansionary periods, the best performing models are EGARCH followed by NAGARCH and the Greer-mean combination in terms of directional accuracy. However, directional value is close to the value attained by EGARCH (0.021) for DCC (0.020) and the Greer forecast combinations (0.020).

Table 4.8 – Recession and Expansion periods directional accuracy (DA) and direction forecast value (DV) of volatility. The holdout period is 250 observations.

	Recession			Expansion		
Single models	DA	Test of equal DA	DV	DA	Test of equal DA	DV
EGARCH	0.554**	0.93	0.145	**0.538*****	1.00	**0.021**
GJR-GARCH	0.552**	0.89	0.148	0.494	0.00	0.018
NAGARCH	0.549**	0.83	0.144	0.522***	0.20	0.019
APGARCH	0.544**	0.72	**0.151**	0.503	0.01	0.019
DCC	0.544**	0.72	0.090	0.507	0.01	0.020
FIGARCH	0.496	0.09	0.099	0.489	0.00	0.019
GARCH	0.462	0.01	0.129	0.437	0.00	0.015
IGARCH	0.462	0.01	0.128	0.433	0.00	0.015
BEKK	0.462	0.01	0.128	0.434	0.00	0.015
Combinations of forecasts						
Mean	0.528	0.42	0.143	0.480	0.00	0.018
Median	0.509	0.19	0.146	0.473	0.00	0.016
H-MSE	**0.557*****	1.00	0.139	0.501	0.00	0.019
H-HMSE	0.544**	0.72	0.142	0.505	0.01	0.019
HR	0.538*	0.60	0.143	0.484	0.00	0.018
EHR	0.520	0.31	0.143	0.478	0.00	0.018
Greer	0.536*	0.56	0.133	0.507	0.01	0.020
Greer-mean	0.538*	0.60	0.142	0.512*	0.04	0.020

Notes: Out-of-sample period: 02/01/2003 − 29/12/2017. Holdout=250 obs. The MSE $\times$ 10^6 (see also Blaskowitz and Herwartz (2009)). GARCH models are described in Table 4.1. H − MSE and H − HMSE denote the hierarchical procedure by Costantini and Pappalardo (2010) based on MSE and HMSE, respectively. HR and EHR indicate hit/success rate and exponential of hit/success rates methods of combination (equations (4.4) and (4.5), respectively). Greer is the method of combination based on equation (4.6). The column 'DA' indicates the proportion of corrected predictions. *,** and *** indicate the significance of DA by the test of Pesaran and Timmermann (1992) at 10%, 5% and 1%, respectively. The highest DA value is reported in bold. The column 'Test of equal DA' reports the p-value of the test of equal proportions by Eberhardt and Fligner (1977) of DA for the pairwise comparisons among the best single model and all the other models and forecast combinations. The null hypothesis states that two models have equal DA, while the alternative one they do not. The column 'DV' reports the mean of directional forecast value. The highest DV value is reported in bold. The results for DV have been multiplied by 10^3. The results are presented for recession and expansion periods as dated by NBER.

The results for the volatility strategies, reported in Table 4.9, are similar to the benchmark case (Table 4.5) with the best performance being observed for the Hierarchical-HMSE forecast combination.

Table 4.9 – Evaluation of strategies. The holdout period is 250 observations.

Panel A: Strategy 1				
Single models	Return	SR	Omega	Sortino
APAGARCH	-1.22	-0.07	0.97	-0.01
EGARCH	-11.71	-0.68	0.86	-0.04
FIGARCH	4.22	0.24	1.03	0.01
GARCH	-4.96	-0.29	0.93	-0.02
GJR-GARCH	5.28	0.31	1.05	0.01
IGARCH	-8.21	-0.48	0.90	-0.03
NAGARCH	-7.26	-0.42	0.91	-0.03
DCC	-8.48	-0.49	0.89	-0.03
BEKK	-7.86	-0.46	0.90	-0.03
Methods of combination of forecasts				
Mean	5.28	0.31	1.05	0.01
Median	-132.49	-8.77	0.13	-0.44
Hierarchical-MSE	-9.23	-0.54	0.89	-0.04
Hierarchical-HMSE	28.61	1.67	1.37	0.09
HR	-3.80	-0.22	0.94	-0.02
EHR	-9.40	-0.55	0.88	-0.04
Greer (2005)	5.28	0.31	1.05	0.01
Greer (2005)-mean	-107.68	-6.78	0.23	-0.36
Panel B: Strategy 2				
Single models	Return	SR	Omega	Sortino
APAGARCH	2.03	0.14	0.96	-0.01
EGARCH	-3.21	-0.24	0.85	-0.05
FIGARCH	4.75	0.35	1.04	0.01
GARCH	0.16	0.01	0.92	-0.03
GJR-GARCH	5.28	0.31	1.04	0.01
IGARCH	-1.46	-0.11	0.88	-0.04
NAGARCH	-0.99	-0.07	0.89	-0.04
DCC	-1.60	-0.11	0.89	-0.03
BEKK	-1.29	-0.10	0.88	-0.04
Methods of combination of forecasts				
Mean	5.28	0.31	1.04	0.01
Median	-63.60	-5.29	0.13	-0.44
H − MSE	-1.97	-0.14	0.88	-0.04
H − HMSE	16.95	1.55	1.35	0.09
HR	0.74	0.05	0.93	-0.02
EHR	-2.06	-0.16	0.88	-0.04
Greer (2005)	5.28	0.31	1.04	0.01
Greer (2005)-mean	-51.20	-4.14	0.23	-0.36

Notes:Out-of-sample period: 02/01/2003 − 29/12/2017. Holdout=250 obs. The MSE $\times 10^6$ (see also Blaskowitz and Herwartz (2009)). GARCH models are described in Table 4.1. H − MSE and H − HMSE denote the hierarchical procedure by Costantini and Pappalardo (2010) based on MSE and HMSE, respectively. HR and EHR indicate hit/success rate and exponential of hit/success rates methods of combination (equations ((4.4)) and ((4.5)), respectively). Greer is the method of combination based on equation ((4.6)). The column Return is the annualized return, column SR is the annualized Sharpe Ratio as in equation (4.9), the column Omega is the Omega ratio as in equation (4.11), where risk-free rate is the 1-Month Treasury Bill Rate and the column Sortino is the Sortino ratio as in equation (5.16), where risk-free rate is the 1-Month Treasury Bill Rate. Annualized returns are reported in percentages.

Overall, the results presented in this section show that the forecast combination methods applied here are highly competitive against the best performer among the GARCH models (EGARCH) in terms of DA and DV. This is particularly true when a holdout period of 500 observations is considered. However, the combinations of forecasts attained some gains in accuracy even for a shorter holdout period, especially for DA in case of bear and recession periods, but not only. Therefore, when there is uncertainty about what model to be chosen, given the equivalent directional forecast performance, the forecast combination methods may be preferable (Armstrong (2001)).

4.4 Conclusion

This chapter aims to evaluate the performance of several GARCH model and methods of forecast combinations for future volatility. The empirical literature on volatility models and forecasting is vast. Univariate and multivariate GARCH models have been used for the prediction of volatility of stock returns, with a special focus on US data. Within this literature, the evaluation of the forecasting models has been mainly carried out by using standard accuracy measures, such as MAE and MSE. Further, those studies exploring the potential of the combination of the forecasts among the GARCH models have also used standard accuracy measures.

This chapter takes a further step and focuses on directional volatility forecasts. In doing so, the chapter uses the directional accuracy and directional forecast value measures proposed by Blaskowitz and Herwartz (2009, 2011). This is the major contribution of the chapter. The importance of direction of the forecasts is largely documented in the forecasting literature, and relates to the economic value of volatility forecasts. Being able to accurately forecast the volatility direction, volatility trading rules can be devised aiming at enhancing the profitability of investors.

Second, the chapter tests for the equivalence of the directional performance of the forecasting models using the test by Eberhardt and Fligner (1977)) based on the equality of proportions. In this way, one is able to establish whether there exists the best performer in terms of correct prediction of direction. This is the second added value of the chapter.

Third, the chapter not only looks at the forecasting ability of the single GARCH models, but it also evaluates the performance of four classes (eight schemes) of combination of forecasts. The first class contains the simple mean and median combinations and the second class contains the methods by Costantini and Pappalardo (2010) and Costantini et al. (2016a) based on a hierarchical encompassing-MSE procedure. The third and fourth classes are directional accuracy forecast combinations and are based on a weighted average forecast scheme for directional change (Greer (2005)) and the DA successes, respectively.

Lastly, this chapter assesses the volatility forecasts with respect to their economic value via constructing volatility investment strategies. This chapter uses two different strategies and several measures for evaluation.

The empirical analysis is conducted using daily data for the S&P 500 index over the period 3/1/2000-29/12/2017 and the realized variance is used as a proxy for the true unobservable volatility. One-step ahead forecasts are generated using a rolling window scheme and two different holdout periods of 500 and 250 observations, respectively. The bear and bull markets and recession and expansion periods are also considered in the analysis. The main results show that forecast combinations applied here are highly competitive against the best performer GARCH model in terms of both DA and DV, especially when a holdout period of 500 observations is considered. However, the combinations of forecasts attained some gains in accuracy even for a shorter holdout period. These results seem to suggest that when there is uncertainty about what model to be chosen, the methods of combination may be preferable in case of directional forecast equivalence among models and forecast combinations. Among the forecast combination methods, the Greer combinations that are based on DA should be preferred followed by the hierarchical methods. Moreover, simple combinations that work well when the evaluation measure is based on MSE do not necessarily perform well when DA is considered. Further, our results from the bear and bull markets and recessions and expansions indicate that models perform better in bad times. Finally, our volatility investment strategies point towards superiority of the H-HMSE combination method, followed by the simple mean and the Greer combination method.

Taking into account the instability in the performance of single forecasts and the uncertainty over the choice of a single model, our findings are valuable to both investors and financial institutions. They provide evidence in favor of the use of combination forecasting methodologies to improve directional forecasting accuracy and economic gains.

4.5 Appendix.

This section reports the results for the MSE based on a holdout period of 250 and 500 observations, respectively. The rolling-window is of 750 observations.

Table 4.10 – MSE-based model ranking. The holdout period is 250 observations.

Single models	MSE	Combinations of forecasts	MSE
APGARCH	**0.030**	Mean	0.036
EGARCH	0.032	Median	0.040
FIGARCH	0.032	H-MSE	**0.030**
GARCH	0.033	H-HMSE	**0.030**
GJR-GARCH	0.042	HR	0.036
IGARCH	0.042	EHR	0.036
NAGARCH	0.042	Greer	0.033
DCC	0.047	Greer-mean	0.033
BEKK	0.047		

Notes: This table reports the MSE of one-step ahead forecasts. Out of sample period: $02/01/2003 - 29/12/2017$. 250 observations are used from the out-of-sample period as the initial holdout period. The MSE $\times 10^6$ (see also Blaskowitz and Herwartz (2009)). GARCH model are described in Table 4.1. $H - MSE$ and $H - HMSE$ denote the hierarchical procedure by Costantini and Pappalardo (2010) based on MSE and HMSE, respectively. HR and EHR indicate hit/success rate and exponential of hit/success rates methods of combination (equations (4.4) and (4.5), respectively). Greer (2005) is the method of combination based on equation (4.6). The lowest MSE is reported in bold.

Table 4.11 – MSE-based model ranking. The holdout period is 500 observations.

Single models	MSE	Combinations of forecasts	MSE
APGARCH	0.032	Mean	0.036
EGARCH	0.034	Median	0.038
FIGARCH	0.034	H-MSE	**0.031**
GARCH	0.036	H-HMSE	0.036
GJR-GARCH	0.045	HR	0.032
IGARCH	0.045	EHR	0.036
NAGARCH	0.045	Greer	0.036
DCC	0.048	Greer-mean	0.039
BEKK	0.050		

Notes: This table reports the MSE of one-step ahead forecasts. Out of sample period: $02/01/2003 - 29/12/2017$. 500 observations are used from the out-of-sample period as the initial holdout period. For other details, see notes table 4.10.

Chapter 5

Multivariate Volatility Models and Risk Parity Portfolio

5.1 Introduction

Chapters 3 and 4 of this thesis were concerned with volatility models of a single asset. More specifically, the forecasting ability of various univariate models, specifications and their combinations in terms of magnitude of error and directional accuracy was examined. However, to better understand the linkages and interactions among markets it is important to study multivariate models too, which will help us estimate and forecast the variance-covariance matrix. The diagonal elements of the matrix contain the variances of the variables and the off-diagonal elements contain the covariances between all possible pairs of assets. Covariance matrix has many applications as for example in the modern portfolio theory following Markowitz (1959) or in the Capital Asset Pricing Model (CAPM) (Bollerslev et al. (1988); Ng (1991); Tse and Tsui (2002)).

Precise estimation of the covariance matrix is the first step towards the construction of a portfolio. Stage two, is asset allocation which is commonly expressed as a maximization problem of a utility function. The result should be a fully diversified portfolio. However, it seems like that is hardly right since the portfolio remains exposed to risk from the highly volatile assets since it allocates more weight to those assets (Qian (2016)). This realization became more apparent during and after the recent financial crisis of 2008 where risk-based approaches for portfolio construction, such as, the risk parity (or methods that focus on the tails of the distribution) started gaining popularity. Risk-based portfolios aim at allocating risk rather than capital by considering the risk contributions from each of the assets in the portfolio. The risk Parity (RP) approach tries to alleviate the heavy contribution by considering the risk contribution to the overall portfolio risk. In other words, RP is a risk-based approach that aims at the marginal risk contributions for every asset in a portfolio to be equal. In that way the exposure of the portfolio to an asset class is limited and therefore, the impact of big losses.

Risk parity has many applications and it is a popular strategy among alternative investments like hedge, mutual and pension funds. The first RP fund was the Allweather portfolio launched by Bridgewaters in 1996. However, RP funds gained popularity after the sub-prime mortgage crisis of 2008 as investors regard them as a safer strategy compared to common portfolios. Similarly, pension funds need to ensure that at the end they will be able to fund their liabilities and pay the pensions hence, RP adjusts for the risk by targeting a specific level. In the literature there are many studies which apply the risk parity approach to construct optimal portfolios such as Platanakis et al. (2019), Roncalli and Weisang (2016), and Qian (2005). The portfolios were a mix of traditional asset classes and hedge funds (Qian (2012)). We believe that this is the first work that applies the risk parity approach to a portfolio of only hedge funds.

More recently, this approach has been applied to built alternative indexes that can be used as benchmark. On August 2018 the S&P Risk Parity Indices were launched and on August, 17 2017 HFR launched a series of six, investable through HFR Asset Management, equally-weighted, annually rebalanced Risk Parity indices. They are "(...) designed to reflect the performance of the universe of managers that employ a portfolio allocation strategy based on targeting risk levels across the various components of an investment portfolio" (HFR (2019)). It has been estimated that RP funds' worth now is at at $500 billion (Bhansali and Harris (2018)). RP portfolios tend to outperform traditional asset allocation optimization techniques (Costa and Kwon (2020); Bessler et al. (2021)).

The fact that RP hedge funds have started to gain wide acceptance while remaining a somewhat enigmatic strategy enhances the need for better measurement and benchmarking of their performance. One of the characteristics of hedge funds is that the availability of data on commercial databases depends on the manager's discretion [1]. Further, due to the lack of regulations, hedge funds returns suffer from an number of biases and it is very hard to find a suitable benchmark. Given the lack of an appropriate benchmark for RP Hedge funds, before the launch of the 6 HFR Investable RP Indices (HFR RP Vol 10, HFR RP Vol 10 Institutional, HFR RP Vol 12, HFR RP Vol 12 Institutional, HFR RP Vol 15 and HFR RP Vol 15 Institutional), this study will explore whether the created RP portfolios of hedge funds with pre-specified volatility targets can outperform the benchmark Indices. To the best of our knowledge this is the first study to compare a risk parity portfolio of hedge funds with the right benchmark. The categorization of HFR RP indices in terms of the volatility target will guide our analysis too. More specifically, this study aims to construct a risk parity portfolio consisting of hedge funds, obtained from the BarclayHedge hedge fund database for the period from January, 1994 to August, 2015. This portfolio is designed in an attempt to outperform the HFR Risk parity

[1] Hedge funds report voluntarily to commercial databases (Joenväärä et al. (2019a)).

Indices out-of-sample. The constructed Fund of hedge funds (FoHFs) consists of the best twenty five performing hedge funds in terms of alpha and beta investment performance measures. Then, the covariance matrix is estimated using the 7-factor model of Fung and Hsieh (2004) and then utilized to extract the weights for the risk parity portfolio. Our results suggest our constructed portfolio have a better performance than the benchmark Indices while the superiority of alpha or SR measure remains open.

This chapter is organized as follows, first there is a literature review around hedge funds, risk parity and HFR, followed by an empirical application using hedge funds from BarclayHedge database.

5.2 Literature Review

5.2.1 Multivariate volatility models

Asset allocation in portfolio management fundamentally relies on multivariate modeling. Multivariate models are prone to suffer from two serious issues. The first one is the curse of dimensionality which renders the estimation of the models infeasible in large dimensions. The model needs to be flexible but this requires many parameters and as the dimensions of the covariance matrix increases the number of parameters also increases rapidly. The second issue concerns the incertitude about imposing positive definiteness of the conditional covariance matrix (Silvennoinen and Teräsvirta (2009b)). Another common complication arises from the fact that some specifications that have been suggested as a solution to the two aforementioned issues produce parameters with no direct interpretation (Silvennoinen and Teräsvirta (2009b)). The literature on multivariate volatility models is vast, see Chapter 2, section 2. One type of model seems to overcome both problems is the factor volatility models.

Multivariate factor GARCH and stochastic volatility models were first analyzed by Diebold and Nerlove (1989) and Engle et al. (1990). "Factor models are motivated by economic theory" (Silvennoinen and Teräsvirta (2009b), p.207). Various assets tend to move together as a reaction to new information. Therefore, identifying common underlying factors in the series under study that drive the return series is crucial for volatility estimation and forecasting. Factor models are successful at decreasing the number of parameters as long as the number of factors is relative smaller than the number of assets (Silvennoinen and Teräsvirta (2009b)). Recent year research has shown that fund returns are depending on Fung and Hsieh (2004) factors and this type of model is well-fitted for forecasting returns.

Engle et al. (1990) assumed that the conditional variance is modeled through K number of factors,

not necessarily uncorrelated. The model can be written as:

$$\Omega_t = C + \sum_{k=1}^{K} w_k w_k' f_{k,t} \tag{5.1}$$

where C is an $N \times N$ positive semi-definite matrix, w_k, $k = 1, ..., K$, are linearly independent $N \times 1$ vectors of factor weights, and the $f_{k,t}$ are the factors which follow a first-order GARCH (Silvennoinen and Teräsvirta (2009b), p.207). The model can be estimated using a two-step procedure. Diebold and Nerlove (1989) also contributed to the literature by proposing a factor model which uses a stochastic volatility model instead of a GARCH one. Having said that, imposing uncorrelated factors is more appropriate for financial data (Cueto et al. (2020)). An additional gain from adopting this assumption is that it is possible to describe the dynamics of the series using less number of factors. The intuition behind this argument is that if the engaged factors are uncorrelated then, they are truly different components that drive the returns. Within this framework, the original observed series contained in r_t are assumed to be linked to unobserved, uncorrelated variables, or factors, z_t, which usually follow a GARCH process through a linear, invertible transformation W: $r_t = Wz_t$. There various specifications proposed for the W matrix. One simple case is to consider an r-factor multivariate GARCH model

$$\epsilon_t = Bf_t + u_t \tag{5.2}$$

where B is a $k \times r$ matrix of full-column rank of so-called factor loadings, f_t is a $r \times 1$ vector containing the common factors and u_t is a $k \times 1$ vector of errors. It is assumed that f_t and u_t have conditional covariances matrices given by Σ_t and Λ_t, respectively. Λ_t is diagonal. Finally, and f_t are assumed to be uncorrelated. The preceding assumptions imply that the conditional covariance matrix of ϵ_t is given by (Franses and Van Dijk (2000), p.203):

$$\Omega_t = B\Lambda_t B' + \Sigma_t \tag{5.3}$$

The above model exists with or without a constant term. A famous illustration is the CAPM model where a portfolio's return is represented by the following regression model:

$$r_t^P = \alpha + \beta R_t + \epsilon_t \tag{5.4}$$

In that case the *beta* coefficient is given as $\beta = \dfrac{Cov(r_t^P, R_t)}{Var(R_t)}$ and R_t is the market return. Extensions of the basic model have been put forward so the return of the portfolio depends on multiple risk factors

$$r_t^P = \alpha + \beta_1 R_{1t} + .. + \beta_n R_{nt} + \epsilon t \tag{5.5}$$

where R_i with $i = 1, .., n$. This model can be viewed as weighted sum of the individual asset sensitiv-

ities. The betas assigned to each individual risk factor can be derived from the covariances between the portfolio return and the risk factors and the risk factors' variances [2].

That represents the simplest, more basic, type of factor models. In the literature it has been suggested the idea of sparse factor model that is based on the assumption of a few weak factors that are allowed to affect only a subset of the available data. It is worth mentioning that the factors considered could either be the market model which is basically the single-index covariance matrix of Sharpe (1964) , or industry factors (for example, as defined by Fama and French (1997)), other factors that are not industry-based (for example, Asness et al. (2013)) or a common ARCH-Feature (Engle and Susmel (1993a)).

5.2.2 Asset Allocation

After estimating the moments of assets' returns using models presented on the section above, these estimates can be used in asset allocation in portfolio management. The section below describes first the classic portfolio allocation before introducing, the more recent, risk-based approaches. Markowitz presented a comprehensive portfolio theory in the following series of publications, Markowitz (1952), Markowitz (1956) and Markowitz (1959), forming the famous, Markowitz portfolio theory. Markowitz portfolio theory is based on the assumption that investors care simply about the mean and the variance (or standard deviation) of the returns of their portfolio and want to maximize the portfolio's returns (for a given level of risk) over a single period (single period mean-variance analysis). In other words, Markowitz simplified the investor's problem to a risk return analysis. Later on, the capital market theory of Sharpe (1964) complemented Markowitz's work.

More accurately investors want to maximize their expected risk premium which can be viewed as the compensation for bearing the risk associated with the expected returns they have been promised by investing on those risky assets. Classic portfolio theory assumes that investors are risk averse. This implies that for every two assets with equal expected returns, the investor will always opt for the asset with the lower risk (Brown (2012)).

A critical extension of this analysis is that an investor can reduce the portfolio risk through diversification. In this way every investors' objective is to keep adding securities in their portfolios, which are negatively correlated, preferably, so that the total risk of the portfolio can be diminished to the level of the systematic risk; that is the whole amount of risk that cannot be eliminated through diversification. Assets from the same country are usually positively correlated so there is need for securities

[2] That is: $\beta_i = \frac{Cov(r_t^P, R_{it})}{Var(R_{it})}$

from the international market. This forms the notable theoretical "Global (market) portfolio" that all investors hold. Considering these assumptions the investor's decision upon adding an asset to the portfolio depends on its covariance with all the other assets already in the portfolio. Sharpe (1964) expressed this condition in the Capital Asset Pricing Model (CAPM) as a linear regression of the asset's return and the return of the Market portfolio. Under this framework risk is expressed as the comovement of the individual asset with the market portfolio, i.e, asset's systematic risk. That is the beta coefficient and it is calculated as "the portion of an individual asset's total variance that is attributable to the variability of the total market portfolio"(Brown (2012), page 20). Nevertheless, investors' risk preferences will determine their final position (Brown (2012)).

However, the above analysis is based on a number of unrealistic assumptions. One of them concerns the investor's ability to borrow and lend unlimited amounts at the risk-free rate. While lending any amount at the risk-free rate may be true, borrowing at the risk-free rate is inaccurate. Usually, the borrowing rate available to the investors is always higher than the risk-free lending rate. What is more, the risk-free asset can only be viewed riskless in the absence of inflation or for a short term period where the inflation risk can be neglected. In all the other cases nominal holdings are exposed to the uncertainty of need to be reinvested at an unknown real interest rate. In that case the risk-free asset needs to be replaced by inflation-indexed bonds where their long-term real payments are known but there is uncertainty about their short term value (Campbell and Viceira (2001)).

The investors' desire to optimize returns with a given level of risk, conventionally, represented as a liner a combination of mean and variance that depend on the amount invested at each asset. A naive assumption is to assign to every asset a weight equal to $1/N$. This is a very simplistic approach and it does not involve any estimation and it completely disregards the given dataset. Although, due to its simplicity, it is associated with lower-costs and it is mainly used a benchmark (DeMiguel et al. (2007)). Another simplistic approach is to employ strategically-weighted portfolios; see Bessler and Wolff (2015) for more information.

Despite, the linear relationship highlighted earlier between mean and risk in practice investors seek to maximize their utility. The most commonly-used utility function, for describing investors behavior, is one that demonstrates diminishing marginal utility of wealth and depends only on the expected return and the risk (Brown (2012)). To comply with Markowitz portfolio theory investors need to be risk-averse thus a standard concave utility function that depends on wealth W, is suitable. Then the maximization problem becomes:

$$maxE_tU(W_{t+1}) \tag{5.6}$$

subject to $W_{t+1} = (1 + R_{p,t+1})W_t$ (Campbell and Viceira (2001)). A methodology that has gained a

lot of attention is restricting the moments of the portfolio. In such a way, the mean-variance portfolio restricts expected returns so that they are identical for all assets. An alternate interpretation is that the investor ignores the expected returns (DeMiguel et al. (2007)). This is how the well-know minimum-variance portfolio can be derived. Its objective is to minimize the variance of returns subject to the weights summing to one (budget constraint) and the main advantage of this strategy is that it only requires the estimation of the sample covariance matrix since the mean is disregarded. That is often desirable as the estimation of the mean is often associated with higher estimation risk.

As in any minimization problem, more constraints may be introduced to the minimum-variance portfolio. One of them is to exclude short sales so that the weights are restricted to be positive. The short sales can be completely forbidden or restricted to a certain level. Nevertheless, introducing short sales constraints on the weights can be applied to any optimization problem like, to the mean-variance, to the minimum-variance or as well as to the Bayesian constrained portfolio. Jagannathan and Ma (2003) showed that imposing restrictions is equivalent to restricting the return moments. Namely, short sale constraints on the mean-variance portfolio (sample based one) is like shrinking the expected return toward its average whereas for the minimum-variance portfolio is analogous of shrinking the variance-covariance matrix (DeMiguel et al. (2007)). Then Jagannathan and Ma (2003) compared the performance and found that a long-only portfolio has almost an equal performance to portfolios which are based on shrank covariance using factor models or shrinkage estimators. Further restrictions that one can impose are the norm constraints, upper and lower boundaries which could also be proportional to the risky assets. However, it is also possible to have equal target-weight constraints (upper and lower) which also "force diversification" based on a constant (Grauer and Shen (2000), p.1258). Then the budget constraint may be relaxed and allow for borrowing.

Another stream of literature has focused on combining the existing techniques. For instance, it has been proposed to combine equal weights policy and the (constrained) minimum-variance portfolio or the mean-variance and the minimum-variance portfolio. Generally this is the idea of combining portfolios. That is achieved by first estimating the moments and then apply the shrinkage procedures directly to the portfolio weights. The mixture of two portfolios will be (DeMiguel et al. (2007)):

$$x^s = cx^c + dx^d, \; s.t \; \mathbf{1}_N^T x^s = 1 \tag{5.7}$$

where x^c and x^d are two reference portfolios chosen by the investor. However, empirically it has been found that the above optimization problems often produce extreme or "corner" solutions that contradicts investors' inclination towards diversification. That motivated Black and Litterman (1990) to develop their model. Their main idea was to combine the CAPM with the mean-variance portfolio theory in a manner that investors' (and portfolio managers) views can enter the problem. Hence,

a more diversified portfolio can be formed Black and Litterman (1990). In that way the expected returns combine market equilibrium and investor views. However, the original paper does not specify clearly how to derive the subjective return estimates and there is some confusion among researchers as to what the inputs should be. Bessler et al. proposed to employ "sample means as subjective return estimates (...)" and historic forecast errors to measure the reliability of these views (Bessler et al. (2017), p.9).

A limitation of all these approaches is that, inevitably, they are associated with estimation error from the inputs; the mean and the covariance of the returns. Thus, it may be better to rely on a statistical or a financial model, such as factor model, or constrain the portfolio weights (Grauer and Hakansson (1995); Grauer and Shen (2000); Ledoit and Wolf (2004a); Kourtis et al. (2012)). For example see also Frost and Savarino (1986) and Jorion (1986) for shrinkage estimators such as one that is a weighted average of the sample mean and the scaled shrinkage target. Finally, it is believed that matrix estimation based on unobservable factors, could reduce the estimation error.

Finally, it is worth mentioning that all those views express the interests of a myopic investor who at the end of each period seeks to maximize their wealth for the next one. Brown (2012) reviews how the investment life cycle affects the optimization problem.

5.2.3 Risk - based portfolio allocation approaches

The classic mean-variance approach has received a lot of criticism for high estimation errors, high portfolio variance, and finally for not achieving true diversification. To secure a more robust solution, risk-based or "smart beta" approaches have been proposed that they only rely on the estimation of the covariance matrix. Then various portfolios can be created such as equally-weighted, minimum variance, equally risk contribution and maximum diversification. Given the lower risk tolerance of investors after the global financial crisis of 2008, "smart beta" approaches have gained popularity. In risk budgeting the risk contribution of each asset class is set to be equal to a pre-defined value chosen by the manager of the portfolio. A special case is the equal risk contribution, known as risk parity where each class contributes the same amount of risk to the final portfolio variance. Risk Parity is an alternative to the traditional asset capital allocation as it seeks to allocate risk rather than assets. There are many measures available to quantify risk but the most common one is the volatility of returns. However, Value-at-Risk and other downside risk measures may be preferred. There are a few studies that have focused on proposing alternative risk measures which are also concentrated on the tails of the returns distribution (Bawa (1975); Fishburn (1977); Harlow (1991); Morton et al. (2006)).

RP portfolios tend to outperform traditional portfolio optimization techniques in terms of risk-

adjusted measures and portfolio turnover. That is because in classic asset allocation, such as the 60/40 portfolio (60% allocated to stocks and the remaining 40% to bonds), the risk contribution of equity tends to be higher than the one of bonds in terms of the total portfolio risk resulting in a non truly "balanced" portfolio (Qian (2016)). In contrast, risk parity portfolios target balanced risk contribution from stocks and bonds and potentially any other asset class (Qian (2016)). For hedge funds RP considers four main different components: equities, credit, interest rates and commodities, and attempts to spread risk evenly across the asset classes (HFR). In addition, Maillard et al. (2010) showed that the volatility of a risk budgeting portfolio is located between the volatility of minimum-variance portfolio and the equally-weighted portfolio.

On the other hand, Asness et al. (2012b) and Anderson et al. (2012) argue that only leveraged risk parity portfolios can outperform classic portfolios in terms of return. Actually, the idea that leverage portfolio with lower level of variance is attributed to Black et al. (1972). Leverage allows the hedge fund manager to tailor the expected returns and risk. There are many instruments that create leverage although, the most commonly-used is through borrowing of derivatives. Derivatives based leverage is the most adopted approach among hedge funds managers and investors. To control further for the downside risk risk parity funds apply volatility targeting. This technique seeks to target a constant level of volatility to reduce the effect of volatility fluctuations. This risk-based methodology, on one hand, is believed to be more robust and produce stable returns on the long term which leads eventually to better risk-adjusted performance. On the other hand, Rudin and Marr (2016) among others have argued that in this way investors' views are disregarded thus they propose a modification to correct for this. Finally, alternative assets like hedge funds have some unique characteristics that hinder the benefits of diversification (for example, non normal returns, higher estimation errors) and therefore a more careful approach based on equal risk contribution is required to achieve the perks of diversification (Platanakis et al. (2019)).

For the reasons stated above in modern risk parity portfolios volatility targets are part of the optimization problem. Volatility targeting can control for the fluctuations of the volatility in a more adept way than the minimum variance approach. For the hedge funds, this implies that their main strategy is Risk Parity and then they are diversified in terms of their volatility target. The fund manager targets annualized volatility level. However, the actual or realized volatility level for longer or shorter periods may be significantly higher or lower depending on market conditions from the forecasted/targeted volatility. Risk Parity allows investors to target specific levels of risk and to divide that risk equally rather than focusing on the specific dollar amounts invested in each component. Risk Parity funds are classified into 3 main volatility targets: 10%, 12% and 15%.

- Funds with a volatility target of 10% or less are classified as volatility target 10%.

- Funds with a volatility target of $> 10\%$ and $< 15\%$ are classified as volatility target 12%.
- Funds with a volatility target of 15% or greater are classified as volatility target 15%.

The 10% target is regarded as moderate volatility whereas, higher volatility approaches are those who target the 15% volatility target.

Assume we have n assets and their covariance is given by a symmetric matrix Σ. Risk parity aims at the marginal risk contributions for every asset in a portfolio to the total risk to be equal. That is defined as:

$$w_i \frac{\partial f}{\partial w_i} = w_j \frac{\partial f}{\partial w_j} \tag{5.8}$$

where w is the weight. Under the assumption of $f(w)$ being the total risk and f being the standard deviation of the portfolio, so that $f(w) = \sigma(w) = \sqrt{w'\Sigma w}$, then it holds that:

$$\Sigma_{i=1}^{n} w_i \frac{\partial \sigma}{\partial w_i} = \Sigma_{i=1}^{n} w_i (\frac{\Sigma w}{\sqrt{w'\Sigma w}})_i = \frac{1}{\sqrt{w'\Sigma w}} \Sigma_{i=1}^{n} w_i (\Sigma w)_i = \frac{1}{\sqrt{w'\Sigma w}} w'\Sigma w = \sigma(w) \tag{5.9}$$

So, we can rewrite equation 5.8 as:

$$w_i (\frac{\Sigma w}{\sqrt{w'\Sigma w}})_i = w_j (\frac{\Sigma w}{\sqrt{w'\Sigma w}})_j => w_i (\Sigma w)_i = w_j (\Sigma w)_j \tag{5.10}$$

The risk contribution can be standardized by the portfolio risk so that:

$$\omega_i \frac{(\Sigma \omega)_i}{\omega'\Sigma\omega} \tag{5.11}$$

where $(\Sigma\omega)_i$ is the i^{th} row of the $n \times 1$ vector $\Sigma\omega$ and $(\omega'\Sigma\omega)^{1/2}$ is the individual volatility. The portfolio weights are restricted to be positive and to sum to one. Then each asset's risk contribution ($\frac{\partial \sigma(x)}{\partial \omega_i}$) may be set be equal to a target of risk contributors, $b = (b_1, ..., b_n)'$ with $\sum_{i=1}^{n} b_i = 1$ and $b_i > 0$. Qian (2006) proposed the equal risk contribution. That can be viewed as a special case were $b = 1/N$. In that case the risk is the same for all the assets (see also Bruder and Roncalli (2012); Maillard et al. (2010); Roncalli (2013)). Then, the optimization problem (with short-sales constraint) in this case is given as:

$$minimize_{x>=0} [\frac{1}{2} x^T \Sigma x - \sum_{i=1}^{N} b_i log(x_i)] \tag{5.12}$$

where $w = x/(1^T x)$ and $b = 1/N * (volatility - target)$. Risk parity portfolios can be viewed as a special case of Markowitz portfolio optimization problems. In case where there are only two assets or in the more general case where the correlations between the assets are the same, the resulting risk parity

weights are inversely proportional to asset volatility: $\omega_i = \frac{\frac{1}{\sigma_i}}{\sum_{j=1}^{n}(\frac{1}{\sigma_j})}$. Thus, the higher the volatility of a component the lower its weight in the Risk Parity portfolio. Applying Risk Parity is making it possible to define the risk of the portfolio as the sum of the risk contributions. For the successful implementation of Risk Parity, careful selection of asset classes is critical (Chaves et al. (2011)). More specifically, while using many assets from various asset classes is important what should matter the most is to choose asset classes that capture quite uncorrelated risk premiums from a balanced mix of various risks. Qian (2016) highlighted that at least three types of risks should be considered; That is equity risk, interest rate risk, and inflation risk. So, correctly classification of asset classes to the right groups of risk premiums is very important. Further, for a better performance asset classes can be decomposed to sub-asset classes and Risk Parity approach can then be applied across asset classes, but also within asset classes (Qian (2016)).

In short Risk Parity is a portfolio allocation strategy that targets risk levels across the various components of an investment portfolio and divides that risk equally in order to achieve true diversification (Qian (2013)). "The goal of risk parity investing is to earn the same level of return with less volatility and risk, or to realize better returns with an equal amount of risk and volatility (versus traditional asset allocation strategies) "(HFR (2019)).

5.2.4 Hedge Funds

Hedge funds are alternative investments available to investors who do not want to settle for the risk and the returns of common portfolios. Hedge funds have great investment flexibility as they can invest in any asset class but also in various financial instruments and can take both, long and short positions. What also characterizes hedge funds is that they are illiquid (this affects the times where they can accept new capital) and that they are managed actively. Traditional absolute return hedge funds attempt to profit regardless of the general movements of the market (bearish or bullish). Hedge funds are known for being non transparent and for the lack of regulation. This feature along with the high fees that they charge ($\simeq 2\%$ of assets under management (AuM)) makes them accessible only to relatively wealthy investors. Although uncorrelated with other traditional asset classes hedge funds can face negative consequences from herding among other hedge funds (Bali et al. (2013)). Most of the funds report end-of the month Net Asset Values (NAV). After the recent global financial crisis action has been taken to impose stricter regulation to hedge funds.

Hedge funds' managers self-report to databases to make the funds' performance available to potential investors and academics. The main commercial databases are: BarclayHedge, EurekaHedge, HFR,

Morningstar (which has acquired the CISDM, MAR, Altvest, and the MSCI Barra's database), and the Lipper TASS (former Tremont TASS) and additionally the eVestment and the Preqin. However, data acquired from the various databases requires special treatment. Joenväärä et al. (2014) and Joenväärä et al. (2019a) expose the main biases that arise in the most popular databases and attempt to show how these biases are responsible for the differences observed in the hedge funds' performance.

The main biases deemed are the survivorship bias, backfill bias and the bias occurring from missing AuM values. More specifically, the survivorship bias concerns the funds that have stop reporting information or ceased operations. Then the funds are categorized as surviving funds and defunct funds. The attrition rate is used to measure the magnitude of the survivorship bias and it is defined as the ratio of the defunct ratio to the total number of funds each year. Lower attrition rate signifies lower survivor bias as there are less defunct funds included in the database. Also it is associated with better performance since defunct funds are worse performing that the surviving ones. Further, this higher performance of lower attrition rate is conditional to the fund strategy. Following, backfill bias also known as "instant history bias" concerns the historical record of a fund before the date that joined the database. It has been calculated that the average backfill period is 31 months so, by excluding the first 31 months of each funds data history it is possible to control for this kind of bias. Other studies use a period of 12, 24, 25 or 27 months (see Joenväärä et al. (2014) for a discussion). An alternative approach to correct for this type of bias is to discard all the observations before the listing date however this might not be possible as not all databases report when that day is. Finally, there have been identified two other types of biases graveyard bias [3] and incubation bias [4](Joenväärä et al. (2019a)). It is considered as good practice to control for these bias when discussing the effect that other biases have.

What is more, the bias related to missing AuM values is prone to affect the value-weighted performance. Joenväärä et al. (2014) after confirming their hypothesis that higher a number of missing observations leads to higher value-weighted returns they suggest that it is preferable to fill the missing vales (for example, by using the last observation) than dropping those observations from the sample. Finally, there is also a bias arising from the fact not all funds being represented by a commercial databases. However, based on studies about the total number of funds it can be said that the union of the funds reporting to a commercial database is a representative sample of the whole hedge fund universe. An also important finding is that funds which select to report to fewer databases are often associated with lower returns while, the opposite is true for the funds which volunteer to report in more than the average number of databases (Joenväärä et al. (2019a)). Other fund characteristics such as share restrictions (i.g lockup, notice and redemption period), compensation structure (for ex-

[3] Graveyard bias arises when the track record of a fund is discarded from the databases (Bhardwaj et al. (2014)).
[4]Incubation bias arises because funds' pre-release returns are only included in the database if they are favorable.

ample, incentive fees and high-water marks), investment style and domicile can also affect the fund's performance to a different degree according to Joenväärä et al. (2014). Finally, Joenväärä et al. (2019a) confirmed that performance persistence vanishes and it is primely leaded by small funds.

Commercial databases is the main source of information for hedge funds. Another information tool about hedge funds and their holdings is the U.S. Securities and Exchange Commission (SEC). Since 2013 all big institutional investment managers who exercise investment discretion over $100 million (on the last trading day of any month during any calendar year) are obliged to file a quarterly form in which they disclose all their portfolio holdings to the SEC. Managers generally should file the first Form 13F for the December quarter of the calendar year during which the fund first reaches the filing threshold and they should continue submitting at least for the following quarters of the same year even if their market values fall below the threshold. The managers should report shares of a foreign issuer only if those shares are traded on a U.S. exchange and should only report their long positions. Finally, managers can request for confidential treatment to prevent or delay their holdings being publicly available. Mutual and pension funds are also required to report their holdings.

Several databases provide investable hedge fund indices. These are very attractive to investors since they offer a direct way to measure hedge fund performance and gain expose to a specific hedge fund (Fung and Hsieh (2002); Dor et al. (2006)). The same applies to Funds of Hedge Funds (FoHFs) which are designed to closely track a broader index. The investors do not have to do the fund selection themselves, the risks are lower and the evaluation is easier. However, broadly based benchmarks, tend to "(...) mask interesting performance characteristics of hedge funds" (Fung and Hsieh (2002)). What is more, since all commercial databases suffer from a number of biases so do their indices.

There are many measures that attempt to assess hedge funds' performance. Hedge fund performance measures are an indicator for investment evaluation and/or fund selection. The first application will be explored in a separate section of this chapter. For the purposes of fund selection, the most common choices are alpha, beta or Sharpe Ratio measures and in some more rare cases the Information Ratio. The motivation is that these measures are better predictors of future performance than raw returns. Performance persistence has been researched a lot in the literature; see for example Kosowski et al. (2007), Boyson (2008), Fung et al. (2008), Ammann et al. (2013). Jensen's alpha is the most popular indicator of investment approach and it can give an indication whether the good perforce come from the manager's skills. For hedge funds the Fung and Hsieh (2004) (F-H hereafter) alpha is used which is a measure of risk that compares the portfolio to the market as whole. F-H seven-factor model includes three Trend-Following Risk Factors (Bond Trend-Following Factor, Currency Trend-Following Factor, Commodity Trend-Following Factor), two Equity-oriented Risk Factors (Equity Market Fac-

tor, The Size Spread Factor) and two Bond-oriented Risk Factors (The Bond Market Factor, The Credit Spread Factor). The beta, refers to the coefficient of the factors. Beta captures the risk that cannot be eliminated by diversification. It gives an indication of the additional risk introduced to the portfolio by a new investment. What is more, some criticize the use of standard measures for hedge funds since its returns are not normal and are prone to manipulation (Amin and Kat (2003), Goetzmann et al. (2007)). However, some researchers believe that performance persistence can not be guaranteed while some believe that managers have no skills and that they are just lucky. Nevertheless, that last argument is still questionable since luck can favor skilled and unskilled funds.

