Charles University in Prague

Faculty of Social Sciences
Institute of Economic Studies

MASTER THESIS

Dynamic Portfolio Optimization During Financial Crisis Using Daily Data and High-frequency Data

Author: **Bc. František Čech**

Supervisor: **PhDr. Jozef Baruník, Ph.D.**

Academic Year: **2012/2013**

Declaration of Authorship

The author hereby declares that he compiled this thesis independently, using only the listed resources and literature.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis document in whole or in part.

Prague, January 7, 2013

Signature

Acknowledgments

I am very grateful to my supervisor, PhDr. Jozef Baruník, Ph.D. for his help, valuable suggestions, comments and encouragement that helped me to complete this thesis. I would also like to thank Matteo Bonato who provided me with the WAR codes.

Abstract

This thesis focuses on variance-covariance matrix modeling and forecasting. Majority of existing research evaluates covariance forecasts by statistical criteria. Our main contribution is economic comparison of parametric and non-parametric approaches of covariance matrix modeling. Parametric approach relies on RiskMetrics and Dynamic Conditional Correlation GARCH models that are applied on daily data. In the second approach, estimates of variance-covariance matrix are directly obtained from the high-frequency data by non-parametric techniques Realized Covariation and Multivariate Realized Kernels. These estimates are further modeled by Heterogeneous and Wishart Autoregression. Moreover, our contribution arises from the use of dataset that covers period of financial crisis. Portfolio of assets that is dynamically optimized consists of two highly liquid assets - Light Crude NYMEX and Gold COMEX, and of European asset represented by DAX index. Forecast evaluation results indicate better economic performance of models estimated on daily data. However, we found out that data synchronization procedure is the main driver of the results.

JEL Classification	C32, C58, G11, G17
Keywords	daily data, high-frequency data, DCC-GARCH model, RiskMetrics model, heterogeneous autoregressive model, Wishart autoregressive model, economic forecast evaluation

Author's e-mail	fero.cech@gmail.com
Supervisor's e-mail	barunik@utia.cas.cz

Abstrakt

Táto práca sa zameriava na modelovanie a prognózovanie variančnej-kovariančnej matice. Väčšina súčasného výskumu vyhodnocuje prognózy kovariančnej matice pomocou štatistických kritérií. Naším hlavným prínosom je ekonomické porovnanie parametrických a neparametrických prístupov k modelovaniu kovariančnej matice. Parametrický prístup je v práci zastúpený modelmi RiskMetrics a Dynamic Conditional Correlation GARCH, ktoré sú odhadnuté na denných dátach. V neparametrickom prístupe sú odhady variančnej-kovariančnej matice získané priamo z vysokofrekvenčných dát pomocou metód Realized Covariation a Multivariate Realized Kernels. Tieto odhady sú ďalej modelované pomocou heterogénnej a Wishartovej autoregresie. Ďalším prínosom tejto práce je použitie dát z obdobia finančnej krízy. Portfólio aktív, ktoré je dynamicky optimalizované, pozostáva z dvoch vysokolikvidných aktív - Light Crude NYMEX (ropa) a Gold COMEX (zlato), a európskeho aktíva zastúpeného DAX indexom. Výsledky ekonomického porovnania prognóz kovariančnej matice naznačujú lepšiu výkonnosť modelov odhadnutých na denných dátach. Zistili sme však, že hlavnou príčinou získania daných výsledkov je proces synchronizácie dát.

Klasifikace C32, C58, G11, G17

Klíčová slova denné dáta, vysokofrekvenčné dáta, DCC-GARCH model, RiskMetrics model, heterogénny autoregresný model, Wishartov autoregresný model, ekonomické ohodnotenie predpovedí

E-mail autora fero.cech@gmail.com

E-mail vedoucího práce barunik@utia.cas.cz

Contents

List of Tables

List of Figures

Acronyms

CL Light Crude NYMEX

DA DAX index

DCC Dynamic Conditional Correlation

GARCH Generalized Autoregressive Conditional Heteroscedasticity

GC Gold COMEX

GMVP Global Minimum Variance Portfolio

HAR Heterogeneous Autoregression

MRK Multivariate Realized Kernels

RCOV Realized Covariation

RMSFE Root Mean Square Forecasting Error

VaR Value-at-Risk

WAR Wishart Autoregression

Master Thesis Proposal

Author	Bc. František Čech
Supervisor	PhDr.Jozef Baruník, Ph.D.
Proposed topic	Dynamic portfolio optimization during financial crisis using daily data and high frequency data

Topic characteristics The aim of the thesis is to find optimal portfolio consisting of four assets (gold futures, euro-dollar futures, crude oil futures and stock index) using two different approaches. Risk knowledge of financial assets is a key issue of portfolios managers. Risk of the individual assets could be described by the volatility of the asset. While volatility can serve as a clue for selecting assets in a portfolio, it does not count with dependencies between assets. If one is interested in correlation structure of many assets, variance-covariance (VCV) matrix is of particular interest. In this thesis dynamic VCV matrix will be estimated using two different data sets. One will contain daily data and will be estimated by Dynamic Conditional Correlation GARCH. The second one will use high frequency data for estimation of realized variance and covariance. Based on the VCV matrices estimates optimal weights of the portfolio will be calculated.

Hypotheses

1. Dynamic Conditional Correlation GARCH estimated on daily data serves as a best estimator of variance covariance matrix used for dynamic portfolio optimization.

2. Realized measures using heterogeneous autoregression provide us with accurate estimates of variance covariance matrix used for dynamic portfolio optimization.
3. Correlations among assets evolve in time and changed dramatically during financial crisis.

Methodology Thesis will be divided into two parts – analysis of daily data and high frequency data. Robert Engel's Dynamic Condition Correlation GARCH will be used for the daily data analysis. DCC-GARCH is a special version of multivariate GARCH models used for dynamic conditional correlations estimation among multiple time series. Model can be estimated in two stages. First one involves univariate GARCH estimation. Second one uses standardized residuals from the first stage for conditional correlations estimation. For the high frequency data realized measures will be used. Consequently, daily realized variance and covariance will be calculated from the high frequency data. Realized measures could be further estimated using simple OLS regression resulting in Heterogeneous Autoregression (HAR). Having the VCV matrices estimated optimal weights of a portfolio will be calculated in order to construct minimum variance portfolio.

Outline

1. Introduction
2. Data
3. Methodology
   3.1. Dynamic Conditional Correlation GARCH
   3.2. Realized measures
4. Portfolio optimization
5. Results – comparison of performance of DCC-GARCH and realized measures
6. Conclusion

Core bibliography

1. ANDERSEN, TORBEN G., BOLLERSLEV TIM, DIEBOLD FRANCIS X. & LABYS PAUL: "Modeling and Forecasting Realized Volatility." *Econometrica* ,Vol. 71, no. 2 (2003): 579-625.
2. ANDERSEN, TORBEN G., BOLLERSLEV TIM, CHRISTOFFERSEN PETER & DIEBOLD FRANCIS X.: "Practical Volatility and Correlation Modeling for Financial Market Risk Management." In *The Risks of Financial Institutions*, by Mark Carey and René M. Stulz, 513-548. University of Chicago Press, 2007.
3. BARNDORFF-NIELSEN, OLE E.& NEIL SHEPHARD: "Econometric Analysis of Relaized Covariation: High Frequency Based Covariance, Regression, and Correlation in Financial Economics." *Econometrica* ,Vol. 72, no. 3 (2004): 885-925.
4. CHIRIAC, ROXANA & VALERI VOEV: "Modelling and Forecasting Multivariate Reralized Volatility." *Journal of Applied Econometrics*, 2011: 922-947.
5. ENGLE, ROBERT: "Dynamic Conditional Correlation - A Simple Class of Multivariate GARCH Models." *Forthcoming Journal of Business and Economic Statistics 2002*, 2002.
6. ENGLE, ROBERT F & SHEPPARD KEVIN: "Theoretical and Empirical Properties of Dynamic Conditional Correlation Multivariate GARCH." *NBER Working Paper Series*, 2001: Working Paper 8554.
7. VOEV, VALERI. „On the Economic Evaluation of Volatility Forecasts.“ *CREATES Research Paper 2009-56 .*

Author

Supervisor

Chapter 1

Introduction

Volatility modeling is one of the key issues in the area of financial econometrics. The risk of individual financial instruments is crucial for asset pricing, portfolio and risk management. Besides volatility of individual assets knowledge of covariance and correlation structure is of great importance. Accurate forecasts of variance-covariance matrices are particularly important in asset allocation and portfolio management.

Nature of the financial data with dependencies in higher moments of the daily return series motivated the work of *Engle (1982)* and later *Bollerslev (1986)*. They have developed a new family of parametric univariate conditionally heteroscedastic models represented by widely used Generalized Autoregressive Conditional Heteroscedasticity (GARCH). In the late eighties and nineties numerous multivariate extensions of the GARCH were created. Among all of them let us mention Constant Conditional Correlation GARCH of *Bollerslev (1990)* further generalized by *Engle (2002)* and BEKK model of *Engle* & *Kroner (1995)*. Multivariate GARCH (MGARCH) models are popular in the literature although they suffer from curse of dimensionality problem. Detailed information about MGARCH specifications can be found in *Bauwens et al. (2006)* for example.

Increased availability of high-frequency data in the last decade resulted in development of the new non-parametric approach of treating volatility, which is an interesting alternative to traditional MGARCH models. Model-free estimator called "realized volatility" that makes volatility observable is proposed in *Andersen et al. (2001)*. Most influential works providing rigorous theoretical background of the concept of realized volatility is *Andersen et al. (2003)* and *Barndorff-Nielsen* & *Shephard (2004)*. In *Barndorff-Nielsen* & *Shephard (2004)* theory of realized volatility is completed with "realized covariation". Estimates of variance-covariance matrix that are obtained by realized covariation method do not have to be necessarily positive semi-definite due to market microstructure noise. Therefore *Barndorff-Nielsen et al. (2011)* introduced Multivariate Realized Kernels estimator guaranteeing the positive semi-definiteness of the variance-covariance matrix.

Once the covariance matrix is estimated from the high-frequency data it needs to be further modeled. There is still ongoing research dedicated to the entire covariance matrices modeling. From the already established methods let us mention Wishart Autoregression (WAR) of *Gourieroux et al. (2009)* with numerous extensions presented in *Bonato et al. (2009)* and *Bonato et al. (2012)*. The use of Cholesky factors further estimated by Vector Autoregressive Fractionally Integrated Moving Average (VARFIMA), Heterogeneous Autoregresion (HAR) or WAR-HAR can be found in *Chiriac* & *Voev (2011)*.

Selection of the assets included in the portfolio that is dynamically optimized is crucial for research. Majority of researchers (*Andersen et al. (2001)*, *Andersen et al. (2003)*, *Bonato et al. (2009)*, *Voev (2007)* among others) concentrate on instruments traded mostly on the United States market (S&P 500 index or U.S. treasury bills) and evaluate forecasting performance generally by statistical criteria. However, our main contribution is that we include the European asset in portfolio and we evaluate covariance forecasts mostly by economic criteria. Economic evaluation of volatility forecasts is of great importance especially for financial practitioners because it provides direct financial evaluation of their decisions.

In our work covariance matrix forecasts used for dynamic portfolio optimization are obtained from two "return based" models represented by

RiskMetrics and DCC-GARCH and four "covariance based" models that include HAR, Cholesky-HAR, WAR and diagonal WAR. Moreover, we evaluate accuracy of these forecasts by one statistical (Root Mean Square Forecasting Error) and three economic criteria (Global Minimum Variance Portfolio, Mean-Variance optimization, Value-at-Risk). Our findings indicate better performance of covariance based models according to Root Mean Square Forecasting Error and Global Minimum Variance Portfolio criteria. On the other hand covariance based models are outperformed by RiskMetrics and DCC-GARCH for Mean-Variance optimization and Value-at-Risk forecast evaluation methods.

The rest of the thesis is structured as follows. We provide theoretical background for models in Chapter 2. Chapter 3 describes daily and high-frequency datasets. Chapter 4 presents results of out-of-sample forecast evaluation. In Chapter 5 we discuss our results. Chapter 6 concludes the thesis.

Chapter 2

Methodology

Riskiness of the financial instruments is crucial for asset allocation. Generally, we have some assets and we are interested in modeling and forecasting their volatility and correlations. In order to capture dependencies among assets we need variance-covariance (VCV) matrix. The general unconditional VCV matrix has the following form

$$VCV = \begin{pmatrix} \sigma_{11}^2 & cov(1,2) & \cdots & cov(1,n) \\ cov(2,1) & \sigma_{22}^2 & \cdots & cov(2,n) \\ \vdots & \vdots & \ddots & \vdots \\ cov(n,1) & cov(n,2) & \cdots & \sigma_{nn}^2 \end{pmatrix}.$$

On the main diagonal it has variances, off-diagonal elements are covariance and by definition is symmetric. Nowadays we know that volatility and correlations are not constant over time but they are time-varying. Therefore unconditional covariances are not sufficient from the asset allocation point of view and the dynamics of the asset dependencies need to be modeled dynamically.

In this chapter the theoretical background of dynamic covariance matrix modeling and forecasting together with forecasts evaluation methods are presented.

The first method of obtaining variance-covariance matrix forecasts is based on traditional time-series models that use daily returns for estimation. The models that we use are Exponentially Weighted Moving Average with parameter set to

RiskMetrics standards presented in section 2.1 and Dynamic Conditional Correlation GARCH described in section 2.2.

The next method relies on the use of high-frequency data. From the intraday closing prices the covariance matrices are directly calculated and afterwards modeled. Realized Covariation and Multivariate Realized Kernels presented in sections 2.3.1 and 2.3.2 are used for covariance matrix estimation. These estimates are further modeled by HAR and WAR model specifications described in sections 2.4 and 2.5.

In the end of the chapter we describe techniques of variance-covariance matrix forecasts evaluation.

2.1. RiskMetrics

Exponentially Weighted Moving Average (EWMA) that uses RiskMetrics is set as the benchmark for all competing models in this thesis. It has the simplest form among used models and it is also easy to implement even if the portfolio is composed of higher number of assets. Due to its simplicity RiskMetrics is commonly set as benchmark model and it is also widely used in financial practice.

Under RiskMetrics methodology[1] we consider a $N \times 1$ vector of returns r_t for $t = 1, \dots, T$ such that

$$r_t | \mathcal{F}_{t-1} \sim N(\mu_t, \sigma_t^2) \tag{1}$$

where μ_t is conditional mean and σ_t^2 stands for conditional variance of daily returns. Moreover if we assume $\mu_t = 0$, conditional variance has the form

$$\sigma^2 = (1 - \lambda) \sum_{t=1}^{T} \lambda^{t-1} r_t^2 \tag{2}$$

with $\lambda \epsilon (0; 1)$ representing decay factor. According to RiskMetrics standards we set decay factor to 0.94.

Equation (2) can be rewritten in recursive form, which will be used for estimation.

[1] for detailed information see *J.P.Morgan* & *Reuters (1996)*

$$\sigma_t^2 = \lambda\sigma_{t-1}^2 + (1-\lambda)r_{t-1}^2 \tag{3}$$

Extension from univariate to multivariate processes and thus to modeling not only variance but also covariance is straightforward. We modify equation (2) into

$$\sigma_{i,j} = (1-\lambda)\sum_{t=1}^{T}\lambda^{t-1}r_i r_j \tag{4}$$

with recursive form

$$\sigma_{i,j,t} = \lambda\sigma_{i,j,t-1} + (1-\lambda)r_{i,t-1}r_{j,t-1} \tag{5}$$

where expression $\sigma_{i,j,t}$ denotes covariance between assets i and j in time t. Note that for $i = j$ equations (2) and (5) are the same.

Nice feature of easily employable estimation remains valid also for EWMA forecasting. Having all relevant information at time t, one-step ahead covariance forecast has form

$$\sigma_{i,j,t+1} = \lambda\sigma_{i,j,t} + (1-\lambda)r_{i,t}r_{j,t} \quad . \tag{6}$$

Simplicity of the RiskMetrics might be the limiting factor of time-varying covariance matrix modeling. In the literature more accurate models that capture dynamic correlation and covariance structure of the assets have been developed. Most commonly used in literature is Dynamic Conditional Correlation GARCH of *Engle (2002)*.