Hedge funds have an extensive list of strategies that can adopt. A non-exhaustive list of possible strategies are event-driven, equity hedge, equity market neutral, dedicated short bias, convertible arbitrage, fixed income arbitrage, global macro, emerging markets, managed futures, multi strategy and risk parity (Bali et al. (2013)). Then a hedge fund may also choose a sub-strategy which gives a more specific direction or target. "Risk Parity is an investment approach that is: Long-only, Systematic, Multi-asset including Stocks, Bonds, and Commodities, Focused on contribution to risk across the various components of the investment portfolio" (HFR (2019)). Risk parity considers four different components: equities, credit, interest rates and commodities, and attempts to spread risk evenly across the asset classes (HFR (2019)). Risk Parity allows investors to target specific levels of risk and to divide that risk equally rather than focusing on the specific dollar amounts invested in each component. Risk parity hedge funds have attracted more attention after the financial crisis of 2008. The first risk parity fund was the Allweather portfolio launched by Bridgewaters in 1996. A non-exhausted list of more Risk Parity funds includes: the AQR, the Invesco, the Putman, the Wealfront, the Bridgewaters, the JP Morgan, the Credit Swiss, the Resolve, the Natixis, the SSGA's. The main strategy of the Fund is defined as the risk parity. Then the risk parity funds are diversified in terms of their volatility target.

Risk parity investments are not restricted geographically and they can invest in a very wide range of financial instruments since the regulations are more relaxed. The Wealthfront Fund uses a broad range of asset classes including but not limited to global developed and emerging market equities, global developed and emerging markets fixed income, real estate investment trusts (REITs) and commodities. Also, the funds invests directly or indirectly in money market instruments such as: short-term U.S. Government securities, U.S. Treasury securities, money market mutual fund shares, and cash and cash equivalents. What is more, the Allweather account focus on the right allocation of stocks, bonds and commodities and the Aquila Capital Fund mainly invests across four asset classes: equities, government bonds, interest rates and commodities. Finally, the Clifton Group's Global Balanced Risk (GBR) aims at 10% volatility by investing in equities, government bonds, corporate bonds, and

commodities as well as in volatility and uses Futures (Exchange traded futures) to gain leverage.

Alternative assets like hedge funds have some unique characteristics and they face different types of risks. Further, the fact that hedge fund managers self-report to the databases implies that they may be tempted to act opportunistically on the way they choose to report asset values and this complicates more the calculation of risks and returns for hedge funds. Therefore, a more careful approach is required when selecting the right benchmark. A mismatch between the portfolio and the benchmark can distort the evaluation process (Qian (2016)). Capital-based (value weighted benchmark or 60/40 portfolio) or asset based benchmarks are no longer suitable for risk-based portfolios (Asness et al. (2012b); Bhansali et al. (2012)). Also the widely used S&P 500 index is not relevant and the same applies to S&P 500 bond index. The main reason is the risk concentration. Transitional investments "tend to have low risk-adjusted returns over time" in comparison to more "robust risk-adjusted returns" of risk-based approaches like risk parity (Alonso and Qian (2013), p.3). "Risk parity strategies are often measured against both absolute and relative return benchmarks. Talk about serving two masters!" (Qian (2016), p.116-117). Some tried to create their own benchmarks by combining for example, primary risk premiums and applying the RP strategy. What is more, in the literature various factor models have been constructed with traditional and/or alternative components of assets' returns that were linked to (observable) investable and non investable indices. The relatively new option-based benchmarks like the S&P Risk Parity Indices are using futures contracts from different asset classes and are attempting to capture the broadest class of hedge funds investments thus, they can be used as a benchmark for some investors but not for hedge funds and funds of hedge funds. "Agarwal and Naik (2004) find that the option-based indices explain a significant fraction of hedge fund return variation" (Aragon (2007)). However, having as a benchmark an investable hedge fund index that matched the exact strategy is the best. HFR Risk Parity Indices fully reflect the performance of risk parity investments for the targeted risk level.

5.2.5 Portfolio Performance Evaluation Measures

Portfolio evaluation is critical for the investment decision-making process. Portfolio evaluation involves calculating the expected return, the portfolio risk and risk-adjusted ratios like the Sharpe ratio (Sharpe (1966)). Most measures utilize only the first two moments of the returns' distribution. It is also common practice to compare the portfolio's performance measures to a benchmark portfolio such as the simple equally-weighted portfolio or to an index. Relative returns allow for the comparison of funds' returns with the benchmark's returns. In hedge funds multiple statistics are used in an attempt to capture the differences in management style and hedge fund style (Agarwal and Naik (2004)). What also dictates the need for multiple measures is the fact that hedge funds are alternative

investments and they experience some unique characteristics such as their returns do not follow the normal distribution. Finally, inevitably, all statistics suffer from drawbacks since they try to summarize loads of information in one single number. Thus, in theory calculating more statistics would give more accurate and robust results. However, in practice it has been indicated that most of the commonly-used statistics tend to be highly correlated with each other (Pfingsten et al. (2004); Pedersen and Rudholm-Alfvin (2003); Eling and Schuhmacher (2006); Eling and Schuhmacher (2007)). As a result the ranking resulting from the most sophisticated measures it is not much different from the one dictated by the Sharpe ratio even for hedge funds. Usually, considering only the first two moments of the distribution for performance evaluation is adequate.

Sharpe Ratio (SR) is very popular and sufficient measure for investments with returns that follow the normal distribution. It represents the risk premium return earned per unit of total risk.

$$Sharpe\ Ratio = \frac{E(R_p) - r_f}{\sigma} \tag{5.13}$$

where $E(R_p)$ is the portfolio returns and σ is the portfolio returns' volatility, r_f is the risk-free rate (3-month US T-Bill). However, Sharpe ratio suffers from estimation error through its moments. Further, it is prone to manipulation and intransparency has always been an issue for hedge funds (Goltz and Schröder (2010)). Finally, upside and downside risk are regarded the same which is not a realistic assumption. Thus, various extensions have been suggested in the literature to adjust for the above limitations. Risk can be substituted by other risk-measures such as Value-at-Risk (VaR). So replacing volatility with VaR it gives the Modified Sharpe ratio:

$$Modified\ Sharpe\ Ratio = \frac{E(R_p) - r_f}{VaR} \tag{5.14}$$

where VaR is the 1-step ahead VaR at 5%. What is more, the generalized Sharpe ratio (GSR) (Hodges, 1998) introduces utility function into the formula. This version falls back to the original statistic when returns are normal and if utility function is assumed to be exponential. Finally, the Adjusted Sharpe Ratio (ASR) uses a Taylor series expansion to account for higher moments of the returns distribution and the Conditional Sharpe Ratio, C-SR, proposed by Agarwal and Naik (2004) uses Expected Shortfall (ES) instead of the VaR.

When two Sharpe ratios are compared from two different strategies, it may be needed, to perform hypothesis testing. In more details a pairwise comparison of the two rations with respect to the two alternative portfolio strategies may be performed. In that case the null hypothesis is of equal Sharpe ratios. Jobson and Korkie (1981) proposed a test which can be performed based on the following t-statistic for two portfolios, {i,n} is given as: $\hat{z}_{jk} = \frac{\hat{\sigma_n}\hat{\mu_i} - \hat{\sigma_i}\hat{\mu_n}}{\sqrt{\hat{\theta}}}$ which asymptotically follows the

standard normal distribution. This holds under the assumption that returns are IID over time with normal distribution (DeMiguel et al. (2007); Jobson and Korkie (1981)). Later, Memmel (2003)) proposed a correction for the test above.

If an investor invests in multiple risky assets then the Sharpe ratio alone is not adequate for decision-making process (Eling and Schuhmacher (2007)). Evaluation measures that also include the market factor such as the Treynor or Jensen alpha are also necessary. In that way the use of financial market indices allows for evaluation of how much return can be attributed to the fund manager's skills. The Jensen's alpha measure can be calculated from a single (CAPM) or multi-factor regression (Fung and Hsieh (2004)). The main difference with the measures seen so far is that there is a different risk-free rate for each point in time rather than depending on the average of the risk free rate. Treynor (1965) formulated the first composite performance evaluation measure that combines risk and returns while introducing the market component and later Treynor and Black (1973) introduced the Appraisal Ratio

$$T_i = \frac{E(R_p) - r_f}{\beta_i} \tag{5.15}$$

It is independent of the investor's risk aversion and it is a risk-adjusted rate, where risk is beta and it reflects the return achieved per systemic unit of risk. It accounts for the correlation between portfolio returns and market returns, namely, the beta factor (β). The main difference between the Sharpe and the Treynor ratio is that the former considers the total risk and not just the systemic risk (ignoring the diversification potentials of the portfolio).

Another broad category of performance evaluation measures are the measures based on Lower Partial Moments (LPM) that only consider deviations that fall below an ex-ante defined threshold which usually is the risk free rate. These measures are also known as downside deviation measures. Downsides risk is defined as the risk below a threshold. The motivation for this adjustment is that upside risk is not a thread to the portfolio thus, it should be disregarded. Measures that fall into that category are the Sortino ratio by Sortino and Van Der Meer (1991) and the Omega ratio by Keating and Shadwick (2002). Sortino ratio is very similar to Sharpe ratio but, it only considers the downside risk (standard deviation).

$$Sortino = \frac{E(R_p) - r_f}{\sqrt{E[(r_f - R_p)_+^2]}} \tag{5.16}$$

where $E(R_p)$ is the portfolio returns and σ is the portfolio returns volatility, r_f is the risk-free rate (3-month US T-Bill), $\sqrt{E[(r_f - R_p)_+^2]}$ is the excess returns of the portfolio over the threshold. This approach may be more suitable to asymmetric return distributions. However, some argue that since it recognizes risk only the downside volatility it may not give the full perception of risk. Thus, Sortino

et al. (1999a) and Sortino et al. (1999b) suggested the Upside Potential Ratio (UPR) extension that considers the "upside potential" of portfolio too. Moreover, just like the Sharpe ratio, the Sortino ratio is prone to manipulation and estimation error due to its distribution moments inputs. Finally, it is susceptible to small sample problems.

Since hedge fund returns are usually not normal it is recommended to use the entire distribution for evaluation purposes. Omega measure considers the whole distribution and does not rely on distributional assumptions. The Omega ratio considers returns above and below the risk-free rate and is calculated as a ratio of the cumulative probability of an investment's outcome above the risk-free rate, to the cumulative probability of an investment's outcome below the risk-free rate. It can be seen as a probability-weighted ratio of gains to losses relative to the risk-free rate. It works for any utility function.

$$Omega = \frac{E[(R_p - r_f)^+]}{E[(r_f - R_p)_+]} \tag{5.17}$$

where $E(R_p)$ is the portfolio returns and r_f is the risk-free rate (3-month US T-Bill). Kazemi et al. (2004), suggested extensions of the Omega ratio. In this category we also find the kappa measure of Kaplan and Knowles (2004) that incorporates LPM of the first and third order.

As mentioned above the main measures are positively correlated to each other (Brown (2012)). In more details the Treynor and Jensen measures are connected with a positive relationship between the performance and the risk involved. Some argue that the Sharpe and the Sortino ratios are superior since they do not take into account the benchmark market portfolio. There is a number of requirements that a benchmark needs to fulfill such as to include asset classes that match the investor's risk preferences and investment needs and it is not guaranteed that an appropriate one exists for every case. Furthermore, both the Treynor and the Sharpe ratio have been criticized for ranking the portfolios in relative terms (not absolute). Finally, Omega has a low correlation with the Sharpe ratio but it tends to correlates more as returns are closer to the normal distribution and higher moments become less significant (Keating and Shadwick (2002)).

A quite different measure that has been introduced is the Manipulation Proof Performance Measure of Goetzmann et al. (2007) which is given as:

$$MPPM\ \hat{\Theta} \equiv \frac{1}{(1-p)\Delta_t} \ln(\frac{1}{T}\sum_{t=1}^{T}[(1+r_t)/(1+r_f)]^{1-p}) \tag{5.18}$$

where $\hat{\Theta}$ is an estimate of the portfolio's premium return after adjusting for risk, T is the total number of observations, and Δ_t is the length of time between observations and p is a coefficient for

risk-aversion ($p = 2$). Another measure is the Certainty-equivalent (CEQ) return which calculates how much risk-free rate an investor would demand in return of deciding not to invest in a risky portfolio (DeMiguel et al. (2007)). Finally, portfolios may be evaluated in terms of the return-loss measure. This measure, uses the naive, equally-weighted portfolio as the benchmark, to capture the additional return required by a strategy (k) to perform as good as the benchmark in term of the Sharpe Ratio (DeMiguel et al. (2007)).

5.3 Data and Methodology

Data

The data were downloaded from BarclayHedge database for the period from January 1994 to August 2015. To account for the survivorship bias, both defunct funds, from the graveyard database, and live funds were included. Some characteristics of the BarclayHedge database are that it has has the largest CTA and US funds (US-based) coverage, it does not include the listing date, it is survivorship bias-free since 1994, it is popular mutual fund database, and half of the lockup observations are missing (Joenväärä et al. (2014); Joenväärä et al. (2019a)).

The aim was to select and keep funds which match the opportunity set of the risk parity hedge funds and eventually the HFR risk parity index. Thus, the funds were filtered according to their main investment style. The main asset classes that risk parity hedge funds invest on are stocks, bonds, commodities, interest rates, but it can also include, emerging markets equity and fixed income, real estate investments trust, currency, inflation, volatility, ETFs. It is a long only, systematic, multi-asset investment approach. Also, it is known that this strategy depends on leverage through investments in derivatives to reach higher returns. The BarclayHedge database recognizes 96 categories. Obvious short-selling strategies, were filtered out. The table below gives an overview the strategies used in the risk parity portfolio. This is the based on the classification of the Hedge Funds strategies proposed by Joenvaara, Kosowski and Tolonen (2014). Thus the 96 recognized primary strategies by BarclayHedge database are mapped to 12 strategy categories. It can be seen that from those, only five were kept and the main contribution was made by long only funds.

Table 5.1 – Mapped Strategies used in the Risk Parity Portfolio

Mapped Strategy	% Keep
Long only	34.00%
Multi-Strategy	11.22%
Others	20.41%
Relative Value	19.39%
Sector	14.97%
SUM	100%

Note: This table reports the percentage contribution of each strategy to the risk parity portfolio. This is the based on the classification of the Hedge Funds strategies proposed by Joenvaara, Kosowski and Tolonen (2014). Thus the 96 recognized primary strategies by BarclayHedge database are mapped to 12 strategy categories.

Then after, filtering for the main strategy, the funds were filtered based on the evaluation criteria of HFR thus, all funds that were not reporting daily or in net-of-all fees or in USD dollars were omitted. Finally, a series of standard filters in the hedge fund literature were applied. To deal with the backfill bias, following Fung and Hsieh (2000), Joenväärä et al. (2019c) and Kosowski et al. (2007), the first twelve months from each fund's reporting history are omitted since BarclayHedge fails to report the listing day of the funds and therefore, Jagannathan et al. (2010) and **?**, method cannot be applied. Finally, all the funds with average assets under management (AUM) of less than 5 million dollars were filtered out. The initial dataset comprises 6,489 live and 16,478 graveyard funds forming a total of 22,967 unique funds. Once all the filters being applied the investor's opportunity set was comprised of 1625 funds in total; 672 live and 953 funds from the graveyard database.

Table 5.2 – Fund Characteristics

	All Funds			Active Funds			Graveyard Funds		
	Mean	Median	St.dev	Mean	Median	St.dev	Mean	Median	St.dev
Return Volatility	3.86	3.10	3.29	4.21	3.41	3.62	3.78	3.02	3.20
Funds per period	401.38	476.50	214.52	187.98	113.00	187.73	213.40	223.00	116.66
Redemption Notice	38.24	30.00	39.13	33.79	30.00	37.76	41.12	30.00	39.74
Lock Up	10.45	0.00	19.09	10.74	0.00	21.82	10.27	0.00	17.15
High Watermark	0.72	1.00	0.45	0.67	1.00	0.47	0.76	1.00	0.43
Management Fee	1.38	1.50	0.51	1.34	1.50	0.54	1.41	1.50	0.50
Performance Fee	14.63	20.00	8.55	11.86	15.00	9.04	16.59	20.00	7.60
Age	9.21	9.00	5.48	6.28	5.00	4.89	11.28	11.00	4.91
AUM	146.30	40.16	456.16	129.07	45.75	246.84	126.63	33.33	462.49
Leverage	0.78	1.00	0.41	0.85	1.00	0.36	0.74	1.00	0.44
Count	1625			672			953		

Note: This table reports the time-series average of the respective cross-sectional statistic for all the funds used in the formation of the risk parity portfolio for a given characteristic and separately for the active and the graveyard funds. The characteristics are the lock up period (in month, $\times 10^{-3}$), a dummy variable for hight watermark, management fee (in %), performance fee (in %), age (in years), assets under management (AUM, in million dollars, $\times 10^{-6}$), a dummy variable for leverage, return volatility (24-month volatility of monthly returns, in %). The last row simply reports the number of all funds, the number of active and the number of graveyard funds. The sample runs from January 1995 to August 2015.

Table 5.2 reports descriptive statistics for the post-filtering sample of hedge funds for the period from January 1994 to August 2015. Each column tabulates the time-series mean of a particular cross-sectional descriptive statistic (mean, median or standard deviation) for a given characteristic. The characteristics examined, consist of the number of funds per period, the 24-month volatility of returns, the fund's redemption notice and lock up period, a dummy variable for high watermark, the fund's management and performance, fees (in percentages), age calculated from the end of the sample since fund's inception day, AUM (in millions), a dummy variable for leverage. Descriptives are tabulated separately for the full sample, as well as for the live and graveyard subsamples.

The last row reports the number of all funds, of active and graveyard funds in the sample. The mean of the returns volatility is 3.86. However, active funds have a higher value of 4.21 led by a slightly higher volatility of volatility. The average AUM value of all the funds included in the portfolio is 146.30$ m while there are no significant differences observed between active and graveyard funds. High watermark and leverage were dummy variables, thus the median is always equal to 1. Grave-

yard funds, have on average 213.4 funds per period as opposed to 187.89 active funds per period. Finally, the average age of all funds is 9.21 with a standard deviation of 5.48 whereas graveyard funds have an average age of 11.28 and active funds, only, 6.28.

Hedge Funds Research (HFR)[5]

Hedge Fund Research, Inc. hereafter HFR was established on 1992. It specializes on the hedge fund industry, on hedge fund indices and their performance. HFR is survivorship bias-free from 1994 and exhibits a consistently high coverage of return, AuM and fund characteristics information. Also, Joenväärä et al. (2019a) stated that if a researcher is interested in a single database for research purposes then this should be the HFR database. HFR has constructed over 200 indices which have been used in many studies such as Hitaj and Zambruno (2016), Hitaj et al. (2019), Carhart et al. (2014), Boigner and Gadzinski (2015). The HFR indices can Investable (ticker ending with "I") or non-investable (ticker ending with "X").

On August, 17 2017 HFR launched a series of six, investable through HFR Asset Management, equally-weighted, annually rebalanced Risk Parity indices. These are the:

- HFR RT Vol 10
- HFR RT Vol 10 Institutional
- HFR RT Vol 12
- HFR RT Vol 12 Institutional
- HFR RT Vol 15
- HFR RT Vol 15 Institutional

The HFR risk parity Institutional Indices represent the performance of the largest managers in the risk parity fund industry. These funds must have assets under management of USD $\$500MM$ or greater at time of inclusion in the index. The first two indices which are targeting the 10% volatility date back to 2003 while, the ones that aim at the 12% volatility have data available from 2005 onwards. Finally, the last two indices have data only from 2015 onwards. The performance of the HFR Risk Parity indices prior to 2017 reflects the strategy and asset grouping classification of the constituent funds as of January 2017. The indices reflect annual rebalance but they are also rebalanced when a new fund is added.

[5] All the information about the construction of the Indices is retrieved from the HFR: `https://www.hedgefundresearch.com/` [Accessed on 20 June 2019]

The number of constituents is approximately 25 funds and the evaluation criteria for inclusion in the HFR risk parity indices are the following: to report monthly returns, to report net of all fees returns and to report assets in USD. The Risk parity funds are derived from the global risk parity funds universe and classified according to their volatility target. Then the relevant funds are filtered to create the strategy universe for a given target volatility relying on the following formula :

$$\delta(freq-12)\cdot\delta(fees)\cdot\delta(ISO-USD)\cdot(Vol-Vol_{Tgt})\cdot H(AUM_{t_o-1}-AUM\_min)\neq 0 \quad (5.19)$$

where $H(x)$ is a step function

$$H(x)=\begin{cases}1, x\geq 0\\ 0, x<0\end{cases} \quad (5.20)$$

and

$$\delta(x)=\begin{cases}1, x=0\\ 0, x\neq 0\end{cases} \quad (5.21)$$

- $freq$ is the reporting frequency (12=month, 4=quarterly)
- $fees$ is the returns net of all fees (o=yes, 1=0)
- ISO is the reporting currency
- Vol_{Tgt} is the volatility target of the fund: 10%,12% and 15%
- AUM_{t0-1} is the fund assets under management in USD at t_0
- t_0-1 corresponds to 1 months before the rebalance day t_0 (i.e the December prior to the January rebalance)
- AUM_min corresponds to the minimum AUM required for qualification ($500 million USD for the HFR Risk parity Institutional Indices and 0 for the rest)

In cases where a manager lists multiple funds with the same or similar investment profile HFR chooses only the most representative fund to be included to the HFR risk parity index. All of the 6 indices are equally-weighted. Then, the index is rebalanced annually the first business day of the calendar year and then evaluated 1 month prior to the rebalance day. At the time of rebalance all index constituents will be equally-weighted. The index is calculated 3 times per month. The HFR risk parity indices are first published as estimates but they are not finalized until two months have elapsed. HFR reports net of all fees in USD and the available time series contain monthly data. What is more, an index

constituent may be removed from the index if it fails to meet the evaluation criteria on or before the rebalance date or for other reasons related to diligence. In those cases, its weight will be equally assigned to the rest constituents. However, a fund may be added to the index between rebalance dates if it meets the evaluation criteria and again in that case the weights will be equally allocated to all the constituents (old and new).

The Index NAV is 1000 at inception ("$t = 0$"). The NAV changes are driven by the index performance which is defined as the percentage change in the value of the index from a previous date "$t - 1$" to current date "t". At rebalancing time "t_0", constituents are equally-weighted so that the NAV of the index is defined as:

$$NAV_{t_0} = NAV_{t_0-1} \times (1 + ROR_{t_0}) \tag{5.22}$$

The NAV of the index at any other times "t" is computed:

$$NAV_t = NAV_{t-1} \times (1 + ROR_t) \tag{5.23}$$

where ROR_t is the percentage change in the total value of the index from "$t-1$" to "t". If a constituent k ceases to be a constituent of the index at the given t' before the rebalance date its weight is then equally distributed to the remaining constituents in the index.

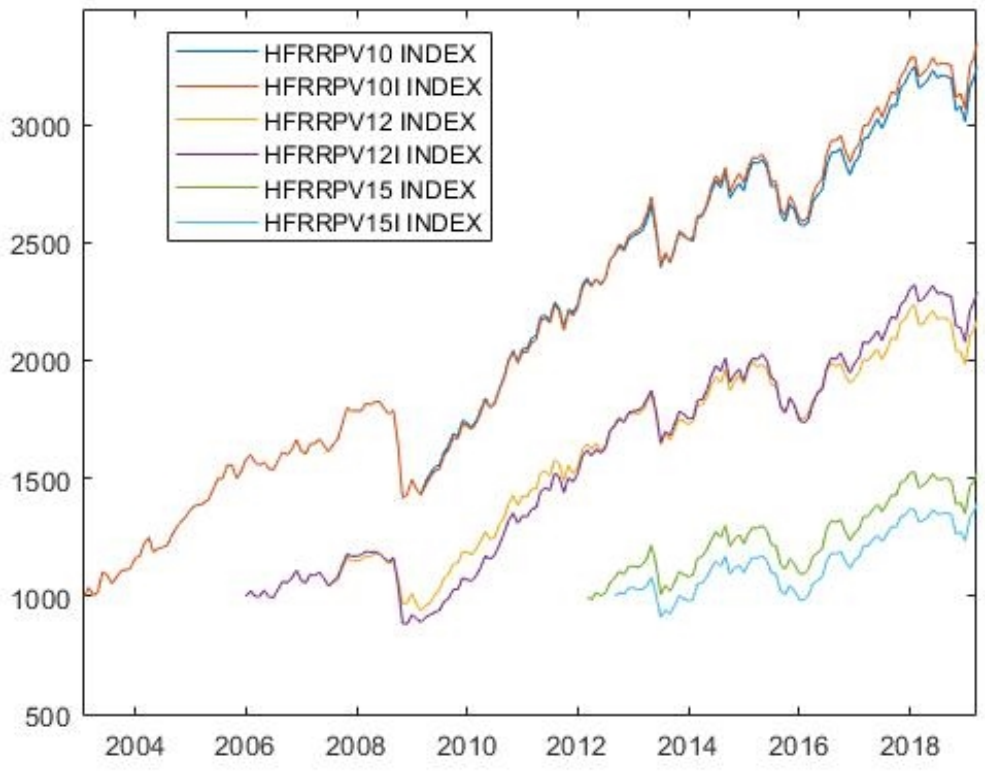


Figure 5.1 – HFR RP Indices (closing) monthly prices graph

Methodology

Our intention is to create a Fund of Hedge Funds(FoHFs) portfolio that follows the risk parity approach, in a way that it mimics the HFR RP indices. In a way our portfolio can be viewed as an easy investment where potential investors can "(...) gain exposure to a certain style or strategy without engaging in the selection of individual funds" (Dor et al. (2006)). To replicate the Indices from the broad selection of funds included only a fraction should be used to construct the FoHFs. An important decision is about the number of funds included. This decision was motivated by the number of hedge funds included to construct the HFR risk parity indices which is known to be around 25 but also from the literature. Joenväärä et al. (2019b) found that the majority of FoHFs hold between 16 to 30 hedge funds in their portfolios while Aiken et al. (2014) collecting data from the SEC EDGAR system found that FoHFs have in holdings 26.7 funds and on average FoHFs own 23.12 unique hedge funds. The number of funds is a trade off between a better parameter estimation and empirical fit. So, we chose to include 25 hedge funds in our portfolio.

Given the luck of not consensus in the literature as which is the best measure for the purposes of this study, hedge fund selection was performed based on statistical measure Sharpe Ratio (Capocci (2007); Manser and Schmid (2009)) and on Jensen's alpha (Capocci et al. (2005); Ammann et al. (2013); Boyson (2008); Fung et al. (2008)). In more details, the risk-adjusted measure, Sharpe ratio, was calculated and the Sharpe 25 funds with the highest ratio were selected. As for the second criterion, the portfolio's alpha as estimated using the Fung and Hsieh's (2004) 7-factor model and again the 25 funds with the higher alpha were selected. This is how the two portfolios were created namely, the alpha-sorted and Sharpe Ratio (SR)-sorted portfolios. Our results suggest that the question as to which measure should be preferred remains open.

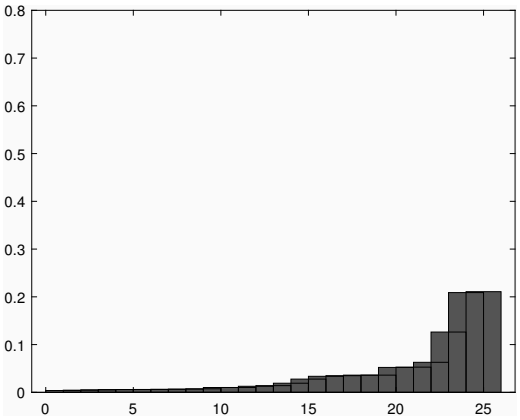
However, we should also keep in mind that the past cannot always mimic the future. Another factor that can effectively improve portfolio selection is accurate estimation of the covariance matrix of returns. Recent years of research have proven that fund returns are depending on Fung and Hsieh (2004) factors. So, the covariance matrix was calculated using F-H seven-factor model. To further improve estimation and forecasting accuracy, we imposed the restriction, on the funds included in the calculation of the covariance matrix to have at least 60 observations at the last day of the iteration period (see Joenväärä et al. (2019b)). A rolling forecasting scheme was adopted. The in-sample period for the 10% volatility target has 116 observations leaving the 132 observations from September 2004 to August 2015 for the out-of-sample. For the 12% volatility target, to match the data available from HFR, the in-sample period is 140 observations leaving the 108 observations from September 2006 to August 2015 for the out-of-sample. The index has less data available. For the HFRI15V and HFRI15VI there are even less data available. The weights were determined through the optimization

problem.

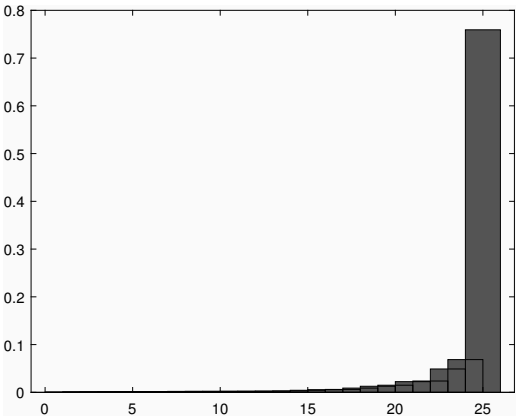
$$minimize_{x>=0}[\frac{1}{2}x^T\Sigma x - \sum_{i=1}^{N} b_i log(x_i)] \tag{5.24}$$

where $w = x/(1^T x)$ and $b = 1/25 * (volatility - target)$. Weights were fixed for 12 months.

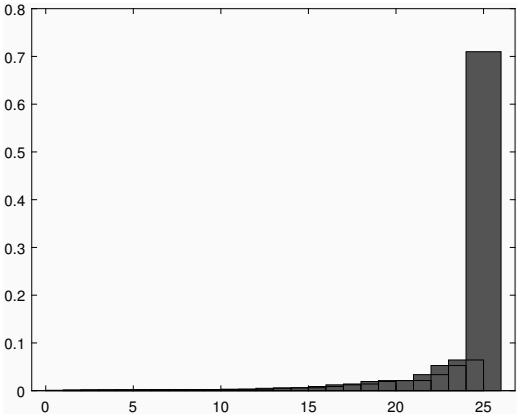
5.4 Empirical Results



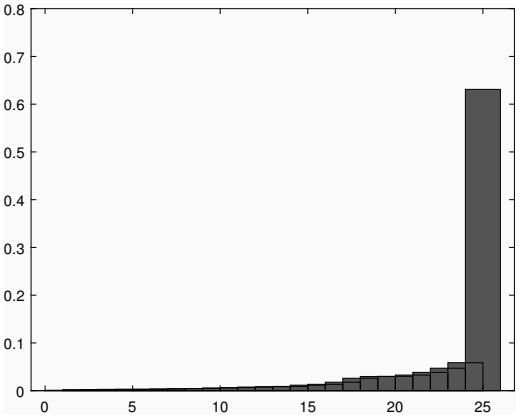
(a) Iter.



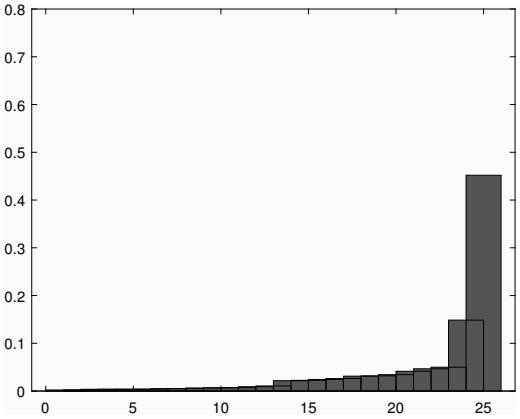
(b) Iter.



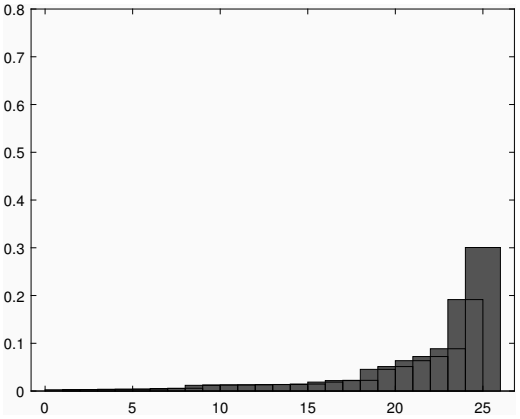
(c) Iter.



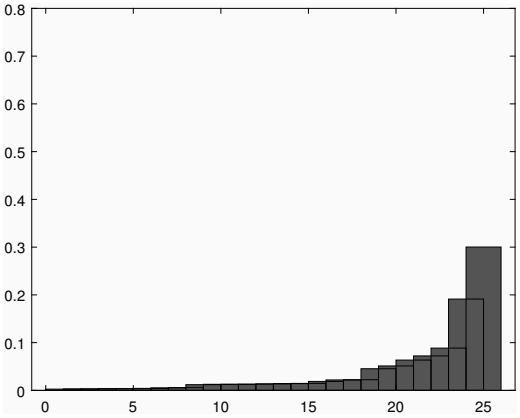
(d) Iter.



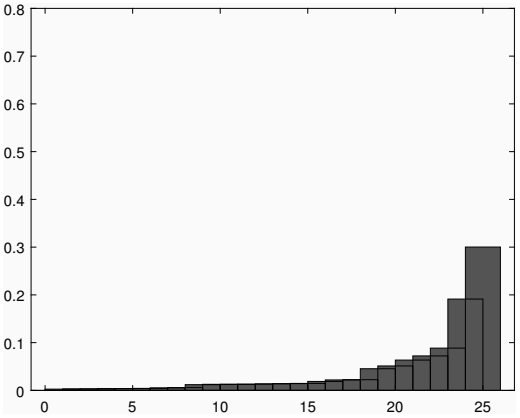
(e) Iter.



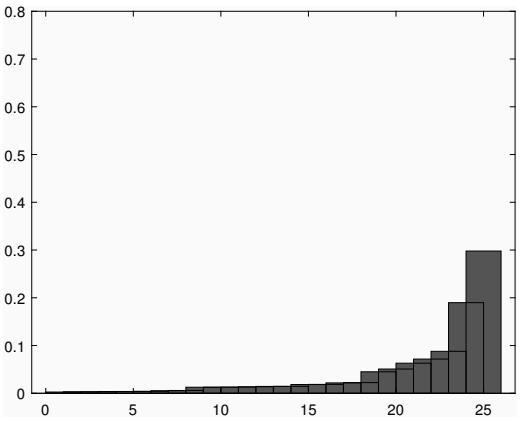
(f) Iter.



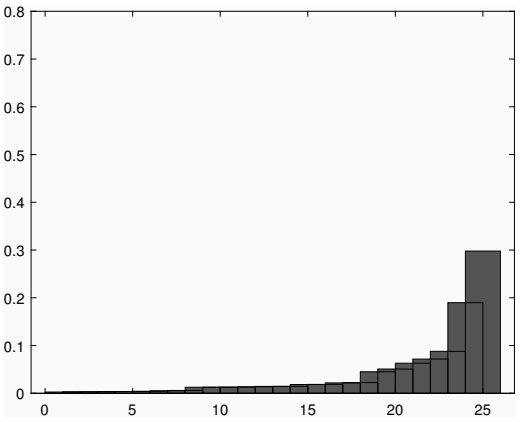
(g) Iter.



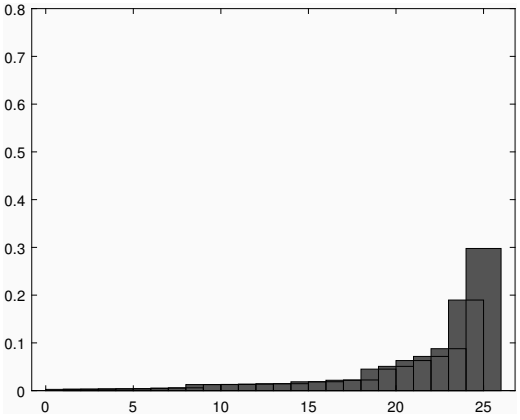
(h) Iter.



(i) Iter.



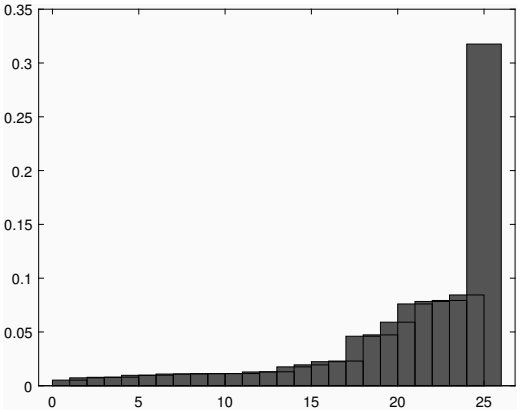
(j) Iter.



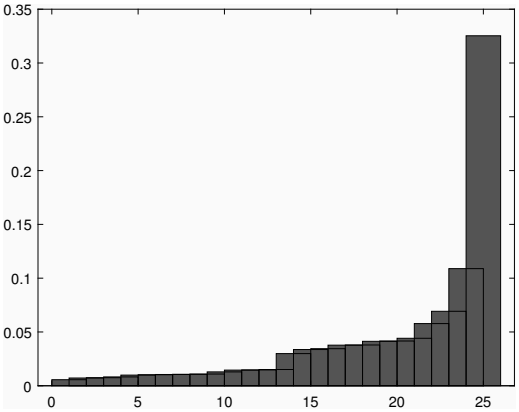
(k) Iter.

Figure 5.2 – Relative weights for the Alpha-sorted portfolio. Yearly rebalancing, 10% volatility targeting.

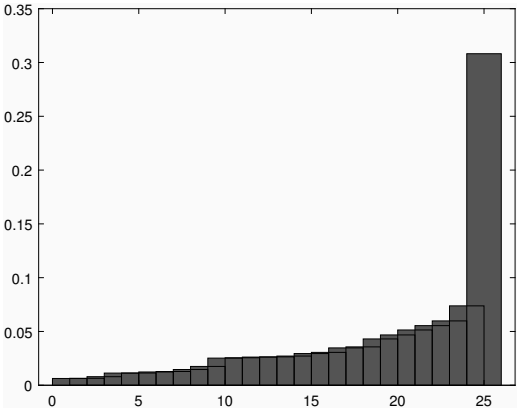
Notes: This figure reports the sorted, relative weights (sum to 1) for each iteration for the Alpha-sorted portfolio. Yearly rebalancing, 10% volatility targeting. Out-of-sample period: September 2004 to August 2015.



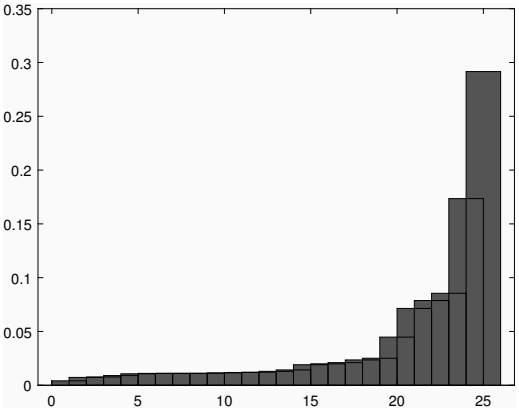
(a) Iter.



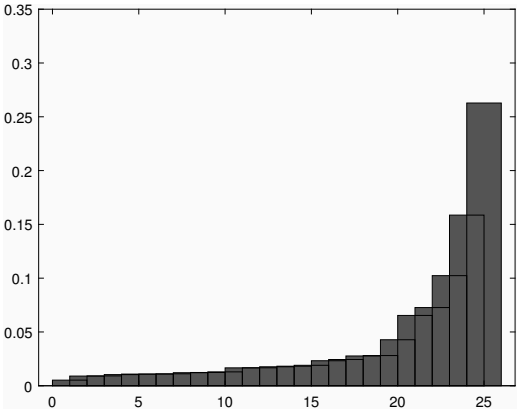
(b) Iter.



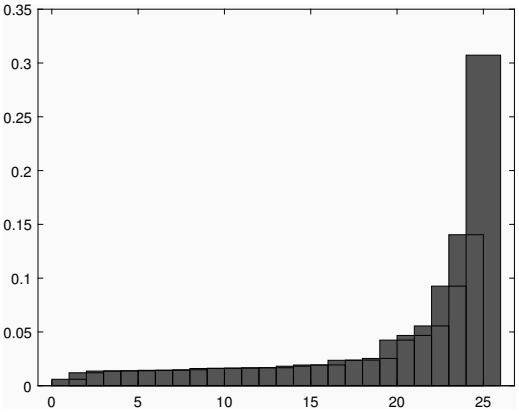
(c) Iter.



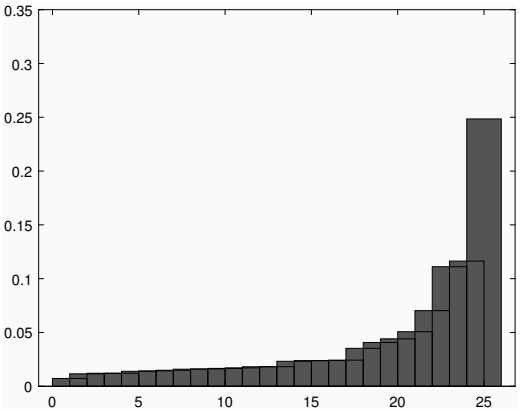
(d) Iter.



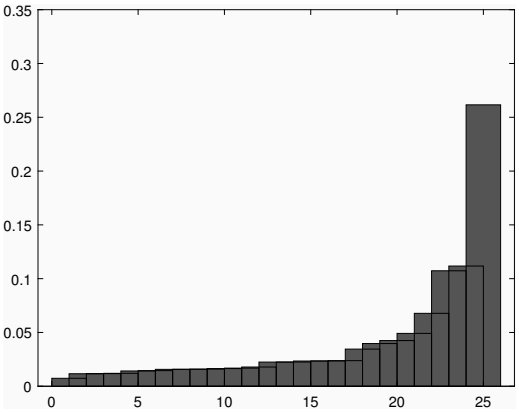
(e) Iter.



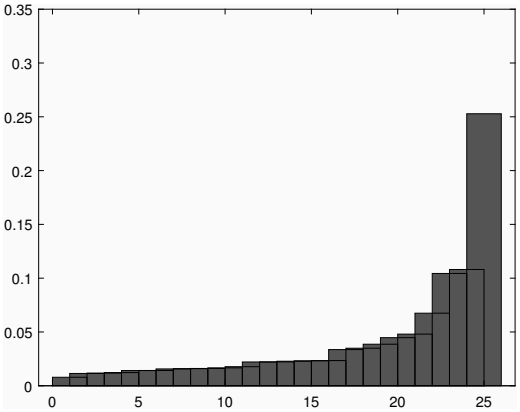
(f) Iter.