2.2. Dynamic Conditional Correlation GARCH

The Dynamic Condition Correlation (DCC) GARCH model is a special version of multivariate GARCH models which allow us to examine dynamic conditional correlations among multiple time series. It was introduced by *Engle (2002)* and can be seen as generalization of *Bollerslev (1990)*'s Constant Conditional Correlation GARCH model. DCC-GARCH is a two stage estimator. In the first stage the univariate GARCH model is estimated for each time series. Second stage works with standardized residuals from the first step that are used to conditional correlations estimation.

Similar to RiskMetrics, let us denote r_t as a daily return such that

$$r_t|\mathcal{F}_{t-1} \sim N(0, H_t) \tag{7}$$

and

$$H_t = D_t R_t D_t \tag{8}$$

Matrices D_t and R_t are diagonal matrix of conditional time varying standard deviations and conditional correlation matrix respectively. D_t is $N \times N$ matrix with conditional standard deviations on the main diagonal and zeros elsewhere. Formally elements of D_t are defined as follows:

$$d_{ij,t} = h_{ij,t}^{-1/2} \text{ for } i = j$$

and

$$d_{ij,t} = 0 \text{ for } i \neq j,\ i,j = 1, \dots, N.$$

Conditional standard deviations h_{it} are obtained from univariate GARCH models

$$h_{i,t} = \omega_i + \sum_{p=1}^{P_i} \alpha_{i,p} r_{i,t-p}^2 + \sum_{q=1}^{Q_i} \beta_{i,q} h_{i,t-q}.$$

R_t is $N \times N$ time varying correlation matrix and has the form

$$R_t = Q_t^{*-1} Q_t Q_t^{*-1} \tag{9}$$

where

$$Q_t = \left(1 - \sum_{m=1}^{M} \alpha_m - \sum_{n=1}^{N} \beta_n\right) \bar{Q} + \sum_{m=1}^{M} \mathrm{A}_m \left(\varepsilon_{t-m} \varepsilon_{t-m}^T\right) + \sum_{n=1}^{N} \mathrm{B}_n\, Q_{t-n} \tag{10}$$

and

$$Q_t^* = \begin{pmatrix} q_{11}^{1/2} & 0 & \cdots & 0 \\ 0 & q_{22}^{1/2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & q_{kk}^{1/2} \end{pmatrix}. \tag{11}$$

$\bar{Q}$ in equation (10) is the unconditional covariance matrix of the standardized residuals from the first stage estimation, $\alpha_m \geq 0$, $\beta_m \geq 0$ and $\sum_{m=1}^{M} \alpha_m + \sum_{n=1}^{N} \beta_n < 1$ and the elements of the R_t are of the form $\rho_{ij,t} = \frac{q_{ij,t}}{\sqrt{q_{ii,t} q_{jj,t}}}$.

The way the model is formulated ensures that the variance-covariance matrix H_t is positive definite. Assumption of normally distributed returns allows us to use Maximum Likelihood estimator. If the assumption of normality is not valid we can still use Quasi-Maximum Likelihood estimator (*Engle (2002)*, *Engle & Sheppard (2001)*).

The log-likelihood can be expressed as

$$L = -\frac{1}{2}\sum_{t=1}^{T}(n\,log(2\pi) + log|H_t| + r_t^T H_t^{-1} r_t)$$

$$L = -\frac{1}{2}\sum_{t=1}^{T}(n\,log(2\pi) + log|D_t R_t D_t| + r_t^T D_t^{-1} R_t^{-1} D_t^{-1} r_t)$$

$$L = -\frac{1}{2}\sum_{t=1}^{T}(n\,log(2\pi) + 2log|D_t| + log|R_t| + \varepsilon_t^T R_t^{-1} \varepsilon_t) \quad (12)$$

$$L = -\frac{1}{2}\sum_{t=1}^{T}(n\,log(2\pi) + 2log|D_t| + r_t^T D_t^{-1} D_t^{-1} r_t - \varepsilon_t^T \varepsilon_t + log|R_t| + \varepsilon_t^T R_t^{-1} \varepsilon_t)$$

Last equation in (12) can be further decomposed to volatility and correlation part. We denote the parameters in D_t as θ and the additional parameters in R_t as ϕ. The volatility part consists of terms containing D_t and correlation part is composed of terms containing R_t. The decomposed log-likelihood has the following form

$$L(\theta, \phi) = L_V(\theta) + L_C(\theta, \phi). \quad (13)$$

Volatility part of the equation is

$$L_V(\theta) = -\frac{1}{2}\sum_{t=1}^{T}(n\,log(2\pi) + log|D_t|^2 + r_t^T D_t^{-2} r_t) \quad (14)$$

and could be rewritten as

$$L_V(\theta) = -\frac{1}{2}\sum_{t=1}^{T}\sum_{i=1}^{n}\left(log(2\pi) + log|h_{i,t}| + \frac{r_{i,t}^2}{h_{i,t}}\right). \quad (15)$$

The equation (15) is simply the sum of individual GARCH log-likelihoods. Generally it is possible to use any GARCH(p,q) process. For the sake of simplicity we will illustrate GARCH(1,1) which is the most widely used process. The conditional variance of GARCH(1,1) process is given by $h_{it}^2 = \omega_i + \alpha_i r_{it-1}^2 + \beta_i h_{it-1}^2$ where $\omega_i > 0; \alpha_i \geq 0;\ \beta_i \geq 0$ and $\alpha_i + \beta_i < 0$.

Our goal is to find

$$\hat{\theta} = \arg\max\{L_V(\theta)\} \tag{16}$$

and it can be done by jointly maximizing equation (15) by separately maximizing each N terms.

The correlation component of equation (13) is

$$L_C(\theta, \phi) = -\frac{1}{2}\sum_{t=1}^{T}(\,log|R_t| + \varepsilon_t^T R_t^{-1}\varepsilon_t - \varepsilon_t^T\varepsilon_t) \tag{17}$$

and can be maximized with respect to ϕ by using $\hat{\theta}$ from the first stage. Formally we are looking for a solution of maximization problem

$$\max_{\phi}\{L_C(\hat{\theta}, \phi)\}. \tag{18}$$

Engle & *Sheppard (2001)* formulated some reasonable regularity condition under which our maximum likelihood estimator is consistent and also asymptotically normally distributed.

Having all available information at time t, one-step ahead forecast in the DCC-GARCH framework can be obtained using following equations:

$$H_{t+1} = D_{t+1}R_{t+1}D_{t+1} \tag{19}$$

where elements of D_{t+1} are square roots of forecasts of univariate GARCH processes

$$h_{i,t+1} = \omega_i + \alpha_i r_{i,t}^2 + \beta_i h_{i,t} \tag{20}$$

and dynamics in R_{t+1} is described by equation

$$R_{t+1} = Q_{t+1}^{*-1}Q_{t+1}Q_{t+1}^{*-1} \tag{21}$$

with

$$Q_{t+1} = (1 - \alpha - \beta)\bar{Q} + \mathrm{A}(\varepsilon_t\varepsilon_t^T) + \mathrm{B}Q_t \tag{22}$$

RiskMetrics and DCC-GARCH are models that require daily data for estimation. However, technological progress in last decade enable us to collect data at higher than daily frequencies. Techniques of covariance matrix estimating and modeling with the use of high-frequency data are presented in next sections.

2.3. Realized Measures

Realized measures as such were first introduced in *Andersen et al. (2001)* where the whole new concept of using high frequency data for volatility calculation was used. Realized volatility can be characterized as a non-parametric model-free estimator. It can be easily computed as a sum of intraday squared returns. Despite the easy construction the theory behind is deep, based on quadratic variation. The main advantage of realized volatility is that it can be seen as observable. In contrast, parametric models such as GARCH treat the volatility as latent. Concept of the realized volatility was further developed in *Andersen et al. (2003)* and *Barndorff-Nielsen & Shephard (2004)* where we can find general framework for treating realized measures.

Similar to *Andersen et al. (2003)* we consider an *n*-dimensional price process defined on a complete probability space, $(\Omega, \mathcal{F}, P)$, evolving in continuous time over the interval $[0, T]$, where T denotes a positive integer. We further consider an information filtration, i.e., in increasing family of σ-fields, $(\mathcal{F}_t)_{t\in[0,T]} \subseteq \mathcal{F}$, which satisfies the usual conditions of P-completeness and right continuity. Finally, we assume that the asset prices through time t, including the relevant state variables, are included in the information set $\mathcal{F}_t$.