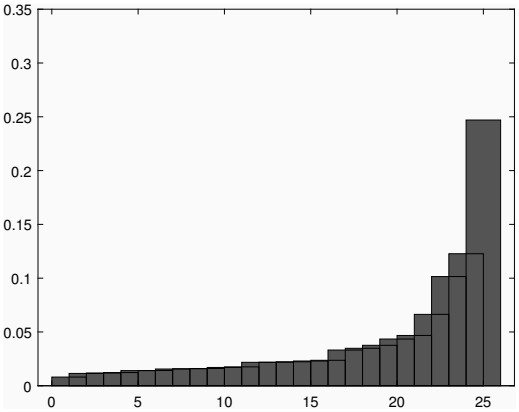
(g) Iter.



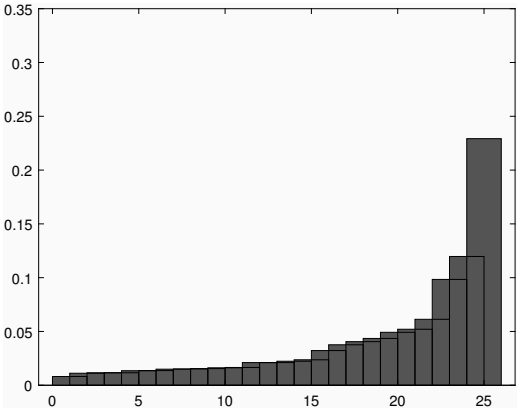
(h) Iter.



(i) Iter.



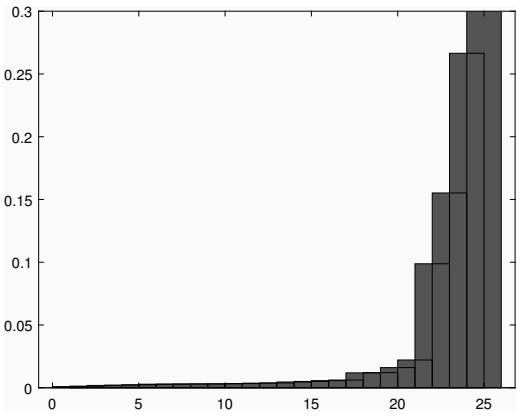
(j) Iter.



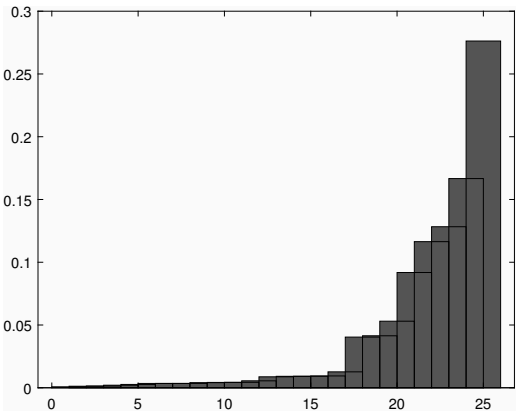
(k) Iter.

Figure 5.4 – Relative weights for the SR-sorted portfolio. Yearly rebalancing, 10% volatility targeting.

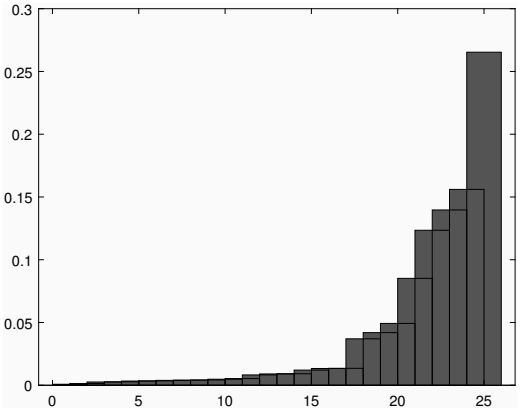
Notes: This figure reports the sorted, relative weights (sum to 1) for each iteration for the Sharpe Ratio (SR)-sorted portfolio. For the creation of the portfolio the first 116 observations were used to determine the first set of weights. Yearly rebalancing, 10% volatility targeting. Out-of-sample period: September 2004 to August 2015.



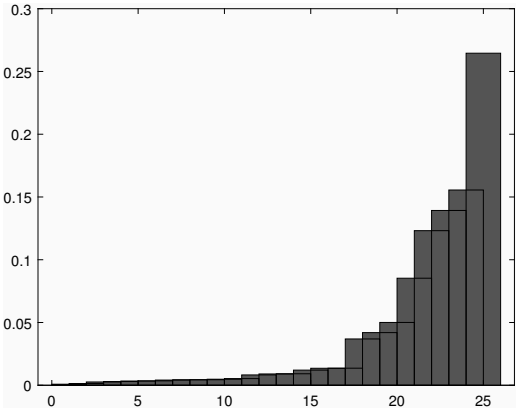
(a) Iter.



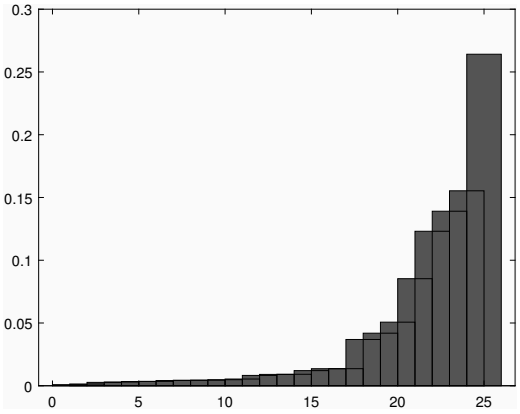
(b) Iter.



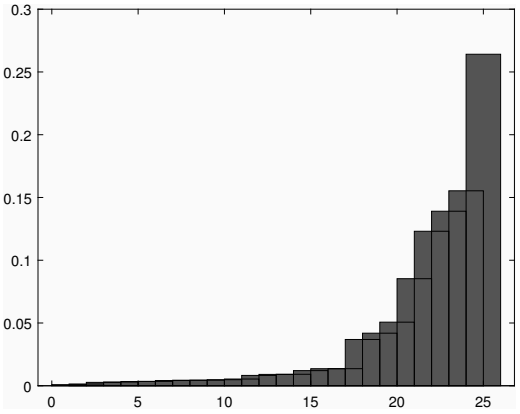
(c) Iter.



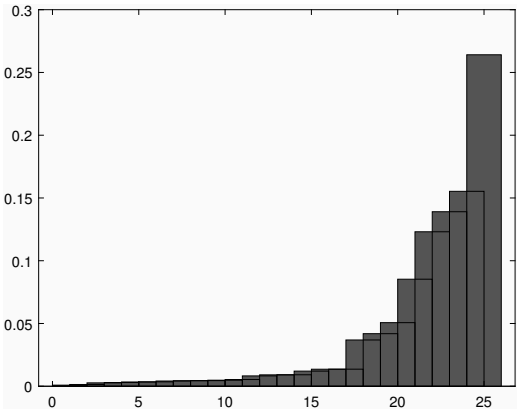
(d) Iter.



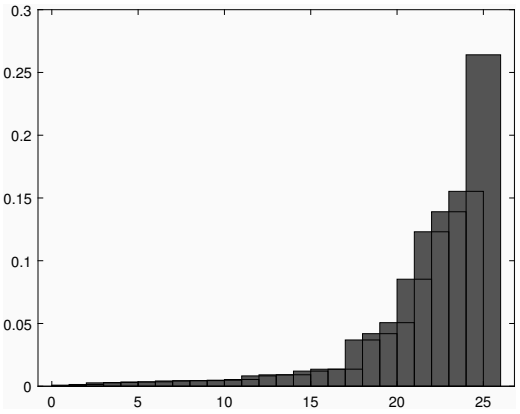
(e) Iter.



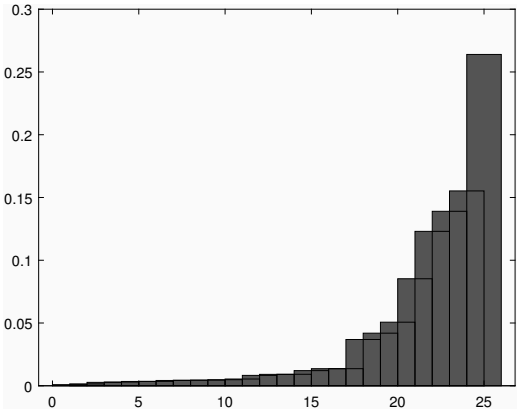
(f) Iter.



(g) Iter.



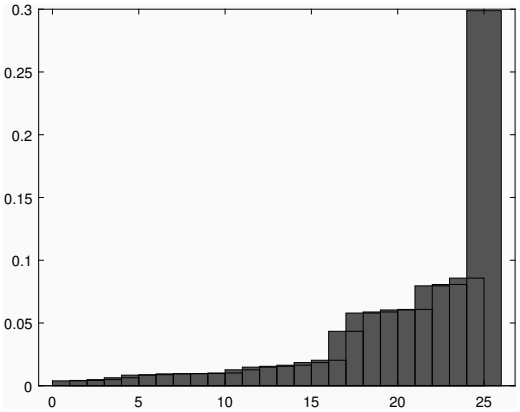
(h) Iter.



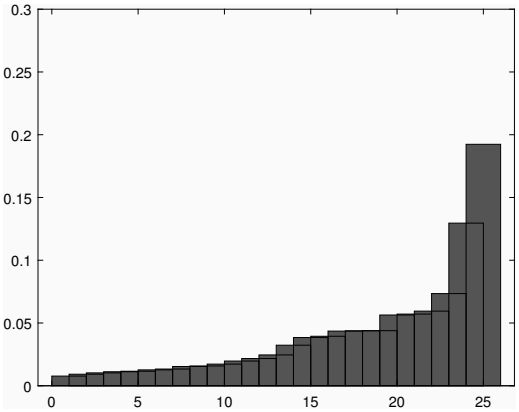
(i) Iter.

Figure 5.6 – Relative weights for the Alpha-sorted portfolio. Yearly rebalancing, 12% volatility targeting.

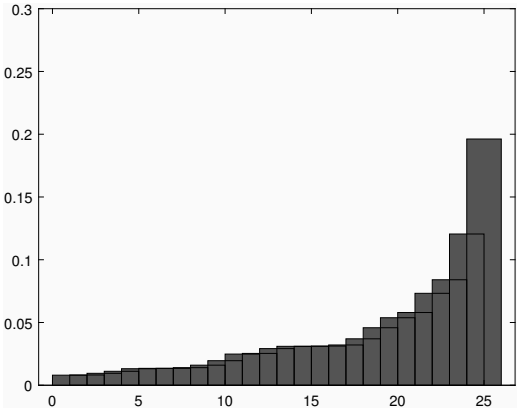
Notes: This figure reports the sorted, relative weights (sum to 1) for each iteration for the Alpha-sorted portfolio. Yearly rebalancing, 12% volatility targeting. Out-of-sample period: September 2006 to August 2015.



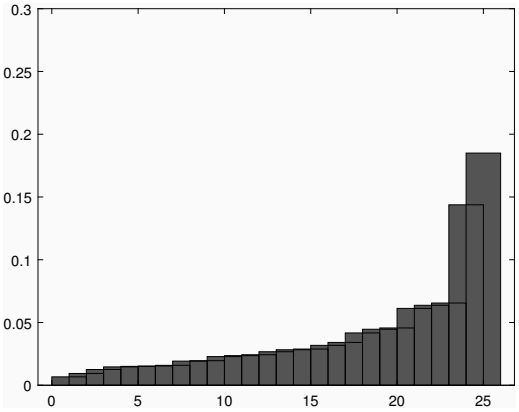
(a) Iter.



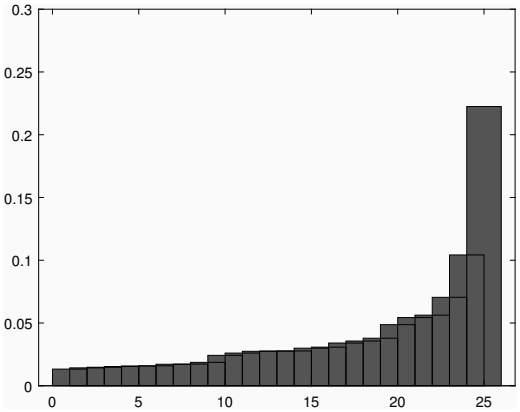
(b) Iter.



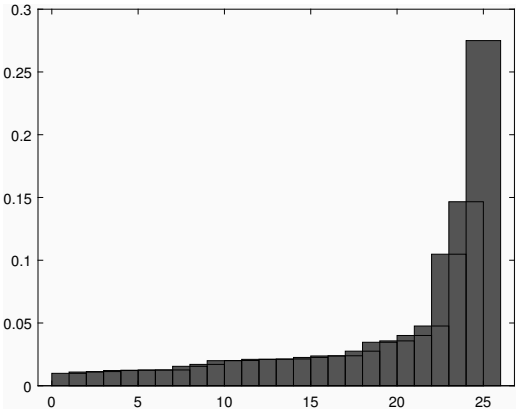
(c) Iter.



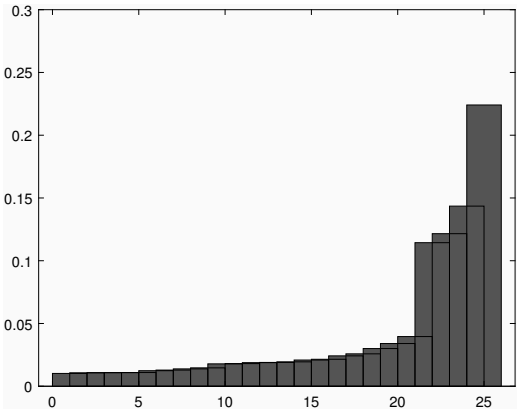
(d) Iter.



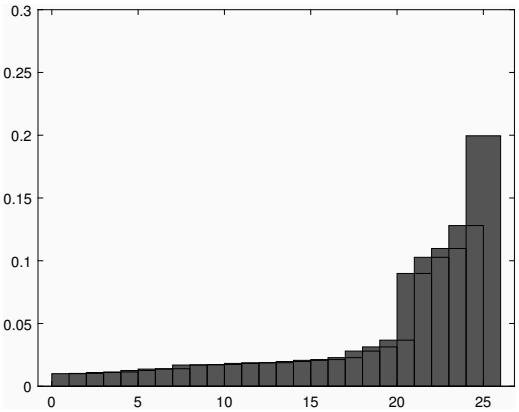
(e) Iter.



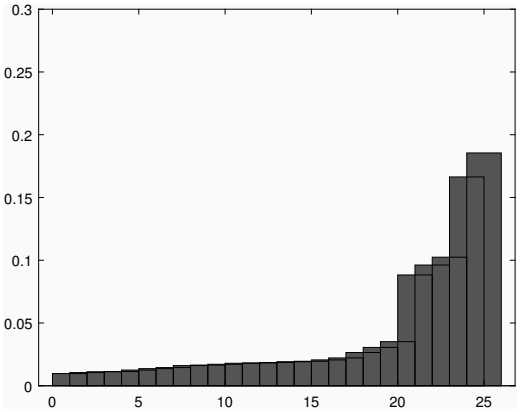
(f) Iter.



(g) Iter.



(h) Iter.



(i) Iter.

Figure 5.8 – Relative weights for the SR-sorted portfolio. Yearly rebalancing, 12% volatility targeting.

Notes: This figure reports the sorted, relative weights (sum to 1) for each iteration for the Sharpe Ratio (SR)-sorted portfolio. Yearly rebalancing, 12% volatility targeting. Out-of-sample period: September 2006 to August 2015.

We consider a set of pre-determined intervention times where the portfolio is rebalanced. These times are at annual frequency. For the 10% volatility targeting portfolios there are 11 iteration periods in the out-of-sample period while for the 12% targeting portfolio, there are 9. Figures 5.2-5.8 report the sorted weights of the all four RP portfolios for each iteration period. The weights are expressed as relative weights that sum to 1. From the plots above it can been seen that weights are homogeneous. It can also be observed that the weights decomposition is not concentrated in a few assets but rather is spread out. For the 10% target the maximum position held is a weight of 0.8 for the alpha-sorted portfolio and 0.35 for the SR-sorted portfolio. However, for the majority of the assets the allocated weight is around 0.3. Similarly, for the 12% targeted portfolios good diversification was also achieved with the maximum weight being 0.3. Finally, the SR-sorted portfolios appear to be more spread out than alpha-sorted portfolios.

Table 5.3 – Number of overlapping funds for Alpha-sorted portfolio at each period

Iterations	1	2	3	4	5	6	7	8	9	10	11
1	-	16	10	6	7	6	3	2	1	0	1
2	16	-	15	9	10	7	4	2	1	1	1
3	10	15	-	13	11	7	6	2	1	2	1
4	6	9	13	-	15	8	7	3	1	2	1
5	7	10	11	15	-	13	8	6	2	4	1
6	6	7	7	8	13	-	17	14	9	4	2
7	3	4	6	7	8	17	-	19	12	7	4
8	2	2	2	3	6	14	19	-	13	6	4
9	1	1	1	1	2	9	12	13	-	15	11
10	0	1	2	2	4	4	7	6	15	-	16
11	1	1	1	1	1	2	4	4	11	16	-

Note: This table reports for the Alpha-sorted portfolio, the number of funds in each period that are repeated at the next period. Yearly rebalancing, 10% volatility targeting. Out-of-sample period: September 2004 to August 2015.

Table 5.4 – Number of overlapping funds for SR-sorted portfolio at each period

Iterations	1	2	3	4	5	6	7	8	9	10	11
1	-	16	8	5	3	2	2	1	1	1	1
2	16	-	15	7	5	4	3	1	2	1	1
3	8	15	-	11	7	6	5	3	4	1	1
4	5	7	11	-	16	12	9	6	7	4	3
5	3	5	7	16	-	17	14	11	9	9	8
6	2	4	6	12	17	-	20	16	13	11	10
7	2	3	5	9	14	20	-	19	15	12	12
8	1	1	3	6	11	16	19	-	21	17	16
9	1	2	4	7	9	13	15	21	-	17	16
10	1	1	1	4	9	11	12	17	17	-	19
11	1	1	1	3	8	10	12	16	16	19	-

Note: This table reports for the Sharpe ratio (SR)-sorted portfolio, the number of funds in each period that are repeated at the next period. Yearly rebalancing, 10% volatility targeting. Out-of-sample period: September 2004 to August 2015.

Table 5.5 – Number of overlapping funds for Alpha-sorted portfolio at each period

Iterations	1	2	3	4	5	6	7	8	9
1	-	14	12	7	7	5	3	4	2
2	14	-	16	7	7	5	3	4	2
3	12	16	-	10	8	6	3	5	2
4	7	7	10	-	17	13	11	5	2
5	7	7	8	17	-	20	13	8	2
6	5	5	6	13	20	-	14	8	3
7	3	3	3	11	13	14	-	15	10
8	4	4	5	5	8	8	15	-	16
9	2	2	2	2	2	3	10	16	-

Note: This table reports for the Alpha-sorted portfolio, the number of funds in each period that are repeated at the next period. Yearly rebalancing, 12% volatility targeting. Out-of-sample period: September 2006 to August 2015.

Table 5.6 – Number of overlapping funds for SR-sorted portfolio at each period

Iterations	1	2	3	4	5	6	7	8	9
1	-	15	12	7	6	6	6	6	5
2	15	-	18	10	6	4	2	2	3
3	12	18	-	12	8	6	4	4	4
4	7	10	12	-	17	14	10	9	8
5	6	6	8	17	-	20	16	14	12
6	6	4	6	14	20	-	19	16	14
7	6	2	4	10	16	19	-	22	17
8	6	2	4	9	14	16	22	-	20
9	5	3	4	8	12	14	17	20	-

Note: This table reports for the Sharpe ratio (SR)-sorted portfolio, the number of funds in each period that are repeated at the next period. Yearly rebalancing, 12% volatility targeting. Out-of-sample period: September 2006 to August 2015.

Tables 5.3 - 5.6 report the number of overlapping funds at each iteration for all portfolios. Yearly our portfolios are reconstructed so we want to observe the number of funds overlapping at each iteration in terms of all the rest iterations. Looking at the Tables 5.3 - 5.6 it is observed that the overlapping for each portfolio and for each volatility target is quite low at each iteration period. So our strategy works and the yearly rebalancing is effective. The majority of the number of funds that overlap is less than 20. The highest overlapping is observed at the SR-sorted portfolio for 12% targeted volatility. So far major differences have not been observed between alpha-sorted and SR-sorted portfolios.

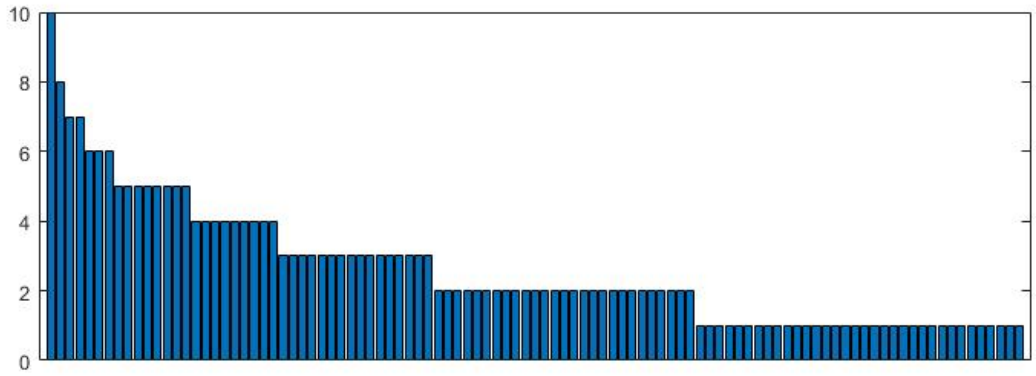


Figure 5.10 – Histogram of all funds used in the Alpha-sorted portfolio. Yearly rebalance, 10% volatility targeting

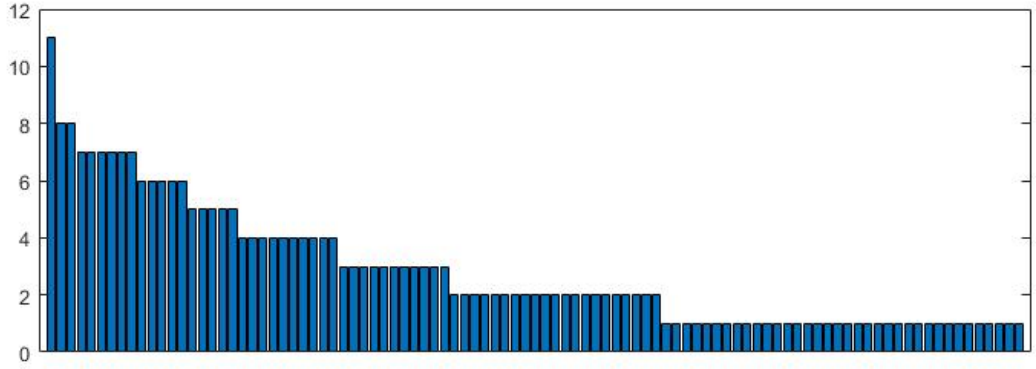


Figure 5.11 – Histogram of all funds used in the SR-sorted portfolio. Yearly rebalance, 10% volatility targeting

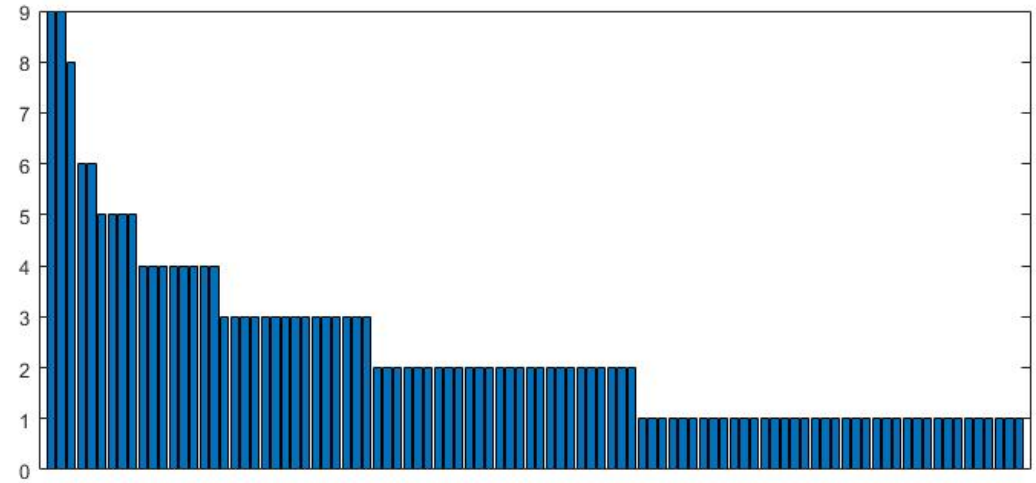


Figure 5.12 – Histogram of all funds used in the Alpha-sorted portfolio. Yearly rebalance, 12% volatility targeting

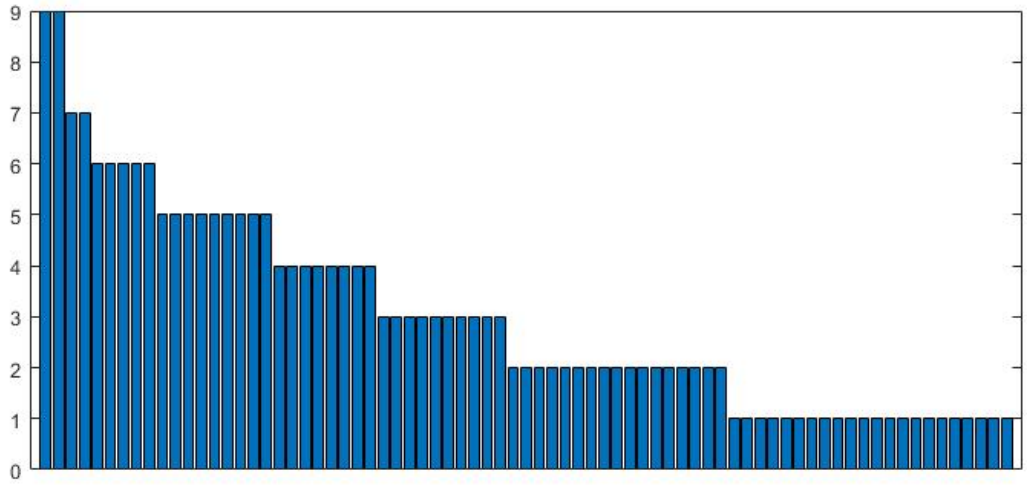


Figure 5.13 – Histogram of all funds used in the SR-sorted portfolio. Yearly rebalance, 12% volatility targeting

To have a better overview of the distribution of the funds used at each iteration for each portfolio the sorted histogram of all funds used in the alpha-sorted and SR-sorted portfolios for each volatility target was plotted. From Figures 5.10 - 5.13 it is confirmed that funds interchange within the portfolios. For the alpha-sorted portfolio at 10% volatility targeting there are 114 different funds and for the SR-sorted portfolio there are 97 out of the maximum 275 potential unique portfolios. Interestingly, the highest number of a specific fund appearing in the portfolio is 10 and 11 times respectively, for the alpha-sorted and the SR-sorted portfolio. This is in compliance with the findings observed from the tables above. Similarly, for the 12% volatility target portfolios out of the maximum of 225 unique funds, the constructed portfolio is comprised of 90 and 70 distinct funds respectively. The maximum number of a fund being represented in the portfolio is 9 times for both 12% portfolios.

Finally, the performance of our constructed risk parity FoHFs is evaluated in respect to the relevant HFRI Indices. Tables 5.7 and 5.8 report the evaluation measures for the constructed portfolios alpha-sorted and SR-sorted as well as the benchmarks (the HFR Risk Parity Index and the HFR Risk Parity Institutional Index) in respect to the volatility target. The performance measures calculated are the annualized (geometric) average return, the annualized average volatility, the Fung and Hsieh (2004) alpha measure, the Sharpe, Omega, Sortino, modified Sharpe (with VaR) and the Appraisal ratio and the the Manipulation Proof Performance Measure reported in basis points (bps). The portfolios are held unchanged for 1 year. The analysis is for the out of sample period.

Table 5.7 – Evaluation measures for Alpha-sorted, SR-sorted portfolios and HFRI10 and HFRI10I Indices

	Alpha-sorted	SR-sorted	HFRI10V Index	HFR10VI Index
Return	4%	11%	7%	7%
Volatility	10%	10%	8.82%	8.83%
Alpha	0.21	0.79	0.44	0.45
Sharpe Ratio	0.07	0.42	0.17	0.18
Omega	1.27	3.61	1.65	1.66
Sortino	0.07	0.51	0.14	0.14
Modified Sharpe Ratio	0.04	0.20	0.01	0.01
MPPM	0.001	0.001	0.001	0.001
Treynor Ratio	0.01	0.03	0.01	0.01

Note: This table reports the evaluation measures for the two constructed portfolios, Alpha-sorted and Sharpe Ratio (SR) as well as the two benchmarks, the HFR Risk Parity Vol 10 Index and the HFR Risk Parity Vol 10 Institutional Index. The performance measures examined are the annualized geometric mean return, the annualized average volatility the Fung and Hsieh (2004) alpha, The Sharpe ratio (Sharpe (1966)), Omega (Keating and Shadwick (2002)), and the Sortino ratio (Sortino et al. (1999b)), the modified Sharpe ratio with VaR, the Manipulation Proof Performance Measure of Goetzmann et al. (2007) and the Treynor and Black (1973) Appraisal Ratio in basis points. Yearly rebalancing, 10% volatility targeting. The sample runs from September 2004 to August 2015.

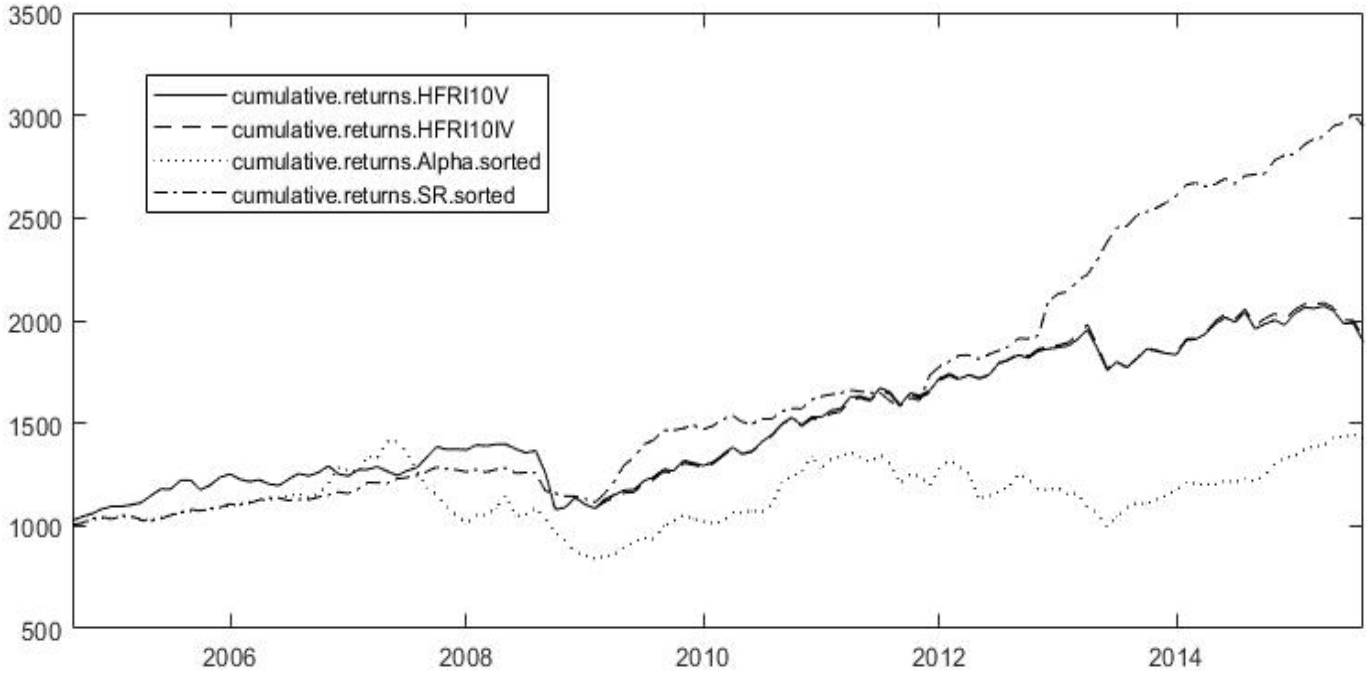


Figure 5.14 – Cumulative returns for the two constructed portfolios, Alpha and Sharpe Ratio (SR) as well as the two benchmarks, the HFR Risk Parity Vol 10 Index and the HFR Risk Parity Vol 12 Institutional Index. Initial investment 1000$.

Table 5.8 – Evaluation measures for Alpha-sorted, SR-sorted portfolios and HFRI12 and HFRI12I Indices

	Alpha-sorted	SR-sorted	HFRI12V Index	HFR12VI Index
Return	11%	6%	6%	6%
Volatility	12%	12%	9.65%	10.78%
Alpha	0.76	0.37	0.35	0.39
Sharpe Ratio	0.28	0.17	0.16	0.15
Omega	2.08	1.57	1.51	1.50
Sortino	0.28	0.15	0.15	0.12
Modified Sharpe Ratio	0.14	0.09	0.09	0.08
MPPM	0.002	0.001	0.002	0.002
Treynor Ratio	0.04	0.01	0.01	0.01

Note: This table reports the evaluation measures for the two constructed portfolios, Alpha and Sharpe Ratio (SR) as well as the two benchmarks, the HFR Risk Parity Vol 12 Index and the HFR Risk Parity Vol 12 Institutional Index. The performance measures examined are the annualized geometric mean return, the annualized average volatility the Fung and Hsieh (2004) alpha, The Sharpe ratio (Sharpe (1966)), Omega (Keating and Shadwick (2002)), and the Sortino ratio (Sortino et al. (1999b)), the modified Sharpe ratio with VaR, the Manipulation Proof Performance Measure of Goetzmann et al. (2007) and the Treynor and Black (1973) Appraisal Ratio in basis points. Yearly rebalancing, 12% volatility targeting. The sample runs from September 2006 to August 2015.

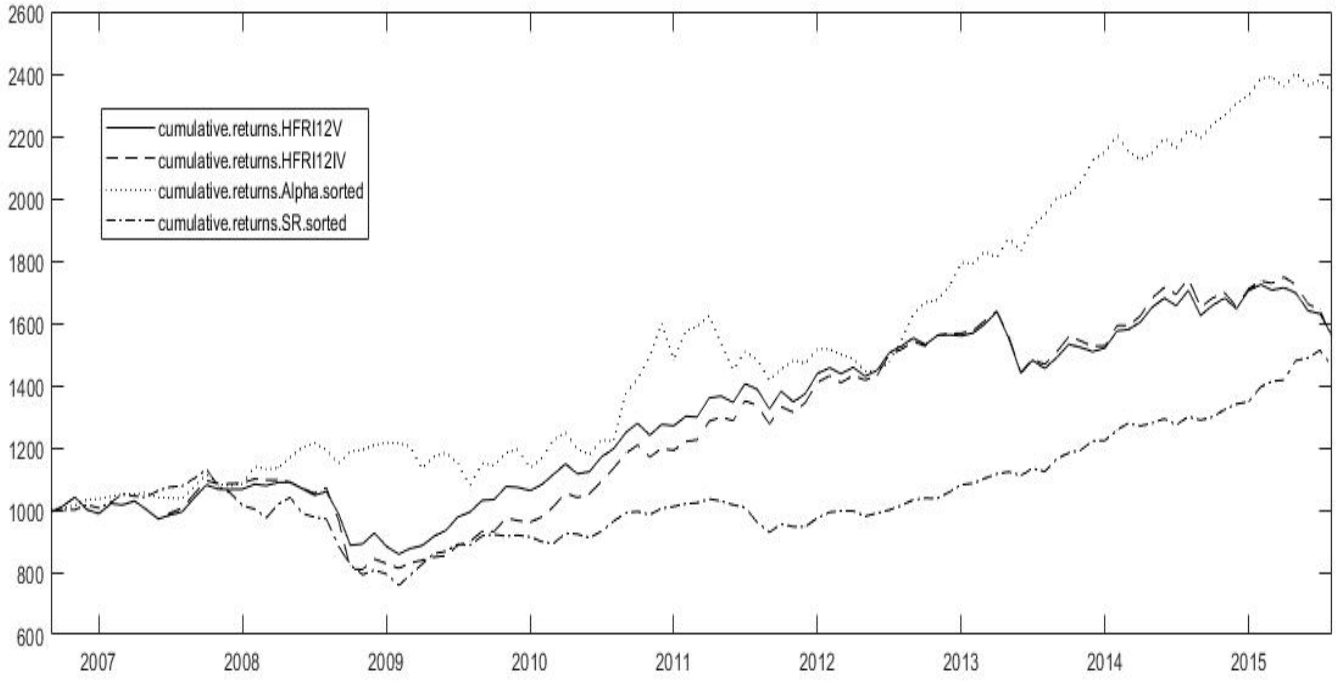


Figure 5.15 – Cumulative returns for the two constructed portfolios, Alpha and Sharpe Ratio (SR) as well as the two benchmarks, the HFR Risk Parity Vol 12 Index and the HFR Risk Parity Vol 12 Institutional Index. Initial investment 1000$.

Table 5.7 has the summary results for the 10% target volatility. First of all observe that our portfolios reached the target of 10% while the benchmark indices (HFRI10V and HFRI10VI) missed it, on average, by roughly 1.2%. SR-sorted portfolio managed to achieve a higher average return of 11% which is 4% more than the indices and 7% more than the alpha-sorted portfolio. SR-sorted portfolio's superior performance is highlighted by all the others risk-adjusted measures since they also rank it as "best". The only exception is with the Manipulation Proof Performance Measure, where all investments report the same value 0.001. As for the alpha-sorted portfolio it seems to be performing worse than the SR portfolio but also compared to the HFR indices. However, in terms of the modified Sharpe ratio the alpha-sorted portfolio outperforms the indices and, in term of the Treynor ratio the three investments report the same number (0.01). The two indices have a very similar performance. The only differences observed are with the Sharpe ratio and the Omega ratio where the Institutional index outperforms the standard one by 1 bp.

Similarly, Table 5.8 presents the summary results for the 12% target volatility portfolios. Our portfolios again match the target exact where as, the indices miss it by roughly 2% on average. However, when the volatility target was set at 12%, the alpha-sorted portfolio seems to be outperforming the SR-sorted portfolio and the benchmark Indices. The alpha-sorted portfolio has an annualized average return of 11% while SR portfolio and both of the indices have only 6%. Alpha-sorted portfolio seems to outperform the SR-sorted one and the benchmark Indices in all measures. The only exception is

the MPPM measure where alpha-sorted portfolio has the same value with HFRI12V index (0.09). Our findings also suggest SR-sorted portfolio has a performance similar to the indices. All three investments report the same value for Treynor ratio. The SR-sorted portfolio (0.37) outperform the HFRI12V (0.35) but not the HFRI12IV (0.39) in terms of alpha, and it has the same value with the HFRI12V (0.15,0.09) and greater than the HFRI12IV (0.12,0.09) in terms of Sortino and the modified Sharpe ratio, respectively. Nevertheless, SR-sorted portfolio is performing better than both Indices in terms of Sharpe and Omega ratio but its performance is worse in terms of the MPPM. Again, both Indices have an almost identical performance, although HFRI12V seems to outperform HFRI12IV for Sharpe, Omega, Sortino, modified Sharpe ratio by one basis point (bp).

To asses the portfolio's performance overtime, Figures 5.14 5.15 report the cumulative returns for the constructed portfolios for each month of the out-of-sample period for all portfolios respectively, for each volatility target. The assumption is of an initial investment of 1000$ and we observe how much it grows by the end of the out-of-sample period to almost 3000$. That is the highest value and it could be achieved holding the SR-sorted portfolio with a 10% volatility target. It can also be observed that the two Indices almost move together. Even though alpha-sorted portfolio has an inferior performance overall, it has a peak at the end of 2007. It is seen that by the end of the sample in 2013 the difference between the SR-sorted portfolio and the indices is expanding, while, alpha-sorted portfolio moves closer to the indices. Figure 5.15, reports the cumulative returns plot for risk parity portfolios at 12% volatility target. Alpha-sorted portfolio cumulative returns are always higher than in the SR-sorted portfolio or the indices for the whole out-of-sample period. Indices seem to move in sync, though their gap grows between 2009 and 2012 years. SR-sorted portfolio is quite close to the indices at the begging of the sample and the distance is increased after 2009. However, the plots seem to tend to meet again closer to 2016.

For both volatility targets, we observe that at least one of our constructed portfolios, outperform the HFRI RP Indices. Under the 10% target case, that is the SR-sorted portfolio, while under the 12% target one, that is the alpha-sorted. We do not seem to observe any significant differences in the performance of the HFR indices.

5.5 Conclusion

The focus is on the risk-based approach, namely Risk Parity (RP) to asset allocation and fund of hedge funds creation. We construct RP portfolios consisting of Hedge Funds (HF) which are competing against, the relatively new, Hedge Funds Research (HFR) RP Indices in an attempt to outperform them. The RP portfolios consist of the best 25 funds, and their performance is evaluated by typical

hedge fund evaluation measures. The funds consisting our RP portfolios were sorted based on risk-adjusted Sharpe Ratio and on Fung and Hsieh (2004)' alpha measure.

Our empirical results suggest that for both volatility targets (10% and 12%), we observe that at least one of our constructed portfolios, managed to outperform the HFRI Indices in the out-of-sample period. Under the 10% target case, that is the SR-sorted portfolio, while under the 12% target, that is the alpha-sorted portfolio. So, a suggestion as to which fund selection measure leads to a superior portfolio cannot be made. What is more, we do not seem to observe any significant differences in the performance of the HFR indices, given the volatility target. Further, commonly-used risk-adjusted measures for evaluation like the Sharpe Ratio seem to be adequate and they produce results similar to more sophisticated measures. Finally, the common misconception that indices are always the best investment for investors was not confirmed in this study.

Chapter 6

Density Volatility Forecasts

6.1 Introduction

All previous chapters have focused on point volatility forecasts using historical observations. This final empirical chapter is addressing two further topics of volatility forecasting namely, implied volatility from option prices and density forecasts. In principle, option-implied forecasts would be expected to dominate forecasts from historical models since the former are inherently forward looking and importantly they already encompass the information included in historical models. However, econometric concerns such as the presence of a time-varying risk premium and the relative illiquidity of option contracts in the distribution's tails suggest that the relative performance of option-implied over historical forecasts constitutes an empirical question. Further, a more recent stream of literature has suggested that density forecasts can have a better forecasting accuracy than point forecasts (Niu and Harvey (2022)). Takings these arguments into consideration, this study aims at addressing both issues with one research question by exploring the dynamics of Implied Volatility (IV) and historical volatility in a density framework. More specifically, it explores when, and under what circumstances, risk transformations of risk-neutral densities can outperform historical densities. Then the forecasts are evaluated in terms of Goodness-of-Fit (GoF) tests.