For the rest of the section we define continuously compounded return over $[t-h, t]$, $[t-h, t]$ denoting time interval such that $0 \leq h \leq t \leq T$, as the difference between log-prices at time t and $t-h$ as

$$r_{t,h} = p_t - p_{t-h} \tag{23}$$

Using previous equation, the cumulative return process from $t = 0$ to T can be written as

$$r_t \equiv r_{t,t} = p_t - p_0. \tag{24}$$

Let us turn to the concept of semi-martingales. Process Y is a semi-martingale if it can be decomposed into a drift term and a local martingale. We can rewrite it as $Y = A + M$, where A is a finite variation process (drift term) and M is a local martingale. If we want the above mentioned decomposition, the canonical decomposition, to be unique, we need to use special semi-martingales. According to *Back (1991)* we have to impose various weak regularity conditions to ensure semi-martingales to be special. The main characteristic of the special semi-martingales is that we assume finite variation process from the canonical decomposition to be predictable, hence the value of the predictable process at time t is known just before t. For more information on semi-martingales see *Protter (2004)* or *Delbaen & Schachermayer (1994)*.

Special semi-martingales are crucial for further work using quadratic variation theory. *Andersen et al. (2003)* use martingales to generally describe asset return process in Proposition 1. According to that proposition, price process can be decomposed into finite variation and predictable mean component and a local martingale, therefore price process can be seen as a semi-martingale process. In addition, both finite variation and local martingale components have continuous and jump part with jumps ensuring the price process to be arbitrage free in the following way. If there is a predictable jump in the price, which generates the arbitrage opportunity, the no-arbitrage condition is ensured by the simultaneous jump, large enough to overweight the jump in the price, in the local martingale in opposite direction.

Using definition of the return process and the Proposition 1 from *Andersen et al. (2003)* the return process is a special semi-martingale and can be further uniquely decomposed to predictable and integrable mean component and a local martingale. From the Proposition 2 in *Andersen et al. (2003)* and the fact that return process is a special semi-martingale, return process has a corresponding quadratic variation process. According to *Barndorff-Nielsen & Shephard (2004)* quadratic variation (QV) process can be defined as

$$[r,r]_t = plim_{M\to\infty} \sum_{j=0}^{M-1} \left[r_{t_{j+1}} - r_{t_j}\right]\left[r_{t_{j+1}} - r_{t_j}\right]', \qquad (25)$$

where $plim$ stands for probability limit. Definition of QV implies

$$RCOV_t \xrightarrow{p} [r,r]_t - [r,r]_{t-1} \qquad (26)$$

for all semi-martingales $[r,r]_t$ and for M going to infinity, that realized covariation consistently estimates increments of quadratic variation process. The more detailed description of relationship between quadratic variation and the conditional covariance matrix is in Theorem 1 in *Andersen et al. (2003)*. In short, the Theorem 1 states that conditional variance-covariance matrix of returns at time t, at given information set available at the same time, is equal to sum of expected value of difference in quadratic variation at time $t+h$ and t, variance of difference between finite variation components at time $t+h$ and t, and expected value of product of finite variation components at time $t+h$ and difference in local martingales from $t+h$ to t.

Following Corollary 1 in *Andersen et al. (2003)*, if we assume that the mean process is pre-determined and independent from the innovation process, conditional variance-covariance matrix of returns might be simplified to the following form

$$Cov(r_{t+h,t} | \mathcal{F}_t) = E([r,r]_{t+h} - [r,r]_t | \mathcal{F}_t). \tag{27}$$

The ex-post realized quadratic variation at time $t+h$ is also an unbiased estimator of conditional variance-covariance matrix of returns at time t. Finally the return distribution is described in Proposition 3 and Theorem 2 of *Andersen et al. (2003)*. For the daily returns it is easy to show that they are normally distributed with daily realized quadratic variation determining the distribution.

2.3.1. Construction of Realized Volatility and Covariation

The previous section describes theory which allows us to use realized variance and covariance. Now we are proceeding to a practical construction of realized volatility and covariation estimators.

First let us use definition from *Hautsch (2011)*. If we have n high-frequency intervals of length Δ, where $\Delta = \frac{1}{n}$, realized variance has the form[2]

$$RV^n := \sum_{k=1}^{n} \left(p_{k\Delta} - p_{(k-1)\Delta}\right)^2 := \sum_{k=1}^{n} r_{k\Delta,n}^2 .$$

Suppose now we have i asset and $t = 1, \dots, T$ days. Realized variance of asset i on day t is defined as

[2] Notation here is slightly different from that in *Hautsch (2011)*.

$$RV_{i,t} = \sum_{k=1}^{n} \left(p_{i,t-1+k\Delta} - p_{i,t-1+(k-1)\Delta}\right)^2 = \sum_{k=1}^{n} r_{i,t-1+k\Delta}^2 \quad , \tag{28}$$

where $p_{i,t-1+j\Delta} - p_{i,t-1+(j-1)\Delta}$ are intraday returns of day t. In a similar way we can construct realized volatility, realized covariance and realized correlation. Realized volatility of asset i at time t, i.e. $RVOL_{i,t}$, is constructed as a square root of realized variance

$$RVOL_{i,t} = \sqrt{RV_{i,t}} \, . \tag{29}$$

Realized covariation of asset i and j, i.e. $RCOV_{i,j,t}$, takes form

$$\begin{aligned} RCOV_{i,j,t} &= \sum_{k=1}^{n} \left(p_{i,t-1+k\Delta} - p_{i,t-1+(k-1)\Delta}\right)\left(p_{j,t-1+k\Delta} - p_{j,t-1+(k-1)\Delta}\right) = \\ &= \sum_{k=1}^{n} r_{i,t-1+k\Delta} r_{j,t-1+k\Delta} \end{aligned} \tag{30}$$

RCOV computation require synchronized data. In a real world we observe asynchronous trading activity for different assets, so some synchronization scheme such as Refresh time in *Barndorff-Nielsen et al. (2011)* have to be employed. Generally variance-covariance matrix can be of dimension $N \times N$. However, *Andersen et al. (2003)* pointed out that it is positive semi-definite as long as number of assets is lower or equal to number of intraday observations. In our case it is therefore of the following form

$$RCOV_t = \sum_{k=1}^{n} \boldsymbol{r}_{t-1+k\Delta} \boldsymbol{r}'_{t-1+k\Delta} \, , \tag{31}$$

where $\boldsymbol{r}_{t-1+k\Delta}$is $m \times 1$ vector composed of difference of price vectors $\boldsymbol{p}_{t-1+k\Delta}$and $\boldsymbol{p}_{t-1+(k-1)\Delta}$. Logarithmic price vector $\boldsymbol{p_t}$ is defined as $\boldsymbol{p_t} = \left(p_{1,t}, \dots, p_{m,t}\right)^T$.
Finally, realized correlation between asset i and j, $RC_{i,j,t}$, can be computed as a ratio of covariance between assets i and j, and product of volatilities of assets i and j. Mathematically speaking

$$RC_{i,j,t} = \frac{RCOV_{i,j,t}}{RVOL_{i,t} RVOL_{j,t}} \, . \tag{32}$$

Under some suitable conditions such as no market-microstructure noise or absence of jumps in prices realized volatility is consistent, unbiased and efficient volatility estimator. However real-world data rarely possess previous properties therefore positive semi-definiteness of covariance matrix estimates is not guaranteed. Recently, the new estimator, Multivariate Realized Kernels, that guarantees covariance matrix to be PSD was introduced in *Barndorff-Nielsen et al. (2011)*.

2.3.2. Multivariate Realized Kernels

The Multivariate Realized Kernels (MRK) estimator proposed in *Barndorff-Nielsen et al. (2011)* is the second type of covariance estimator we use in the thesis. The MRK estimates are guaranteed to be positive-semidefinite and the estimator can also deal with non-synchronous trading. If the high-frequency returns are synchronized the positive semi-definite MRK covariances are defined as follows

$$RCOV_t^{MRK} = \sum_{h=-n}^{n} k\left(\frac{h}{H}\right)\Gamma_h \,, \tag{33}$$

where Γ_h stands for h-th realized autocovariance and k is a non-stochastic weight function. Γ_h is defined in the following way

$$\Gamma_h = \sum_{\mathrm{j=j+1}}^{\mathrm{n}} x_j x_{j-h}^{'}, \tag{34}$$

for $h \geq 0$, and

$$\Gamma_h = \Gamma_{-h}^{'}, \tag{35}$$

for $h < 0$.