The relative forecasting performance of option-implied and historical densities has been found to depend on a number of factors such as the forecasting horizon and the sample period. Densities from options produce risk-neutral densities (RND) that are misspecified. Risk transformations that incorporate the risk-premium requested by risk-averse investors are necessary. The calibrated densities are then called real-world densities (RWD). Then, those risk transformations of risk-neutral densities typically work best for longer forecasting horizons, while historical models tend to provide more accurate forecasts for shorter horizons (Shackleton et al. (2010)).

In other words, our aim is to understand the optimal way to determine ex-ante when and why real-

world densities should be preferred to historical models when forecasting. Even though the methods have been studied extensively there are not many financial applications. This kind of research has been conducted only on US stock indices and a handful of US stocks. Therefore, the first contribution of this study comes from expanding the analysis on European stock indices and options written on them. Options written on US and on European assets have some fundamental differences such as in their liquidity which can affect the performance of the models and result to different findings. So it is important to evaluate the sensitivity of the main findings in the literature to different markets. We are testing the forecasting accuracy on multiple forecasting horizons from 1 day to bimonthly (56 days). Another contribution comes from examining the performance of the density forecasts at various regions of the distribution (intervals); A performance evaluation analysis which was not performed by most studies in the field including the seminal paper of Shackleton et al. (2010).

Early research on density forecast evaluation did not include multiple forecasting horizons. Anagnou et al. (2005) compare the accuracy of option-based and historical-based density forecasts at predicting the actual price-return distribution of some financial assets. They used option-implied density forecasts and applied both a simple mean shift and a more rigorous utility-based transformation. They found no significant difference between option-implied and historical density forecasts. They considered both currency and index future contracts. Later, Rompolis (2010) focused on S&P 500 index, and found that his method for RND works best. Shackleton et al. (2010) were the first to examine different forecasting horizons for the US market. Høg and Tsiaras (2011) compared density forecasts for RND, RWD and historical densities for various forecast horizons and concluded that for crude-oil, the non-parametric adjustments of RND outperform their parametric calibration counterparts, RND and historical density forecasts. Kostakis et al. (2011) in a similar exercise and Shackleton et al. (2010) also supported the conclusion that RWD outperform historical forecasts for the US market.

Furthermore, Huang and Wang (2012) only compare the density prediction performance of alternative GARCH models for US and UK benchmark indices for various forecasting horizons. They found that jump models are the best performing ones. However, the non-normal distribution assumption does not offer any significant improvement on forecasting accuracy. Then, Yun (2014) examines again the predictive ability of RND, RWD, historical densities for jump and RV models for short and medium forecasting horizon using the S&P 500 index. He found mixed results depending on the horizon and the evaluation measure. However, one important finding is that the time-varying jump risk premia models are able to forecast densities better across all forecast horizons in terms of the log likelihood. For 1-step ahead, these models marginally outperform the RV models. These conclusions are almost opposite to Shackleton et al. (2010). They also found that the risk-adjusted real-world model matters.

On the other hand, Fan et al. (2018) are comparing option-implied density forecasts (RND and RWD) to historical densities but across four forecast horizons for individual stocks from DJIA using high-frequency returns and daily prices. They found that simple option-implied models outperform historical ones. During the same year Crisóstomo and Couso (2018) found that using a new composite indicator, the integrated forecast score, RNDs outperform historical-based predictions in terms of information content. However, they compared only RND with historical densities for 1-step ahead (they did not consider RWD). Their analysis was conducted on IBEX 35 index. However, depending on different evaluation measures they found conflicting results. For instance, GARCH-based GJR-FHS delivers the most consistent forecasts across the entire density range.

The analysis aboves highlights the lack of consensus in the literature and the need to explore the topic more. The rest of the chapter is organized as following; first the framework for risk-neutral, real world volatility density and historical densities models is presented. Then the tests for density forecast evaluation are introduced in section 6.2.4. Section 6.3 explains the data and methodology used. In Section 6.4 our empirical results are presented. Finally, the chapter finishes with the conclusion which summarizes the key points.

6.2 Literature Review

6.2.1 Risk-Neutral & Real World Volatility Density

"A density forecast of the realization of a random variable at some future time is an estimate of the probability distribution of the possible future values of that variable" (Tay and Wallis (2000), p.235). Density forecasts provide an estimate of the entire future distribution of returns and as such they are considered more accurate than point forecasts. Density forecasts have also been very popular in banks. What is also important for commercial and central banks is to have up-to-date information of the market conditions. So, central banks have been also, heavily relying on options markets to obtain forward-looking investors' expectations considering that option-based densities can quickly reflect changes in market conditions.

Banz and Miller (1978) and Breeden and Litzenberger (1978) were the first to explore options to extract risk-neutral density forecasts. The first step to get risk-neutral densities is to use an option-pricing model. The most famous one is the Black-Scholes (B-S) model. "The Black–Scholes lognormal densities are defined by using the implied volatility of the nearest-the-money, nearest-to-expiry contract, for which $|S - K|$ is nearer zero than for all other contemporaneous strikes" (Fan et al. (2018), p.90). Implied volatility is derived from priced options on stocks and it is the volatility that equates the theoretical price after using a model, such as the Black-Scholes (B-S), to the market price (Bollerslev

(2009)). From the B-S call option-pricing formula it is possible to derive the ex-ante IV. However, B-S model has received a lot of criticism due to its limitations. The B-S model predicts that the implied volatility curve is flat when plotted against varying strike prices and it assumes that risk-free rate of return and volatility remain constant for the option's duration. However, the empirically observed volatility smile contradicts these assumptions.

For a more accurate estimation that overcomes both above limitations of the B-S model Heston (1993) developed the Heston Model. This model belongs to stochastic volatility models that assume uncertain volatility behavior and thus they are more realistic. An additional advantage of the Heston model is that it provides a closed-form solution for pricing options. The dynamics of the model for the price and the volatility are given by the following formulas:

$$dp/p = \sqrt{V}dW_1$$

$$dV = \kappa(\theta - V)dt + \xi\sqrt{V}dQ_2 \tag{6.1}$$

where θ is long-run variance, κ is the rate at which V (variance of the asset price) reverts to θ, ξ is the volatility of the volatility, W_1 is the Brownian motion of the price and W_2 is the Brownian motion of the variance with their correlation being ρ. Bates (1996) suggested an extension of the price dynamics that includes jumps which is believed to increase the probability that the realized and the theoretical option prices match. Heston model is using the characteristic functions of option values rather than the density functions. More specifically, it is based on the characteristic functions of the logarithmic prices there are easier to work with and provides the closed-form solution. Then, the characteristic function can be transformed to densities using the Fourier Inversion theorem. Accordingly, the RND is given by

$$g(X) = e^{rT}\frac{\vartheta^2 c}{\vartheta X^2} \tag{6.2}$$

$$c(X) = e^{-rT}\int_X^{\inf}(x - X)g(x)dx \tag{6.3}$$

$$c(X) = e^{-rT}E^Q[max(S_T - X, 0)] \tag{6.4}$$

$$F = E^Q[S_T] \tag{6.5}$$

Liu et al. (2007). Then using the probabilistic approach of Heston the value of a European call option is given by:

$$c(P_0, K) = e^{-rt}(p_0 P_1(p_0, K) - K P_2(p_0, K) \tag{6.6}$$

where r is the risk-free rate, $P_2(p_0, K)$ is the risk-neutral probability that the option expires in-the-money and $P_1(p_0, K)$ is a probability for the same event when a different measure is applied (Shackleton et al. (2010)). A further advantage of the Heston model is that options with different expiry dates can be used to extract the risk-neutral densities. Other parametric specifications include the Mixture of two lognormal densities, the Generalized Beta of the second kind (GB2) and the Normal Inverse Gaussian (NIG) (Anagnou et al. (2005)). Finally Madan et al. (1998) suggested the purely discontinuous Variance Gamma (VG) (Crisóstomo and Couso (2018)). Risk-neutral density probabilities can be derived from non-parametric and parametric approaches. For the nonparametric the main categories are: implied trees, maximum-entropy methods, curve fitting methods and regularization methods (for more details see: Rubinstein (1994), Campa et al. (1998), Aıt-Sahalia et al. (2001), Anagnou et al. (2005)).

However, a serious weakness is that such option-implied CDFs, as the ones that have been presented above, give an accurate estimation only under the risk-neutrality assumption. In reality, investors are risk-averse and demand a compensation for bearing extra risk. Risk-neutral-densities do not incorporate any risk-premium and thus they tend to be biased. "Options only reveal information associated with a risk-neutral environment, and transitioning to a non-martingale reality is the ultimate objective (although the calibration of this risk-premium parameter will be discussed at a later point)" (Papadaki (2011), p.143). Risk-neutral density estimates need to be transited to the real world that is to move from Q-measure to P-measure. Finally, the comparison between risk-neutral and real-world densities can reveal more information about investors' risk preferences and its relationship with option prices (de Vincent-Humphreys and Noss (2012)).

Two main approaches have been proposed in the literature to transform the risk-neutral densities to the real-world. The first way requires to estimate the RND, then use the historical data to estimate the empirical density and the final step is to compare them in order to infer the risk-premium (Anagnou et al. (2005)). However, this method is not motivated by the theory and could be inconsistent with the observed properties of the density. The second approach is to produce time-varying risk-adjusted densities by specifying parametric forms for the risk-aversion functions, and then combine them with the RND (Anagnou et al. (2005)). It based on the assumption that risk preferences are assumed to be stationary and well-behaving (Anagnou et al. (2005)). In more detail, the second approach to transform the densities is to calibrate the risk premium required by investors by relying on economic theory by using a utility function (Bliss and Panigirtzoglou (2004)). Bliss and Panigirtzoglou (2004) idea is that "if risk-neutral PDFs do not forecast accurately, we may infer that the difference between

the risk-neutral and the accurate or objective forecast arises from the risk aversion of the representative agent. We can then use this difference to infer the degree of risk aversion of the representative investor" (p.408). Bliss and Panigirtzoglou (2004) estimate a single parameter utility function, where Liu et al. (2007) use a two-parameter utility function. Finally, Shackleton et al. (2010) provide two additional transformations, "one assumes that affine (jump-) diffusion price dynamics apply in the real world by incorporating appropriate risk-premia, and the other applies a nonparametric calibration function" (p.2679). To transform the risk-neutral density presented above the following formula can be applied:

$$dp/p = \eta_1 V dt + \sqrt{V} d\bar{W}_1 dV = [\eta_2 V + \kappa(\theta - V)]dt + \xi\sqrt{V} d\bar{W}_2 \tag{6.7}$$

that depends on parameters η_1 and η_2 (Shackleton et al. (2010)). An alternative approach to utility-based transformations is to treat the risk-neutral densities as misspecified and attempt to properly calibrate them into real-world densities using only statistical criteria. Finally, a nonparametric calibration function can be estimated which relies only on past observations. To achieve that we rely on the fact that correctly specified forecasts are independent uniform U(1,0) variates. Then we can enhance the risk-neutral densities accordingly based on their deviations from the theoretical uniform distribution to generate more accurate real-world densities. "The nonparametric transformation calibration function for a selected horizon (h) and a selected risk-neutral method, is determined by a set of $n - h$ cumulative, risk-neutral probabilities"

$$u_{t+h} = F_{Q,s,T}(p_{s+h}|\Theta_s),\ 1 \leq s \geq h - s \tag{6.8}$$

with T (years) matching h (trading days). Here $F_{Q,s,T}$ is the cumulative distribution function (CDF) of the stock price p_{s+h}, with Θ_s a vector of density parameters (Fan et al. (2018), p.86). At time t the risk-neutral CDF of p_{s+h}, denoted $F_{Q,T}$, is transformed into the real-world CDF of p_{s+h}, denoted $F_{P,T}$, by the equation:

$$F_{P,T}(x) = \hat{C}_T(F_{Q,T}(x)) \tag{6.9}$$

where $\hat{C}_T$ is the empirical calibration function, which is the estimated real-world CDF of the risk-neutral PIT. A Gaussian kernel with bandwidth B is used to obtain the kernel density and CDF

6.2.2 Historical Density

Even though risk transformations give better estimates, the best approach to transform risk-neutral to real-world densities is yet to be found. To overcome this limitation it has also been suggested in the literature to directly forecast real-world volatility densities. In that case "we only need to assume the process that drives the volatility of that asset (discrete-time model) and then use bootstrap techniques to obtain the predictive volatility density according to different assumptions for the error distribution" (Papadaki (2011), p.157). Those models are the well-known GARCH models and their extensions

(such as GJR) under different distributional assumptions (such as Normal, Student-T, GED or Skew-t distributions). A better approximation of the error distribution leads to more accurate forecasts. Historical models predict future densities based on past observations. Thus, they are simpler to implement but the window size is very important since it can affect the forecasting ability especially in time of financial crisis. The returns and volatility are given as:

$$r_t = \mu + u_t u \sim N(o, \sigma^2) \tag{6.10}$$

$$\sigma_t^2 = \alpha + \beta_1 * u_{t-1}^2 + \sigma_{t-1}^2 \tag{6.11}$$

and for GJR:

$$\sigma_t^2 = \alpha + \beta_1 * u_{t-1}^2 + \sigma_{t-1}^2 + \gamma * I_t * u_{t-1}^2 \tag{6.12}$$

where $I_t = 1$ if $u_t^2 < 0$ or 0 otherwise. First the parameters are calibrated based on past information (return series at t-1) by using a specific GARCH model to forecast the one-step ahead conditional volatility. To do the density forecast, "all that is required is to simulate a sufficiently large number of variance forecasts for time t by taking n independent random draws from the error distribution and adding them to the volatility equation" (Papadaki (2011), p.167).

6.2.3 Other Methods

Since both approaches come with drawbacks and strengths but also with many choices for the selected specification averaging has also been popular in the literature such as the Thick Model Averaging (Granger and Jeon (2004)). More specifically, when averaging for risk-neutral densities, then the approach refers to not having to stick to a single option-pricing model while for real-world ones, the averaging refers to not have to select a single errors' distribution. The selection is performed on the basis of the BIC (Schwartz Bayesian Information Criterion) and/or MSE (Mean Squared Error).
Finally, the theory has also suggested that a mixture of densities can also be used,

$$f_{mix,T}(x) = a f_{P,T}(x) + (1 - a) f_{ARCH,T} \tag{6.13}$$

where $f_{P,T}(x)$ is the "transformed", risk-neutral densities and $f_{ARCH,T}$ is the historical densities (Shackleton et al. (2010)). The coefficients, add to 1, and are determined by the previous performance Finally, Rompolis (2010) suggested the maximum entropy method which incorporates better all the available information provided by the market.

6.2.4 Density Forecast evaluation

The evaluation of Density Forecasts, is not much different than the one applied of point models. In this context however, there are three main types of measures the log likelihood ratio where can be used to sort the densities, the Goodness-of-Fit (GoF) tests where it is to evaluate whether the densities give well-specified forecasts and finally the tests of equal accuracy for pairwise comparisons. In density forecasts, the objective is to choose parameter a such as:

$$a^*(p(y)) = argmin_{a \in A} \int L(a, y)p(y)dy \tag{6.14}$$

where p(y) is the true density. Then the expected loss with respect to the true data generating process is:

$$E[L(a^*, y)] = \int L(a^*, y)f(y)dy \tag{6.15}$$

where $f(y)$ is the density forecast Diebold et al. (1998a). See also Diebold and Lopez (1996), Diebold et al. (1998b), Diebold et al. (1998a). Similarly to the point forecasts we can use the MSE, the RMSE, the MAE loss functions and the Wilcoxon's signed-rank test (Rompolis (2010);Huang (2012)). An alternative measure is the root mean integrated squared error (RMISE).

$$RMISE = \sqrt{E[\int (\hat{f}_Q(x) - f_Q(x))^2]} \tag{6.16}$$

where $E(\cdot)$ is the expectation. The RMISE metric can be decomposed to the root integrated squared bias (RISB) and the root integrated variance (RIV) (Rompolis (2010)).

Log Likelihood (LR)

The most popular way to rank density predictions is in terms of their out-of-sample accuracy as measured by the Likelihood ratio (LR).

$$L_m = \sum_{y=1}^{j} \log(f_{m,t}(p_{t+1})) \tag{6.17}$$

where $f_{t,m}$ is the ex-ante density forecast and p_{t+1} is the ex-post realization at $t+1$ (Shackleton et al. (2010)). For each model, LR assigns a loss score and then we sum these scores over the sample. The higher the LR of the model, the higher the chances that the model generates densities that are closer to the true density. Another measure is the continuous ranked probability score (CRPS) which it is similar to the LR but considers the whole distribution. It uses the difference between the actual realization and all other probabilistic outcomes.

$$CRPS_t = \int_{-\infty}^{+\infty} [CDF^m(x) - CDF^r(x)]^2 dx \tag{6.18}$$

where CDF^m and CDF^r is the CDF of the model and the realization, respectively. $CDF^r(x) = 0$ if $x_t* < x$ and $CDF^r(x) = 1$ if $x_t* \geq x$. Crisóstomo and Couso (2018) proposed a transformation of the above measure to consider return deviations instead of index points. Finally, we can also use sequential likelihoods proposed by Johannes et al. (2009) that provide more insight into the times that the models fails (Yun (2014)).

$$LR_t^{(m1,m2)} = \sum_{s=1}^{t} \ln(\frac{f^{m1}(y_s|I_{s-1})}{f^{m2}(y_s|I_{s-1})}),\ t = 1, ...T \tag{6.19}$$

Goodness-of-Fit (GoF)

Ranking the densities forecasts based on the loss function or on the out-of-sample LR is not always enough. What really matters in density forecasts is that the models provide well-specified forecasts. That is why the relevant literature has proposed several Goodness-of-Fit (GoF) tests for density forecast evaluation. However, diagnostics cannot evaluate the model's accuracy and/or the error magnitude.

A key property of correctly specified forecasts, is that the Probability Integral Transform (PIT) series are independent uniform U(0, 1) variables. PIT is the quantile of the ex-ante distribution evaluated at the actual realization. That is better expressed by performing a simple transformation proposed by Berkowitz (2001).

$$z_t = \Phi^{-1}(PIT_t) \tag{6.20}$$

where

$$PIT_t = \int_{-\infty}^{X_{t+1}} f_t(x)dx \tag{6.21}$$

where $f_t(x)$ is the forecast generated by a model at time t and X_{t+1} is the actual realization of the random variable. Diebold et al. (1998a) tested the uniformity and independence hypothesis separately. To test the assumption about the distribution there are two well-known tests the Kolmogorov–Smirnov (KS) and the Jarque–Bera (JB). The first ones is using the following t-statistic:

$$KS = \sup_{0 \leq PIT \leq 1} |\tilde{C}(PIT) - PIT| \tag{6.22}$$

where $\tilde{C}(PIT) = \frac{1}{j-i+1} \sum_{y=i+1}^{j+1} S(PIT - PIT_t)$ and $S(x) = 1$ if $x \leq 0$ and $S(x) = 0$ if $x < 0$ Shackleton et al. (2010). In other words it uses the maximum difference between the sample and a theoretical CDF. Now to test the hypothesis of no autocorrelation, the general framework is that we examine the autocorrelation, of the z_t series, in the mean,variance, kurtosis and skewness (4 moments). In more details, a test has been suggested that it is based on regressing $(z_t - \bar{z}_t)$ on 6, 12, and 24 lags.

The proposed test follows the x^2 distribution under the null. Another approach is to use the test of Berkowitz (2001) for autocorrelation which is given as:

$$LR1 = (L(\hat{\mu}, \hat{\sigma}^2, 0) - L(\hat{\mu}, \hat{\sigma}^2, \hat{\rho})) \sim x^2(1) \tag{6.23}$$

Then μ, σ^2 and ρ can be estimated via maximum likelihood. Finally, Berkowitz (2001) proposed a test that can simultaneously test for uniformity and independence. The test is based on estimating the following model:

$$z_t - \mu = \rho(z_{t-1} - \mu) + \epsilon_t \tag{6.24}$$

$\epsilon_t \sim n.i.i.dN(0, \sigma^2)$. The null hypothesis is that z_t should be i.i.d N(0,1) and the alternative that the series are stationary and follow a Gaussian AR(1) using a standard likelihood ratio test (LR3). Another way to view this is having the likelihood of a restricted and an unrestricted model. In other words, under the null hypothesis that the forecast is correctly specified z_t should be i.i.d N(0,1) (Høg and Tsiaras (2011)). The test is given as:

$$LR3 = (L(0, 1, 0) - L(\hat{\mu}, \hat{\sigma}^2, \hat{\rho})) \sim x^2(3) \tag{6.25}$$

where the first term is the maximum log-likelihoods (MLL) for the null and the second term is the MLL under the alternative hypotheses. However, if the series z_t have $\mu = 0$ and $\sigma^2 = 1$ but their higher moments imply that the distribution is not normal then the test will fail to reject the null. An alternative test is the Hong and Li (2005) that tests for the two hypothesis but also it takes into consideration the estimation uncertainty and can be used to compare non-nested models (Yun (2014)). Alternative tests include the Kuiper statistic, the Watson statistic, and the Anderson-Darling test Anagnou et al. (2005).

Christoffersen (1998), Engle and Manganelli (2004) and later on Patton (2006) focus on measures to evaluate the forecasting densities on certain regions. The null is that the forecasting density is correctly specified in the specific region. One condition for that is that the probability according to a model ($p_{i,t}$) and the true probability ($\pi_{i,t}$) that the realization of the random variable X_t under investigation will fall into the i'th region are equal. We can define: $H_{i,t} = 1$ if $X_t \in R_{i,t}$ and $H_{i,t} = 0$ otherwise, where $R_{o,t}, ...R_{K,t}$ are disjoint regions covering the support, S of $X_t : S = R_{o,t}U....UR_{K,t}$. So the second condition under the null is that $H_{i,t} \sim i.d\ Bernouli(p_i)$. Following Patton (2006)'s transformation we can re-write:

$$\pi_{i,t} = \pi_i(Z_{it}, \beta_i, p_{it}) = \Lambda(Z_{it}\beta_i + \ln[\frac{p_{it}}{1 - p_{it}}]) \tag{6.26}$$

This corresponds to test for $\beta_i = 0$, all 4 elements. Z_{it} includes a constant which is to check whether

the model has the correct proportion of hits, and three variables that count the number of hits in that region in the last 1, 5, and 10 observations. So a likelihood ratio test can be conducted: $LR = -2(LL_0 - LL_1) \sim x^2(4)$ where LL_0 and LL_1 is the restricted and unrestricted log-likelihood of the model, respectively (Høg and Tsiaras (2011)).

Pairwise comparisons

Similarly to point forecasts, is essential to test also the statistical significance of the forecasting accuracy among the models by carrying out pairwise comparisons. The null hypothesis is always the same, the competitive models have equal expected likelihood. The test once again is using the difference between the log-likelihoods and it assumes that the differences are uncorrelated. Following Amisano and Giacomini (2007) The t-statistic is:

$$AG = \frac{\bar{d}}{s_d\sqrt{j-i+1}} \tag{6.27}$$

where

$$d_t = \log(f_{m,t}(p_{y+1})) - \log(f_{n,t}(p_{y+1})), i \leq t \leq j \tag{6.28}$$

where $f_{m,t}$ is the density forecast. A Newey–West adjustment corrects for the autocorrelation in overlapping forecasts (Fan et al. (2018)).

Other methods

A density-specific measure is a criterion of the expected logarithmic score (lnS) which is given as:

$$lnS = \frac{1}{T-M}\sum_{i=1}^{T-M} \ln f_t(\hat{V}_t) \tag{6.29}$$

where $f_t(\hat{V}_t)$ is the predictive density (Meng and Liu (2019)). The preferred model maximizes this measure. Finally, another category is the tail tests or scoring rules which are "based on the distance between the probability forecast P_t^f formulated at time $t-1$ for an event at time t, and a binary variable R_t which assumes value of 1 if the event occurs, and value of 0 otherwise" (Anagnou et al. (2005)). A possible test is the Brier's quadratic probability score.

$$B_n = 1/n\sum_t 2(P_t^f - R_t)^2 \tag{6.30}$$

The lower the score, the more accurate the forecast is. Finally, it is also possible to test the statistically significance of this test (Anagnou et al. (2005)).

6.3 Data & Methodology

6.3.1 Data description

For the empirical application we use options written on European indices for the period from October 2006 to November 2016. Precisely, data from CAC 40 index for France, DAX for Germany, FTSE MIB index for Italy, AEX for Netherlands, MINI IBEX 35 for Spain, OMXS 30 index for Sweden, SSMI for Switzerland, and FTSE 100 index for UK were considered

Table 6.1 summarizes the number of options, the moneyness and the maturity of the option contracts before filtering. On the table it can been seen that the average number of options per day varies among the indices with data from Spain having the highest and Sweden having the lowest number of contracts. The total number of puts and calls seems to be close, but on every dataset the number of put contracts is a bit higher than the number of calls with the exception of Netherlands. Almost all out-of-the-money (OTM) options of the contracts seem to have average to high maturity except from the options written on Netherlands stock index. The average maturity is measured between one to six months while high maturity is considered anything higher than 6 months. In more details stock indices from France, Germany, Italy, Spain, Switzerland and UK have the majority of their contracts with a duration with more than 6 months. However, options written on the Swedish stock index have between and and six months to expire.

Table 6.1 – Summary Statistics for European Options

Panel A: France				
	Total Number	Average Options/day	Max/day	Min/day
Call	1414174	409	807	78
Put	1440784	417	807	93
Subtotal	2854958			
Moneyness/Maturity	< 1 m	1-6 m	> 6 m	
OTM put	3%	11%	31%	
OTM call	4%	11%	30%	
Subtotal	11%	26%	63%	
Panel B: Germany				
	Total Number	Average Options/day	Max/day	Min/day
Call	2771110	734	1298	277
Put	2771252	734	1298	277
Subtotal	5542362			
Moneyness/Maturity	< 1 m	1-6 m	> 6 m	
OTM put	6%	13%	23%	
OTM call	7%	14%	23%	
Subtotal	19%	32%	49%	
Panel C: Italy				
	Total Number	Average Options/day	Max/day	Min/day
Call	922151	362	503	142
Put	978695	384	508	147
Subtotal	1900846			
Moneyness/Maturity	< 1 m	1-6 m	> 6 m	
OTM put	5%	10%	29%	
OTM call	6%	10%	28%	
Subtotal	17%	23%	60%	

Table 6.2 – Summary Statistics for European Options - Cont'd

Panel D: Netherlands				
	Total Number	Average Options/day	Max/day	Min/day
Call	848067	250	581	12
Put	847938	250	581	12
Subtotal	1696005			
Moneyness/Maturity	< 1 m	1-6 m	> 6 m	
OTM put	10%	12%	15%	
OTM call	12%	14%	15%	
Subtotal	38%	30%	32%	
Panel E: Spain				
	Total Number	Average Options/day	Max/day	Min/day
Call	1882727	736	1599	88
Put	1927363	753	1572	104
Subtotal	3810090			
Moneyness/Maturity	< 1 m	1-6 m	> 6 m	
OTM put	2%	9%	34%	
OTM call	2%	9%	34%	
Subtotal	7%	20%	73%	
Panel F: Sweden				
	Total Number	Average Options/day	Max/day	Min/day
Call	434558	176	675	28
Put	538597	218	675	34
Subtotal	973155			
Moneyness/Maturity	< 1 m	1-6 m	> 6 m	
OTM put	11%	23%	12%	
OTM call	11%	19%	10%	
Subtotal	31%	46%	23%	

Table 6.3 – Summary Statistics for European Options - Cont'd (2)

Panel G: Switzerland				
	Total Number	Average Options/day	Max/day	Min/day
Call	1864226	494	896	196
Put	1864291	494	896	196
Subtotal	3728517			
Moneyness/Maturity	< 1 m	1-6 m	> 6 m	
OTM put	5%	17%	23%	
OTM call	5%	18%	23%	
Subtotal	15%	38%	47%	
Panel H: UK				
	Total Number	Average Options/day	Max/day	Min/day
Call	1921849	516	1628	67
Put	1979850	532	1616	100
Subtotal	3901699			
Moneyness/Maturity	< 1 m	1-6 m	> 6 m	
OTM put	4%	17%	26%	
OTM call	4%	16%	25%	
Subtotal	11%	37%	52%	

Notes: This table reports the summary statistics for all European options written on futures from 2006 to 2016. This is about for the whole dataset before data clearing.

Then, we checked our data and any options where their delta is not in included in the interval [-1,1] was deleted. Any contracts that had less than a week (7 days) to expiration were also excluded (Shackleton et al. (2010)). Options are allowed to have different maturities since we are using the Heston model. For the purposes of this study, the data from the OTM calls written on futures are used. That means that only options where the absolute value of delta was less than 0.5. A straightforward conversion using the put and call parity was performed to calculate the call price from the relevant put price.

6.3.2 Methodology

To construct the RND forecasts the Heston model was used. The first step was to estimate the 5 parameters of Heston model namely, the initial variance, the long-term variance, the speed of mean reversion, the volatility of volatility and the correlation. They were calibrated by minimizing the average of the squared pricing errors.

$$1/N \sum_{i=1}^{N} (CALL(K_i) - CALL_{GB2}(K_i|\theta_{GB2}))^2 \tag{6.31}$$

where $CALL_{GB2}(K_i|\theta_{GB2})$ is the theoretical price given be a GB2 density with parameters vector θ_{GB2} (Høg and Tsiaras (2011)). We tried various approaches to estimate the parameters. For instance as initial values, on the optimization, we used the parameters found in Shackleton et al. (2010), some "random" arbitrary starting points, the median of all previous observations and the previous fit. What we found is that the parameters from the various approaches did not have significant differences and for most datasets the best fit as measured by the mean square error (MSE) was given by using Shackleton et al. (2010) parameters so we proceeded with these parameters for our first optimization.

Following the estimation of the parameters, non-overlapping RND forecasts were produced for daily, weekly, biweekly, monthly and bimonthly forecasting horizons considering 1, 5, 10, 20 and 40 trading days respectively. To estimate the RND for each horizon the PDF and the CDF were calculated based on the Heston model. We used the Breeden and Litzenberger (1978) density approximation. The PDF was calculated numerically as the second derivative of the call function and CDF was calculated numerically as the first derivative of the call function.

For the reasons outlined above, RNDs cannot fully reflect reality so non-parametric recalibration of density forecasts were also employed following Shackleton et al. (2010) and Høg and Tsiaras (2011) to construct RWD forecasts. To create our new series the following transformation was used: $z_t = \Phi^{-1}(u_t)$, where Φ is the cumulative distribution function of the standard normal density (Høg and Tsiaras (2011)). Our window length (number of past periods) at which non-parametric calibration was estimated was the whole sample. For the purposes of comparison, historical density forecasts were also generated. We used the GARCH and the GJR models assuming a Normal and a Student-T distribution for each model. For any horizon longer than 1-day ahead simulated returns were used as described in section 6.2.2. Our rolling window was 5 years.

To evaluate and compare the density forecasts produced by the various approaches and models we focused on evaluating whether the forecasts are misspecified or not using the Goodness-of-Fit (GoF) tests. It is necessary here to clarify exactly what is meant by misspecified. We started by creating the probability integral transform (PIT) series that transform the realization of the data generating process with respect to the density forecast. Then to test whether the forecasts are correctly specified, all we need to do is to test whether those PIT series are i.i.d. N(0, 1). These hypotheses can be tested separately or jointly. The Kolmogorov-Smirnov (KS) test tests if the series are N(0, 1). To test for the independence hypothesis the Lagrange Multiplier (LM) test was used which corresponds to the null of no auto-correlation in the first four moments of the series using the 6, 12, and 24 lags

of the standardized (squared standardized) residuals. To test the joint hypothesis of normality and independence the LR3 tests of Berkowitz (2001) was used. Nevertheless, the previous tests focus on the whole distribution. To test whether the distribution is correctly specified in different regions of the distribution the "hit tests" of Patton (2006) were used. This is useful if we want to concentrate on specific parts of the distributions as for instance in the tails. Finally, the descriptive statistics of the z_t series, where $z_t = \Phi^{-1}(PIT)$ were calculated. If the forecasts are well specified the series should follow the standard normal distribution.

6.4 Results

Table 6.4 gives the p-values for the KS1 test for all density forecasts for all forecasting horizons. Under the null hypothesis the PIT series follow a normal distribution. First of all, what we observe is that RND forecasts have the highest rate of rejections at any level of statistical significance compared to all the other methods considered namely, the RWD forecasts and the historical densities forecasts. This finding was expected as the RND method gives mis-specified forecasts. That was also found by Shackleton et al. (2010) and Fan et al. (2018). However, it can also been see that as the horizon increases the results for RND seem to improve. For 1, 5 or 10 days ahead, the p-value is 0 for all datasets and for 14 days ahead forecasts the p-values are 0 except for Italy which is 0.08, for Sweden which is 0.01 and for Netherlands which is 0.05. By contrast, for all datasets, for 2-month ahead RND forecasts we fail to reject the null hypothesis at 1% significance level.

Moving on to RWD forecasts it can been seen again that the longer the forecasting horizon the better the fit we get which may reflect low power in small samples as there are fewer observations. Crisóstomo and Couso (2018) also found an improvement as the horizon increased for non-parametric calibrations. Also we get a noticeable improvement as instead of rejecting 28 times the null hypothesis with RNDs we only reject it 8 times at 1%. At 5% significance level, we had 31 rejections under RND forecasts versus 14 after the non-parametric calibration to RWD. In more details, for 1-week, 1 and 2 months ahead, RWD forecast fails to pass the test at 5% only for Sweden for 1-month ahead. For the shorter horizons, at 5% only Spain and Germany pass the test for 1 day ahead and Germany for 1-week ahead too.

Now moving to the historical simulations (HS) it can been seen that they should also be preferred over RND forecasts and seemingly over RWD forecasts. Under the GARCH-N model at 1% significant level the null hypothesis of normality is rejected only for 1-day ahead for Spain, Sweden, France, Switzerland and Netherlands. At 5% it seems like it is for all countries except Italy. However, compared to RWD forecasts moving to longer horizons means that the null hypothesis is only rejected for Sweden at 5% significance level for 1 and 2 months ahead. Changing the distributional assumption

on the GARCH model to Student-T gives slightly worse results. Now for all datasets for 1-day ahead forecasts the null hypothesis is rejected. What is more, at 5% now except from Sweden also for UK at 1-month ahead and for Germany at 2-month ahead forecasts the null hypothesis is rejected.

Focusing on the GJR model under the Normal distribution it seems to be the specification that works the best as it has the lowest rate of rejections. Also any rejections only exist on the daily forecasts. In more details, Germany is the exception where the null hypothesis does not get rejected at any level. Finally, for GJR-T the results seem to be better than GARCH-T but not GARCH-N where GARCH-N gives slightly better results for daily forecast only for a few countries. The good performance of historical methods was also documented in Crisóstomo and Couso (2018) were "(...) all historical methods pass the KS test at 5% level" (p.596). However, the KS tests for uniformity under the i.i.d. assumption rather than testing for the two hypothesis jointly as the LR3 test. So, it is important to evaluate our density forecasts in terms of the LR3 test as well.

Table 6.4 – KS1 test results

RND	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
ITALY	0.00	0.00	0.00	**0.08**	0.01
SPAIN	0.00	0.00	0.00	0.00	**0.10**
UK	0.00	0.00	0.00	0.00	**0.06**
SWEDEN	0.00	0.00	0.00	0.01	**0.51**
FRANCE	0.00	0.00	0.00	0.00	**0.13**
GERMANY	0.00	0.00	0.00	0.00	**0.11**
SWITZERLAND	0.00	0.00	0.00	0.00	**0.05**
NETHERLANDS	0.00	0.00	0.00	**0.05**	**0.29**
RWD	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
ITALY	0.01	0.01	**0.47**	**0.12**	**0.99**
SPAIN	**0.06**	0.01	**0.89**	**0.36**	**0.67**
UK	0.00	0.00	**0.87**	**0.23**	**0.23**
SWEDEN	0.00	0.00	**0.64**	0.04	**0.89**
FRANCE	0.00	0.01	**0.70**	**0.07**	**0.77**
GERMANY	**0.34**	**0.07**	**0.18**	**0.18**	**0.48**
SWITZERLAND	0.00	0.01	**0.17**	**0.74**	**0.43**
NETHERLANDS	0.00	0.00	**0.07**	**0.12**	**0.5**5
HS (GARCH -N)	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
ITALY	**0.06**	**0.49**	**0.46**	**0.65**	**0.32**
SPAIN	0.00	**0.84**	**0.62**	**0.47**	**0.18**
UK	0.03	**0.75**	**0.59**	**0.07**	**0.16**
SWEDEN	0.00	**0.16**	**0.18**	0.04	0.01
FRANCE	0.00	**0.60**	**0.20**	**0.13**	**0.09**
GERMANY	0.02	**0.46**	**0.08**	**0.28**	**0.07**
SWITZERLAND	0.00	**0.70**	**0.64**	**0.26**	**0.33**
NETHERLANDS	0.00	**0.78**	**0.25**	**0.36**	**0.14**
HS(GARCH -T)	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
ITALY	0.00	**0.12**	**0.11**	**0.31**	**0.23**
SPAIN	0.00	**0.81**	**0.44**	**0.19**	**0.05**
UK	0.00	**0.25**	**0.50**	0.02	**0.07**
SWEDEN	0.00	**0.18**	**0.19**	0.02	0.00
FRANCE	0.00	**0.19**	**0.09**	0.04	0.03
GERMANY	0.00	**0.46**	**0.06**	**0.25**	0.02
SWITZERLAND	0.00	**0.53**	**0.50**	**0.29**	**0.37**

NETHERLANDS	0.00	**0.64**	**0.09**	**0.28**	**0.10**
HS(GJR-N)	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
ITALY	0.01	**0.60**	**0.69**	**0.99**	**0.56**
SPAIN	0.00	**0.39**	**0.32**	**0.98**	**0.34**
UK	0.03	**1.00**	**0.70**	**0.45**	**0.81**
SWEDEN	0.00	**0.56**	**0.77**	**0.12**	**0.22**
FRANCE	0.00	**0.73**	**0.38**	**0.63**	**0.69**
GERMANY	**0.10**	**0.59**	**0.63**	**0.95**	**0.49**
SWITZERLAND	0.00	**0.98**	**0.96**	**0.86**	**0.89**
NETHERLANDS	0.00	**0.89**	**0.90**	**0.86**	**0.78**
HS (GJR-T)	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
ITALY	0.01	**0.15**	**0.20**	**0.4**1	**0.27**
SPAIN	0.03	**0.92**	**0.74**	**0.37**	**0.20**
UK	0.01	**0.60**	**0.61**	**0.16**	**0.13**
SWEDEN	0.00	**0.53**	**0.34**	**0.06**	**0.07**
FRANCE	0.00	**0.38**	**0.12**	**0.21**	0.03
GERMANY	0.00	**0.54**	**0.07**	**0.36**	**0.10**
SWITZERLAND	0.00	**0.89**	**0.8**7	**0.67**	**0.92**
NETHERLANDS	0.00	**0.75**	**0.21**	**0.37**	**0.21**

Notes: This table shows the p-values for the Kolmogorov-Smirnov (KS) test for all European indices, for all horizons, for the risk-neutral density (RND) forecasts, the real world density (RWD) forecasts and the historical density forecasts using the GARCH-N, GARCH-T, GJR-N and GJR-T models. The statistically significant results at the 5% level are reported in bold.

Table 6.5 gives the p-values for the LR3 test for all density forecasts for all forecasting horizons. The null hypothesis is that of independence and normality. Once again we see the highest rate of rejections for the RND forecasts for all horizons and for all datasets. The only exceptions where RND forecasts pass the test at any significance level are for 1 and 2 months ahead for Sweden and Netherlands. Also for UK and France only for bimonthly forecast at 1% and 5%, Germany for bimonthly forecast at 1% and Switzerland only for monthly forecast at any level of significance. Fan et al. (2018) found that none of the US stocks passed the test under the RND method as densities are misspecified and have higher variances than the RWDs.

Moving to RWD forecasts an improvement can be seen especially for longer horizons. Shackleton et al. (2010) suggested this could be attributed to the lower power of the test in smaller samples. However, for Sweden and Netherlands now the null hypothesis is rejected at 5% for monthly forecasts. Nevertheless, moving from RND to RWD it means that there are two non rejections at 1% for 1-day

ahead, and 7 out of 8 for for 10-day ahead compared to zero for both horizons.

Moving to historical density forecasts, an improvement is clearly observed and it seems like GARCH-N models produce well specified forecasts for all horizons, but the daily one. For 1-day ahead forecasts the null hypothesis is rejected for all but Italy at 1%. For longer horizons it is only rejected for Sweden and UK but for the latter only for 1-month ahead at 5%. Assuming a different distribution for the GARCH model did not seem to improve the results. Now all daily forecasts are misspecified and UK, Sweden, Germany and France also have more rejections in the longer horizons.

LR3 tests seems to indicate that GJR model is is the best performing model. In particular GJR-N and GJR-T give almost identical results. All daily forecasts seems to be misspecified except for GJR-N for Italy where the p-value is 0.01, for UK were the p-value is 0.03 and for Germany were the p-value is 0.10. For GJR-T the the only exceptions for 1-day ahead are Italy, Spain and UK where we fail to reject the null at 1%. Then GJR-T has no rejections for any horizon for any country at 5% except France for 2-month ahead where the p-value is 0.03. However, GJR-N has no rejections for any level for significance for weekly, biweekly, monthly and bimonthly forecasts for any dataset. On the contrary, Huang and Wang (2012) stated that "non-normal models provide more accurate density prediction than the normal model (...)" (p.167).

To summarize, based on the LR3 tests, overall, we can observe that moving from the RND to RWD improves the results especially for the longer horizons. Still, the historical density forecasts have a clearly better fit and the GJR seems to be the best performing model. However, results needs to be interpreted cautiously as the LR3 test may fail to detect non-normality in the series in their third and fourth moments, as it does not directly tests for normality. So, it is important to complement our analysis with the KS test (Crisóstomo and Couso (2018)).

Table 6.5 – LR3 test results

RND	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
ITALY	0.00	0.00	0.00	0.00	0.00
SPAIN	0.00	0.00	0.00	0.00	0.00
UK	0.00	0.00	0.00	0.00	**0.06**
SWEDEN	0.00	0.00	0.00	**0.11**	**0.39**
FRANCE	0.00	0.00	0.00	0.01	**0.09**
GERMANY	0.00	0.00	0.00	0.00	0.02
SWITZERLAND	0.00	0.00	0.00	**0.48**	0.00
NETHERLANDS	0.00	0.00	0.00	**0.81**	**0.27**
RWD	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
ITALY	0.00	0.00	**0.06**	**0.04**	0.00
SPAIN	0.00	0.00	**0.96**	**0.16**	**0.87**
UK	0.00	0.00	**0.06**	0.03	**0.20**
SWEDEN	0.00	0.00	0.03	0.00	**0.09**
FRANCE	0.00	0.00	0.03	**0.22**	0.03
GERMANY	0.03	0.00	0.01	0.01	0.02
SWITZERLAND	0.00	0.00	**0.08**	**0.82**	**0.49**
NETHERLANDS	**0.13**	0.00	0.00	0.01	**0.13**
HS (GARCH -N)	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
ITALY	0.01	**0.51**	**0.59**	**0.43**	**0.61**
SPAIN	0.00	**0.51**	**0.13**	**0.09**	**0.33**
UK	0.00	**0.06**	**0.70**	0.04	**0.08**
SWEDEN	0.00	**0.06**	0.03	0.01	0.03
FRANCE	0.00	**0.50**	**0.35**	**0.28**	**0.21**
GERMANY	0.00	**0.28**	**0.06**	**0.11**	**0.19**
SWITZERLAND	0.00	**0.52**	**0.67**	**0.31**	**0.23**
NETHERLANDS	0.00	**0.63**	**0.17**	**0.15**	**0.25**
HS (GARCH -T)	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
ITALY	0.00	**0.21**	**0.23**	**0.14**	**0.24**
SPAIN	0.00	**0.41**	**0.05**	0.03	**0.09**
UK	0.00	0.03	**0.7**5	0.01	0.01
SWEDEN	0.00	0.03	0.01	0.01	0.01
FRANCE	0.00	**0.13**	0.02	0.01	0.01
GERMANY	0.00	**0.30**	0.02	0.04	**0.08**
SWITZERLAND	0.00	**0.57**	**0.62**	**0.30**	**0.24**

NETHERLANDS	0.00	**0.41**	**0.06**	**0.06**	**0.0**8
HS(GJR-N)	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
ITALY	0.00	**0.87**	**0.97**	**0.90**	**0.95**
SPAIN	0.00	**0.31**	**0.44**	**0.27**	0.00
UK	0.01	**0.16**	**0.84**	**0.62**	**0.60**
SWEDEN	0.00	**0.34**	**0.59**	**0.12**	**0.35**
FRANCE	0.00	**0.63**	**0.83**	**0.75**	**0.72**
GERMANY	0.00	**0.40**	**0.39**	**0.69**	**0.71**
SWITZERLAND	0.00	**0.67**	**0.90**	**0.56**	**0.36**
NETHERLANDS	0.00	**0.70**	**0.73**	**0.86**	**0.93**
HS (GJR-T)	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
ITALY	0.00	**0.44**	**0.56**	**0.41**	**0.59**
SPAIN	0.00	**0.51**	**0.29**	**0.12**	**0.34**
UK	0.00	**0.11**	**0.85**	**0.11**	**0.15**
SWEDEN	0.00	**0.21**	**0.17**	**0.05**	**0.12**
FRANCE	0.00	**0.40**	**0.14**	**0.10**	**0.05**
GERMANY	0.00	**0.70**	**0.15**	**0.27**	**0.39**
SWITZERLAND	0.00	**0.84**	**0.92**	**0.63**	**0.37**
NETHERLANDS	0.00	**0.76**	**0.36**	**0.39**	**0.38**

Notes: This table shows the p-values for the Likelihood ratio (LR3) test of Berkowitz (2001) for all European indices, for all horizons, for the risk-neutral density (RND) forecasts, the real world density (RWD) forecasts and the historical density forecasts using the GARCH-N, GARCH-T, GJR-N and GJR-T models. The statistically significant results at the 5% level are reported in bold.

Tables 6.6 - 6.11 give the p-values for the LM autocorrelation test for all methods for all horizons for each dataset. The first four moments are being tested using 6,12 and 24 lags. Focusing on RND forecasts we can see that as the forecasting horizon increases the results improve for all datasets. For 1 day and 1 week ahead for the vast majority of the cases the null hypothesis is rejected. However, there are fewer rejections on the 3rd and 4th lags and that for Germany the weekly forecasts seem to pass the test at 1%. For the longer horizons Italy has 12 rejections, Spain has 6, UK has 23, Sweden has 7, France has 15, Germany and Netherlands have 1 and Switzerland has 0 at 5%.

Table 6.6 presents the same test results for RWD forecasts. For the first two horizons there does not seem to be an improvement for most markets and the RWD forecasts seem to fail to pass the test. Also slightly less rejections on the first moments are being observed. Looking at the last three forecasting horizons we see that Italy has 12 rejections, Spain has 7, Uk has 17, Sweden and Netherlands have 15, France has 16, Germany has 8, and Switzerland has 7 rejections. So overall by moving from RND to RWD did not achieve any improvement but some deterioration. However, it is worth mentioning

that for the 2-month ahead forecasts we had 6 rejections relying on RWD compared to 10 at 5% for RND, overall.

On the other hand, historical simulations seems to produce better forecasts for all horizons for each dataset. Table 6.8 - 6.11 present the results. Using the GARCH-N model it gives only a few rejections for each dataset for the shortest horizon. But Switzerland and France pass the test for all moments, lags and horizons at 5% and 10% level. As for the longer horizons the results are very satisfactory as well. Italy has 5 rejections, Spain has 3, UK and Netherlands have 1, Sweden has 0 and Germany has 2 at 5% overall. Similar are the results for GARCH-T, GJR-N and GJR-T. But now France and Germany have no rejections for GARCH-T. In more details, using GARCH-T means that Italy has 6 rejections, Spain has 3, UK, Switzerland and Netherlands have 1, and Sweden has 0 at 5% for weekly, biweekly, monthly and bimonthly forecasts. While using GJR-N or GJR-T will result on only having one dataset (Switzerland) with no rejections at 5% for all moments and each lag for all horizons. What is more, for these models the overall rejection rate is also very low and improves as horizons increases. For instance using the GJR-N specification results in France, Sweden and UK to have no rejections at 5% for any horizon longer than 1-day and under GJR-T specification that is true only for France and Sweden.

Table 6.6 – Autocorrelation test results-RND

	MOMENTS	LAGS	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
ITALY	1	6	0.00	0.02	0.03	**0.16**	**0.8**6
	1	12	0.00	0.00	0.00	**0.45**	**0.18**
	1	24	0.00	0.00	0.00	**0.75**	**0.53**
	2	6	0.00	0.03	0.00	**0.11**	**0.68**
	2	12	0.00	0.00	0.00	**0.06**	**0.45**
	2	24	0.00	0.01	0.00	**0.36**	**0.77**
	3	6	0.00	**0.99**	0.00	**0.10**	**0.73**
	3	12	0.00	**0.76**	0.00	**0.06**	**0.38**
	3	24	0.00	**1.00**	0.00	**0.44**	**0.84**
	4	6	0.00	**1.00**	0.00	**0.14**	**0.71**
	4	12	0.00	**1.00**	0.00	**0.06**	**0.39**
	4	24	0.00	**1.00**	0.00	**0.38**	**0.82**
SPAIN	1	6	0.02	0.02	**0.70**	**0.24**	**0.43**
	1	12	0.00	0.03	**0.64**	**0.48**	**0.70**
	1	24	0.00	0.01	**0.32**	**0.50**	**0.70**
	2	6	0.00	0.00	0.01	**0.58**	0.00
	2	12	0.00	0.00	0.04	**0.68**	**0.36**
	2	24	0.00	0.00	0.00	**0.84**	**0.46**
	3	6	**0.08**	0.00	**0.62**	**0.38**	**0.29**
	3	12	0.00	0.00	**0.89**	**0.55**	**0.29**
	3	24	0.00	0.00	**0.58**	**0.36**	**0.35**
	4	6	0.00	0.00	**0.48**	**0.96**	**0.01**
	4	12	0.00	0.00	**0.87**	**0.97**	**0.32**
	4	24	0.00	0.00	0.02	**1.00**	**0.23**
UK	1	6	**0.43**	0.00	**0.19**	0.01	**0.20**
	1	12	**0.67**	0.00	**0.11**	0.00	**0.49**
	1	24	0.02	0.00	**0.22**	0.01	**0.73**
	2	6	0.00	0.00	0.00	0.00	0.00
	2	12	0.00	0.00	0.00	0.00	0.02
	2	24	0.00	**0.05**	0.00	0.00	**0.37**
	3	6	0.00	0.00	**0.18**	0.01	0.01
	3	12	0.00	0.04	**0.43**	0.00	0.03
	3	24	0.00	**0.55**	0.00	0.00	**0.65**
	4	6	0.00	0.00	**0.25**	0.04	0.00

	4	12	0.00	0.08	**0.72**	0.00	0.03
	4	24	0.00	**0.73**	0.00	0.00	**0.21**
SWEDEN	1	6	0.02	0.00	0.00	**0.39**	**0.67**
	1	12	0.00	0.00	0.00	0.00	**0.13**
	1	24	0.00	0.00	0.00	**0.41**	**0.34**
	2	6	0.00	0.00	**0.15**	**0.82**	**0.59**
	2	12	0.00	0.00	**0.64**	**0.59**	**0.55**
	2	24	0.00	0.00	0.00	**0.05**	**0.89**
	3	6	**0.66**	**0.32**	**0.98**	**1.00**	**0.49**
	3	12	0.00	**0.77**	**1.00**	**1.00**	**0.49**
	3	24	0.00	**1.00**	0.00	**0.57**	**0.69**
	4	6	0.00	**0.99**	**1.00**	**1.00**	**0.85**
	4	12	0.00	**1.00**	**1.00**	**1.00**	**0.94**
	4	24	0.00	**1.00**	0.00	**0.92**	**0.50**
FRANCE	1	6	0.02	0.02	**0.24**	**0.10**	**0.18**
	1	12	0.00	**0.24**	**0.12**	**0.14**	**0.67**
	1	24	0.00	**0.30**	**0.16**	**0.60**	**0.87**
	2	6	0.00	0.00	0.00	0.00	**0.11**
	2	12	0.00	0.00	0.00	0.00	**0.26**
	2	24	0.00	0.00	0.00	0.00	**0.59**
	3	6	**1.00**	**0.88**	0.00	0.03	**0.09**
	3	12	**1.00**	**0.99**	0.00	0.02	**0.37**
	3	24	**1.00**	**0.42**	0.00	**0.18**	**0.57**
	4	6	**1.00**	**1.00**	0.00	**0.06**	**0.36**
	4	12	**1.00**	**1.00**	0.00	**0.05**	**0.82**
	4	24	**1.00**	**1.00**	0.00	0.03	**0.79**
GERMANY	1	6	0.00	0.02	**0.14**	**0.45**	**0.08**
	1	12	0.00	0.01	**0.20**	**0.15**	**0.25**
	1	24	0.00	0.03	0.00	**0.60**	**0.41**
	2	6	0.00	**0.96**	**1.00**	**0.99**	**0.87**
	2	12	0.00	**1.00**	**1.00**	**0.98**	**0.96**
	2	24	0.00	**1.00**	**1.00**	**1.00**	**0.31**
	3	6	0.00	**1.00**	**1.00**	**1.00**	**0.98**
	3	12	0.00	**1.00**	**1.00**	**1.00**	**0.99**
	3	24	0.00	**1.00**	**1.00**	**1.00**	**0.30**
	4	6	0.01	**1.00**	**1.00**	**1.00**	**1.00**

	4	12	**0.05**	**1.00**	**1.00**	**1.00**	**1.00**
	4	24	**0.58**	**1.00**	**1.00**	**1.00**	**0.33**
SWITZERLAND	1	6	0.00	0.03	**0.95**	**0.86**	**0.09**
	1	12	0.00	0.00	**0.88**	**0.96**	**0.36**
	1	24	0.00	0.00	**1.00**	**1.00**	**0.41**
	2	6	0.00	**0.77**	**1.00**	**1.00**	**0.29**
	2	12	0.00	0.02	**1.00**	**1.00**	**0.31**
	2	24	0.00	**0.31**	**1.00**	**1.00**	**0.41**
	3	6	0.00	**1.00**	**1.00**	**1.00**	**0.16**
	3	12	0.00	**0.99**	**1.00**	**1.00**	**0.56**
	3	24	0.00	**1.00**	**1.00**	**1.00**	**0.41**
	4	6	0.00	**1.00**	**1.00**	**1.00**	**0.62**
	4	12	0.00	**1.00**	**1.00**	**1.00**	**0.48**
	4	24	0.00	**1.00**	**1.00**	**1.00**	**0.41**
NETHERLANDS	1	6	0.01	0.00	**0.38**	0.02	**0.16**
	1	12	0.01	0.00	**0.10**	**0.05**	**0.45**
	1	24	0.00	0.00	0.00	**0.15**	**0.34**
	2	6	0.00	0.00	**0.98**	0.00	0.03
	2	12	0.00	0.00	**0.90**	0.02	**0.11**
	2	24	0.00	0.00	**0.19**	**0.16**	**0.54**
	3	6	**0.51**	0.00	**1.00**	**0.37**	0.02
	3	12	0.00	0.00	**1.00**	**0.82**	**0.26**
	3	24	0.00	0.00	**1.00**	**1.00**	**0.07**
	4	6	**0.99**	0.00	**1.00**	**0.74**	**0.06**
	4	12	0.00	0.00	**1.00**	**0.99**	**0.42**
	4	24	0.00	0.00	**1.00**	**1.00**	**0.91**

Notes: This table shows the p-values for the Lagrange Multiplier (LM) test of no autocorrelation for all European indices, for all horizons, for the risk-neutral density (RND) forecasts. The statistically significant results at the 5% level are reported in bold.

Table 6.7 – Autocorrelation test results - RWD

	MOMENTS	LAGS	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
ITALY	1	6	0.00	**0.65**	**0.24**	**0.65**	**0.87**
	1	12	0.00	**0.16**	0.00	**0.88**	**0.12**
	1	24	0.01	0.03	0.02	**0.98**	**0.54**
	2	6	0.00	0.00	0.00	0.00	**0.57**
	2	12	0.00	0.00	0.00	0.00	**0.89**
	2	24	0.00	0.00	0.00	**0.07**	**1.00**
	3	6	0.00	0.04	**0.76**	**0.20**	**0.80**
	3	12	0.00	0.02	**0.11**	**0.55**	**0.46**
	3	24	0.00	0.00	0.00	**0.89**	**0.8**9
	4	6	0.00	0.00	0.00	0.00	**0.63**
	4	12	0.00	0.00	0.00	0.01	**0.91**
	4	24	0.00	0.00	0.00	**0.16**	**0.99**
SPAIN	1	6	**0.35**	**0.8**1	**0.87**	**0.25**	**0.46**
	1	12	**0.67**	**0.75**	**0.74**	**0.28**	**0.80**
	1	24	**0.46**	**0.55**	**0.31**	**0.48**	**0.63**
	2	6	0.00	0.00	0.00	**0.05**	0.03
	2	12	0.00	0.00	0.00	0.02	**0.40**
	2	24	0.00	0.00	0.00	**0.27**	**0.61**
	3	6	0.02	0.02	**0.35**	**0.36**	**0.56**
	3	12	**0.05**	**0.07**	**0.54**	**0.52**	**0.60**
	3	24	0.01	**0.10**	**0.30**	**0.07**	**0.25**
	4	6	0.00	0.00	0.01	**0.33**	0.03
	4	12	0.00	0.00	**0.11**	0.00	**0.44**
	4	24	0.00	0.01	0.00	0.02	**0.16**
UK	1	6	**0.38**	0.00	**0.13**	**0.10**	**0.50**
	1	12	**0.69**	0.00	**0.08**	0.04	**0.83**
	1	24	**0.32**	0.00	**0.17**	**0.08**	**0.91**
	2	6	0.00	0.00	0.00	0.00	0.00
	2	12	0.00	0.00	0.00	0.00	0.01
	2	24	0.00	0.00	0.00	0.00	**0.44**
	3	6	**0.36**	0.00	**0.07**	**0.05**	**0.12**
	3	12	0.03	0.00	0.04	**0.05**	**0.35**
	3	24	0.00	0.00	0.00	**0.07**	**0.81**
	4	6	0.00	0.00	0.00	0.00	0.00

	4	12	0.00	0.00	0.00	0.00	**0.10**
	4	24	0.00	0.00	0.00	0.00	**0.75**
SWEDEN	1	6	0.04	0.00	**0.32**	**0.66**	**0.63**
	1	12	0.01	0.01	**0.12**	**0.42**	**0.1**1
	1	24	0.01	0.01	0.03	**0.95**	**0.34**
	2	6	0.00	0.00	0.00	0.01	**0.53**
	2	12	0.00	0.00	0.00	0.01	**0.28**
	2	24	0.00	0.00	0.00	**0.08**	**0.62**
	3	6	**0.07**	0.00	0.00	**0.31**	**0.68**
	3	12	0.02	0.00	0.00	0.02	**0.49**
	3	24	0.00	0.00	0.00	**0.13**	**0.66**
	4	6	0.00	0.00	0.00	**0.06**	**0.33**
	4	12	0.00	0.00	0.00	0.01	**0.31**
	4	24	0.00	0.00	0.00	0.03	**0.73**
FRANCE	1	6	**0.24**	0.03	**0.80**	**0.89**	**0.08**
	1	12	**0.30**	**0.20**	**0.64**	**0.58**	**0.25**
	1	24	**0.38**	**0.14**	**0.66**	**0.86**	**0.55**
	2	6	0.00	0.00	0.00	0.00	0.01
	2	12	0.00	0.00	0.00	0.00	**0.05**
	2	24	0.00	0.00	0.00	0.00	**0.37**
	3	6	**0.18**	**0.05**	**0.38**	0.04	**0.08**
	3	12	**0.06**	0.04	**0.31**	0.02	**0.18**
	3	24	0.01	0.00	0.00	**0.18**	**0.52**
	4	6	0.00	0.00	0.00	0.00	**0.08**
	4	12	0.00	0.00	0.00	0.00	**0.21**
	4	24	0.00	0.00	0.00	0.00	**0.42**
GERMANY	1	6	0.00	**0.05**	**0.06**	**0.34**	**0.19**
	1	12	0.00	0.01	0.02	**0.23**	**0.61**
	1	24	0.03	0.03	0.01	**0.50**	**0.61**
	2	6	0.00	0.00	0.00	**0.90**	**0.45**
	2	12	0.00	0.00	0.00	**0.48**	**0.52**
	2	24	0.00	0.00	0.01	**0.59**	**0.77**
	3	6	0.04	0.02	**0.05**	**0.37**	**0.59**
	3	12	0.00	0.01	**0.16**	**0.05**	**0.71**
	3	24	0.00	0.01	0.01	**0.32**	**0.44**
	4	6	0.00	0.00	**0.05**	**0.99**	**0.74**

4	12	0.00	0.00	**0.21**	**0.80**	**0.91**
4	24	0.00	0.00	**0.30**	**0.69**	**0.90**
SWITZERLAND 1	6	0.01	**0.08**	**0.30**	**0.36**	**0.47**
1	12	**0.07**	0.04	**0.39**	**0.60**	**0.21**
1	24	0.02	**0.08**	**0.81**	**0.97**	**0.41**
2	6	0.00	0.00	0.00	**0.81**	0.03
2	12	0.00	0.00	0.00	**0.31**	**0.4**1
2	24	0.00	0.00	0.01	**0.86**	**0.41**
3	6	0.00	**0.07**	**0.21**	**0.38**	**0.13**
3	12	0.00	0.01	**0.08**	**0.52**	**0.41**
3	24	0.00	0.00	**0.08**	**0.79**	**0.41**
4	6	0.00	0.00	0.01	**0.76**	**0.13**
4	12	0.00	0.00	0.00	**0.60**	**0.76**
4	24	0.00	0.00	0.04	**0.97**	**0.41**
NETHERLANDS 1	6	**0.09**	**0.11**	**0.52**	**0.44**	**0.23**
1	12	**0.20**	0.01	**0.36**	**0.50**	**0.47**
1	24	**0.10**	0.00	**0.2**1	**0.75**	**0.25**
2	6	0.00	0.00	0.00	0.00	0.02
2	12	0.00	0.00	0.00	0.00	**0.17**
2	24	0.00	0.00	0.00	0.00	**0.43**
3	6	0.00	0.00	**0.45**	**0.28**	0.02
3	12	0.03	0.00	**0.34**	**0.49**	**0.26**
3	24	0.00	0.00	0.09	**0.68**	**0.11**
4	6	0.00	0.00	0.00	0.00	0.00
4	12	0.00	0.00	0.00	0.00	**0.05**
4	24	0.00	0.00	0.00	0.02	**0.25**

Notes: This table shows the p-values for the Lagrange Multiplier (LM) test of no autocorrelation, for all European indices, for all horizons for the real world density (RWD) forecasts. The statistically significant results at the 5% level are reported in bold.

Table 6.8 – Autocorrelation test results-Historical Simulation GARCH-N

	MOMENTS	LAGS	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
ITALY	1	6	0.00	**0.41**	**0.82**	**0.41**	**0.60**
	1	12	0.02	**0.84**	**0.58**	**0.41**	**0.93**
	1	24	**0.07**	**0.62**	**0.81**	**0.74**	**0.97**
	2	6	**0.08**	**0.25**	**0.40**	**0.08**	**0.39**
	2	12	**0.15**	0.00	**0.37**	**0.27**	**0.54**
	2	24	**0.11**	**0.05**	**0.48**	**0.21**	**0.97**
	3	6	0.00	**0.38**	**0.98**	0.02	**0.96**
	3	12	0.02	**0.28**	**0.98**	**0.13**	**0.89**
	3	24	**0.16**	**0.44**	**0.94**	**0.14**	**0.97**
	4	6	0.02	**0.61**	**0.68**	0.01	**0.92**
	4	12	0.00	0.03	**0.77**	**0.11**	**0.83**
	4	24	0.01	**0.16**	**0.89**	0.03	**0.97**
SPAIN	1	6	**0.05**	0.01	**0.39**	**0.95**	**0.76**
	1	12	0.03	**0.16**	**0.80**	**0.94**	**0.98**
	1	24	0.00	**0.50**	**0.91**	**0.93**	**1.00**
	2	6	0.00	**0.85**	**0.74**	**0.17**	0.04
	2	12	0.00	**0.65**	**0.52**	**0.44**	**0.46**
	2	24	0.00	**0.61**	**0.60**	**0.47**	**1.00**
	3	6	**0.11**	**0.26**	**0.99**	**0.21**	**0.94**
	3	12	**0.11**	**0.55**	**0.84**	**0.52**	**0.63**
	3	24	**0.11**	**0.71**	**0.99**	**0.57**	**1.00**
	4	6	**0.29**	**0.93**	**1.00**	**0.13**	0.02
	4	12	**0.29**	**0.95**	**1.00**	**0.46**	**0.47**
	4	24	**0.34**	**0.95**	**1.00**	**0.93**	**1.00**
UK	1	6	**0.71**	0.02	**0.52**	**0.29**	**0.62**
	1	12	**0.93**	**0.11**	**0.60**	**0.40**	**0.56**
	1	24	**0.90**	**0.15**	**0.96**	**0.11**	**0.79**
	2	6	0.01	**0.35**	**0.70**	**1.00**	**0.53**
	2	12	0.01	**0.31**	**0.94**	**0.95**	**0.99**
	2	24	0.00	**0.15**	**0.75**	**0.85**	**0.78**
	3	6	**0.92**	**0.22**	**0.81**	**0.95**	**0.30**
	3	12	**0.91**	**0.62**	**0.99**	**1.00**	**0.79**
	3	24	**0.97**	**0.50**	**0.96**	**0.78**	**0.84**
	4	6	**0.56**	**0.99**	**0.81**	**0.98**	**0.30**

	4	12	0.00	**1.00**	**0.99**	**0.99**	**1.00**
	4	24	0.01	**1.00**	**0.99**	**1.00**	**0.82**
SWEDEN	1	6	0.02	**0.44**	**0.07**	**0.24**	**0.62**
	1	12	**0.06**	**0.87**	**0.09**	**0.56**	**0.40**
	1	24	**0.14**	**0.67**	**0.47**	**0.38**	**0.74**
	2	6	0.01	**0.31**	**0.62**	**0.62**	**0.06**
	2	12	0.00	**0.14**	**0.74**	**0.75**	**0.26**
	2	24	0.01	**0.11**	**0.73**	**0.83**	**0.74**
	3	6	**0.47**	**0.96**	**0.85**	**0.85**	**0.73**
	3	12	**0.37**	**0.99**	**0.98**	**0.98**	**0.09**
	3	24	**0.58**	**0.77**	**1.00**	**0.99**	**0.44**
	4	6	**0.08**	**0.99**	**0.94**	**0.93**	**0.20**
	4	12	0.00	**0.86**	**1.00**	**0.98**	**0.21**
	4	24	0.04	**0.45**	**1.00**	**1.00**	**0.43**
FRANCE	1	6	**0.16**	**0.73**	**0.84**	**0.56**	**0.96**
	1	12	**0.51**	**0.90**	**0.65**	**0.76**	**0.58**
	1	24	**0.71**	**0.93**	**0.96**	**0.73**	**0.26**
	2	6	**0.17**	**0.52**	**0.64**	**0.86**	**0.39**
	2	12	**0.15**	**0.07**	**0.74**	**0.98**	**0.9**9
	2	24	**0.63**	**0.06**	**0.87**	**1.00**	**0.75**
	3	6	**0.51**	**0.81**	**1.00**	**0.98**	**0.97**
	3	12	**0.89**	**0.80**	**0.99**	**1.00**	**1.00**
	3	24	**0.99**	**0.78**	**1.00**	**1.00**	**0.78**
	4	6	**0.97**	**0.77**	**0.99**	**1.00**	**0.90**
	4	12	**0.64**	**0.61**	**1.00**	**1.00**	**1.00**
	4	24	**0.88**	**0.95**	**1.00**	**1.00**	**0.93**
GERMANY	1	6	**0.16**	**0.55**	**0.25**	**0.47**	**0.74**
	1	12	**0.50**	**0.77**	**0.12**	**0.45**	**0.92**
	1	24	**0.61**	**0.77**	**0.73**	**0.67**	**0.54**
	2	6	**0.04**	**0.78**	**0.24**	**0.97**	**0.99**
	2	12	**0.14**	**0.76**	**0.41**	**0.97**	**0.98**
	2	24	**0.23**	**0.14**	**0.65**	**0.96**	**0.61**
	3	6	**0.35**	**0.63**	**0.85**	**0.87**	**0.73**
	3	12	**0.11**	**0.95**	**0.31**	**0.88**	**0.9**8
	3	24	**0.15**	0.03	**0.40**	**0.92**	**0.88**
	4	6	0.04	**0.99**	**0.65**	**0.97**	**1.00**

	4	12	0.00	**1.00**	**0.89**	**0.88**	**1.00**
	4	24	0.01	0.00	**0.84**	**0.99**	**0.89**
SWITZERLAND	1	6	**0.40**	**0.09**	**0.66**	**0.60**	**0.89**
	1	12	**0.32**	**0.28**	**0.70**	**0.66**	**0.80**
	1	24	**0.23**	**0.72**	**0.50**	**0.86**	**1.00**
	2	6	**0.66**	**0.93**	**0.59**	0.01	**0.88**
	2	12	**0.73**	**0.58**	**0.88**	**0.10**	**0.67**
	2	24	**0.91**	**0.80**	**0.58**	**0.45**	**1.00**
	3	6	**0.18**	**0.35**	**0.96**	**0.48**	**1.00**
	3	12	**0.38**	**0.22**	**1.00**	**0.82**	**0.27**
	3	24	**0.53**	**0.61**	**0.99**	**0.95**	**1.00**
	4	6	**0.95**	**0.85**	**0.94**	**0.08**	**1.00**
	4	12	**0.21**	**0.90**	**1.00**	**0.57**	**0.20**
	4	24	**0.69**	**0.99**	**1.00**	**0.93**	**1.00**
NETHERLANDS	1	6	**0.06**	**0.8**8	**0.82**	**0.85**	**0.67**
	1	12	**0.42**	**0.89**	**0.79**	**0.73**	**0.40**
	1	24	**0.45**	**0.73**	**0.96**	**0.89**	**0.45**
	2	6	0.00	**0.80**	**0.40**	**0.97**	**0.48**
	2	12	0.00	**0.87**	**0.75**	**0.98**	**0.98**
	2	24	0.00	**0.24**	**0.90**	**0.97**	**0.85**
	3	6	**0.30**	**0.97**	**0.96**	**0.98**	**0.61**
	3	12	**0.08**	**1.00**	**0.65**	**0.99**	**0.94**
	3	24	**0.20**	**0.10**	**0.66**	**1.00**	**0.40**
	4	6	0.00	**0.98**	**0.59**	**0.99**	**0.54**
	4	12	0.00	**1.00**	**0.52**	**0.99**	**1.00**
	4	24	0.03	0.00	**0.57**	**1.00**	**0.69**

Notes: This table shows the p-values for the Lagrange Multiplier (LM) test of no autocorrelation, for all European indices, for all horizons for the historical density forecasts using the GARCH-N model. The statistically significant results at the 5% level are reported in bold.

Table 6.9 – Autocorrelation test results-Historical Simulation GARCH-T

	MOMENTS	LAGS	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
ITALY	1	6	0.01	**0.39**	**0.81**	**0.44**	**0.58**
	1	12	0.02	**0.81**	**0.5**4	**0.42**	**0.92**
	1	24	**0.07**	**0.60**	**0.81**	**0.74**	**0.97**
	2	6	**0.16**	**0.21**	**0.44**	**0.14**	**0.43**
	2	12	**0.42**	0.00	**0.49**	**0.34**	**0.50**
	2	24	**0.21**	0.04	**0.48**	**0.29**	**0.97**
	3	6	0.01	**0.42**	**0.96**	0.04	**0.93**
	3	12	**0.05**	**0.36**	**0.96**	**0.18**	**0.81**
	3	24	**0.20**	**0.43**	**0.90**	**0.16**	**0.97**
	4	6	0.02	**0.49**	**0.70**	0.01	**0.88**
	4	12	0.01	0.02	**0.84**	**0.14**	**0.78**
	4	24	0.00	**0.09**	**0.82**	0.03	**0.97**
SPAIN	1	6	0.01	0.01	**0.35**	**0.95**	**0.80**
	1	12	0.01	**0.15**	**0.78**	**0.95**	**0.95**
	1	24	0.00	**0.49**	**0.91**	**0.94**	**1.00**
	2	6	0.00	**0.86**	**0.42**	**0.12**	0.02
	2	12	0.00	**0.67**	**0.45**	**0.18**	**0.33**
	2	24	0.00	**0.61**	**0.64**	**0.69**	**1.00**
	3	6	**0.19**	**0.22**	**0.87**	**0.23**	**0.73**
	3	12	**0.11**	**0.55**	**0.54**	**0.47**	**0.98**
	3	24	**0.06**	**0.70**	**0.87**	**0.56**	**1.00**
	4	6	**0.17**	**0.89**	**0.98**	**0.11**	0.00
	4	12	**0.06**	**0.91**	**0.90**	**0.30**	**0.18**
	4	24	0.04	**0.87**	**0.98**	**0.89**	**1.00**
UK	1	6	**0.65**	0.02	**0.52**	**0.29**	**0.64**
	1	12	**0.89**	**0.13**	**0.65**	**0.39**	**0.49**
	1	24	**0.86**	**0.13**	**0.96**	**0.10**	**0.60**
	2	6	0.00	**0.22**	**0.60**	**1.00**	**0.47**
	2	12	0.00	**0.19**	**0.88**	**0.93**	**0.99**
	2	24	0.00	**0.08**	**0.61**	**0.79**	**0.76**
	3	6	**0.87**	**0.12**	**0.96**	**0.92**	**0.26**
	3	12	**0.92**	**0.41**	**0.98**	**0.98**	**0.74**
	3	24	**0.96**	**0.33**	**0.93**	**0.66**	**0.74**
	4	6	**0.19**	**0.94**	**0.99**	**0.98**	**0.17**

	4	12	0.00	**0.98**	**0.99**	**0.99**	**1.00**
	4	24	0.00	**0.95**	**0.99**	**0.99**	**0.68**
SWEDEN	1	6	0.02	**0.44**	**0.10**	**0.20**	**0.53**
	1	12	**0.08**	**0.86**	**0.09**	**0.52**	**0.38**
	1	24	**0.13**	**0.66**	**0.45**	**0.41**	**0.67**
	2	6	0.00	**0.24**	**0.64**	**0.44**	**0.07**
	2	12	0.00	**0.13**	**0.69**	**0.69**	**0.26**
	2	24	0.00	**0.11**	**0.78**	**0.78**	**0.86**
	3	6	**0.28**	**0.94**	**0.81**	**0.72**	**0.62**
	3	12	**0.26**	**0.98**	**0.97**	**0.98**	**0.16**
	3	24	**0.35**	**0.58**	**1.00**	**0.97**	**0.71**
	4	6	0.00	**0.98**	**0.89**	**0.89**	**0.12**
	4	12	0.00	**0.86**	**0.99**	**0.97**	**0.31**
	4	24	0.00	**0.30**	**1.00**	**1.00**	**0.78**
FRANCE	1	6	**0.14**	**0.61**	**0.85**	**0.50**	**0.88**
	1	12	**0.48**	**0.86**	**0.67**	**0.70**	**0.44**
	1	24	**0.65**	**0.91**	**0.97**	**0.62**	**0.22**
	2	6	**0.06**	**0.57**	**0.57**	**0.89**	**0.49**
	2	12	**0.05**	**0.17**	**0.42**	**0.96**	**0.95**
	2	24	**0.40**	**0.14**	**0.76**	**0.99**	**0.47**
	3	6	**0.40**	**0.80**	**1.00**	**0.95**	**0.96**
	3	12	**0.77**	**0.82**	**0.93**	**1.00**	**1.00**
	3	24	**0.96**	**0.70**	**1.00**	**1.00**	**0.53**
	4	6	**0.75**	**0.67**	**0.95**	**1.00**	**0.88**
	4	12	**0.08**	**0.47**	**0.99**	**1.00**	**1.00**
	4	24	**0.40**	**0.75**	**1.00**	**1.00**	**0.79**
GERMANY	1	6	**0.14**	**0.56**	**0.17**	**0.49**	**0.74**
	1	12	**0.48**	**0.74**	**0.09**	**0.45**	**0.97**
	1	24	**0.57**	**0.79**	**0.68**	**0.72**	**0.56**
	2	6	**0.05**	**0.73**	**0.32**	**0.82**	**0.94**
	2	12	**0.18**	**0.73**	**0.34**	**0.97**	**0.89**
	2	24	**0.23**	**0.29**	**0.69**	**0.88**	**0.76**
	3	6	**0.66**	**0.68**	**0.74**	**0.74**	**0.63**
	3	12	**0.60**	**0.94**	**0.17**	**0.92**	**0.83**
	3	24	**0.70**	**0.29**	**0.26**	**0.81**	**0.96**
	4	6	**0.29**	**0.99**	**0.61**	**0.93**	**0.99**

	4	12	**0.32**	**1.00**	**0.71**	**0.96**	**0.99**
	4	24	**0.24**	0.02	**0.75**	**0.98**	**0.98**
SWITZERLAND	1	6	**0.52**	**0.07**	**0.67**	**0.54**	**0.75**
	1	12	**0.43**	**0.25**	**0.68**	**0.59**	**0.79**
	1	24	**0.31**	**0.68**	**0.50**	**0.84**	**1.00**
	2	6	**0.56**	**0.96**	**0.66**	0.01	**0.87**
	2	12	**0.75**	**0.61**	**0.90**	**0.15**	**0.64**
	2	24	**0.91**	**0.80**	**0.56**	**0.38**	**1.00**
	3	6	**0.17**	**0.25**	**0.94**	**0.20**	**0.99**
	3	12	**0.47**	**0.22**	**0.99**	**0.54**	**0.24**
	3	24	**0.44**	**0.46**	**0.92**	**0.84**	**1.00**
	4	6	**0.80**	**0.80**	**0.85**	0.03	**1.00**
	4	12	**0.48**	**0.79**	**0.99**	**0.35**	**0.17**
	4	24	**0.83**	**0.93**	**0.91**	**0.84**	**1.00**
NETHERLANDS	1	6	0.04	**0.92**	**0.72**	**0.85**	**0.54**
	1	12	**0.29**	**0.86**	**0.76**	**0.73**	**0.52**
	1	24	**0.36**	**0.76**	**0.96**	**0.93**	**0.48**
	2	6	0.02	**0.67**	**0.42**	**0.97**	**0.44**
	2	12	**0.11**	**0.81**	**0.73**	**1.00**	**0.98**
	2	24	**0.14**	**0.32**	**0.94**	**0.98**	**0.76**
	3	6	**0.30**	**0.95**	**0.96**	**0.99**	**0.45**
	3	12	**0.53**	**1.00**	**0.62**	**1.00**	**0.96**
	3	24	**0.78**	**0.49**	**0.72**	**1.00**	**0.39**
	4	6	**0.07**	**0.96**	**0.58**	**0.99**	**0.45**
	4	12	**0.26**	**0.99**	**0.39**	**1.00**	**1.00**
	4	24	**0.17**	0.00	**0.46**	**1.00**	**0.59**

Notes: This table shows the p-values for the Lagrange Multiplier (LM) test of no autocorrelation, for all European indices, for all horizons for the historical density forecasts using the GARCH-T model. The statistically significant results at the 5% level are reported in bold.

Table 6.10 – Autocorrelation test results-Historical Simulation GJR-N

	MOMENTS	LAGS	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
ITALY	1	6	0.01	**0.47**	**0.82**	**0.58**	**0.58**
	1	12	0.02	**0.89**	**0.67**	**0.53**	**0.91**
	1	24	**0.08**	**0.72**	**0.85**	**0.70**	**0.97**
	2	6	**0.42**	**0.30**	**0.39**	**0.02**	**0.53**
	2	12	**0.75**	**0.04**	**0.19**	**0.10**	**0.72**
	2	24	**0.86**	**0.30**	**0.31**	**0.35**	**0.97**
	3	6	**0.20**	**0.24**	**0.97**	**0.29**	**0.83**
	3	12	**0.47**	**0.47**	**0.99**	**0.48**	**0.98**
	3	24	**0.34**	**0.78**	**0.96**	**0.57**	**0.97**
	4	6	**0.62**	**0.52**	**0.48**	0.01	**0.66**
	4	12	**0.82**	**0.60**	**0.51**	**0.15**	**0.51**
	4	24	**0.84**	**0.92**	**0.54**	**0.23**	**0.97**
SPAIN	1	6	0.02	0.01	**0.50**	**0.93**	**0.77**
	1	12	0.02	**0.14**	**0.88**	**0.97**	**0.99**
	1	24	0.00	**0.48**	**0.96**	**0.96**	**1.00**
	2	6	**0.08**	**0.48**	**0.81**	**0.60**	**0.79**
	2	12	0.01	**0.47**	**0.73**	**0.90**	**0.84**
	2	24	0.00	**0.45**	**0.76**	**0.44**	**1.00**
	3	6	**0.09**	**0.07**	**0.95**	**0.33**	**0.87**
	3	12	**0.25**	**0.25**	**0.72**	**0.68**	**0.97**
	3	24	0.02	**0.40**	**0.94**	**0.50**	**1.00**
	4	6	**0.92**	**0.59**	**0.99**	**0.82**	**1.00**
	4	12	**0.93**	**0.70**	**1.00**	**0.94**	**0.96**
	4	24	**0.81**	**0.70**	**1.00**	**0.80**	**1.00**
UK	1	6	**0.59**	**0.05**	**0.73**	**0.37**	**0.47**
	1	12	**0.92**	**0.19**	**0.80**	**0.49**	**0.52**
	1	24	**0.95**	**0.22**	**0.98**	**0.16**	**0.84**
	2	6	0.00	**0.35**	**0.60**	**0.97**	**0.25**
	2	12	0.00	**0.22**	**0.88**	**0.72**	**0.55**
	2	24	0.00	**0.17**	**0.60**	**0.81**	**0.98**
	3	6	**0.37**	**0.13**	**0.95**	**0.94**	**0.24**
	3	12	**0.52**	**0.47**	**0.90**	**1.00**	**0.61**
	3	24	**0.43**	**0.67**	**0.91**	**0.42**	**0.95**
	4	6	**0.10**	**0.93**	**0.81**	**0.99**	**0.79**

	4	12	0.02	**0.98**	**0.98**	**0.96**	**0.94**
	4	24	0.00	**0.99**	**0.63**	**0.99**	**0.96**
SWEDEN	1	6	0.03	**0.73**	**0.11**	**0.15**	**0.32**
	1	12	**0.09**	**0.98**	**0.11**	**0.58**	**0.33**
	1	24	**0.20**	**0.84**	**0.45**	**0.63**	**0.76**
	2	6	0.02	**0.28**	**0.42**	**0.37**	**0.21**
	2	12	0.00	**0.15**	**0.49**	**0.64**	**0.51**
	2	24	**0.07**	**0.20**	**0.55**	**0.82**	**0.44**
	3	6	**0.86**	**0.76**	**0.92**	**0.67**	**0.95**
	3	12	0.02	**0.86**	**0.99**	**0.99**	**0.09**
	3	24	**0.24**	**0.61**	**1.00**	**0.89**	**0.24**
	4	6	**0.99**	**0.95**	**0.87**	**0.86**	**0.33**
	4	12	**0.58**	**0.60**	**0.97**	**0.97**	**0.73**
	4	24	**0.99**	**0.58**	**0.99**	**1.00**	**0.20**
FRANCE	1	6	**0.10**	**0.73**	**0.82**	**0.57**	**0.92**
	1	12	**0.39**	**0.90**	**0.69**	**0.74**	**0.62**
	1	24	**0.65**	**0.94**	**0.98**	**0.67**	**0.30**
	2	6	0.03	**0.56**	**0.81**	**0.72**	**0.07**
	2	12	0.01	**0.17**	**0.30**	**0.87**	**0.77**
	2	24	0.02	**0.06**	**0.32**	**0.99**	**0.43**
	3	6	**0.22**	**1.00**	**0.99**	**0.83**	**0.86**
	3	12	**0.45**	**0.99**	**0.84**	**0.99**	**0.92**
	3	24	**0.55**	**0.73**	**0.96**	**1.00**	**0.28**
	4	6	**0.16**	**0.97**	**0.91**	**0.99**	**0.28**
	4	12	**0.07**	**0.92**	**0.89**	**1.00**	**1.00**
	4	24	**0.06**	**0.47**	**0.98**	**1.00**	**0.30**
GERMANY	1	6	**0.05**	**0.61**	**0.19**	**0.73**	**0.58**
	1	12	**0.23**	**0.84**	**0.11**	**0.47**	**0.97**
	1	24	**0.39**	**0.86**	**0.71**	**0.72**	**0.51**
	2	6	**0.05**	**0.82**	**0.23**	**0.93**	**0.45**
	2	12	**0.22**	**0.70**	**0.17**	**0.94**	**0.95**
	2	24	**0.06**	**0.41**	**0.28**	**0.89**	**0.60**
	3	6	**0.08**	**0.34**	**0.19**	**0.80**	**0.09**
	3	12	0.03	**0.80**	0.00	**0.88**	**0.91**
	3	24	0.02	**0.19**	0.02	**0.67**	**0.79**
	4	6	**0.84**	**0.99**	**0.10**	**0.96**	**0.85**

	4	12	**0.89**	**0.99**	0.01	**0.88**	**0.97**
	4	24	**0.94**	0.04	0.01	**0.99**	**0.63**
SWITZERLAND	1	6	**0.62**	**0.06**	**0.66**	**0.54**	**0.81**
	1	12	**0.48**	**0.25**	**0.74**	**0.62**	**0.89**
	1	24	**0.38**	**0.70**	**0.59**	**0.90**	**1.00**
	2	6	**0.14**	**0.79**	**0.47**	**0.06**	**0.28**
	2	12	**0.45**	**0.59**	**0.73**	**0.39**	**0.46**
	2	24	**0.59**	**0.73**	**0.38**	**0.36**	**1.00**
	3	6	**0.92**	**0.14**	**0.98**	**0.43**	**0.96**
	3	12	**0.88**	**0.27**	**1.00**	**0.58**	**0.45**
	3	24	**0.81**	**0.54**	**0.87**	**0.99**	**1.00**
	4	6	**0.62**	**0.80**	**0.85**	**0.12**	**0.80**
	4	12	**0.91**	**0.52**	**1.00**	**0.64**	**0.63**
	4	24	**0.86**	**0.60**	**0.86**	**0.86**	**1.00**
NETHERLANDS	1	6	0.02	**0.94**	**0.78**	**0.94**	**0.46**
	1	12	**0.24**	**0.92**	**0.83**	**0.84**	**0.47**
	1	24	**0.35**	**0.81**	**0.98**	**0.96**	**0.58**
	2	6	0.00	**0.57**	**0.28**	**0.87**	**0.25**
	2	12	0.00	**0.68**	**0.74**	**0.97**	**0.75**
	2	24	0.00	**0.50**	**0.86**	**0.84**	**0.99**
	3	6	0.04	**0.99**	**0.87**	**0.93**	**0.22**
	3	12	0.00	**1.00**	**0.90**	**0.99**	**0.56**
	3	24	0.03	**0.75**	**0.74**	**1.00**	**0.58**
	4	6	**0.91**	**0.80**	**0.49**	**0.95**	**0.39**
	4	12	**0.53**	**0.92**	**0.83**	**0.98**	**0.93**
	4	24	**0.97**	0.03	**0.60**	**1.00**	**0.99**

Notes: This table shows the p-values for the Lagrange Multiplier (LM) test of no autocorrelation, for all European indices, for all horizons for the historical density forecasts using the GJR-N model. The statistically significant results at the 5% level are reported in bold.