The kernel function we use for estimation is a Parzen function defined as

$$k(x) = \begin{cases} 1 - 6x^2 + 6x^3 & 0 \leq x \leq 1/2 \\ 2(1-x)^3 & 1/2 \leq x \leq 1. \\ 0 & x > 1 \end{cases} \tag{36}$$

Having described Realized Covariation and Multivariate Realized Kernels, techniques that are designed to obtain variance-covariance matrix estimates, we now turn to description of Heterogeneous Autoregression and Wishart Autoregression. These models use VCV matrix estimates for further modeling and forecasting.

2.4. Heterogeneous Autoregressive model

The Heterogeneous Autoregressive model (HAR) presented here was proposed by *Corsi (2009)*. It represents a new approach to volatility modeling using high-frequency data and realized volatility. The work of *Corsi (2009)* was inspired by Heterogeneous Market Hypothesis presented in *Müller et al. (1993)*. The Heterogeneous Market Hypothesis states that heterogeneity among investors arises due to the different investment strategies and time horizons. In heterogeneous market we can divide investors to a two categories according to their time horizons. Short term investors are represented by speculative traders, whose positions are likely to be closed at the end of day. Long term investors are represented by central banks, commercial organizations and pension funds which trade less frequently but in larger volumes. *Müller et al. (1993)* claim that volatility is empirically positively correlated with market presence and volume. On the contrary, in the homogeneous market, volatility is supposed to be negatively correlated with market presence and activity. Therefore different trading strategies of market participants in heterogeneous market create the volatility. *Corsi (2009)* suggests three primary volatility components: short term traders with trading horizon of one day, medium term investors with weekly trading frequency and long term investors who trade at monthly basis.

HAR is of the AR-type class of models. It tries to combine a short term and long term volatility together. Although it is not truly long memory model such as ARFIMA or FIGARCH, it is still able to capture long memory of volatility.

Now let us turn to the model. We consider the following standard continuous price process

$$dp(t) = \mu(t)dt + \sigma(t)dW(t)\,, \tag{37}$$

where $p(t)$ is a logarithmic price process, $\mu(t)$ is a cadlág[3] finite variation process, $W(t)$ is a Brownian motion and $\sigma(t)$is a stochastic process independent of $W(t)$. Integrated volatility associated with day t is defined as a square root of the integral of instantaneous variance over a one trading day $1d$,

$$\sigma_t^{(d)} = \sqrt{\int_{t-1d}^{t} \sigma^2(\omega)d(\omega)}\,. \tag{38}$$

Theory described in section 2.3.1 allows us to construct an estimator of realized volatility defined as

$$RVOL_t^{(d)} = \sqrt{\sum_{j=O}^{M-1} r_{t-j.\Delta}^2}\,, \tag{39}$$

where t denotes day, j stands for the time within day t, $\Delta= 1d/M$ is the number of observation during one day and $r_{t-j.\Delta} = p_{t-j.\Delta} - p_{t-(j-1).\Delta}$ represents intraday returns sampled at frequency Δ. $RVOL_t^{(d)}$ is referred as a daily volatility. For the HAR model we also need volatilities over longer time period. Namely we use weekly and monthly one. Weekly, $RVOL_t^{(w)}$, is the average of last five working days volatilities and monthly, $RVOL_t^{(m)}$, is the average of the volatilities of last 22 working days, i.e.

$$RVOL_t^{(w)} = \frac{1}{5}\left(RVOL_t^{(d)} + RVOL_{t-1d}^{(d)} + \cdots + RVOL_{t-4d}^{(d)}\right) \tag{40}$$

$$RVOL_t^{(m)} = \frac{1}{22}\left(RVOL_t^{(d)} + RVOL_{t-1d}^{(d)} + \cdots + RVOL_{t-21d}^{(d)}\right) \tag{41}$$

Next, we define the latent partial volatility $\tilde{\sigma}_t^{(\cdot)}$. In the market each participant is generating only certain volatility component. These should be divided into three categories: daily, weekly and monthly denoted as $\tilde{\sigma}_t^{(d)}$, $\tilde{\sigma}_t^{(w)}$ and $\tilde{\sigma}_t^{(m)}$ respectively. In addition, the highest frequency volatility component determines the market volatility, i.e. $\tilde{\sigma}_t^{(d)} = \sigma_t^{(d)}$. The daily and weekly components depend on past daily and weekly realized volatility respectively and the expectation of

[3] right continuous process with left limits

the next-period volatility of the longer-term. The monthly component depends only on its past realizations. The model is then characterized by three equations

$$\tilde{\sigma}_{t+1m}^{(m)} = c^{(m)} + \phi^{(m)} RVOL_t^{(m)} + \tilde{\omega}_{t+1m}^{(m)} \quad (42)$$

$$\tilde{\sigma}_{t+1w}^{(w)} = c^{(w)} + \phi^{(w)} RVOL_t^{(w)} + \gamma^{(w)} E_t\left[\tilde{\sigma}_{t+1m}^{(m)}\right] + \tilde{\omega}_{t+1w}^{(w)} \quad (43)$$

$$\tilde{\sigma}_{t+1d}^{(d)} = c^{(d)} + \phi^{(d)} RVOL_t^{(d)} + \gamma^{(d)} E_t\left[\tilde{\sigma}_{t+1w}^{(w)}\right] + \tilde{\omega}_{t+1d}^{(d)} \quad (44)$$

where $\tilde{\omega}_{t+1m}^{(d)}$, $\tilde{\omega}_{t+1m}^{(w)}$ and $\tilde{\omega}_{t+1m}^{(m)}$ contemporaneously and serially independent zero mean innovations. Moreover, to ensure positivity of partial volatilities error terms have appropriately truncated left tail.

By substituting equation (42) into the (43) and then transformed (43) into (44) and using fact that $\tilde{\sigma}_t^{(d)} = \sigma_t^{(d)}$ we get

$$\sigma_{t+1d}^{(d)} = c + \beta^{(d)} RVOL_t^{(d)} + \beta^{(w)} RVOL_t^{(w)} + \beta^{(m)} RVOL_t^{(m)} + \tilde{\omega}_{t+1d}^{(d)}, \quad (45)$$

Moreover, ex-post $\sigma_{t+1d}^{(d)}$ can be written as a sum of realized volatility and innovations term that count for latent daily volatility measurement and estimation error.

$$\sigma_{t+1d}^{(d)} = RVOL_t^{(d)} + \omega_{t+1d}^{(d)}, \quad (46)$$

Finally, substituting equation (46) into the equation (45) we arrive at

$$RVOL_{t+1d}^{(d)} = c + \beta^{(d)} RVOL_t^{(d)} + \beta^{(w)} RVOL_t^{(w)} + \beta^{(m)} RVOL_t^{(m)} + \omega_{t+1d}, \quad (47)$$

where $\omega_{t+1d} = \tilde{\omega}_{t+1d}^{(d)} - \omega_{t+1d}^{(d)}$. Equation (47) represents simple AR-type model of realized volatility which can be estimated using the ordinary least squares (OLS).

One-step ahead forecast using Heterogeneous Autoregression is computed as follows

$$RVOL_{t+1} = c + \beta^{(d)} RVOL_t^{(d)} + \beta^{(w)} RVOL_t^{(w)} + \beta^{(m)} RVOL_t^{(m)} \quad (48)$$

If the Heterogeneous Autoregressive model is applied on the elements of variance-covariance matrix estimates positive semi-definiteness of the covariance forecasts is not guaranteed. In order to ensure positivity of the VCV matrix forecasts we use Wishart Autoregressive model of *Gourieroux et al. (2009)*.

2.5. Wishart Autoregressive model

Gourieroux et al. (2009) introduced new model for modeling dynamics of realized covariance by conditional Wishart distribution. The Wishart Autoregressive model (WAR) is an interesting alternative to the classical multivariate volatility models as it reduces number of parameters[4] from $\left[\frac{n(n+1)}{2}\right]^2$ to $1+\frac{n(n+1)}{2}+n^2$ and it does not require any further restrictions to produce positive definite and symmetric covariance matrix.