Table 6.11 – Autocorrelation test results-Historical Simulation GJR-T

	MOMENTS	LAGS	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
ITALY	1	6	0.01	**0.43**	**0.81**	**0.61**	**0.55**
	1	12	0.02	**0.87**	**0.64**	**0.54**	**0.90**
	1	24	**0.09**	**0.68**	**0.83**	**0.79**	**0.97**
	2	6	**0.23**	**0.28**	**0.40**	**0.09**	**0.30**
	2	12	**0.60**	0.02	**0.28**	**0.18**	**0.52**
	2	24	**0.61**	**0.17**	**0.29**	**0.43**	**0.97**
	3	6	**0.13**	**0.25**	**0.93**	**0.32**	**0.75**
	3	12	**0.30**	**0.45**	**0.97**	**0.46**	**0.86**
	3	24	**0.21**	**0.69**	**0.95**	**0.44**	**0.97**
	4	6	**0.39**	**0.40**	**0.54**	0.04	**0.47**
	4	12	**0.70**	**0.29**	**0.70**	**0.27**	**0.57**
	4	24	**0.22**	**0.59**	**0.48**	**0.24**	**0.97**
SPAIN	1	6	0.01	0.01	**0.47**	**0.94**	**0.81**
	1	12	0.01	**0.14**	**0.85**	**0.98**	**0.99**
	1	24	0.00	**0.49**	**0.96**	**0.97**	**1.00**
	2	6	**0.75**	**0.51**	**0.56**	**0.26**	**0.33**
	2	12	**0.12**	**0.53**	**0.66**	**0.58**	**0.69**
	2	24	**0.14**	**0.46**	**0.69**	**0.46**	**1.00**
	3	6	**0.19**	**0.08**	**0.83**	**0.30**	**0.79**
	3	12	**0.31**	**0.31**	**0.62**	**0.71**	**0.73**
	3	24	0.01	**0.48**	**0.88**	**0.56**	**1.00**
	4	6	**0.91**	**0.54**	**0.97**	**0.36**	**0.77**
	4	12	**0.82**	**0.71**	**0.95**	**0.74**	**0.96**
	4	24	**0.69**	**0.64**	**0.97**	**0.77**	**1.00**
UK	1	6	**0.50**	**0.05**	**0.67**	**0.35**	**0.47**
	1	12	**0.87**	**0.21**	**0.78**	**0.45**	**0.45**
	1	24	**0.92**	**0.21**	**0.98**	**0.14**	**0.64**
	2	6	0.00	**0.28**	**0.54**	**0.99**	**0.27**
	2	12	0.00	**0.18**	**0.83**	**0.83**	**0.71**
	2	24	0.00	**0.11**	**0.56**	**0.71**	**0.99**
	3	6	**0.40**	**0.08**	**0.94**	**0.87**	**0.22**
	3	12	**0.65**	**0.36**	**0.89**	**0.97**	**0.52**
	3	24	**0.73**	**0.60**	**0.94**	**0.26**	**0.98**
	4	6	0.01	**0.90**	**0.80**	**0.99**	**0.58**

	4	12	0.00	**0.94**	**0.98**	**0.98**	**0.93**
	4	24	0.00	**0.94**	**0.75**	**0.93**	**1.00**
SWEDEN	1	6	0.03	**0.64**	**0.12**	**0.14**	**0.34**
	1	12	**0.09**	**0.95**	**0.10**	**0.44**	**0.36**
	1	24	**0.21**	**0.76**	**0.42**	**0.61**	**0.72**
	2	6	0.00	**0.29**	**0.53**	**0.26**	**0.10**
	2	12	0.00	**0.18**	**0.58**	**0.51**	**0.33**
	2	24	0.00	**0.22**	**0.75**	**0.73**	**0.63**
	3	6	**0.56**	**0.83**	**0.89**	**0.46**	**0.90**
	3	12	0.03	**0.85**	**0.98**	**0.98**	**0.05**
	3	24	**0.07**	**0.43**	**1.00**	**0.93**	**0.37**
	4	6	0.02	**0.97**	**0.79**	**0.70**	0.03
	4	12	0.00	**0.87**	**0.95**	**0.82**	**0.40**
	4	24	**0.07**	**0.70**	**0.97**	**0.99**	**0.31**
FRANCE	1	6	**0.09**	**0.70**	**0.87**	**0.49**	**0.78**
	1	12	**0.37**	**0.89**	**0.74**	**0.71**	**0.52**
	1	24	**0.61**	**0.95**	**0.99**	**0.56**	**0.19**
	2	6	0.01	**0.75**	**0.76**	**0.82**	**0.07**
	2	12	0.00	**0.67**	**0.13**	**0.78**	**0.35**
	2	24	0.00	**0.29**	**0.23**	**0.93**	**0.28**
	3	6	**0.22**	**1.00**	**0.99**	**0.68**	**0.85**
	3	12	**0.38**	**0.99**	**0.70**	**0.97**	**0.75**
	3	24	**0.36**	**0.74**	**0.97**	**0.99**	**0.16**
	4	6	**0.20**	**0.98**	**0.90**	**0.98**	**0.43**
	4	12	0.00	**0.97**	**0.78**	**1.00**	**1.00**
	4	24	0.01	**0.63**	**0.97**	**1.00**	**0.17**
GERMANY	1	6	**0.05**	**0.57**	**0.12**	**0.69**	**0.49**
	1	12	**0.24**	**0.77**	**0.08**	**0.50**	**0.99**
	1	24	**0.37**	**0.86**	**0.64**	**0.77**	**0.50**
	2	6	0.01	**0.82**	**0.51**	**0.81**	**0.09**
	2	12	0.04	**0.77**	**0.11**	**0.93**	**0.95**
	2	24	0.00	**0.46**	**0.36**	**0.73**	**0.70**
	3	6	**0.12**	**0.58**	**0.15**	**0.69**	0.01
	3	12	**0.11**	**0.92**	0.00	**0.89**	**0.72**
	3	24	0.02	**0.63**	0.01	**0.47**	**0.67**
	4	6	**0.31**	**1.00**	**0.08**	**0.93**	0.03

	4	12	**0.55**	**0.99**	0.00	**0.95**	**0.99**
	4	24	**0.05**	**0.22**	0.00	**0.95**	**0.64**
SWITZERLAND	1	6	**0.71**	**0.06**	**0.62**	**0.49**	**0.72**
	1	12	**0.60**	**0.25**	**0.69**	**0.58**	**0.89**
	1	24	**0.50**	**0.67**	**0.58**	**0.87**	**1.00**
	2	6	**0.14**	**0.78**	**0.54**	**0.09**	**0.37**
	2	12	**0.42**	**0.47**	**0.76**	**0.50**	**0.64**
	2	24	**0.59**	**0.73**	**0.40**	**0.47**	**1.00**
	3	6	**0.87**	**0.17**	**0.98**	**0.32**	**0.95**
	3	12	**0.76**	**0.34**	**1.00**	**0.48**	**0.50**
	3	24	**0.69**	**0.53**	**0.81**	**0.96**	**1.00**
	4	6	**0.41**	**0.80**	**0.80**	**0.12**	**0.88**
	4	12	**0.73**	**0.47**	**0.99**	**0.63**	**0.66**
	4	24	**0.50**	**0.60**	**0.67**	**0.76**	**1.00**
NETHERLANDS	1	6	0.02	**0.95**	**0.72**	**0.93**	**0.35**
	1	12	**0.19**	**0.84**	**0.79**	**0.81**	**0.56**
	1	24	**0.37**	**0.82**	**0.96**	**0.97**	**0.57**
	2	6	0.00	**0.57**	**0.53**	**0.86**	**0.20**
	2	12	0.02	**0.65**	**0.86**	**0.95**	**0.63**
	2	24	0.00	**0.49**	**0.94**	**0.81**	**0.98**
	3	6	**0.22**	**0.99**	**0.90**	**0.95**	**0.19**
	3	12	**0.44**	**1.00**	**0.89**	**0.96**	**0.48**
	3	24	**0.59**	**0.93**	**0.83**	**0.99**	**0.52**
	4	6	**0.41**	**0.77**	**0.50**	**0.91**	**0.15**
	4	12	**0.86**	**0.76**	**0.85**	**0.95**	**0.91**
	4	24	**0.07**	0.01	**0.87**	**1.00**	**0.95**

Notes: This table shows the p-values for the Lagrange Multiplier (LM) test of no autocorrelation, for all European indices, for all horizons for the historical density forecasts using the GJR-T model. The statistically significant results at the 5% level are reported in bold.

Tables 6.12 - 6.17 report the p-values for the Hit tests. For some sections of the distribution there are no observations or not enough observations to perform the test so we mark them as "-". For RND as expected from the statistics from the whole distribution, the forecasts are misspecified and do not pass the test for the whole distribution or parts of it in the vast majority of the cases. However, we see again an improvement as the horizon increases. It is hard to say if the fit improves at a certain part of the distribution as the results are quite mix for each dataset and for each horizon. Moving to the RWD the forecasts it seems to have a marginal improvement especially for biweekly and monthly

forecasts. For example UK did not pass the test for the whole distribution or parts of it for 1-day and 1-week ahead at any significance level but did 4 times for RWD forecasts for (0.1-0.3) and (0.7-0.9).

Tables 6.14 - 6.17 show the results for the historical simulation forecasts. For GARCH-N there are on average 4 rejections, for GARCH-T there are 5, for GJR-N there are 3 and for GJR-T there are 4 at 5% significance level. So it seems like GJR-N is the best model.

Table 6.12 – Hit tests-RND

ITALY	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	0.00	0.00	0.00	0.00	0.00
(0.1,0.3)	0.00	0.00	0.01	0.02	0.60
(0.3, 0.7)	0.00	0.00	0.00	0.00	0.04
(0.7,0.9)	0.00	0.00	**0.07**	**0.07**	-
(0.9,1)	0.00	0.00	0.00	0.01	0.02
SPAIN	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	0.00	0.00	-	**1.00**	-
(0.1,0.3)	0.00	0.00	0.00	**0.38**	0.19
(0.3, 0.7)	0.00	0.00	0.00	0.00	0.00
(0.7,0.9)	0.00	0.00	**0.08**	0.03	**0.35**
(0.9,1)	0.00	0.00	**1.00**	**1.00**	-
UK	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	0.00	0.00	0.00	0.02	**0.14**
(0.1,0.3)	0.00	0.00	**0.14**	0.03	**0.56**
(0.3, 0.7)	0.00	0.00	0.00	0.00	**0.26**
(0.7,0.9)	0.00	0.00	0.03	0.04	**0.37**
(0.9,1)	0.00	0.00	-	0.01	**0.07**
SWEDEN	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	0.00	0.00	0.00	**0.05**	**0.50**
(0.1,0.3)	0.00	0.00	0.00	0.01	**0.31**
(0.3, 0.7)	0.00	0.00	0.00	0.00	**0.13**
(0.7,0.9)	0.00	0.00	0.01	0.02	**0.27**
(0.9,1)	0.00	0.00	0.00	**0.31**	**0.57**
FRANCE	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	0.00	0.00	0.00	0.03	-
(0.1,0.3)	0.00	0.00	0.00	**0.15**	**0.48**
(0.3, 0.7)	0.00	0.00	0.00	0.00	**0.99**
(0.7,0.9)	0.00	0.00	0.03	0.01	**0.56**
(0.9,1)	0.00	0.00	0.00	0.00	**0.06**
GERMANY	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	0.00	0.00	**0.05**	0.01	**1.00**
(0.1,0.3)	0.00	**0.05**	**0.07**	**0.43**	**0.14**
(0.3, 0.7)	0.00	0.00	0.00	0.00	0.00
(0.7,0.9)	0.00	**0.11**	0.00	**0.21**	**0.22**

(0.9,1)	0.00	0.00	0.00	0.01	-
SWITZERLAND	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	0.00	**1.00**	**1.00**	0.04	0.04
(0.1,0.3)	0.00	0.00	0.00	0.00	-
(0.3, 0.7)	0.00	0.00	0.00	0.00	0.00
(0.7,0.9)	0.00	0.00	0.00	**0.29**	**0.78**
(0.9,1)	1.00	0.00	0.00	**0.07**	**1.00**
NETHERLANDS	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	0.00	0.00	0.00	0.03	-
(0.1,0.3)	0.00	0.00	0.01	**0.25**	0.04
(0.3, 0.7)	0.00	0.00	0.00	0.00	**0.25**
(0.7,0.9)	0.00	0.00	0.00	0.06	**0.60**
(0.9,1)	0.00	0.00	0.00	**0.34**	-

Notes: This table shows the p-values for the Hit tests of Patton (2006), for all European indices, for all horizons, for the risk-neutral density (RND) forecasts. The statistically significant results at the 5% level are reported in bold.

Table 6.13 – Hit tests-RWD

ITALY	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	0.00	0.00	0.00	0.00	0.00
(0.1,0.3)	0.00	0.00	0.01	0.02	0.60
(0.3, 0.7)	0.00	0.00	0.00	0.00	0.04
(0.7,0.9)	0.00	0.00	**0.07**	**0.07**	-
(0.9,1)	0.00	0.00	0.00	0.01	0.02
SPAIN	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	0.00	0.00	**0.10**	**0.60**	**0.65**
(0.1,0.3)	**0.11**	**0.39**	**0.58**	**0.47**	0.01
(0.3, 0.7)	0.00	0.00	**0.83**	**0.27**	**0.13**
(0.7,0.9)	**0.05**	**0.55**	**0.63**	**0.43**	**0.21**
(0.9,1)	0.00	0.00	**0.09**	**0.24**	-
UK	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	0.00	0.00	**0.43**	0.00	**0.90**
(0.1,0.3)	0.02	**0.57**	**0.70**	**0.33**	**0.85**
(0.3, 0.7)	0.00	0.00	0.01	**0.32**	**0.45**
(0.7,0.9)	**0.23**	**0.08**	**0.88**	**0.16**	**0.28**
(0.9,1)	0.00	0.00	**0.44**	0.01	**0.19**
SWEDEN	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	0.00	0.00	0.00	**0.13**	-
(0.1,0.3)	0.00	**0.08**	**0.10**	**0.49**	**0.76**
(0.3, 0.7)	0.00	0.00	0.00	**0.09**	**0.84**
(0.7,0.9)	0.00	**0.10**	**0.34**	**0.64**	**0.49**
(0.9,1)	0.00	0.00	0.00	0.02	**0.57**
FRANCE	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	0.00	0.00	0.02	0.03	-
(0.1,0.3)	0.00	**0.99**	**0.24**	**0.42**	**0.30**
(0.3, 0.7)	0.00	0.00	0.01	0.00	**0.08**
(0.7,0.9)	0.00	0.01	**0.37**	0.04	**0.49**
(0.9,1)	0.00	0.00	0.00	**0.42**	-
GERMANY	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	0.00	0.00	**0.12**	**0.37**	**1.00**
(0.1,0.3)	**0.30**	**0.68**	**0.38**	**0.95**	**0.91**
(0.3, 0.7)	0.00	0.00	0.00	**0.22**	**0.25**
(0.7,0.9)	**0.37**	**0.51**	**0.08**	**0.45**	**0.17**

(0.9,1)	0.00	0.01	**0.08**	**0.07**	**1.00**
SWITZERLAND	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	0.00	0.00	**0.16**	**0.26**	0.03
(0.1,0.3)	**0.24**	**0.77**	**0.24**	**0.35**	**0.28**
(0.3, 0.7)	0.00	0.00	0.02	**0.31**	**0.93**
(0.7,0.9)	0.01	**0.81**	**0.58**	**0.88**	**0.29**
(0.9,1)	0.00	0.01	**0.19**	**0.73**	-
NETHERLANDS	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	0.00	0.00	0.00	0.01	-
(0.1,0.3)	0.00	**0.59**	**0.15**	**0.11**	**0.11**
(0.3, 0.7)	0.00	0.00	0.01	0.00	**0.10**
(0.7,0.9)	0.00	0.00	**0.20**	**0.46**	**0.69**
(0.9,1)	0.00	0.00	0.00	**0.07**	-

Notes: This table shows the p-values for the Hit tests of Patton (2006), for all European indices, for all horizons, for the real world density (RWD) forecasts. The statistically significant results at the 5% level are reported in bold.

Table 6.14 – Hit tests-Historical Simulation-GARCH-N

ITALY	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	**0.38**	**0.79**	**0.40**	**0.78**	0.01
(0.1,0.3)	0.01	**0.55**	**0.99**	**0.92**	-
(0.3, 0.7)	0.00	**0.53**	**0.93**	**0.84**	**0.18**
(0.7,0.9)	0.00	**0.70**	**0.53**	**0.93**	**0.61**
(0.9,1)	0.01	**0.89**	**0.40**	-	**1.00**
SPAIN	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	0.02	**0.29**	**0.28**	-	-
(0.1,0.3)	**0.08**	**0.50**	**0.11**	**0.49**	-
(0.3, 0.7)	0.00	**0.80**	**0.27**	**0.66**	0.04
(0.7,0.9)	**0.51**	**0.50**	**0.73**	**0.70**	-
(0.9,1)	0.00	**0.22**	**0.66**	**1.00**	**1.00**
UK	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	**0.36**	**0.81**	0.02	**0.24**	**0.24**
(0.1,0.3)	**0.60**	**0.88**	**0.65**	**0.17**	**0.98**
(0.3, 0.7)	0.00	**0.25**	**0.65**	**0.35**	**0.55**
(0.7,0.9)	**0.23**	**0.13**	**0.82**	**0.06**	**0.09**
(0.9,1)	0.00	**0.59**	**0.20**	**0.24**	**1.00**
SWEDEN	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	0.01	**0.23**	**0.20**	**0.67**	**0.21**
(0.1,0.3)	**0.80**	**0.88**	**0.22**	**0.56**	**0.46**
(0.3, 0.7)	0.00	**0.26**	0.03	**0.11**	**0.83**
(0.7,0.9)	**0.71**	**0.33**	**0.82**	**0.29**	-
(0.9,1)	0.00	0.02	0.00	0.01	**1.00**
FRANCE	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	**0.05**	**0.64**	**0.54**	**0.29**	**0.48**
(0.1,0.3)	**0.31**	**0.56**	**0.34**	0.04	**0.67**
(0.3, 0.7)	0.00	**0.34**	**0.89**	**0.44**	**0.38**
(0.7,0.9)	0.03	**0.18**	**0.42**	**0.79**	**0.33**
(0.9,1)	0.00	**0.46**	**0.05**	**1.00**	-
GERMANY	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	0.02	**0.15**	**0.55**	**0.34**	**0.74**
(0.1,0.3)	**0.90**	**0.86**	**0.65**	**0.86**	**0.91**
(0.3, 0.7)	0.00	**0.93**	**0.55**	**0.36**	**0.24**
(0.7,0.9)	**0.18**	**0.51**	**0.13**	**0.48**	**0.21**

(0.9,1)	0.00	**0.17**	-	-	**1.00**
SWITZERLAND	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	**0.05**	**1.00**	**0.75**	-	**0.21**
(0.1,0.3)	**0.33**	**0.86**	**0.24**	**0.29**	0.03
(0.3, 0.7)	0.00	**0.20**	**0.36**	**0.35**	**0.13**
(0.7,0.9)	**0.27**	**0.91**	**0.48**	**0.74**	**0.35**
(0.9,1)	0.00	**0.28**	**0.37**	-	-
NETHERLANDS	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	0.00	**0.99**	**0.33**	**0.29**	**0.09**
(0.1,0.3)	0.04	**0.61**	**0.47**	**0.19**	**0.97**
(0.3, 0.7)	0.00	**1.00**	**0.49**	**0.24**	**0.82**
(0.7,0.9)	**0.36**	**0.73**	**0.22**	**0.55**	0.02
(0.9,1)	0.00	**0.82**	**0.06**	**0.26**	-

Notes: This table shows the p-values for the Hit tests of Patton (2006), for all European indices, for all horizons, for the historical density forecasts using the GARCH-N model. The statistically significant results at the 5% level are reported in bold.

Table 6.15 – Hit tests-Historical Simulation-GARCH-T

ITALY	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	0.00	**0.14**	**0.32**	**0.49**	0.01
(0.1,0.3)	0.01	**0.57**	**0.46**	**0.32**	**0.28**
(0.3, 0.7)	0.00	**0.45**	**0.47**	**0.21**	**0.21**
(0.7,0.9)	0.03	**0.86**	**0.65**	**0.92**	**0.79**
(0.9,1)	**0.26**	**0.72**	**0.42**	-	**1.00**
SPAIN	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	0.00	**0.29**	**0.22**	**0.50**	-
(0.1,0.3)	**0.87**	**0.68**	**0.09**	**0.18**	0.03
(0.3, 0.7)	0.00	**0.82**	**0.27**	**0.15**	**0.61**
(0.7,0.9)	**0.43**	**0.84**	**0.89**	**0.87**	-
(0.9,1)	0.00	**0.19**	**0.46**	-	**1.00**
UK	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	0.00	**0.92**	0.00	**0.24**	**0.37**
(0.1,0.3)	**0.40**	**0.98**	**0.58**	**0.35**	**0.84**
(0.3, 0.7)	0.00	**0.28**	**0.88**	**0.44**	**0.73**
(0.7,0.9)	**0.62**	**0.30**	**0.68**	0.01	**0.17**
(0.9,1)	**0.07**	**0.52**	**0.42**	**0.07**	**1.00**
SWEDEN	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	0.00	0.04	**0.20**	**0.67**	**0.21**
(0.1,0.3)	**0.24**	**0.74**	**0.05**	**0.56**	**0.34**
(0.3, 0.7)	0.00	**0.77**	0.02	**0.07**	**0.92**
(0.7,0.9)	**0.30**	**0.08**	**0.73**	**0.49**	-
(0.9,1)	0.00	0.03	0.00	0.00	**1.00**
FRANCE	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	0.00	**0.15**	0.02	**0.46**	**0.74**
(0.1,0.3)	**0.08**	**0.56**	**0.84**	0.03	**0.58**
(0.3, 0.7)	0.00	**0.19**	**0.87**	**0.91**	**0.98**
(0.7,0.9)	**0.78**	**0.21**	**0.39**	**0.79**	**0.30**
(0.9,1)	0.00	**0.20**	0.04	**1.00**	-
GERMANY	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	0.00	**0.14**	**0.45**	**0.35**	**0.22**
(0.1,0.3)	0.04	**0.95**	**0.91**	**0.63**	**0.20**
(0.3, 0.7)	0.00	**0.78**	**0.32**	**0.57**	**0.46**
(0.7,0.9)	**0.87**	**0.51**	**0.22**	**0.60**	**0.16**

(0.9,1)	0.01	**0.22**	-	**1.00**	**1.00**
SWITZERLAND	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	0.00	**0.78**	**0.87**	**0.18**	**0.70**
(0.1,0.3)	0.02	**0.39**	**0.19**	0.03	**0.05**
(0.3, 0.7)	0.00	**0.30**	**0.31**	**0.56**	**0.13**
(0.7,0.9)	**0.65**	**0.96**	**0.40**	**0.28**	**0.35**
(0.9,1)	0.00	**0.66**	**0.33**	-	-
NETHERLANDS	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	0.00	**0.61**	**0.13**	**0.09**	**0.06**
(0.1,0.3)	**0.05**	**0.68**	**0.83**	**0.25**	**0.82**
(0.3, 0.7)	0.00	**0.49**	**0.57**	**0.09**	**1.00**
(0.7,0.9)	**0.73**	**0.85**	**0.59**	**0.60**	**0.70**
(0.9,1)	0.00	**0.91**	0.00	-	**1.00**

Notes: This table shows the p-values for the Hit tests of Patton (2006), for all European indices, for all horizons, for the historical density forecasts using the GARCH-T model. The statistically significant results at the 5% level are reported in bold.

Table 6.16 – Hit tests-Historical Simulation-GJR-N

ITALY	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	**0.42**	**0.34**	**0.33**	**0.59**	0.02
(0.1,0.3)	0.00	**0.40**	**0.53**	**0.59**	**0.56**
(0.3, 0.7)	0.00	**0.39**	**0.79**	**0.73**	**0.47**
(0.7,0.9)	**0.41**	**0.31**	**0.09**	**0.76**	**0.87**
(0.9,1)	**0.37**	**0.60**	**0.48**	**0.71**	**0.72**
SPAIN	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	0.00	**0.50**	**0.29**	-	-
(0.1,0.3)	**0.08**	**0.34**	0.01	**0.20**	-
(0.3, 0.7)	0.00	**0.98**	**0.55**	**0.79**	0.03
(0.7,0.9)	**0.09**	**0.12**	**0.45**	**0.79**	**0.05**
(0.9,1)	**0.09**	**0.31**	**0.52**	-	**1.00**
UK	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	**0.18**	**0.42**	0.00	**0.73**	**0.17**
(0.1,0.3)	**0.64**	**0.96**	**0.65**	**0.18**	**0.27**
(0.3, 0.7)	**0.05**	**0.14**	**0.73**	**0.85**	**0.10**
(0.7,0.9)	**0.36**	**0.20**	**0.21**	**0.07**	**0.27**
(0.9,1)	**0.46**	**0.96**	**0.71**	**0.30**	-
SWEDEN	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	0.00	**0.50**	**0.39**	**0.92**	**0.56**
(0.1,0.3)	**0.20**	**0.99**	**0.57**	**0.82**	**0.72**
(0.3, 0.7)	**0.05**	**0.73**	0.02	**0.13**	**0.69**
(0.7,0.9)	**0.32**	**0.11**	**0.88**	**0.37**	**0.65**
(0.9,1)	**0.08**	**0.42**	**0.19**	**0.48**	**0.36**
FRANCE	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	0.01	**0.96**	**0.40**	**0.54**	**0.25**
(0.1,0.3)	**0.44**	**0.35**	**0.88**	**0.83**	**0.89**
(0.3, 0.7)	0.00	**0.38**	**0.98**	**0.69**	**0.80**
(0.7,0.9)	0.04	**0.49**	**0.24**	**0.71**	**0.06**
(0.9,1)	0.00	**0.54**	**0.42**	-	**0.61**
GERMANY	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	0.02	**0.24**	**0.82**	**0.51**	-
(0.1,0.3)	**0.89**	**0.61**	**0.68**	**0.86**	**0.39**
(0.3, 0.7)	**0.06**	**0.46**	**0.21**	**0.62**	**0.50**
(0.7,0.9)	**0.23**	**0.08**	**0.19**	**0.80**	**0.31**

(0.9,1)	0.00	**0.13**	**0.09**	-	-
SWITZERLAND	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	0.00	**0.97**	**0.92**	-	**1.00**
(0.1,0.3)	**0.48**	**0.85**	**0.46**	**0.43**	**0.05**
(0.3, 0.7)	0.00	**0.12**	**0.73**	**0.89**	**0.26**
(0.7,0.9)	**0.14**	**0.56**	**0.19**	**0.97**	**0.91**
(0.9,1)	**0.05**	**0.50**	**0.43**	**0.93**	-
NETHERLANDS	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	0.00	**0.61**	**0.13**	**0.09**	**0.06**
(0.1,0.3)	**0.05**	**0.68**	**0.83**	**0.25**	**0.82**
(0.3, 0.7)	0.00	**0.49**	**0.57**	**0.09**	**1.00**
(0.7,0.9)	**0.73**	**0.85**	**0.59**	**0.60**	**0.70**
(0.9,1)	0.00	**0.91**	0.00	-	**1.00**

Notes: This table shows the p-values for the Hit tests of Patton (2006), for all European indices, for all horizons, for the historical density forecasts using the GJR-N model. The statistically significant results at the 5% level are reported in bold.

Table 6.17 – Hit tests-Historical Simulation-GJR-T

ITALY	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	0.00	**0.18**	**0.19**	**0.99**	0.01
(0.1,0.3)	**0.04**	**0.79**	**0.53**	**0.62**	**0.28**
(0.3, 0.7)	0.00	**0.45**	**0.52**	**0.41**	**0.21**
(0.7,0.9)	0.03	**0.61**	**0.23**	**0.55**	**0.54**
(0.9,1)	0.00	**0.68**	**0.40**	**0.71**	**0.89**
SPAIN	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	0.00	**0.27**	**0.16**	**0.96**	-
(0.1,0.3)	**0.11**	**0.81**	**0.07**	**0.62**	**0.46**
(0.3, 0.7)	0.01	**0.90**	**0.36**	**0.70**	**0.18**
(0.7,0.9)	**0.47**	**0.62**	**0.67**	**0.54**	0.03
(0.9,1)	0.00	**0.19**	**0.52**	**1.00**	**1.00**
UK	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	0.01	**0.87**	0.02	**0.55**	**0.54**
(0.1,0.3)	**0.77**	**0.96**	**0.85**	**0.19**	**0.94**
(0.3, 0.7)	0.00	**0.38**	**0.93**	**0.49**	**0.13**
(0.7,0.9)	**0.09**	**0.06**	**0.40**	**0.20**	**0.08**
(0.9,1)	0.00	**0.54**	**0.60**	**0.10**	-
SWEDEN	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	0.00	**0.17**	**0.40**	**0.86**	**0.21**
(0.1,0.3)	**0.64**	**0.95**	0.03	**0.59**	**0.68**
(0.3, 0.7)	0.00	**0.69**	0.03	**0.44**	**0.80**
(0.7,0.9)	0.01	0.01	**0.65**	**0.05**	-
(0.9,1)	0.00	**0.30**	**0.19**	**0.29**	**1.00**
FRANCE	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	0.00	**0.39**	**0.12**	**0.81**	**0.65**
(0.1,0.3)	0.03	**0.45**	**0.88**	**0.28**	**0.56**
(0.3, 0.7)	0.00	**0.07**	**0.65**	**0.79**	**0.92**
(0.7,0.9)	**0.54**	**0.39**	**0.50**	**0.96**	-
(0.9,1)	0.00	**0.66**	**0.40**	**1.00**	**0.26**
GERMANY	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	0.00	**0.18**	**0.80**	**0.76**	**0.28**
(0.1,0.3)	**0.37**	**1.00**	**0.68**	**0.98**	**0.08**
(0.3, 0.7)	0.01	**0.08**	**0.35**	**0.87**	**0.08**
(0.7,0.9)	**0.05**	0.02	**0.72**	**0.95**	**0.14**

(0.9,1)	0.00	0.04	**0.13**	-	-
SWITZERLAND	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	0.00	**0.93**	**0.92**	**0.16**	**1.00**
(0.1,0.3)	0.01	**0.65**	**0.49**	**0.22**	**0.05**
(0.3, 0.7)	0.00	**0.29**	**0.36**	**0.75**	**0.26**
(0.7,0.9)	**0.79**	**0.69**	**0.45**	**0.88**	**0.92**
(0.9,1)	0.00	**0.34**	**0.28**	**0.93**	**1.00**
NETHERLANDS	DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
(0,1)	0.00	**0.74**	**0.11**	**0.33**	**0.07**
(0.1,0.3)	**0.47**	**0.20**	**0.54**	**0.40**	**0.46**
(0.3, 0.7)	0.00	**0.71**	0.04	**0.15**	**0.61**
(0.7,0.9)	**0.09**	**0.42**	**0.08**	**0.41**	**0.64**
(0.9,1)	0.00	**0.71**	**0.06**	**0.28**	-

Notes: This table shows the p-values for the Hit tests of Patton (2006), for all European indices, for all horizons, for the historical density forecasts using the GJR-T model. The statistically significant results at the 5% level are reported in bold.

Finally, we aim to observe directly the mean and the variance of the transformed series in a more straightforward way of confirming the fit of the forecasts. Table 6.18 reports the statistics. In brief it was confirmed what we have observed from the tests. When we move from RND to RWD there is some improvement both in terms of mean and standard deviation for each horizon. But that is not the case for all datasets. For example there is an improvement for Italy for 1 and 2 months ahead moving from (0.37, 2.40) and (0.62, 3.19) to (-0.01, 1.18) and (-0.03, 0.88) respectively. However, for most datasets the mean of RND forecasts is close to zero and the standard deviation close to 1 for longer horizons. In those cases the non-parametric calibration gives slightly worse results as it is the case for Netherlands for 1-month ahead where we go from (0, 0.95) to (0.02, 1.19).

Moving to the historical simulations, a better fit can be seen mostly in terms of standard deviations where now the values are closer to 1. It can be observed that for GARCH-N, GARCH-T, GJR-N and GJR-T we have cases where the standard deviation is equal to 1 meaning that the fit is better. Surprisingly, the perfect fit for the mean (mean=0) can be located only under RND, RWD and GARCH-N. Even though a better fit is observed in terms of standard deviations, in historical densities forecasts, there are isolated cases where the fit for a horizon might get slightly worse. For example, for Netherlands using the GARCH-N specification the statistics are: (-0.03, 0.88), (-0.05, 0.99), (-0.13, 0.99), (-.020, 1.20) and (-0.26, 1.02) whereas for RWD are: (0, 1.02), (-0.05, 0.89), (0.02, 1.20), (0.02, 0.19) and (-0.04, 0.89) for 1, 7, 14, 28 and 56 days ahead respectively. Also, it is observed that for almost all GARCH models the values of the mean increases as the horizon increases. But that is

hardly the case for GJR-N which seems like it gives the best fit.

Table 6.18 – Statistics

RND		DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
ITALY	mean	0.02	0.03	0.04	0.37	0.62
	st.dev	0.53	0.78	1.36	2.40	3.19
SPAIN	mean	0.00	0.03	0.00	0.01	0.02
	st.dev	0.80	0.51	0.56	0.58	0.61
UK	mean	0.01	0.05	-0.01	0.00	0.02
	st.dev	0.51	1.19	0.65	0.78	0.82
SWEDEN	mean	0.03	0.03	0.02	0.06	0.05
	st.dev	0.50	1.12	0.91	1.11	0.90
FRANCE	mean	0.03	0.05	0.04	0.07	0.14
	st.dev	0.53	0.94	0.69	0.83	0.99
GERMANY	mean	0.02	0.04	0.05	0.03	0.04
	st.dev	0.43	0.86	0.91	0.73	0.76
SWITZERLAND	mean	0.01	0.05	0.09	0.15	-0.06
	st.dev	0.48	0.37	0.74	1.03	1.41
NETHERLANDS	mean	0.02	-0.05	0.02	0.00	0.05
	st.dev	0.83	1.34	0.86	0.95	0.98
RWD		DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
ITALY	mean	-0.02	-0.03	-0.03	-0.01	-0.03
	st.dev	1.09	1.23	1.12	1.18	0.88
SPAIN	mean	-0.01	0.00	0.00	0.01	-0.05
	st.dev	1.08	1.22	0.98	0.86	0.96
UK	mean	0.01	-0.03	0.01	0.05	-0.06
	st.dev	1.14	0.87	0.99	1.16	0.86
SWEDEN	mean	0.02	0.00	0.04	0.09	0.02
	st.dev	1.28	0.87	1.07	1.19	0.84
FRANCE	mean	0.00	-0.03	0.03	0.06	0.01
	st.dev	1.24	0.94	1.11	1.11	0.88
GERMANY	mean	0.02	0.01	0.01	0.00	0.02
	st.dev	1.02	0.89	0.89	0.82	0.76
SWITZERLAND	mean	-0.02	-0.07	-0.08	-0.08	-0.13
	st.dev	0.89	1.18	1.10	1.05	1.01
NETHERLANDS	median	0.00	-0.05	0.02	0.02	-0.04
	st.dev	1.02	0.89	1.20	1.19	0.89
HS (GARCH -N)		DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY

ITALY	mean	-0.02	-0.09	-0.11	-0.17	-0.18
	st.dev	0.95	1.03	1.01	1.02	1.06
SPAIN	mean	0.02	-0.01	-0.10	-0.16	-0.16
	st.dev	0.91	0.96	1.08	0.81	0.83
UK	mean	-0.04	-0.07	-0.27	-0.23	-0.30
	st.dev	0.95	1.02	1.41	1.00	1.00
SWEDEN	mean	-0.05	-0.12	-0.17	-0.24	-0.33
	st.dev	0.89	0.97	0.99	0.89	0.89
FRANCE	mean	-0.03	-0.06	-0.11	-0.17	-0.26
	st.dev	0.91	0.97	0.98	0.98	1.03
GERMANY	mean	-0.04	-0.06	-0.16	-0.19	-0.23
	st.dev	0.94	0.96	0.94	1.00	0.88
SWITZERLAND	mean	-0.02	-0.05	-0.10	-0.14	-0.19
	st.dev	0.86	0.96	0.97	0.87	0.80
NETHERLANDS	mean	-0.03	-0.05	-0.13	-0.20	-0.26
	st.dev	0.88	0.99	0.99	1.02	1.02
HS (GARCH -T)		DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
ITALY	mean	0.03	-0.13	-0.17	-0.25	-0.30
	st.dev	0.84	1.03	1.01	1.02	1.07
SPAIN	mean	0.03	-0.07	-0.17	-0.28	-0.33
	st.dev	0.81	0.97	1.06	0.82	0.82
UK	mean	0.05	-0.11	-0.29	-0.29	-0.39
	st.dev	0.85	1.02	1.36	0.99	0.99
SWEDEN	mean	0.05	-0.14	-0.20	-0.28	-0.40
	st.dev	0.81	0.99	1.00	0.91	0.92
FRANCE	mean	0.05	-0.12	-0.20	-0.29	-0.43
	st.dev	0.82	0.99	1.01	0.98	1.03
GERMANY	mean	0.04	-0.08	-0.21	-0.25	-0.33
	st.dev	0.85	0.99	0.98	1.01	0.92
SWITZERLAND	mean	0.02	-0.06	-0.11	-0.16	-0.21
	st.dev	0.78	0.97	0.97	0.88	0.83
NETHERLANDS	mean	0.04	-0.07	-0.17	-0.23	-0.34
	st.dev	0.77	1.01	1.01	1.01	1.03
HS (GJR-N)		DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
ITALY	mean	0.05	-0.03	-0.02	-0.04	0.00
	st.dev	0.95	1.03	1.02	1.01	1.08

SPAIN	mean	0.08	0.06	0.01	-0.03	0.04
	st.dev	0.90	0.95	1.04	0.83	0.86
UK	mean	0.04	0.00	-0.20	-0.10	-0.11
	st.dev	0.96	1.01	1.34	1.00	1.04
SWEDEN	mean	0.04	-0.05	-0.06	-0.11	-0.15
	st.dev	0.89	0.95	0.98	0.87	0.92
FRANCE	mean	0.03	-0.01	-0.04	-0.07	-0.13
	st.dev	0.91	0.96	0.98	0.96	1.02
GERMANY	mean	0.04	0.02	-0.07	-0.07	-0.06
	st.dev	0.94	0.96	0.94	0.98	0.91
SWITZERLAND	mean	0.06	0.02	0.01	0.01	0.02
	st.dev	0.87	0.96	0.97	0.88	0.82
NETHERLANDS	mean	0.04	0.02	-0.04	-0.07	-0.08
	st.dev	0.89	0.98	0.97	0.98	1.01
HS (GJR-T)		DAILY	WEEKLY	BIWEEKLY	MONTHLY	BIMONTHLY
ITALY	mean	-0.02	-0.09	-0.10	-0.16	-0.17
	st.dev	0.86	1.04	1.03	1.01	1.09
SPAIN	mean	-0.03	-0.01	-0.09	-0.18	-0.17
	st.dev	0.83	0.97	1.04	0.84	0.84
UK	mean	-0.01	-0.06	-0.24	-0.21	-0.26
	st.dev	0.88	1.01	1.30	0.99	1.02
SWEDEN	mean	-0.02	-0.10	-0.13	-0.20	-0.29
	st.dev	0.83	0.98	1.01	0.91	0.95
FRANCE	mean	-0.01	-0.09	-0.15	-0.23	-0.35
	st.dev	0.84	1.00	1.02	0.98	1.05
GERMANY	mean	-0.02	-0.04	-0.15	-0.18	-0.23
	st.dev	0.87	0.99	0.99	1.00	0.97
SWITZERLAND	mean	-0.04	-0.01	-0.03	-0.05	-0.06
	st.dev	0.80	0.97	0.97	0.90	0.84
NETHERLANDS	mean	-0.02	-0.03	-0.11	-0.16	-0.24
	st.dev	0.78	1.00	0.99	0.99	1.03

Notes: This table shows the values for the mean and the standard deviation of the transformed PIT series using the the CDF of the standard normal distribution, for all European indices, for all horizons, for the risk-neutral density (RND) forecasts, the real world density (RWD) forecasts and the historical density forecasts using the GARCH-N, GARCH-T, GJR-N and GJR-T models.