Following *Gourieroux et al. (2009)* and *Bonato et al. (2009)*, let us denote realized covariance as Y_t. Y_t follows Wishart process of order 1, denote $W[K,M,\Sigma]$, if it satisfies

$$Y_t = \sum_{k=1}^{K} x_{k,t}x_{k,t}^{'} \ , \tag{49}$$

where processes $x_{k,t}$ for $k=1,\dots,K$ are independent Gaussian $VAR(1)$ processes of dimension n with K degrees of freedom. Moreover $x_{k,t}$ have the same autoregressive parameter matrix M and common innovation variance Σ

$$x_{k,t} = Mx_{k,t-1} + \varepsilon_{k,t} \qquad \varepsilon_{k,t} \sim N(0,\Sigma). \tag{50}$$

Proposition 2 from *Gourieroux et al. (2009)* describes first and second order conditional moments of $WAR(1)$ process. Useful property of this proposition is that we are able to combine equation (49) and (50) which yields

$$Y_t = MY_{t-1}M' + K\Sigma + \eta_t, \tag{51}$$

where η_t is matrix of heteroscedastic error terms with zero conditional mean. Important feature of the model is number of degrees of freedom. According to *Bonato (2009)* and *Chiriac* & *Voev (2011)* matrix Y_t is positive definite if and only if $K > n$. On the other hand if the K is smaller than n, process Y_t has a degenerate Wishart distribution and if the $K < n-1$ there is no density function defined for the variance-covairance distribution.

There are three parameters that need to be estimated in the WAR model:

- matrix of autoregressive parameters M

[4] see *Gourieroux et al. (2009)*, p.168

- symmetric, positive definite innovation covariance matrix Σ
- degrees of freedom K

Estimation procedure follows Proposition 9 from *Gourieroux et al. (2009)* in which if $K > n - 1$

(i) K and Σ are identifiable, and the autoregressive matrix M is identifiable up to its sign.

(ii) $\Sigma^* = \Sigma K$ is first-order identifiable and M is first-order identifiable up to its sign. The degree of freedom K and the innovation covariance matrix Σ are second-order identifiable.

The non-linear least squares using first-order conditional moments are used for the first order identification as they are equivalent to Method of Moments estimation. The estimator is given by

$$\left(\widehat{M}, \widehat{\Sigma}^*\right) = arg \min_{M,\Sigma^*} S^2\,(M, \Sigma^*), \tag{52}$$

where

$$\begin{aligned} S^2(M, \Sigma^*) &= \sum_{t=2}^{T} \sum_{i<j} \left(Y_{ij,t} - \sum_{k=1}^{n} \sum_{l=1}^{n} Y_{kl,t-1} m_{ik} m_{lj} - \sigma_{ij}^* \right)^2 \\ &= \sum_{t=2}^{T} \| vech(Y_t) - vech(M Y_{t-1} M' + \Sigma^*) \|^2 \end{aligned} \tag{53}$$

and $vech$ represents vector of stacked elements of lower/upper triangular matrix Y_t. Equation (52) can be estimated using any software that accounts for heteroscedasticity.

Identification of covariance matrix Σ and degrees of freedom K is obtained from second-order moments. In case of $WAR(1)$, the marginal distribution of the process is centered Wishart distribution. The portfolio's volatility conditional variance has the following form

$$V(\alpha' Y_t \alpha) = \frac{2}{K} [\alpha' \Sigma^*(\infty) \alpha]^2, \tag{54}$$

where

$$\Sigma^*(\infty) = K \Sigma(\infty) \tag{55}$$

and

$$\Sigma(\infty) = M\Sigma(\infty)M' + \Sigma. \tag{56}$$

A consistent estimator of degrees of freedom K is

$$\hat{K}(\alpha) = 2\frac{\left[\alpha'\hat{\Sigma}^*(\infty)\alpha\right]^2}{\hat{V}(\alpha' Y_t \alpha)}. \tag{57}$$

and consistent estimator of covariance matrix Σ is

$$\hat{\Sigma}(\alpha) = \frac{\hat{\Sigma}^*}{\hat{K}(\alpha)}. \tag{58}$$

Detailed estimation procedure can be found in *Gourieroux et al. (2009)* or *Bonato et al. (2009)*.

One-step ahead forecast in WAR framework can be computed according to following equation

$$Y_{t+1} = MY_tM' + K\Sigma. \tag{59}$$

Having defined models that are designed to produce variance-covariance matrix forecasts we now turn to forecast evaluation methods.

2.6. Evaluation of forecasts

Last section is dedicated to statistical and economic evaluation of variance-covariance matrix forecasts. Statistical evaluation employed in this thesis is based on the Root Mean Square Forecasting Error. According to perspective of portfolio optimization we employ three economic criteria. Namely they are Mean-Variance optimization, Global Minimum Variance Portfolio and Value-at-Risk.

2.6.1. Root Mean Square Forecasting Error

Root Mean Square Forecasting Error (RMSFE) as a statistical measure was selected due to the fact that it was used in *Chiriac* & *Voev (2011)* and *Voev (2009)* where they compared model based on daily as well as high-frequency data. We define RMSFE loss function as

$$RMSFE = \frac{1}{T-1}\sum_{t=1}^{T}\left\|\Sigma_{t+1} - \hat{\Sigma}_{t+1|t}\right\|_F, \tag{60}$$

where $\hat{\Sigma}_{t+1|t}$ is a forecast of VCV matrix, Σ_{t+1} is the true VCV matrix proxied by both Realized Covariation and Multivariate Realized Kernel estimates and $\left\|\Sigma_{t+1} - \hat{\Sigma}_{t+1|t}\right\|_F$ denotes a Frobenius norm. According to *Watkins (2002)* the Frobenius norm is defined as

$$\|A\|_F = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}\left|a_{ij}\right|^2}. \tag{61}$$

Having described statistical evaluation of covariance forecasts we now turn to economic evaluation criteria. We begin with Mean-Variance optimization approach followed by Global Minimum Variance Portfolio and Value-at-Risk approaches.

2.6.2. Mean-Variance optimization

Portfolio selection in the Mean-Variance optimization framework dates back to *Markowitz (1952)* and his work Portfolio Selection. We have two possibilities for selecting optimal portfolio. First one involves specifying expected portfolio return and finding optimal weights of portfolio assets so the volatility of the entire portfolio is minimized. In second option investor is trying to maximize expected return of portfolio at given volatility level.

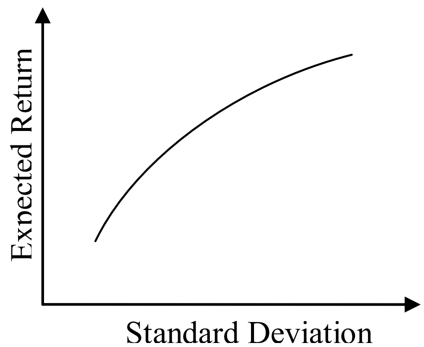
Let us consider a risk averse investor who wants to maximize his utility. Weights of portfolio assets, $w = (w_1, \dots, w_n)'$, maximizing utility of the investor can be found solving following problem

$$\begin{gathered} \min_{w_{t+1}} \; w'_{t+1}\hat{\Sigma}_{t+1|t}w_{t+1} \\ s.t. \;\; l'''w_{t+1} = 1 \\ w'_{t+1}\hat{\mu}_{t+1|t} = \mu_P \end{gathered} \tag{62}$$

where w_{t+1} is vector of assets weights, $\hat{\Sigma}_{t+1|t}$ represents a covariance matrix forecast, l denotes a vector of ones, $\hat{\mu}_{t+1|t}$ is a vector of mean forecasts and μ_P

stands for portfolio return. Detailed analytical procedure of finding such a portfolios is described in *Merton (1972)*. Once we solved equation (62) for different expected portfolio returns or volatility levels, we are able to construct mean-variance efficient frontier.

Figure 2-1: Efficient frontier



Source: Author

2.6.3. Global Minimum Variance Portfolio

Using the mean forecasts is the main drawback of the Mean-Variance optimization. Mean forecasts might be extremely noisy so the portfolio weights and variance can become notably sensitive to changes in assets mean. Solution to this problem is the Global Minimum Variance Portfolio (GMVP) optimization as it does not require use of mean forecasts. GMVP is specified as follows

$$\min_{w_{t+1}} w'_{t+1}\hat{\Sigma}_{t+1|t}w_{t+1} \qquad s.t. \;\; l''' w_{t+1} = 1 \tag{63}$$

According to *Kempf & Memmel (2006)* the solution of the previous minimization problem has the form

$$w_{t+1}^{GMV} = \frac{\hat{\Sigma}_{t+1|t}^{-1} l}{l'\hat{\Sigma}_{t+1|t}^{-1} l} \;, \tag{64}$$

and the expected return variance is calculated as

$$\sigma^{2}{}_{t+1}^{GMV} = w_{t+1}^{GMV\,\prime} \hat{\Sigma}_{t+1|t}\, w_{t+1}^{GMV} = \frac{1}{l' \hat{\Sigma}_{t+1|t}^{-1} l} \,. \qquad (65)$$

2.6.4. Value-at-Risk

Previous methods of evaluating covariance forecasts "only" rank the forecasts according to some criterion. Value-at-Risk approach has the advantage that it quantifies the risk as a single number representing maximum potential loss at given probability level.

Jorion (2007) defines Value-at-Risk as the *"worst loss over a target horizon such that there is a low, prespecified probability that the actual loss will be larger"*. Let the c be the confidence level and L be the loss of the portfolio. Then the VaR is defined as

$$P(L > VaR) \leq 1 - c \,. \qquad (66)$$

If we set confidence level to $c = 0.95$ or $c = 0.99$, then the probability of experiencing loss greater than VaR is less than 5 or 1 percent respectively. Three basic approaches can be used for Value-at-Risk calculation. According to *Jorion (2007)* VaR can be computed using

- *variance-covariance method* - VaR is calculated analytically using forecasts of variance-covariance matrix and assumption that returns are normally distributed
- *historical simulation* - from the historical data we construct return distribution in which the specific quantile represents the VaR
- *Monte Carlo simulation* - we generate N random return processes according to specific distribution. From the generated processes the quantile of the return process is selected and the VaR is calculated.

Our aim is to access the performance of different covariance forecasting models so we stick to the variance-covariance method of VaR calculation. The more detailed description of the method follows.

For variance-covariance method we use parametric approach with standard deviation of the portfolio as the input parameter. The most simple case assume that the returns are standard normally distributed. Having estimated portfolio's standard deviation, the VaR at confidence level c with corresponding quantile of standard normal distribution α calculated over t days with invested amount W is defined as

$$VaR = \alpha\sigma\sqrt{t}\,W. \tag{67}$$

The most frequent confidence levels or cut-off points are 0.95 and 0.99 with corresponding quantiles 1.645 and 2.326 respectively.

VaR as a risk measure is widely used in the banking sector. In the RiskMetrics framework we calculate one-day VaR with a cut-off point set to 0.95. Banking regulation standards introduced by Basel Committee proposed to calculate minimum capital requirements of the banks as 10-day VaR at 99% confidence level.

Backtesting of VaR

Normality of returns distribution might be too strong and restrictive leading to miscalculation of VaR. In order to minimize financial risks we need to test the VaR performance of different covariance forecasts. Importance of backtesting is best described in *Brown (2008)*: *"VaR is only as good as its backtest. When someone shows me a VaR number, I don't ask how it is computed, I ask to see the backtest. If I think I could make money betting either side at 100 to 1 on whether or not a break will occur tomorrow, I disregard the VaR. If anyone argues with me, I challenge them to take my bet over the next year."*

From the numerous backtesting procedures we employ test proposed by *Kupiec (1995)*. Basic idea of unconditional coverage Proportion of Failures (PoF) test is to compare the number of VaR exceedance with total numbers of VaR forecasts. If we denote number of VaR exceptions x and total number of forecasted VaRs N, the PoF ratio $\frac{x}{N}$ should not deviate much from $(1 - confidence\ level)$. To put it more formally we test

$$H_0: \quad p = p^* \tag{68}$$

where $p = (1 - c)$ and $p^* = \frac{x}{N}$. The null hypothesis is tested using likelihood ratio test with test statistic

$$LR = -2ln[(1-p^*)^{N-x}(p^*)^x] + 2ln\left[\left(1-\frac{x}{N}\right)^{N-x}\left(\frac{x}{N}\right)^x\right] \tag{69}$$

Under the H_0, the PoF test is χ^2 distributed with one degree of freedom. In *Kupiec (1995)* the non-rejection regions for different p^* and N can be found.

Chapter 3

Data

Having defined methodology, let us turn to data description. In our work we use daily and high-frequency data. By term high-frequency (HF) we mean data collected at 5-minute interval. To be precise, we use 5-minute closing prices. All data were obtained from Tick Data. Samples which we use cover period from July 1, 2003 until November 30, 2011.

Selection of the financial instruments was done on the basis of European investor´s perspective, whose portfolio includes not only world's most traded assets such as S&P 500 but also local one. The final portfolio consists of highly liquid commodity, "safe haven" investment and European asset. Namely, our portfolio contains Light Crude NYMEX, Gold COMEX and DAX index. To get a better picture of our dataset short description of assets follows.

Light Crude NYMEX (denoted with symbol CL) represents light crude oil futures contracts traded at the New York Mercantile Exchange. Crude oil is the most traded physical commodity worldwide with prices quoted in U.S. dollars (USD). CL contracts are highly liquid. One trade unit of CL is 1 000 barrels. The minimum price fluctuation is one cent per barrel that is 10 USD per contract. Our sample consists of 2503 trading days.

Gold COMEX (denoted with symbol GC) is traded at the metal division, formerly known as Commodity Exchange, of the New York Mercantile Exchange. Gold is considered as a secure investment by many investors, thus safe haven. From the historical point of view, gold preserves its value during politically and economically uncertain times and serves as inflationary hedge. GC contracts are traded in USD. The size of the contract is 100 troy ounces with the minimum price change of ten cents per troy ounce (10 USD per contract). Our sample covers 2485 trading days.

DAX index (denoted with symbol DA) is the representative of European market in our portfolio. DAX is the German stock index, that consists of 30 German most actively traded companies at Frankfurt Stock Exchange. Its stock volume represents approximately 3/4 of German listed stocks. DAX was chosen as a representative of European market due to the prominent role of German economy within Europe. The prices of DAX are quoted in Euro (EUR). Number of days when DAX was traded within our sample is 2148.

3.1. Data processing

Our goal is to examine and model mutual dependencies among chosen assets. Models defined in previous chapter require synchronized dataset for proper estimation - assets have to be traded at the same time. By this constraint several problems arise. First one emerges from different trading hours. Final dataset has to contain only those trading days during which all assets were traded in case of daily data. For HF data, we keep closing prices with the same timestamp depending on the day, hour and minute. Another problem arises from the goal of the thesis to compare forecasts based on daily and HF data. To avoid this problem, days used for estimation must be the same in both datasets. Generally, data are synchronized using following procedure

1. only observation with the same date for all time series are kept
2. in case of HF data we repeat procedure from the first step also for values of hours and minutes

   2.1. we keep only days with more than 69 observations, resulting in at least 6.5 hours of trading

3. once 5-minute data are synchronized we delete those days from daily data that are not included in the final high-frequency dataset.

To illustrate the cleaning procedure, let us explain it using a simple example, supposing we have 3 trading days - November 1 to November 3, 2011. On November 1 we are left with 100 closing prices, November 2 contains 65 observations and on November 3 assets were traded 90 times. On November 2 there were less than 70 trades so we do not use this day for estimation although open and close prices are available and daily returns are possible to calculate. Final dataset would thus contain only days November 1 and November 3, 2011.
1930 days left once the data processing procedure is applied. The more detailed description of the time-series follows.

3.2. High-frequency data

Theoretically, the more data we have, the more information we are able to extract and thus our estimates should be more precise. The dramatic increase of computer performance in the last decade enables us to collect data for all executed trades. However, using tick data for estimation is not effective due to complications arising from their nature. The raw data might be contaminated by market microstructure noise due to bid-ask bounce or might suffer from asynchronous trading. To overcome problems, proper filters and sampling frequency must be chosen. As mentioned at the beginning of the chapter, we use 5-minute closing prices similar to *Andersen et al. (2005)* or *Chiriac* & *Voev (2011)*. Having briefly described concept of using HF data, let us summarize main characteristics of time-series.