6.5 Conclusion

We used data from options written on eight European indices to estimate and forecast densities for 1-day, 1-week, 2-week, 1-month and 2-month ahead. The Heston model was used to get RND forecasts which were then transformed to RWD forecasts through a non-parametric calibration. Also to avoid any issues existing with the estimation of implied densities historical simulations were performed to get historical densities forecasts using the GARCH-N, the GARCH-T, the GJR-N and the GJR-T models.

Our results suggest that RND forecast are misspecified and for that reason should not be preferred. Non-parametric calibration to RWD is necessary but does not give the best fit. From our analysis historical densities seems to give the best fit. This result is contrary to Shackleton et al. (2010) findings. Further, Høg and Tsiaras (2011), Fan et al. (2018) and Crisóstomo and Couso (2018) even though they all focused on different assets, forecast horizons, models and evaluation methods they all somewhat compared implied densities to historical ones and they also found that the former outperforms the latter. One interpretation could be that the lower liquidity in European indices makes implied densities relatively less informative compared to the US ones. Now, comparing between GARCH and GJR models, that were used for the historical densities forecasts, it seems like that GJR was the best according to the GoF tests. What is more, changing the distribution for the GARCH and GJR models does not give significant changes and the signal as to which distribution is best are mixed. Cautiously it could be argued that GJR-N is best for normality and GJR-T for independence. Therefore, we can conclude that Historical models are superior to option-implied densities for forecasts.

Chapter 7

Conclusion

The main aim in this thesis was to evaluate various approaches to forecast volatility. In Chapter 2 the methodological framework was presented. Chapter 3 focuses on univariate volatility models. A plethora of non-parametric, parametric models and forecast combinations were considered. We used various forecast evaluation tools, to evaluate the models' forecasting ability and compare them in the out of sample framework. These tests were the GW, the MCS and the SPA tests. An empirical application was performed using the S&P 500 index returns. We found that there is not a single specification that outperforms the rest but there were a few that performed better than the rest and that is the TGARCH, the log HAR and the mean forecast combination. To robust our results a sensitivity analysis was conducted that was directed to the choice of the loss function, the sample size and the stock index. Further research extend this study by considering different forecast combinations schemes, multivariate volatility models and different approaches on forecast evaluation while focusing on the economic significance for investors and managers as well.

In Chapter 4 we addressed the directional accuracy of the forecasts derived by GARCH models and their combinations. In more details we aimed at establishing whether more sophisticated models are better than more parsimonious ones and evaluating a variety of combinations methods using the DA and the DV measures. To the best of our knowledge this is the first study that uses the single models and sophisticated combinations and evaluates them in terms of both DA and DV. Further, some technical trading rules were applied to measure profitability since DA is more correlated with profitability than the magnitude of error. The analysis was carried out using S&P 500 data. We observed that EGARCH model performed better than the rest in terms of single models and the procedures by Greer (2005) and Costantini and Pappalardo (2010) were the best among the forecast combinations both in terms of DA and DV. The results by the test of Eberhardt and Fligner (1977) gave evidence that combinations can alleviate uncertainty over the choice of a single model. Further, our investment strategies pointed to superiority of the hierarchical encompassing-HMSE combination method. Finally, sensitivity analysis was performed to robust our results. A future research step

could be to develop forecast combination methods based on economic value loss functions, expecting these combination schemes to lead to superior economic gains for an investor.

Chapter 5 focuses on multivariate volatility models. In this chapter a risk parity portfolio was constructed using hedge funds in an attempt to outperform the HFR Risk Parity Indices. One portfolio was constructed based on alpha measure and the second one based on Sharpe ratio. Our empirical results demonstrated that at least one of the constructed portfolios outperformed the HFR Indices for the both volatility targets, 10% and 12%. Following analysis should investigate more thoroughly which method is best for constructing FoHFs and also repeat the analysis for other volatility targets.

The final, empirical, chapter (Chapter 6) explores density forecasts. In more details given the lack of agreement in the literature as to which of these two is generally superior to the other, this study examines under what circumstances, risk transformations of risk-neutral densities can outperform historical densities. In more details, we considered a non-parametric calibration to transform the risk-neutral implied densities to real-world densities. To evaluate our forecasts Goodness-of-Fit tests were used. To the best of our knowledge this is the first study to perform the empirical application on options written on European stock indices. We found that historical densities outperform RNDs and RWDs. Future research should also include other methods for calibration and forecast combinations. What is more, it should also evaluate the models and the methods in terms of pairwise comparisons as well.

Bibliography

Abadir, K. M., Distaso, W., and Žikeš, F. (2014). Design-free estimation of variance matrices. *Journal of Econometrics*, 181(2):165–180.

Agarwal, V. and Naik, N. Y. (2004). Risks and portfolio decisions involving hedge funds. *The Review of Financial Studies*, 17(1):63–98.

Aguilar, O. and West, M. (2000). Bayesian dynamic factor models and portfolio allocation. *Journal of Business & Economic Statistics*, 18(3):338–357.

Aielli, G. P. (2013). Dynamic conditional correlation: On properties and estimation. *Journal of Business & Economic Statistics*, 31(3):282–299.

Aielli, G. P. and Caporin, M. (2015). Dynamic principal components: A new class of multivariate garch models. Working paper 0193. Available from: `https://economia.unipd.it/sites/economia.unipd.it/files/20150194.pdf`.

Aiken, A. L., Clifford, C. P., and Ellis, J. (2014). The value of funds of hedge funds: Evidence from their holdings. *Management Science*, 61(10):2415–2429.

Aiolfi, M., Capistran, C., and Timmermann, A. (2011). Forecast combinations. In Clements, M. and Hendry, D., editors, *The Oxford Handbook of Economic Forecasting*, chapter 11, pages 355–388.

Aıt-Sahalia, Y., Wang, Y., and Yared, F. (2001). Do option markets correctly price the probabilities of movement of the underlying asset? *Journal of Econometrics*, 102(1):67–110.

Alexander, C. (2001). Orthogonal GARCH. *Mastering risk*, pages 21–38.

Alexander, C. and Chibumba, A. (1997). Orthogonal garch: an empirical validation in equities, foreign exchange and interest rates. *School of Mathematical Sciences Discussion Paper, Sussex University.*

Alizadeh, S., Brandt, M. W., and Diebold, F. X. (2002). Range-based estimation of stochastic volatility models. *Journal of Finance*, 57(3):1047–1091.

Alonso, N. and Qian, E. (2013). Risk parity equity strategy with flexible risk targets. *The Journal of Investing*, 22(3):99–106.

Amado, C. and Teräsvirta, T. (2014). Conditional correlation models of autoregressive conditional heteroscedasticity with nonstationary garch equations. *Journal of Business & Economic Statistics*, 32(1):69–87.

Amin, G. S. and Kat, H. M. (2003). Hedge fund performance 1990-2000: Do the "money machines" really add value? *Journal of financial and quantitative analysis*, 23(2):251–274.

Amisano, G. and Giacomini, R. (2007). Comparing density forecasts via weighted likelihood ratio tests. *Journal of Business & Economic Statistics*, 25(2):177–190.

Ammann, M., Huber, O., and Schmid, M. (2013). Hedge fund characteristics and performance persistence. *European Financial Management*, 19(2):209–250.

Anagnou, I., Bedendo, M., Hodges, S., and Tompkins, R. (2005). Forecasting accuracy of implied and garch-based probability density functions. *Review of Futures Markets*, 11(1):41–66.

Andersen, T. G.and Davis, R., Kreiß, J. ., and Mikosch, T. V. (2009). *Handbook of financial time series*. Springer Science & Business Media.

Andersen, T. and Benzoni, L. (2009). *Stochastic volatility*. Working paper series 2009-04, Federal Reserve Bank of Chicago. Available from: `https://www.chicagofed.org/publications/working-papers/2009/wp-04`.

Andersen, T., Bollerslev, T., Christoffersen, P., and Diebold, F. (2005a). Volatility forecasting. NBER Working Papers 11188, National Bureau of Economic Research, Inc. Available from: `https://EconPapers.repec.org/RePEc:nbr:nberwo:11188`.

Andersen, T., Bollerslev, T., Christoffersen, P., and Diebold, F. (2006). Volatility and correlation forecasting. volume 1, chapter 15, pages 777–878. Elsevier, 1 edition.

Andersen, T., Bondarenko, O., and Gonzalez-Perez, M. (2015). Exploring return dynamics via corridor implied volatility. *Review of Financial Studies*, 28(10):2902–2945.

Andersen, T. G., Bollerslev, T., and Diebold, F. X. (2007). Roughing it up: Including jump components in the measurement, modeling and forecasting of return volatility. *Review of Economics and Statistics*, 89(4):701–720.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Ebens, H. (2001a). The distribution of realized stock return volatility. *Journal of financial economics*, 61(1):43–76.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2001b). The distribution of exchange rate volatility. *Journal of the American Statistical Association*, 96(1):42–55.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, 71(2):579–625.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Wu, J. (2005b). A framework for exploring the macroeconomic determinants of systematic risk. *American Economic Review*, 95(2):398–404.

Andersen, T. G., Bollerslev, T., and Meddahi, N. (2011). Realized volatility forecasting and market microstructure noise. *Journal of Econometrics*, 160(1):220–234.

Andersen, T. G., Bondarenko, O., and Gonzalez-Perez, M. T. (2010). *Coherent model-free implied volatility: A corridor fix for high-frequency VIX*. Creates research papers 2011-49, Department of Economics and Business Economics, Aarhus University. Available from: `ftp://ftp.econ.au.dk/creates/rp/11/rp1149.pdf`.

Anderson, H. M., Nam, K., and Vahid, F. (1999). Asymmetric nonlinear smooth transition GARCH models. In *Nonlinear time series analysis of economic and financial data*, pages 191–207. Springer.

Anderson, R. M., Bianchi, S. W., and Goldberg, L. R. (2012). Will my risk parity strategy outperform? *Financial Analysts Journal*, 68(6):75–93.

Aragon, G. O. (2007). Share restrictions and asset pricing: Evidence from the hedge fund industry. *Journal of Financial Economics*, 83(1):33–58.

Armstrong, J. S. (1984). Forecasting by extrapolation: Conclusions from 25 years of research. *Interfaces*, 14(6):52–66.

Armstrong, J. S. (1989). Combining forecasts: The end of the beginning or the beginning of the end? *International Journal of Forecasting*, 5(4):585–588.

Armstrong, J. S. (2001). *Principles of forecasting: A handbook for researchers and practitioners*, volume 30. Springer Science & Business Media.

Asai, M. and McAleer, M. (2004). *Dynamic leverage and threshold effects in stochastic volatility models*. Unpublished manuscript, Faculty of Economics, Tokyo Metropolitan University. Available from: `https://doi.org/10.1080/07474930500243035`.

Asai, M. and McAleer, M. (2005). Dynamic asymmetric leverage in stochastic volatility models. *Econometric Reviews*, 24(3):317–332.

Asai, M. and McAleer, M. (2007). Non-trading day effects in asymmetric conditional and stochastic volatility models. *The Econometrics Journal*, 10(1):113–123.

Asai, M. and McAleer, M. (2015). Forecasting co-volatilities via factor models with asymmetry and long memory in realized covariance. *Journal of Econometrics*, 189(2):251–262.

Asai, M., McAleer, M., and Medeiros, M. C. (2012). Modelling and forecasting noisy realized volatility. *Computational Statistics & Data Analysis*, 56(1):217–230.

Asness, C. S., Frazzini, A., and Pedersen, L. H. (2012a). Leverage aversion and risk parity. *Financial Analysts Journal*, 68(1):47–59.

Asness, C. S., Frazzini, A., and Pedersen, L. H. (2012b). Leverage aversion and risk parity. *Financial Analysts Journal*, 68(1):47–59.

Asness, C. S., Moskowitz, T. J., and Pedersen, L. H. (2013). Value and momentum everywhere. *The Journal of Finance*, 68(3):929–985.

Astatkie, T., Watts, D. G., and Watt, W. E. (1997). Nested threshold autoregressive (neTAR) models. *International Journal of Forecasting*, 13(1).

Attanasio, O. P. (1991). Risk, time-varying second moments and market efficiency. *The Review of Economic Studies*, 58(3):479–494.

Audrino, F. (2014). Forecasting correlations during the late-2000s financial crisis: The short-run component, the long-run component, and structural breaks. *Computational Statistics & Data Analysis*, 76:43–60.

Audrino, F. and Bühlmann, P. (2001). Tree-structured generalized autoregressive conditional heteroscedastic models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(4):727–744.

Awartani, B. and Corradi, V. (2005). Predicting the volatility of the S&P-500 stock index via GARCH models: The role of asymmetries. *International Journal of forecasting*, 21:167–183.

Bacchini, F., Ciammola, A., Iannaccone, R., and Marini, M. (2010). Combining forecasts for a flash estimate of euro area gdp. *Contributi Istat*, 3.

Back, A. D. and Weigend, A. S. (1997). A first application of independent component analysis to extracting structure from stock returns. *International journal of neural systems*, 8(04):473–484.

Baillie, R. T., Bollerslev, T., and Mikkelsen, H. O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 74(1):3–30.

Bali, T., Atilgan, Y., and Demirtas, O. (2013). *Investing in hedge funds: A guide to measuring risk and return characteristics*. Academic Press, 1 edition.

Banz, R. W. and Miller, M. H. (1978). Prices for state-contingent claims: Some estimates and applications. *Journal of Business*, 51(4):653–672.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and Shephard, N. (2011). Multivariate realised kernels: consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading. *Journal of Econometrics*, 162(2):149–169.

Barndorff-Nielsen, O. E. and Shephard, N. (2004). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics*, 2(1):1–48.

Barndorff-Nielsen, O. and Shephard, N. (2002a). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(2):253–280.

Barndorff-Nielsen, O. and Shephard, N. (2002b). Estimating quadratic variation using realized variance. *Journal of Applied econometrics*, 17(5):457–477.

Bates, D. S. (1996). Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options. *The Review of Financial Studies*, 9(1):69–107.

Bates, J. and Granger, C. (1969). The combination of forecasts. *Journal of the Operational Research*, 20:451–468.

Bauwens, L., Hafner, C. M., and Laurent, S. (2012a). *Handbook of volatility models and their applications*, volume 3. John Wiley & Sons.

Bauwens, L. and Otranto, E. (2020). Nonlinearities and regimes in conditional correlations with different dynamics. *Journal of Econometrics*, 217(2):496–522.

Bauwens, L., Storti, G., Violante, F., et al. (2012b). Dynamic conditional correlation models for realized covariance matrices. Core discussion paper 2012/60. Available from: `https://dial.uclouvain.be/pr/boreal/object/boreal:122203`.

Bawa, V. S. (1975). Optimal rules for ordering uncertain prospects. *Journal of Financial Economics*, 2(1):95–121.

Becker, R. and Clements, A. (2008). Are combination forecasts of S&P 500 volatility statistically superior? *International Journal of Forecasting*, 24(1).

Bera, A. K., Higgins, M. L., and Lee, S. (1992). Interaction between autocorrelation and conditional heteroskedasticity: A random-coefficient approach. *Journal of Business and Economic Statistics*, 10(2):133–142.

Berkowitz, J. (2001). Testing density forecasts, with applications to risk management. *Journal of Business and Economic Statistics*, 19(5):465–474.

Bessler, W., Opfer, H., and Wolff, D. (2017). Multi-asset portfolio optimization and out-of-sample performance: An evaluation of black–litterman, mean-variance, and naïve diversification approaches. *The European Journal of Finance*, 23(1):1–30.

Bessler, W., Taushanov, G., and Wolff, D. (2021). Factor investing and asset allocation strategies: A comparison of factor versus sector optimization. *Journal of Asset Management*, 22(6):488–506.

Bessler, W. and Wolff, D. (2015). Do commodities add value in multi-asset portfolios? An out-of-sample analysis for different investment strategies. *Journal of Banking & Finance*, 60:1–20.

Bhansali, V., Davis, J., Rennison, G., Hsu, J., and Li, F. (2012). The risk in risk parity: A factor-based analysis of asset-based risk parity. *The Journal of Investing*, 21(3):102–110.

Bhansali, V. and Harris, L. (2018). Everybody's doing it: Short volatility strategies and shadow financial insurers. *Financial Analysts Journal*, 74(2):12–23.

Bhardwaj, G., Gorton, G. B., and Rouwenhorst, K. G. (2014). Fooling some of the people all of the time: The inefficient performance and persistence of commodity trading advisors. *The Review of Financial Studies*, 27(11):3099–3132.

Billio, M. and Caporin, M. (2009). A generalized dynamic conditional correlation model for portfolio risk evaluation. *Mathematics and Computers in Simulation*, 79(8):2566–2578.

Billio, M., Caporin, M., and Gobbo, M. (2006). Flexible dynamic conditional correlation multivariate garch models for asset allocation. *Applied Financial Economics Letters*, 2(02):123–130.

Black, F. (1976). *Studies of stock price volatility changes*. Proceedings of the 1976 meeting of business and economic statistics section, American statistical association, 177-81.

Black, F., Jensen, M. C., Scholes, M., et al. (1972). The capital asset pricing model: Some empirical tests. *Studies in the theory of capital markets*, 81(3):79–121.

Black, F. and Litterman, R. (1990). Asset allocation: Combining investor views with market equilibrium. Discussion paper, Goldman, Sachs & Co.

Black, F. and Scholes, M. (1973). The valuation of options and corporate liabilities. *Journal of political economy*, 81(3):637–654.

Blair, B., Poon, S. H., and Taylor, S. J. (2001a). The pricing of options and corporate liabilities. *Journal of Political Economy*, 81:637–654.

Blair, B. J., Poon, S.-H., and Taylor, S. J. (2001b). Forecasting S&P 100 volatility: The incremental information content of implied volatilities and high-frequency index returns. *Journal of Econometrics*, 105(1):5–26.

Blaskowitz, O. and Herwartz, H. (2009). Adaptive forecasting of the EURIBOR swap term structure. *Journal of Forecasting*, 28:575–594.

Blaskowitz, O. and Herwartz, H. (2011). On economic evaluation of directional forecasts. *International of Journal of Forecasting*, 27:1058–1065.

Blaskowitz, O. and Herwartz, H. (2014). Testing the value of directional forecasts in the presence of serial correlation. *International Journal of Forecasting*, 30:30–42.

Bliss, R. R. and Panigirtzoglou, N. (2004). Option-implied risk aversion estimates. *The journal of finance*, 59(1):407–446.

Boigner, P. and Gadzinski, G. (2015). Diversification with risk factors and investable hedge fund indices. *Journal of Asset Management*, 16(2):101–116.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307–327.

Bollerslev, T. (1990). Modelling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH model. *The review of economics and statistics*, 72(3):498–505.

Bollerslev, T. (2009). Glossary to ARCH (GARCH). Technical report, CREATES Research Papers 2008-49, School of Economics and Management, University of Aarhus. Available from: `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1263250`.

Bollerslev, T., Engle, R. F., and Wooldridge, J. M. (1988). A capital asset pricing model with time-varying covariances. *Journal of political Economy*, 96(1):116–131.

Bollerslev, T. and Ghysels, E. (1996). Periodic autoregressive conditional heteroskedasticity. *Journal of Business and Economic Statistics*, 14(2):139–151.

Bollerslev, T. and Mikkelsen, H. O. (1996). Modeling and pricing long memory in stock market volatility. *Journal of econometrics*, 73(1):151–184.

Bollerslev, T., Patton, A. J., and Quaedvlieg, R. (2016). Exploiting the errors: A simple approach for improved volatility forecasting. *Journal of Econometrics*, 192(1):1–18.

Boudt, K., Galanos, A., Payseur, S., and Zivot, E. (2019). Multivariate GARCH models for large-scale applications: A survey. In Maddala, G. and Rao, C., editors, *Handbook of statistics*, volume 41, chapter 7, pages 193–242. Elsevier.

Boyson, N. M. (2008). Hedge fund performance persistence: A new approach. *Financial Analysts Journal*, 64(6):27–44.

Breeden, D. T. and Litzenberger, R. H. (1978). Prices of state-contingent claims implicit in option prices. *Journal of business*, 51(4):621–651.

Breidt, F. J., Crato, N., and de Lima, P. (1998). The detection and estimation of long memory in stochastic volatility. *Journal of econometrics*, 83(1-2):325–348.

Briner, B. G. and Connor, G. (2008). How much structure is best? A comparison of market model, factor model and unstructured equity covariance matrices. *Journal of Risk*, 10(4):3–30.

Brito, D., Medeiros, M. C., and Ribeiro, R. (2018). Forecasting large realized covariance matrices: The benefits of factor models and shrinkage. Working paper. Available from: `https://ssrn.com/abstract=3163668`.

Brooks, C. and Persand, G. (2003). Volatility forecasting for risk management. *Journal of forecasting*, 22(1):1–22.

Broto, C. and Ruiz, E. (2004). Estimation methods for stochastic volatility models: A survey. *Journal of Economic Surveys*, 18(5):613–649.

Broto, C. and Ruiz, E. (2006). Unobserved component models with asymmetric conditional variances. *Computational Statistics and Data Analysis*, 50(9):2146–2166.

Brown, R. (2012). *Analysis of investments & management of portfolios*. South Western College Publishing, 10 edition.

Brownlees, C. T. and Gallo, G. M. (2009). Comparison of volatility measures: a risk management perspective. *Journal of Financial Econometrics*, 8(1):29–56.

Bruder, B. and Roncalli, T. (2012). *Managing risk exposures using the risk budgeting approach*. Working paper. Available from: `https://ssrn.com/abstract=2009778`.

Cai, J. (1994). A Markov model of switching-regime ARCH. *Journal of Business & Economic Statistics*, 12(3):309–316.

Callot, L. A., Kock, A. B., and Medeiros, M. C. (2017). Modeling and forecasting large realized covariance matrices and portfolio choice. *Journal of Applied Econometrics*, 32(1):140–158.

Calvet, L. and Fisher, A. (2001). Forecasting multifractal volatility. *Journal of econometrics*, 105(1):27–58.

Calvet, L. E., Fisher, A. J., and Thompson, S. B. (2006). Volatility comovement: a multifrequency approach. *Journal of econometrics*, 131(1-2):179–215.

Campa, J. M., Chang, P. K., and Reider, R. L. (1998). Implied exchange rate distributions: Evidence from OTC option markets. *Journal of international Money and Finance*, 17(1):117–160.

Campbell, J. Y. and Viceira, L. (2001). Who should buy long-term bonds? *American Economic Review*, 91(1):99–127.

Capocci, D., Corhay, A., and Hübner, G. (2005). Hedge fund performance and persistence in bull and bear markets. *The European Journal of Finance*, 11(5):361–392.

Capocci, D. P. (2007). The sustainability of hedge fund performance: New insights. *Journal of Financial Management*, 1(2):59–86.

Caporin, M. and McAleer, M. (2006). Dynamic asymmetric garch. *Journal of Financial Econometrics*, 4:385–412.

Caporin, M. and McAleer, M. (2012a). Do we really need both BEKK and DCC? A tale of two multivariate GARCH models. *Journal of Economic Surveys*, 26(4):736–751.

Caporin, M. and McAleer, M. (2012b). Model selection and testing of conditional and stochastic volatility models. In L. Bauwens, C. H. and Laurent, S., editors, *Handbook of Volatility models and their applications*, chapter 8, pages 199–222.

Cappiello, L., Engle, R., and Sheppard, K. (2006). Asymmetric dynamics in the correlations of global equity and bond returns. *Journal of Financial econometrics*, 4:537–572.

Carhart, M., Cheah, U.-W., De Santis, G., Farrell, H., and Litterman, R. (2014). Exotic beta revisited. *Financial Analysts Journal*, 70(5):24–52.

Catania, L. and Proietti, T. (2020). Forecasting volatility with time-varying leverage and volatility of volatility effects. *International Journal of Forecasting*, 36:1301–1317.

Chan, D., Kohn, R., and Kirby, C. (2006). Multivariate stochastic volatility models with correlated errors. *Econometric Reviews*, 25(2-3):245–274.

Chan, L. K., Karceski, J., and Lakonishok, J. (1999). On portfolio optimization: Forecasting covariances and choosing the risk model. *The review of Financial studies*, 12(5):937–974.

Charles, A. and Darné, O. (2019). The accuracy of asymmetric GARCH model estimation. *International Economics*, 157:179–202.

Chaves, D., Hsu, J., Li, F., and Shakernia, O. (2011). Risk parity portfolio vs. other asset allocation heuristic portfolios. *The Journal of Investing*, 20(1):108–118.

Chen, Q., Costantini, M., and Deschamps, B. (2016). How accurate are professional forecasts in Asia? evidence from ten countries. *International Journal of Forecasting*, 32:154–167.

Chen, X., Ghysels, E., and Wang, F. (2015). Hybrid-GARCH: A generic class of models for volatility predictions using high frequency data. *Statistica Sinica*, 25(2):759–786.

Chen, Y.-C., Rogoff, K. S., and Rossi, B. (2010). Can exchange rates forecast commodity prices? *The Quarterly Journal of Economics*, 125(3):1145–1194.

Chernov, M., Gallant, A. R., Ghysels, E., and Tauchen, G. (2003). Alternative models for stock price dynamics. *Journal of Econometrics*, 116(1-2):225–257.

Chib, S., Nardari, F., and Shephard, N. (2006). Analysis of high dimensional multivariate stochastic volatility models. *Journal of Econometrics*, 134(2):341–371.

Chib, S., Omori, Y., and Asai, M. (2009). Multivariate stochastic volatility. In T. G. Andersen, R. A. Davis, J.-P. K. and Mikosch, T., editors, *Handbook of Financial Time Series*, pages 365–400. Springer.

Chiriac, R. and Voev, V. (2011). Modelling and forecasting multivariate realized volatility. *Journal of Applied Econometrics*, 26:922–947.

Chortareas, G., Jiang, Y., John, Y., and Nankervis, C. (2011). Forecasting exchange rate volatility using high-frequency data: Is the EURO different? *International Journal of Forecasting*, 27(4).

Chou, R. Y., Chou, H., and Liu, N. (2010). Range volatility models and their applications in finance. In *Handbook of quantitative finance and risk management*, pages 1273–1281. Springer.

Christensen, B. J. and Prabhala, N. R. (1998). The relation between implied and realized volatility. *Journal of financial economics*, 50(2):125–150.

Christensen, K., Oomen, R., and Podolskij, M. (2010). Realised quantile-based estimation of the integrated variance. *Journal of Econometrics*, 159(1):74–98.

Christodoulakis, G. A. and Satchell, S. (2000). *Evolving Systems of Financial Returns: Auto-Regressive Conditional Beta*. Working paper 02. Available from: `https://warwick.ac.uk/fac/soc/wbs/subjects/finance/research/wpaperseries/2000/00-97.pdf`.

Christoffersen, P. F. (1998). Evaluating interval forecasts. *International economic review*, 39(4):841–862.

Christoffersen, P. F. and Diebold, F. X. (2000). How relevant is volatility forecasting for financial risk management? *Review of Economics and Statistics*, 82(1):12–22.

Chung, H., Huang, C.-S., and Tseng, T.-C. (2008). Modeling and forecasting of realized volatility based on high-frequency data: Evidence from Taiwan. *International Research Journal of Finance and Economics*, 22(22):178–191.

Clark, T. E. and McCracken, M. (2001). Tests of equal forecast accuracy and encompassing for nested models. *Journal of Econometrics*, 105(1):85–110.

Clark, T. E. and McCracken, M. W. (2009). Improving forecast accuracy by combining recursive and rolling forecasts. *International Economic Review*, 50(2):363–395.

Clark, T. E. and West, K. D. (2007). Approximately normal tests for equal predictive accuracy in nested models. *Journal of econometrics*, 138(1):291–311.

Clemen, R. (1989). Combining forecasts: A review and annotated bibliography. *International Journal of Forecasting*, 5(4):559–583.

Clements, M. and Hendry, D. (1998). *Forecasting economic time series*. Cambridge University Press.

Clements, M. P. and Hendry, D. F. (2008). Economic forecasting in a changing world. *Capitalism and Society*, 3(2):1–20.

Colacito, R., Engle, R. F., and Ghysels, E. (2011). A component model for dynamic correlations. *Journal of Econometrics*, 164(1):45–59.

Conrad, C. and Kleen, O. (2020). Two are better than one: Volatility forecasting using multiplicative component garch-midas models. *Journal of Applied Econometrics*, 35:19–45.

Conrad, C. and Loch, K. (2015). Anticipating long-term stock market volatility. *Journal of Applied Econometrics*, 30(7):1090–1114.

Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, 7:174–196.

Corsi, F., Audrino, F., and Renó, R. (2012). HAR modeling for realized volatility forecasting. In Bauwens, L., Hafner, C., and Laurent, S., editors, *Handbook of Volatility Models and Their Applications*, chapter 15, pages 363–382. John Wiley & Sons, Inc.

Corsi, F., Mittnik, S., Pigorsch, C., and Pigorsch, U. (2008). The volatility of realized volatility. *Econometric Reviews*, 27:46–78.

Corsi, F., Pirino, D., and Renó, R. (2010). Threshold bipower variation and the impact of jumps on volatility forecasting. *Journal of Econometrics*, 159(2):276–288.

Corsi, F. and Reno, R. (2009). *HAR volatility modelling with heterogeneous leverage and jumps*. Working paper,ssrn 1316953. Available from: `https://web.stanford.edu/group/SITE/archive/SITE_2009/segment_1/s1_papers/corsi.pdf`.

Costa, G. and Kwon, R. (2020). A robust framework for risk parity portfolios. *Journal of Asset Management*, 21(5):447–466.

Costantini, M., Cuaresma, J., and Hlouskova, J. (2016a). Forecasting errors, directional accuracy and profitability of currency trading: The case of EUR/USD exchange rate. *Journal of Forecasting*, 35:652–668.

Costantini, M., Cuaresma, J., and Hlouskova, J. (2016b). Forecasting errors, directional accuracy and profitability of currency trading: The case of eur/usd exchange rate. *Journal of Forecasting*, 35:652–668.

Costantini, M., Gunter, U., and Kunst, R. (2017). Forecast combinations in a DSGE-VAR lab. *Journal of Forecasting*, 36(1):305–324.

Costantini, M. and Pappalardo, C. (2010). A hierarchical procedure for the combination of forecasts. *International Journal of Forecasting*, 26:725–743.

Crisóstomo, R. and Couso, L. (2018). Financial density forecasts: A comprehensive comparison of risk-neutral and historical schemes. *Journal of Forecasting*, 37(5):589–603.

Cueto, J. M., Grané, A., and Cascos, I. (2020). Models for expected returns with statistical factors. *Journal of Risk and Financial Management*, 13(12):314.

Cumby, R. E. and Modest, D. M. (1987). Testing for market timing ability: A framework for forecast evaluation. *Journal of Financial Economics*, 19(1):169–189.

Dacorogna, M., Müller, U., Pictet, O., and Olsen, R. (1998). *Modelling short-term volatility with GARCH and HARCH models*. Working paper. Available from: `https://ssrn.com/abstract=36960`.

Danielsson, J. (1998). Multivariate stochastic volatility models: Estimation and a comparison with VGARCH models. *Journal of Empirical Finance*, 5(2):155–173.

Darolles, S., Francq, C., and Laurent, S. (2018). Asymptotics of cholesky garch models and time-varying conditional betas. *Journal of Econometrics*, 204(2):223–247.

Davidson, J. (2004). Moment and memory properties of linear conditional heteroskedasticity models, and a new model. *Journal of Business and Economic Statistics*, 22:16–29.

Davidson, L., Cooper, D., et al. (1976). A simple way of developing a probability distribution of present value. *Journal of Petroleum Technology*, 28(9):1069–1078.

Day, T. E. and Lewis, C. M. (1993). Forecasting futures market volatility. *The Journal of Derivatives*, 1(2):33–50.

de Almeida, D., Hotta, L. K., and Ruiz, E. (2018). MGARCH models: Trade-off between feasibility and flexibility. *International Journal of Forecasting*, 34(1):45–63.

de Vincent-Humphreys, R. and Noss, J. (2012). *Estimating probability distributions of future asset prices: Empirical transformations from option-implied risk-neutral to real-world density functions*. Working paper no. 455, Bank of England. Available from: `https://www.bankofengland.co.uk/-/media/boe/files/working-paper/2012/estimating-probability-distributions-of-future-asset-prices-emperical-transformations.pdf`.

DeMiguel, V., Garlappi, L., and Uppal, R. (2007). Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *The review of Financial studies*, 22(5):1915–1953.

den Hertog, R. G. J. (1994). Pricing of permanent and transitory volatility for U.S. stock returns: A composite GARCH model. *Economic Letters*, 44(4):421–426.

Dendramis, Y., Kapetanios, G., and Tzavalis, E. (2015). Shifts in volatility driven by large stock market shocks. *Journal of Economic Dynamics & Control*, 55:130–147.

Diebold, F. X. (2007). *Elements of forecasting*. Thomson.

Diebold, F. X. (2015). Comparing predictive accuracy, twenty years later: A personal perspective on the use and abuse of diebold–mariano tests. *Journal of Business & Economic Statistics*, 33:1–1.

Diebold, F. X., Gunther, T. A., and Tay, A. S. (1998a). Evaluating density forecasts, with applications to financial risk management. *International Economic Review*, 39(4):863–883.

Diebold, F. X., Hahn, J., and Tay, A. S. (1998b). Real-time multivariate density forecast evaluation and calibration: Monitoring the risk of high-frequency returns on foreign exchange. Nber working papers 6845, National Bureau of Economic Research, Inc. Available from: `https://ideas.repec.org/p/nbr/nberwo/6845.html`.

Diebold, F. X. and Lopez, J. (1996). Forecast evaluation and combination. In Maddala, G. and Rao, C., editors, *Handbook of statistics*, volume 14, chapter 8, pages 241–268. Elsevier.

Diebold, F. X. and Mariano, R. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 13:253–265.

Diebold, F. X. and Nerlove, M. (1989). The dynamics of exchange rate volatility: A multivariate latent factor ARCH model. *Journal of Applied econometrics*, 4(1):1–21.

Dieobold, F. X. (1986). Modelling the persistence of conditional variances: A comment. *Econometric Review*, 5(1):71–74.

Ding, Z. and Engle, R. F. (2001). Large scale conditional covariance matrix modeling, estimation and testing. *Academia Economic Papers*, 29:157–184.

Ding, Z., Granger, C. W. J., and Engle, R. F. (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, 1.

Dor, A. B., Dynkin, L., and Gould, A. (2006). Style analysis and classification of hedge funds. *The Journal of Alternative Investments*, 9(2):10–29.

Dunis, C., Kellard, N. M., and Snaith, S. (2013). Forecasting EUR-USD implied volatility: The case of intraday. *Journal of Banking and Finance*, 37:4943–4957.

Duong, D. and Swanson, N. R. (2015). Empirical evidence on the importance of aggregation, asymmetry, and jumps for volatility prediction. *Journal of Econometrics*, 187(2):606–621.

Durham, G. B. (2007). SV mixture models with application to S&P 500 index returns. *Journal of Financial Economics*, 85(3):822–856.

Eberhardt, K. and Fligner, M. (1977). A comparison of two tests for equality of two proportions. *The American Statistician*, 31(4):151–155.

Eling, M. and Schuhmacher, F. (2006). Performance measurement of hedge fund indices-Does the measure matter? In Haasis HD., Kopfer H., S. J., editor, *Operations Research Proceedings 2005*, volume 2005, pages 205–210. Springer.

Eling, M. and Schuhmacher, F. (2007). Does the choice of performance measure influence the evaluation of hedge funds? *Journal of Banking & Finance*, 31(9):2632–2647.

Elton, E. J. and Gruber, M. J. (1973). Estimating the dependence structure of share prices–implications for portfolio selection. *The Journal of Finance*, 28(5):1203–1232.

Engle, R. (2002a). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics*, 20:339–350.

Engle, R., Granger, C., and Kraft, D. (1984). Combining competing forecasts of inflation using a bivariate arch model. *Journal of economic dynamics and control*, 8:151–165.

Engle, R. and Patton, A. (2001). What good is a volatility model ? *Quantitative Finance*, 1(2):237–245.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50(4):987–1007.

Engle, R. F. (1990a). Discussion. *Review of Finacial Studies*, 3:103–106.

Engle, R. F. (1990b). Stock market volatility and the crash of '87: Discussion. *Review of Financial Studies*, 3(1):103–106.

Engle, R. F. (2002b). New frontiers for ARCH models. *Journal of Applied Econometrics*, 17(5):425–446.

Engle, R. F. (2016). Dynamic conditional beta. *Journal of Financial Econometrics*, 14(4):643–667.

Engle, R. F. and Gallo, G. M. (2006). A multiple indicators model for volatility using intra-daily data. *Journal of Econometrics*, 131(1-2):3–27.

Engle, R. F., Ghysels, E., and Sohn, B. (2008). On the economic sources of stock market volatility. In *AFA 2008 New Orleans Meetings Paper*.

Engle, R. F., Ghysels, E., and Sohn, B. (2013). Stock market volatility and macroeconomic fundamentals. *Review of Economics and Statistics*, 95(3):776–797.

Engle, R. F. and Gonzalez-Rivera, G. (1991). Semiparametric ARCH models. *Journal of Business and Economic Statistics*, 9(4):345–359.

Engle, R. F. and Kelly, B. (2012). Dynamic equicorrelation. *Journal of Business & Economic Statistics*, 30(2):212–228.

Engle, R. F. and Kroner, K. F. (1995). Multivariate simultaneous generalized ARCH. *Econometric theory*, 11(1):122–150.

Engle, R. F. and Lee, G. (1999a). A long-run and short-run component model of stock return volatility. In Engle, R. and White, H., editors, *Cointegration, Causality, and Forecasting: A Festschrift in Honour of Clive WJ Granger*, pages 475–497. Oxford University Press.

Engle, R. F. and Lee, G. G. J. (1999b). A permanent and transitory component model of stock return volatility. In Engle, R. and White, H., editors, *Cointegration, Causality, and Forecasting: A Festschrift in Honor of Clive*, pages 475–497. Oxford University Press.

Engle, R. F., Lilien, D. M., and Robbins, R. P. (1987). Estimating time varying risk premia in the term structure: The arch-m model. *Econometrica*, 55:391–407.

Engle, R. F. and Manganelli, S. (2004). Caviar: Conditional autoregressive value at risk by regression quantiles. *Journal of business & economic statistics*, 22(4):367–381.

Engle, R. F. and Marcucci, J. (2006). A long-run pure variance common features model for the common volatilities of the dow jones. *Journal of Econometrics*, 132(1):7–42.

Engle, R. F. and Mezrich, J. (1996). GARCH for groups: A round-up of recent developments in GARCH techniques for estimating correlation. *RISK*, 9(8):36–40.

Engle, R. F. and Ng, V. K. (1993). Measuring and testing the impact of news on volatility. *The journal of finance*, 48(5):1749–1778.

Engle, R. F., Ng, V. K., and Rothschild, M. (1990). Asset pricing with a factor-ARCH covariance structure: Empirical estimates for treasury bills.

Engle, R. F. and Susmel, R. (1993a). Common volatility in international equity markets. *Journal of Business & Economic Statistics*, 11(2):167–176.

Engle, R. F. and Susmel, R. (1993b). Common volatility in international equity markets. *Journal of Business & Economic Statistics*, 11(2):167–176.

Eugenie, M. H. (2003). *Empirical Studies on Volatility in International Stock Markets*. Springer, 1 edition.

Fama, E. F. and French, K. R. (1997). Industry costs of equity. *Journal of financial economics*, 43(2):153–193.

Fameliti, S. and Skintzi, V.D., . (2020). Predictive ability and economic gains from volatility forecast combinations. *Journal of Forecasting*, 39:200–219.

Fan, J., Li, Y., and Yu, K. (2012). Vast volatility matrix estimation using high-frequency data for portfolio selection. *Journal of the American Statistical Association*, 107(497):412–428.

Fan, J., Liao, Y., and Mincheva, M. (2011). High dimensional covariance matrix estimation in approximate factor models. *Annals of statistics*, 39(6):3320.

Fan, R., Taylor, S. J., and Sandri, M. (2018). Density forecast comparisons for stock prices, obtained from high-frequency returns and daily option prices. *Journal of Futures Markets*, 38(1):83–103.

Figlewski, S. (1997). Forecasting volatility. *Financial markets, institutions & instruments*, 6(1):1–88.

Fishburn, P. C. (1977). Mean-risk analysis with risk associated with below-target returns. *The American Economic Review*, 67(2):116–126.

Fleming, J., Kirby, C., and Ostdiek, B. (2003). The economic value of volatility timing using realized volatility. *Journal of Financial Economics*, 67:473–509.

Fornari, F. and Mele, A. (1997). Sign- and volatility-switching ARCH models: Theory and applications to international stock markets. *Journal of Applied Econometrics*, 12(1):49–65.

Forte, G. and Manera, M. (2006). *Forecasting volatility in Asian and European stock markets with asymmetric GARCH models*. Newfin working paper 3. Available from: `https://boa.unimib.it/retrieve/handle/10281/44657/67016/wp_03_06.pdf`.

Franses, P. H. and Van Dijk, D. (2000). *Non-linear time series models in empirical finance*. Cambridge University Press, 1 edition.

French, K. R., Schwert, G. W., and Stambaugh, R. F. (1987). Expected stock returns and volatility. *Journal of financial Economics*, 19(1):3–29.

Friedman, B. M., Laibson, D. I., and Minsky, H. P. (1989). Economic implications of extraordinary movements in stock prices. *Brookings Papers on Economic Activity*, 1989(2):137–189.

Frost, P. A. and Savarino, J. E. (1986). Portfolio size and estimation risk. *The Journal of Portfolio Management*, 12(4):60–64.

Fuertes, A., Izzeldin, M., and Kalotychou, E. (2009a). On forecasting daily stock volatility: The role of intraday information and market conditions. *International Journal of Forecasting*, 25(2):259–281.

Fuertes, A., Izzeldin, M., and Kalotychou, E. (2009b). On forecasting daily stock volatility: The role of intraday information and market conditions. *International Journal of Forecasting*, 25:259–281.

Fung, W. and Hsieh, D. A. (2000). Performance characteristics of hedge funds and commodity funds: Natural vs. spurious biases. *Journal of Financial and Quantitative analysis*, 35(3):291–307.

Fung, W. and Hsieh, D. A. (2002). Hedge-fund benchmarks: Information content and biases. *Financial Analysts Journal*, 58(1):22–34.

Fung, W. and Hsieh, D. A. (2004). Hedge fund benchmarks: A risk-based approach. *Financial Analysts Journal*, 60(5):65–80.

Fung, W., Hsieh, D. A., Naik, N. Y., and Ramadorai, T. (2008). Hedge funds: Performance, risk, and capital formation. *The Journal of Finance*, 63(4):1777–1803.

Gallant, A. R. (1981). On the bias in flexible functional forms and an essentially unbiased form: The Fourier flexible form. *Journal of Econometrics*, 15(2):211–244.

Gallant, A. R., Hsu, C.-T., and Tauchen, G. (1999). Using daily range data to calibrate volatility diffusions and extract the forward integrated variance. *Review of Economics and Statistics*, 81(4):617–631.

Garman, M. and Klass, M. (1980). On the estimation of security price volatilities from historical data. *Journal of Business*, 53(1):67–78.

Geweke, J. (1986). Modelling the persistence of conditional variances: A comment. *Econometric Review*, 5(1):57–61.

Ghalanos, A. (2018). *Package 'rugarch'*. econometric software. Available from: `http://www.unstarched.net/` [Accessed on 25 October 2018].

Ghysels, E. (2014). Matlab toolbox for mixed sampling frequency data analysis using midas regression models. *Available on MATLAB Central at http://www. mathworks. com/matlabcentral/fileexchange/45150-midas-regression*.

Ghysels, E., Harvey, A. C., and Renault, E. (1996). Stochastic volatility. In Maddala, G. and Rao, C., editors, *Handbook of statistics*, volume 14, chapter 5, pages 119–191. Elsevier.

Ghysels, E., Santa-Clara, P., and Valkanov, R. (2005). There is a risk-return trade-off after all. *Journal of Financial Economics*, 76(3):509–548.

Giacomini, R. and White, H. (2006). Tests of conditional predictive ability. *Econometrica*, 74(6):1545–1578.

Glosten, L. R., Jagannathan, R., and Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance*, 48(5):1779–801.

Gneiting, T. (2011). Making and evaluating point forecasts. *Journal of the American Statistical Association*, 106(494):746–762.

Goetzmann, W., Ingersoll, J., Spiegel, M., and Welch, I. (2007). Portfolio performance manipulation and manipulation-proof performance measures. *The Review of Financial Studies*, 20(5):1503–1546.

Golosnoy, V., Gribisch, B., and Liesenfeld, R. (2012). The conditional autoregressive wishart model for multivariate stock market volatility. *Journal of Econometrics*, 167(1):211–223.

Goltz, F. and Schröder, D. (2010). Hedge fund transparency: where do we stand? *The Journal of Alternative Investments*, 12(4):20–35.

Gonzalez-Rivera, G. (1998). Smooth transition GARCH models. *Studies in Nonlinear Dynamics and Econometrics*, 3(2):61–78.

González-Rivera, G., Lee, T.-H., and Mishra, S. (2004). Forecasting volatility: A reality check based on option pricing, utility function, value-at-risk, and predictive likelihood. *International Journal of forecasting*, 20(4):629–645.

Gouriéroux, C., Jasiak, J., and Sufana, R. (2009). The wishart autoregressive process of multivariate stochastic volatility. *Journal of Econometrics*, 150(2):167–181.

Gourieroux, C. and Monfort, A. (1992). Qualitative threshold ARCH models. *Journal of Econometrics*, 52(1-2):159–199.

Gradojevic, N. and Yang, J. (2006). Non-linear, non-parametric, non-fundamental exchange rate forecasting. *Journal of Forecasting*, 25(4):227–245.

Granger, C. W. and Jeon, Y. (2004). Thick modeling. *Economic Modelling*, 21(2):323–343.

Granger, C. W. and Ramanathan, R. (1984). Improved methods of combining forecasts. *Journal of Forecasting*, 3(2):197–204.

Grauer, R. R. and Hakansson, N. H. (1995). Stein and capm estimators of the means in asset allocation. *International Review of Financial Analysis*, 4(1):35–66.

Grauer, R. R. and Shen, F. C. (2000). Do constraints improve portfolio performance? *Journal of banking & finance*, 24(8):1253–1274.

Greco, T., Biondi-Zoccai, G., Gemma, M., Guérin, C., Zangrillo, A., and Landoni, G. (2015). How to impute study-specific standard deviations in meta-analyses of skewed continuous endpoints? *World Journal of Meta-Analysis*, 3(5):215–224.

Greer, M. (2005). Combination forecasting for directional accuracy: An application to survey interest rate forecasts. *Journal of applied statistics*, 32(6):607–615.

Guégan, D. and Diebolt, J. (1994). Probabilistic properties of the β-ARCH model. *Statistica Sinica*, 4(1):71–87.

Guidolin, A. and Pedio, M. (2021). Forecasting commodity futures returns with stepwise regressions: Do commodity.specific factors help. *Annals of Operations Research*, 299:1317–1356.

Guidolin, M. and Timmermann, A. (2005). Economic implications of bull and bear regimes in UK stock and bond returns. *The Economic Journal*, 115(500):111–143.

Gulay, E. and Emec, H. (2018). Comparison of forecasting performances: Does normalization and variance stabilization method beat GARCH(1,1)-type models? empirical evidence from the stock markets. *Journal of Forecasting*, 79:133–150.

Hafner, C. M. and Franses, P. H. (2009). A generalized dynamic conditional correlation model: Simulation and application to many assets. *Econometric Reviews*, 28(6):612–631.

Hafner, C. M. and Reznikova, O. (2012). On the estimation of dynamic conditional correlation models. *Computational Statistics & Data Analysis*, 56(11):3533–3545.

Hafner, C. M. and Rombouts, J. V. (2007). Semiparametric multivariate volatility models. *Econometric Theory*, 23(2):251–280.

Hafner, C. M., Van Dijk, D., and Hans Franses, P. (2006). Semi-parametric modelling of correlation dynamics. In Terrell, D. and Fomby, T., editors, *Econometric Analysis of Financial and Economic Time Series*, volume 1, pages 59–103. Emerald Group Publishing Limited.

Hagerud, G. E. (1997). A new non-linear garch model. Phd dissertation, EFI Economic Research Institute, Stockholm. Available from: `https://ex.hhs.se/dissertations/221859-FULLTEXT01.pdf`.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57(2):357–384.

Han, Y. (2006). Asset allocation with a high dimensional latent factor stochastic volatility model. *The Review of Financial Studies*, 19(1):237–271.

Hang, Q. (2017). *User Guide of GARCH-MIDAS and DCC-MIDAS MATLAB Programs*. Midas matlab toolbox - repack of mi(xed) da(ta) s(ampling) regressions (midas) written by eric ghysels and collaborators. Available from: `https://ww2.mathworks.cn/matlabcentral/fileexchange/45150-midas-matlab-toolbox`.

Hansen, P. and Lunde, A. (2004). *An unbiased measure of realized variance*. Working paper. Available from: `https://ssrn.com/abstract=524602`.

Hansen, P. and Lunde, A. (2005). A forecast comparison of volatility models: Does anything beat a GARCH(1,1)? *Journal of Applied Econometrics*, 20:873–889.

Hansen, P. R., Lunde, A., and Nason, J. M. (2005). A test for superior predictive ability. *Journal of Business Economics and Statistics*, 23:365–380.

Hansen, P. R., Lunde, A., and Nason, J. M. (2011). The model confidence set. *Econometrica*, 79(2):453–497.

Harlow, W. V. (1991). Asset allocation in a downside-risk framework. *Financial analysts journal*, 47(5):28–40.

Harvey, A. C. (1998). Long memory in stochastic volatility. In Knight, J. and Satchell, S., editors, *Forecasting Volatility in the Financial Markets*, chapter 16, pages 351–363. Elsevier, 1 edition.

Harvey, A. C., Ruiz, E., and Shephard, N. (1994). Multivariate stochastic variance models. *The Review of Economic Studies*, 61(2):247–264.

Harvey, A. C. and Shephard, N. (1996). The estimation of an asymmetric stochastic volatility model for asset returns. *Journal of Business and Economic Statistics*, 14(4):429–434.

Harvey, C. and Whaley, R. (1992). Market volatility prediction and the efficiency of the S&P 100 index option market. *Journal of Financial Economics*, 31:43–74.

Harvey, D., Leybourne, S., and Newbold, P. (1997). Testing the equality of prediction mean squared errors. *International Journal of forecasting*, 13(2):281–291.

Haug, S. and Czado, C. (2007). An exponential continuous-time GARCH process. *Journal of Applied Probability*, 44(4):960–976.

Hautsch, N. and Ou, Y. (2008). *Discrete-time stochastic volatility models and MCMC-based statistical inference*. Sfb 649 discussion paper 2008-063. Available from: `https://ideas.repec.org/p/hum/wpaper/sfb649dp2008-063.html`.

Hentschel, L. (1995). All in the family: Nesting symmetric and asymmetric GARCH models. *Journal of Financial Economics*, 39(1):71–104.

Herwartz, H. (2017). Stock return prediction under GARCH-an empirical assessment. *International Journal of Forecasting*, 33(3):569–580.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The review of financial studies*, 6(2):327–343.

Heston, S. L. and Nandi, S. (2000). A closed-form garch option valuation model. *The review of financial studies*, 13(3):585–625.

HFR (2019). Hedge fund research. https://www.hedgefundresearch.com.

Higgins, M. L. and Bera, A. K. (1992). A class of nonlinear ARCH models. *International Economic Review*, 33(1):137–158.

Hitaj, A., Mercuri, L., and Rroji, E. (2019). Sensitivity analysis of mixed tempered stable parameters with implications in portfolio optimization. *Computational Management Science*, 16(1):1–25.

Hitaj, A. and Zambruno, G. (2016). Are smart beta strategies suitable for hedge fund portfolios? *Review of Financial Economics*, 29:37–51.

Høg, E. and Tsiaras, L. (2011). Density forecasts of crude-oil prices using option-implied and ARCH-type models. *Journal of Futures Markets*, 31(8):727–754.

Hong, Y. and Li, H. (2005). Nonparametric specification testing for continuous-time models with applications to term structure of interest rates. *The Review of Financial Studies*, 18(1):37–84.

Hu, Y.-P. and Tsay, R. S. (2014). Principal volatility component analysis. *Journal of Business & Economic Statistics*, 32(2):153–164.

Huang, A. Y. (2012). Volatility forecasting by quantile regression. *Applied Economics*, 44(4):423–433.

Huang, T.-H. and Wang, Y.-H. (2012). The volatility and density prediction performance of alternative garch models. *Journal of Forecasting*, 31(2):157–171.

Hull, J. and White, A. (1987). The pricing of options with stochastic volatilities. *Journal of Finance*, 42(2):281–300.

Hwang, S. and Satchell, S. E. (2005). GARCH model with cross-sectional volatility: GARCHX models. *Applied Financial Economics*, 15(3):203–216.

Ishihara, T., Omori, Y., and Asai, M. (2016). Matrix exponential stochastic volatility with cross leverage. *Computational Statistics & Data Analysis*, 100:331–350.

Jacquier, E., Polson, N. G., and Rossi, P. E. (2004). Bayesian analysis of stochastic volatility models with fat-tails and correlated errors. *Journal of Econometrics*, 122(1):185–212.

Jagannathan, R. and Ma, T. (2003). Risk reduction in large portfolios: Why imposing the wrong constraints helps. *The Journal of Finance*, 58(4):1651–1683.

Jagannathan, R., Malakhov, A., and Novikov, D. (2010). Do hot hands exist among hedge fund managers? An empirical evaluation. *The Journal of Finance*, 65(1):217–255.

Jeantheau, T. (1998). Strong consistency of estimators for multivariate ARCH models. *Econometric theory*, 14(1):70–86.

Jiang, G. and Tian, Y. (2005). The model-free implied volatility and its information content. *The Review of Financial Studies*, 18(4):1305–1342.

Jiang, Y., Ahmed, S., and Liu, X. (2017). Volatility forecasting in the Chinese commodity futures market with intraday data. *Review of Quantitative Finance and Accounting*, 48.

Jobson, J. D. and Korkie, B. M. (1981). Performance hypothesis testing with the Sharpe and Treynor measures. *The Journal of Finance*, 36(4):889–908.

Joenväärä, J., Kaupila, M., Kosowski, R., and Tolonen, P. (2014). *Hedge fund performance: What do we know?* Cepr discussion paper no. dp13618.

Joenväärä, J., Kaupila, M., Kosowski, R., and Tolonen, P. (2019a). *Hedge Fund Performance: Are Stylized Facts Sensitive to Which Database One Uses?* Working paper. Available from: `https://ssrn.com/abstract=1989410`.

Joenväärä, J., Kaupila, M., Kosowski, R., and Tolonen, P. (2019b). *Hedge fund performance: What do we know?* Cepr discussion paper no. dp13618.

Joenväärä, J., Kosowski, R., and Tolonen, P. (2019c). The effect of investment constraints on hedge fund investor returns. *Journal of Financial and Quantitative Analysis*, 54(4):1539–1571.

Johannes, M. S., Polson, N. G., and Stroud, J. R. (2009). Optimal filtering of jump diffusions: Extracting latent states from asset prices. *The Review of Financial Studies*, 22(7):2759–2799.

Johnson, D. (1998). The robustness of mean and variance approximations in risk analysis. *Journal of the Operational Research Society*, 49(3):253–262.

Jorion, P. (1986). Bayes-stein estimation for portfolio analysis. *Journal of Financial and Quantitative Analysis*, 21(3):279–292.

Jung, Y. C. (2016). Relative performance of vixc vs. garch in predicting realised volatility changes. *Investment Analysts Journal*, 45(sup-1):1–16.

Kalimipalli, M. and Susmel, R. (2007). Switching stochastic volatility and the short-term interest rates. *Journal of Empirical Finance*, 11(3):309–329.

Kambouroudis, D., McMillan, D., and Tsakou, K. (2016). Forecasting stock return volatility: A comparison of GARCH, implied volatility, and realized volatility models. *Journal of Futures Markets*, 36(12).

Kaplan, P. D. and Knowles, J. A. (2004). Kappa: A generalized downside risk-adjusted performance measure. *Journal of Performance Measurement*, 8:42–54.

Karanasos, M., Yfanti, S., and Christopoulos, A. (2021). The long memory heavy process: modeling and forecasting financial volatility. *Annals of Operations Research*, 306:111–130.

Kawakatsu, H. (2006). Matrix exponential GARCH. *Journal of Econometrics*, 134(1):95–128.

Kazemi, H., Schneeweis, T., and Gupta, B. (2004). Omega as a performance measure. *Journal of performance measurement*, 8(3):16–25.

Keating, C. and Shadwick, W. F. (2002). A universal performance measure. *Journal of performance measurement*, 6(3):59–84.

Keefer, D. L. (1994). Certainty equivalents for three-point discrete-distribution approximations. *Management science*, 40(6):760–773.

Keefer, D. L. and Bodily, S. E. (1983). Three-point approximations for continuous random variables. *Management Science*, 29(5):595–609.

Keefer, D. L. and Verdini, W. A. (1993). Better estimation of pert activity time parameters. *Management science*, 39(9):1086–1091.

Khemiri, R. (2011). The smooth transition GARCH model: application to international stock indexes. *Applied Financial Economics*, 21(8):555–562.

Kodres, L. E. (1993). Test of unbiasedness in foreign exchange futures markets: An examination of price limits and conditional heteroskedasticity. *Journal of Business*, 66(3):463–490.

Koehler, P. A. (2009). *Predicting Realized Volatility: The Effects of Extended Hours Trading*. Citeseer.

Konstantinidi, E., G., S., and E. Tzagkaraki, E. (2008). Can the evolution of implied volatility be forecasted? *Journal of Banking and Finance*, 32(11):2401–2411.

Koopman, S. J., Jungbacker, B., and Hol, E. (2005). Forecasting daily variability of the S&P 100 stock index using historical, realised and implied volatility measurements. *Journal of Empirical Finance*, 12(3):445–475.

Koopman, S. J. and Uspensky, E. H. (2002). The stochastic volatility in mean model: Empirical evidence from international stock markets. *Journal of Applied Econometrics*, 17(6):667–689.

Kosowski, R., Naik, N. Y., and Teo, M. (2007). Do hedge funds deliver alpha? A Bayesian and bootstrap analysis. *Journal of Financial Economics*, 84(1):229–264.

Kostakis, A., Panigirtzoglou, N., and Skiadopoulos, G. (2011). Market timing with option-implied distributions: A forward-looking approach. *Management Science*, 57(7):1231–1249.

Kourtis, A., Dotsis, G., and Markellos, R. N. (2012). Parameter uncertainty in portfolio selection: Shrinking the inverse covariance matrix. *Journal of Banking & Finance*, 36(9):2522–2531.

Koutsandreas, D., Spiliotis, E., Petropoulos, F., and Assimakopoulos, V. (2021). On the selection of forecasting accuracy measures. *Journal of the Operational Research Society*, 0:1–18.

Kroner, K. F. and Ng, V. K. (1998). Modeling asymmetric comovements of asset returns. *The review of financial studies*, 11(4):817–844.

Kuan, C.-M. (2002). Lecture on the Markov switching model. *Institute of Economics Academia Sinica*, 8(15):1–30.

Kwan, C., Li, W., and Ng, K. (2005). A multivariate threshold GARCH model with time-varying correlations. *Econometric Reviews*, 29(1):20–38.

Lanne, M. and Saikkonen, P. (2007). A multivariate generalized orthogonal factor GARCH model. *Journal of Business & Economic Statistics*, 25(1):61–75.

Ledoit, O. and Wolf, M. (2003). Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of empirical finance*, 10(5):603–621.

Ledoit, O. and Wolf, M. (2004a). Honey, I shrunk the sample covariance matrix. *Journal of Portfolio Management*, 30(4):110–119.

Ledoit, O. and Wolf, M. (2004b). A well-conditioned estimator for large-dimensional covariance matrices. *Journal of multivariate analysis*, 88(2):365–411.

Leitch, G., . T. J. E. (1995). Professional economic forecasts: Are they worth their costs? *Journal of Forecasting*, 14:143–157.

Leitch, G. and Tanner, J. E. (1991). Economic forecast evaluation: Profits versus the conventional error measures. *The American Economic Review*, pages 580–590.

Li, C. W. and Li, W. K. (1996). On a double-threshold autoregressive heteroscedastic time series model. *Journal of Applied Econometrics*, 11:253–74.

Li, W., Gao, J., Li, K., and Yao, Q. (2016). Modeling multivariate volatilities via latent common factors. *Journal of Business & Economic Statistics*, 34:564–573.

Li, X. and Zakamulin, V. (2020). Stock volatility predictability in bull and bear markets. *Quantitative Finance*, 20(7):1149–1167.

Liu, C. and Maheu, J. M. (2009). Forecasting realized volatility: a bayesian model-averaging approach. *Journal of Applied Econometrics*, 24(5):709–733.

Liu, H.-C. and Hung, J.-C. (2010). Forecasting S&P-100 stock index volatility: The role of volatility asymmetry and distributional assumption in garch models. *Expert Systems with Applications*, 37(7):4928–4934.

Liu, L. and Pan, Z. (2020). Forecasting stock market volatility: The role of technical variables. *Economic Modelling*, 84:55–65.

Liu, L., Patton, A., and Sheppard, K. (2015). Does anything beat 5-minute RV? A comparison of realized measures across multiple asset classes. *Journal of Econometrics*, 187(1):293–311.

Liu, X., Shackleton, M. B., Taylor, S. J., and Xu, X. (2007). Closed-form transformations from risk-neutral to real-world distributions. *Journal of Banking & Finance*, 31(5):1501–1520.

Lu, Y. (2005). *Modeling and forecasting daily stock return volatility with intra-day price fluctuation information*. Phd thesis, Department of Economics, University of Pennsylvania. Available from: `https://repository.upenn.edu/dissertations/AAI3179771`.

Luo, L., Pairote, S., and Chatpatanasiri, R. (2017). GARCH-type forecasting models for volatility of stock market and MCS test. *Communications in Statistics-Simulation and Computation*, 46(7):5303–5312.

Ma, F., Li, Y., Liu, L., and Zhang, Y. (2018). Are low-frequency data really uninformative? A forecasting combination perspective. *The North American Journal of Economics and Finance*, 44:92–108.

Madan, D. B., Carr, P. P., and Chang, E. C. (1998). The variance gamma process and option pricing. *Review of Finance*, 2(1):79–105.

Maheua, J. M. and McCurdycd, T. H. (2011). Do high-frequency measures of volatility improve forecasts of return distributions? *Journal of Econometrics*, 160 (1):69–76.

Maillard, S., Roncalli, T., and Teiletche, J. (2010). The properties of equally weighted risk contribution portfolios. *The Journal of Portfolio Management*, 36(4):60–70.

Makridakis, S. and Winkler, R. L. (1983). Averages of forecasts: Some empirical results. *Management Science*, 29(9):987–996.

Manser, S. and Schmid, M. M. (2009). The performance persistence of equity long/short hedge funds. *Journal of derivatives & hedge funds*, 15(1):51–69.

Marcucci, J. (2005). Forecasting stock market volatility with regime-switching GARCH models. *Studies in Nonlinear Dynamics & Econometrics*, 9:1–55.

Markowitz, H. (1952). Portfolio selection. *The journal of finance*, 7(1):77–91.

Markowitz, H. (1956). The optimization of a quadratic function subject to linear constraints. *Naval research logistics Quarterly*, 3(1-2):111–133.

Markowitz, H. (1959). Portfolio selection: Efficient diversification of investments. *Yale University Press*.

Martens, M. and Poon, S.-H. (2001). Returns synchronization and daily correlation dynamics between international stock markets. *Journal of Banking & Finance*, 25(10):1805–1827.

Martens, M., van Dijk, D., and de Pooter, M. (2004). *Modeling and forecasting S&P 500 volatility: long memory, structural breaks and nonlinearity*. Tinbergen institute discussion papers 04-067/4, Erasmus University Rotterdam. Available from: `http://www.finance.martinsewell.com/stylized-facts/volatility/MartensvanDijkdePooter2004.pdf`.

McAleer, M. and Medeiros, M. (2008). A multiple regime smooth transition heterogeneous autoregressive model for long memory and asymmetries. *Journal of Econometrics*, 147(1):104–119.

Mcmillan, D. and Speight, A. (2012). Daily FX volatility forecasts: Can the GARCH(1,1) model be beaten using high-frequency data? *Journal of Forecasting*, 31(4).

Meddahi, N. (2001). *An eigenfunction approach for volatility modeling*. Cahiers de recherche 2001-29, Universite de Montreal, Departement de sciences economiques. Available from: `http://www.crde.umontreal.ca/cahiers/29-2001-cah.pdf` [Accessed on 20 May 2018].

Meddahi, N. (2002). A theoretical comparison between integrated and realized volatility. *Journal of Applied Econometrics*, 17(5):479–508.

Medeiros, M. and Veiga, A. (2003). Diagnostic checking in a flexible nonlinear time series model. *Journal of Time Series Analysis*, 24(4):461–482.

Meese, R. and Rogoff, K. (1988). Was it real? The exchange rate-interest differential relation over the modern floating-rate period. *Journal of Finance*, 43(4):933–948.

Meligkotsidou, L., Panopoulou, E., Vrontos, I. D., and Vrontos, S. D. (2019). Quantile forecast combinations in realised volatility prediction. *Journal of the Operational Research*, 70(10):1720–1733.

Memmel, C. (2003). Performance hypothesis testing with the Sharpe ratio. *Finance Letters*, 1(1):21–23.

Meng, F. and Liu, L. (2019). Analyzing the economic sources of oil price volatility: An out-of-sample perspective. *Energy*, 177:476–486.

Merton, R. (1991). On market timing and investment performance. i. an equilibrium theory of value for market forecasts. *Journal of Business*, 54:363–406.

Merton, R. C. (1973). Theory of rational option pricing. *The Bell Journal of economics and management science*, 4(1):141–183.

Mincer, J. A. and Zarnowitz, V. (1969). The evaluation of economic forecasts. In *Economic forecasts and expectations: Analysis of forecasting behavior and performance*, pages 3–46. NBER.

Moder, J. J. and Rodgers, E. (1968). Judgment estimates of the moments of PERT type distributions. *Management Science*, 15(2):76–83.

Moosa, I. and Vaz, J. (2015). Directional accuracy, forecasting error and the profitability of currency trading: Model-based evidence. *Applied Economics*, 47(57):6191–6199.

Morales-Arias, L. and Moura, G. V. (2013). Adaptive forecasting of exchange rates with panel data. *International Journal of Forecasting*, 29(3):493–509.

Morgan, J. et al. (1996). *Riskmetrics technical document*. New York RiskMetrics, 4 edition.

Morton, D. P., Popova, E., and Popova, I. (2006). Efficient fund of hedge funds construction under downside risk measures. *Journal of Banking & Finance*, 30(2):503–518.

Mosteller, F. (1946). On some useful "inefficient" statistics. *The Annals of Mathematical Statistics*, 17(4):377–408.

Muller, U. A., Dacorogna, M. M.and Davé, R. D., Olsen, R. B., Puctet, O. V., and von Weizsacker, J. (1997). Volatilities of different time resolutions - analysing the dynamics of market component. *Journal of Empirical Finance*, 4:213–239.

Näsholm, A. and Bunjaku, B. (2010). *Forecasting Volatility-A Comparison Study of Model Based Forecasts and Implied Volatility*. Master's thesis, Lund University. Available from: `http://lup.lub.lu.se/student-papers/record/1627421`.

Nelson, D. B. (1990). Stationarity and persistence in the GARCH(1, 1) mode. *Econometric Theory*, 6(3):318–334.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, 59(2):347–370.

Newey, W. and West, K. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consisten covariance matrix, econametrica, 55, 703-708. *Organizational Behavior and Human Decision Processes*, 39:228–241.

Ng, L. (1991). Tests of the CAPM with time-varying covariances: A multivariate approach. *The Journal of Finance*, 46(4):1507–1521.

Niu, X. and Harvey, N. (2022). Point, interval, and density forecasts: Differences in bias, judgment noise, and overall accuracy. *Futures & Foresight Science*, 4(3-4).

Nowicka-Zagrajek, J. and Werron, A. (2001). Dependence structure of stable R-GARCH processes. *Probability and Mathematical Statistics*, 21(2):371–380.

Omori, Y., Chib, S., Shephard, N., and Nakajima, J. (2007). Stochastic volatility with leverage: Fast and efficient likelihood inference. *Journal of Econometrics*, 140(2):425–449.

Palandri, A. (2009). Sequential conditional correlations: Inference and evaluation. *Journal of Econometrics*, 153(2):122–132.

Palm, F. C. and Zellner, A. (1992). To combine or not to combine? Issues of combining forecasts. *Journal of Forecasting*, 11(8):687–701.

Papadaki, G. (2011). *Model averaging for volatility forecasting, option pricing and asset allocation*. PhD thesis, Imperial College London. Available from: `https://spiral.imperial.ac.uk/bitstream/10044/1/6395/1/Papadaki-G-2011-PhD-Thesis.pdf`.

Parkinson, M. (1980). The extreme value method for estimating the variance of the rate of return. *Journal of Business*, 53(1):61–65.

Patton, A. J. (2006). Modelling asymmetric exchange rate dependence. *International economic review*, 47(2):527–556.

Patton, A. J. (2011). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, 160:246–256.

Patton, A. J. (2018). *Forecast Evaluation and Comparison Theory*. Unpublished presentation at: 1^{st} *QFFE Summer School*, 28-30 may 2018, marseille, france.

Patton, A. J. and Sheppard, K. (2009a). Evaluating volatility and correlation forecasts. In T. G. Andersen, R. A. Davis, J.-P. K. and Mikosch, T., editors, *Handbook of financial time series*, pages 801–838. Springer.

Patton, A. J. and Sheppard, K. (2009b). Optimal combinations of realised volatility estimators. *International Journal of Forecasting*, 25(2):218–238.

Patton, A. J. and Sheppard, K. (2015). Good volatility, bad volatility: Signed jumps and the persistence of volatility. *Review of Economics and Statistics*, 97(3):683–697.

Pearson, E. S. and Tukey, J. W. (1965). Approximate means and standard deviations based on distances between percentage points of frequency curves. *Biometrika*, 52(3/4):533–546.

Pedersen, C. S. and Rudholm-Alfvin, T. (2003). Selecting a risk-adjusted shareholder performance measure. *Journal of asset management*, 4(3):152–172.

Pedersen, R. S. and Rahbek, A. (2014). Multivariate variance targeting in the BEKK–GARCH model. *The Econometrics Journal*, 17(1):24–55.

Pelletier, D. (2006). Regime switching for dynamic correlations. *Journal of econometrics*, 131(1-2):445–473.

Perez-Quiros, G. and Timmermann, A. (2007). Variations in the mean and volatility of stock returns around turning points of the business cycle. In *Forecasting Volatility in the Financial Markets*, pages 333–350. Elsevier.

Pesaran, H. M. and Timmermann, A. (1992). A simple nonparametric test of predictive performance. *Journal of Business & Economic Statistics*, 10(4):461–465.

Pesaran, M. and Pick, A. (2009). Forecast combination across estimation windows. *Journal of Business & Economic Statistics*, 29(2):307–318.

Pesaran, M. H. and Timmerman, A. (2005). *Selection of estimation window for strictly exogenous regressors*. Working paper, Cambridge University and University of California, San Diego. Available from: `http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.153.5149&rep=rep1&type=pdf`.

Pfingsten, A., Wagner, P., and Wolferink, C. (2004). An empirical investigation of the rank correlation between different risk measures. *The Journal of Risk*, 6(4):55–74.

Platanakis, E., Sakkas, A., and Sutcliffe, C. (2019). Harmful diversification: Evidence from alternative investments. *The British Accounting Review*, 51(1):1–23.

Politis, D. N. and Romano, J. P. (1994). The stationary bootstrap. *Journal of the American Statistical association*, 89(428):1303–1313.

Poon, S.-H. (2005). *A practical guide to forecasting financial market volatility*. John Wiley & Sons.

Poon, S.-H. and Granger, C. (2003). Forecasting volatility in financial markets: A review. *Journal of Economic Literature*, 41(2):478–539.

Priestley, M. B. (1988). *Non-linear and non-stationary time series analysis*. London:Academic press.

Qian, E. (2013). Are risk-parity managers at risk parity? *The Journal of Portfolio Management*, 40(1):20–26.

Qian, E. E. (2005). Risk parity portfolios: Efficient portfolios through true diversification. *Panagora Research paper*.

Qian, E. E. (2006). On the financial interpretation of risk contribution: Risk budgets do add up. *Journal of Investment Management*, 4(4):1–11.

Qian, E. E. (2012). Pension liabilities and risk parity. *The Journal of Investing*, 21(3):93–101.

Qian, E. E. (2016). *Risk parity fundamentals*. CRC Press.

Quaedvlieg, R. (2016). *Risk and uncertainty*. PhD thesis, Maastricht University. Available from: `https://higherlogicdownload.s3.amazonaws.com/AMSTAT/7328eb46-c57e-49af-8d4f-0460ca4a6a2c/UploadedImages/Quaedvlieg_thesis.pdf`.

Rabemananjara, R. and Zakoían, J. M. (1993). Threshold ARCH models and asymmetries in volatility. *Journal of Applied Econometrics*, 8(1):31–49.

Rapach, D. and Strauss, J.K., . (2008). Structural breaks and GARCH models of exchange rate volatility. *Journal of Applied Econometrics*, 23(1):65–90.

Rapach, D. E., Strauss, J. K., and Zhou, G. (2010). Out-of-sample equity premium prediction: Combination forecasts and links to the real economy. *The Review of Financial Studies*, 23(2):821–862.

Robinson, P. M. (1991). Testing for strong serial correlation and dynamic conditional heteroskedasticity in multiple regression. *Journal of Econometrics*, 47(1):67–84.

Rompolis, L. S. (2010). Retrieving risk neutral densities from european option prices based on the principle of maximum entropy. *Journal of Empirical Finance*, 17(5):918–937.

Roncalli, T. (2013). *Introduction to risk parity and budgeting*. CRC Press, 1 edition.

Roncalli, T. and Weisang, G. (2016). Risk parity portfolios with risk factors. *Quantitative Finance*, 16(3):377–388.

Rubinstein, M. (1994). Implied binomial trees. *The journal of finance*, 49(3):771–818.

Rudin, A. and Marr, W. M. (2016). Investor views, drawdown-based risk parity, and hedge fund portfolio construction. *The Journal of Alternative Investments*, 19(2):63–69.

Samkharadze, B. (2013). *" Essays on modeling international stock market volatility spillovers using multivariate GARCH models*. PhD thesis, UCL-Université Catholique de Louvain. Available from: `http://hdl.handle.net/2078.1/139184`.

Samuels, J. and Sekkel, R. (2017). Model confidence sets and forecast combination. *International Journal of Forecasting*, 33:48–60.

Santos, A. A. and Moura, G. V. (2014). Dynamic factor multivariate GARCH model. *Computational Statistics & Data Analysis*, 76:606–617.

Santos, A. A. P., Nogales, F. J., and Ruiz, E. (2013). Comparing univariate and multivariate models to forecast portfolio value-at-risk. *Journal of Financial Econometrics*, 23:400–441.

Satchell, S. and Knight, J. (2007). *Forecasting volatility in the financial markets*. Elsevier, 3 edition.

Satchell, S. and Timmermann, A. (1995). An assessment of the economic value of non-linear foreign exchange rate forecasts. *Journal of Forecasting*, 14(6):477–497.

Scharth, M. and Medeiros, M. C. (2009). Asymmetric effects and long memory in the volatility of Dow Jones stocks. *International Journal of Forecasting*, 25(2):304–327.

Schwarz, G. et al. (1978). Estimating the dimension of a model. *The annals of statistics*, 6(2):461–464.

Schwert, G. W. (1989). Why does stock market volatility change over time? *Journal of Finance*, 44(5):1115–1153.

Sentana, E. (1995). Quadratic ARCH models. *Review of Economic Studies*, 62(4):639–661.

Shackleton, M. B., Taylor, S. J., and Yu, P. (2010). A multi-horizon comparison of density forecasts for the S&P 500 using index returns and option prices. *Journal of Banking and Finance*, 34(11):2678–2693.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The journal of finance*, 19(3):425–442.

Sharpe, W. F. (1966). Mutual fund performance. *The Journal of business*, 39(1):119–138.

Shephard, N. (2005). *Stochastic volatility: selected readings*. Oxford University Press on Demand.

Shephard, N. and Sheppard, K. (2010). Realising the future: forecasting with high-frequency-based volatility (HEAVY) models. *Journal of Applied Econometrics*, 25(2):197–231.

Sheppard, K. (2009). *Oxford MFE toolbox*. econometric software, University of Oxford. Available from: `hhttp://www.kevinsheppard.com/MFE_Toolbox`.

Sheppard, K. (2016). *Financial econometrics notes*. Technical report, University of Oxford. Available from: `https://www.kevinsheppard.com/images/4/4c/Financial_Econometrics_2016-2017_tablet.pdf` [Accessed on 25 November 2017].

Silvennoinen, A. and Teräsvirta, T. (2005). Multivariate autoregressive conditional heteroskedasticity with smooth transitions in conditional correlations. Technical report, SSE/EFI Working Paper Series in Economics and Finance. Available from: `https://ideas.repec.org/p/hhs/hastef/0577.html`.

Silvennoinen, A. and Teräsvirta, T. (2009a). Modeling multivariate autoregressive conditional heteroskedasticity with the double smooth transition conditional correlation GARCH model. *Journal of Financial Econometrics*, 7(4):373–411.

Silvennoinen, A. and Teräsvirta, T. (2009b). Multivariate garch models. In T. G. Andersen, R. A. Davis, J.-P. K. and Mikosch, T., editors, *Handbook of financial time series*, pages 201–229. Springer.

Silvennoinen, A. and Teräsvirta, T. (2015). Modeling conditional correlations of asset returns: A smooth transition approach. *Econometric Reviews*, 34(1-2):174–197.

Smith, M. and Pitts, A. (2006). Foreign exchange intervention by the bank of Japan: Bayesian analysis using a bivariate stochastic volatility model. *Econometric Reviews*, 25(2-3):425–451.

So, M., Lam, K., and Li, W. (1998). A stochastic volatility model with Markov switching. *Journal of Business and Economic Statistics*, 16(2):244–253.

So, M., Li, W., and Lam, K. (2002). A threshold stochastic volatility model. *Journal of Forecasting*, 21:473–500.

So, M. K., Li, W., and Lam, K. (1997). Multivariate modelling of the autoregressive random variance process. *Journal of Time Series Analysis*, 18(4):429–446.

Sortino, F. A. and Van Der Meer, R. (1991). Downside risk. *Journal of portfolio Management*, 17(4):27.

Sortino, F. A., Van Der Meer, R., and Plantinga, A. (1999a). The Dutch triangle. *The Journal of Portfolio Management*, 26(1):50–57.

Sortino, F. A., Van der Meer, R., and Plantinga, A. (1999b). The upside potential ratio. *Journal of Performance Measurement*, 4(1):10–15.

Stock, J. H. and Watson, M. W. (2002). Forecasting using principal components from a large number of predictors. *Journal of the American statistical association*, 97(460):1167–1179.

Stock, J. H. and Watson, M. W. (2004). Combination forecasts of output growth in a seven-country data set. *Journal of Forecasting*, 23(6):405–430.

Swanson, N. R. and White, H. (1995). A model-selection approach to assessing the information in the term structure using linear models and artificial neural networks. *Journal of Business & Economic Statistics*, 13(3):265–275.

Swanson, N. R. and White, H. (1997a). Forecasting economic time series using flexible versus fixed specification and linear versus nonlinear econometric models. *International journal of Forecasting*, 13(4):439–461.

Swanson, N. R. and White, H. (1997b). A model selection approach to real-time macroeconomic forecasting using linear models and artificial neural networks. *Review of Economics and Statistics*, 79(4):540–550.

Symitsi, E., Symeonidis, L., Kourtis, A., and Markellos, R. (2018). Covariance forecasting in equity markets. *Journal of Banking & Finance*, 96:153–168.

Tay, A. S. and Wallis, K. F. (2000). Density forecasting: A survey. *Journal of forecasting*, 19(4):235–254.

Taylor, J. W. (2005). Generating volatility forecasts from value at risk estimates. *Management Science*, 51(5):712–725.

Taylor, S. J. (1982). Financial returns modelled by the product of two stochastic processes-A study of the daily sugar prices 1961-75. *Time series analysis: theory and practice*, 6(1):203–226.

Taylor, S. J. (1986). *Modelling Financial Time Series*. New York: John Wiley., 1 edition.

Taylor, S. J. (2004). *Asset Price Dynamics and Prediction*. Princeton, NJ: Princeton University Press.

Taylor, S. J., Yadav, P., and Zhang, Y. (2010). The information content of implied volatilities and model-free volatility expectations: Evidence from options written on individual stocks. *Journal of Banking & Finance*, 34(4):871–881.

Tian, S. and Hamori, S. (2015). Modeling interest rate volatility: A realized approach. *Journal of Banking & Finance*, 61:158–171.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(1):267–288.

Timmer, J. and Weigend, A. S. (1997). Modeling volatility using state space models. *International Journal of Neural Systems*, 8(04):385–398.

Timmermann, A. (2006). Forecast combinations. In *Handbook of economic forecasting*, volume 1, chapter 4, pages 135–196. Elsevier.

Treynor, J. L. (1965). How to rate management of investment funds. *Harvard business review*, 43(1):63–75.

Treynor, J. L. and Black, F. (1973). How to use security analysis to improve portfolio selection. *The journal of business*, 46(1):66–86.

Trucíos, C., Hotta, L. K., and Pereira, P. L. V. (2019). On the robustness of the principal volatility components. *Journal of Empirical Finance*, 52:201–219.

Tse, Y. K. (1998). The conditional heteroscedasticity of the Yen-Dollar exchange rate. *Journal of Applied Econometrics*, 13(1):49–55.

Tse, Y. K. and Tsui, A. K. C. (2002). A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations. *Journal of Business & Economic Statistics*, 20(3):351–362.

Van der Weide, R. (2002). GO-GARCH: a multivariate generalized orthogonal GARCH model. *Journal of Applied Econometrics*, 17(5):549–564.

van Dijk, D. and Franses, P. H. (1999). Modeling multiple regimes in the business cycle. *Macroeconomic Dynamics*, 3(3):311–340.

Vilhelmsson, A. (2006). Garch forecasting performance under different distribution assumptions. *Journal of Forecasting*, 25:561–578.

Violante, F. and Laurent, S. (2012). Volatility forecasts evaluation and comparison. In *Handbook of Volatility Models and Their Applications*, pages 465–486. John Wiley and Sons.

Vrontos, I. D., Dellaportas, P., and Politis, D. N. (2003). A full-factor multivariate GARCH model. *The Econometrics Journal*, 6(2):312–334.

Vrontos, S. D., Galakis, J., and Vrontos, I. D. (2021). Implied volatility directional forecasting: a machine learning approach. *Quantitative Finance*, pages 1–20.

Wang, Y., Wu, C., and Yang, L. (2015). Hedging with futures: Does anything beat the naïve hedging strategy? *Management Science*, 61(12):2870–2889.

Wecker, W. E. (1981). Asymmetric time series. *Journal of the American Statistical Association*, 76(373):16–21.

Welch, I. and Goyal, A. (2008). A comprehensive look at the empirical performance of equity premium prediction. *The Review of Financial Studies*, 21(4):1455–1508.

West, K. (1996). Asymptotic inference about predictive ability. *Econometrica*, pages 1067–1084.

West, K. D. (2006). Forecast evaluation. *Handbook of economic forecasting*, 1:99–134.

White, H. (2000). A reality check for data snooping. *Econometrica*, 68:1097–1126.

Wu, E. H. and Philip, L. (2005). Volatility modelling of multivariate financial time series by using ICA-GARCH models. In Gallagher, M., H. J. and Maire, F., editors, *International Conference on Intelligent Data Engineering and Automated Learning*, volume 3578, pages 571–579. Springer.

Yeh, J.-H., Wang, J.-N., and Kuan, C.-M. (2014). A noise-robust estimator of volatility based on interquantile ranges. *Review of Quantitative Finance and Accounting*, 43(4):751–779.

Yfanti, S. and Karanasos, M. (2021). Financial volatility modeling with option-implied information and important macro-factors. *Journal of the Operational Research Society*, 0(0):1–21.

Yu, J. (2005). On leverage in a stochastic volatility model. *Journal of Econometrics*, 127(2):165–178.

Yu, J. and Meyer, R. (2006). Multivariate stochastic volatility models: Bayesian estimation and model comparison. *Econometric Reviews*, 25(2-3):361–384.

Yun, J. (2014). Out-of-sample density forecasts with affine jump diffusion models. *Journal of Banking & Finance*, 47:74–87.

Zakoian, J.-M. (1994). Threshold heteroskedastic models. *Journal of Economic Dynamics and control*, 18(5):931–955.

Zhang, Y., Lei, L., and Wei, Y. (2020). Forecasting the chinese stock market volatility with international market volatilities: The role of regime switching. *The North American Journal of Economics and Finance*, 52:101145.

Zhang, Y., Taylor, S. J., and Wang, L. (2013). Investigating the information content of the model-free volatility expectation by monte carlo methods. *Journal of Futures Markets*, 33(11):1071–1095.

Zhang, Y., Wang, Y., and Ma, F. (2021). Forecasting us stock market volatility: How to use international volatility information. *Journal of Forecasting*, 40(5):733–768.

Zou, H. (2006). The adaptive lasso and its oracle properties. *Journal of the American statistical association*, 101(476):1418–1429.

Zumbach, G. O. (2007). *The riskmetrics 2006 methodology*. Technical document. Available from: `https://ssrn.com/abstract=1420185`.