Our original dataset covers the period from July 1, 2003 until November 30, 2011. In total we have 2 503 trading days with 541 721 closing prices for CL, 2 485 trading days with 501 660 closing prices for GC and 2 148 trading days with 221 203 closing prices for DA. Substantial increase in trading activity for CL and GC occurred during sample period. The average number of observations per day increases from 176 in 2003, 231 in 2005, to 279 in both 2010 and 2011 in case of crude oil. For gold, observations increased from 185 in 2003, almost 200 in 2005 to 279 in 2010 and 267 in 2011. Since 2007 market for crude oil and gold became real 24/7 global market with only one hour of no trading activity during trading day. Trading hours and number of observation per day remains almost constant for DAX. There are 102 observations per day during period 2003 to 2006. Two more observations occurred in 2007. This increase remained unchanged also in the following years.

By the process of synchronization 573, 555 and 218 days were removed from CL, GC and DA time series respectively. The total number of synchronized closing prices included in our sample is 184 699. Elimination of more than 350 000 data points in case of CL, almost 317 000 and 37 000 in case of GC and DA led to higher overnight returns. The mean of overnight returns is highest for crude oil, -0.0735% (DA: -0.00754%; GC: 0.0409%) with highest positive, 17.98% (DA: 12.99%; GC: 10.57%) and negative, -16.73% (DA: -13.29%; GC: -10.59%) return.

The basic characteristics of the synchronized HF dataset is summarized in the Table 3-1.

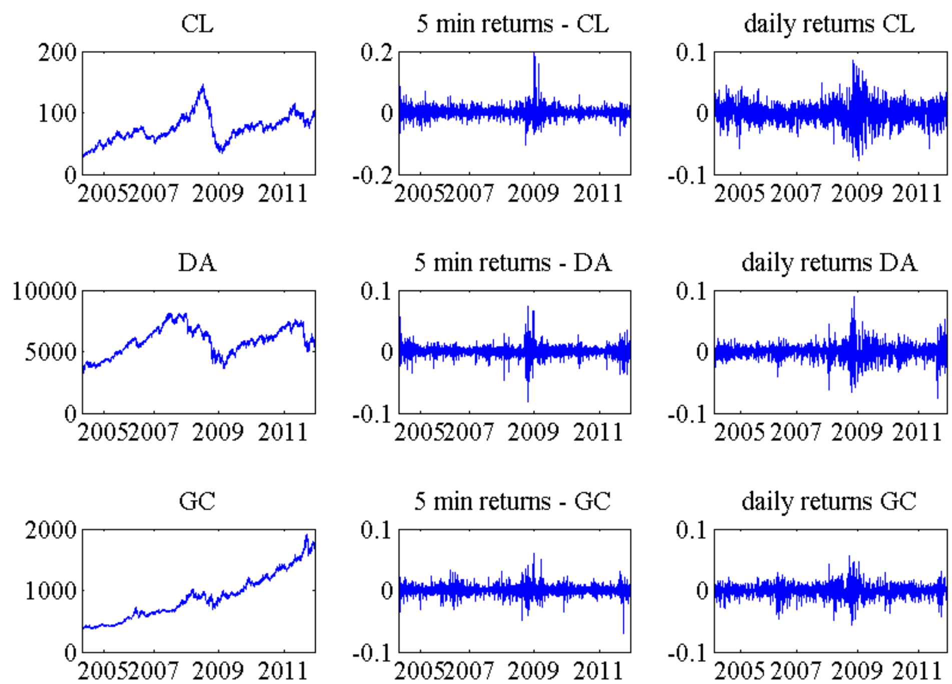
Table 3-1: Descriptive statistics of high-frequency closing prices and returns

	5 min prices			*daily returns*		
	CL	DA	GC	CL	DA	GC
Mean	73.7	5877.1	888.9	-0.0013	-3.5402e-04	-2.7987e-04
Min	26.8	3207.7	342.1	-0.0781	-0.0766	-0.0556
Max	145.1	8143.7	1913.5	0.086	0.0894	0.0559
Std. dev.	22.3	1169.7	372.7	0.0164	0.0119	0.0102
Kurtosis	3.3037	2.1089	2.5686	5.6011	9.4141	6.725
Skeweness	0.5142	-0.1259	0.6152	-0.081	-0.199	-0.3191

Source: Author's computation

Figure 3-1 shows final closing prices series with corresponding returns. The effects of financial crisis are mostly evident in CL and DA. There was a huge drop in value of both assets in 2008 and at the beginning of 2009. Financial crisis did not affect price of gold as much as in case of crude oil and DAX index. During 2008 and first quarter of 2009 price of gold more or less stagnated. In the following period until the end of our sample price almost doubled. If we concentrate more on return series, period of high volatility is clearly observable for all the time series during financial crisis 2008/2009.

Figure 3-1: 5-minute closing prices, 5-minute returns and daily high-frequency returns



Source: Author's computation

3.3. Daily data

Daily data are divided into two groups. In the first group, returns are calculated as a logarithmic difference between closing price at time t and $t-1$. Such data are referred as "close-close" or "CC" in the rest of the work. Second group of daily data consists of returns computed as logarithmic difference between closing price and opening price at time t. These data will be denoted as "open-close" or "OC".

Original and synchronized datasets consist of identical number of days as in case of HF data. Data synchronization procedure significantly impacts values of overnight returns. The highest mean of overnight returns is recorded for DA, 0.0587% (CL: 0.0294%; GC: 0.0409%). Values of maximum positive returns for CL, DA and GC equal 19.24%, 6.78% and 5.73% respectively. The maximum negative returns are -7.62%, -9.11% and -3.51% for CL, DA and GC respectively.

Descriptive statistics of CC and OC prices and returns are in Table 3-2 and Table 3-3.

Table 3-2: Descriptive statistics of close-close prices and returns

	prices CC			*daily returns CC*		
	CL	DA	GC	CL	DA	GC
Mean	72.1	5789.8	859.5	6.1765e-04	2.8616e-04	8.3343e-04
Min	27.2	3235.1	344.4	-0.1294	-0.0809	-0.0764
Max	145.5	8105.7	1902.8	0.1902	0.1187	0.1052
Std. dev.	22.5	1193.1	375.9	0.025	0.0153	0.0134
Kurtosis	3.2833	2.0465	2.6239	6.7504	10.1059	7.3386
Skeweness	0.5353	-0.0603	0.6851	0.1166	0.1449	-0.1919

Source: Author's computation

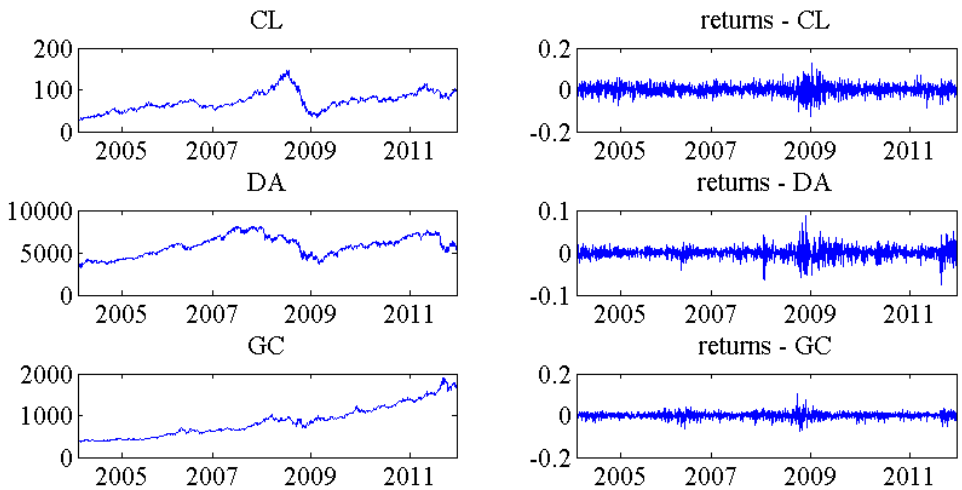
Table 3-3: Descriptive statistics of open-close prices and returns

	prices OC			*daily returns OC*		
	CL	DA	GC	CL	DA	GC
Mean	72.2	5790.6	859.4	3.2695e-04	-3.0130 e-04	4.1722e-04
Min	26.9	3235.1	344.4	-0.1263	-0.0756	-0.0766
Max	145.4	8137.7	1912	0.1269	0.0879	0.1025
Std. dev.	22.5	1192.7	375.8	0.0235	0.012	0.013
Kurtosis	3.2827	2.0439	2.6257	5.0705	8.8081	7.8462
Skeweness	0.5343	-0.0599	0.6847	-0.1863	-0.2023	-0.2755

Source: Author's computation

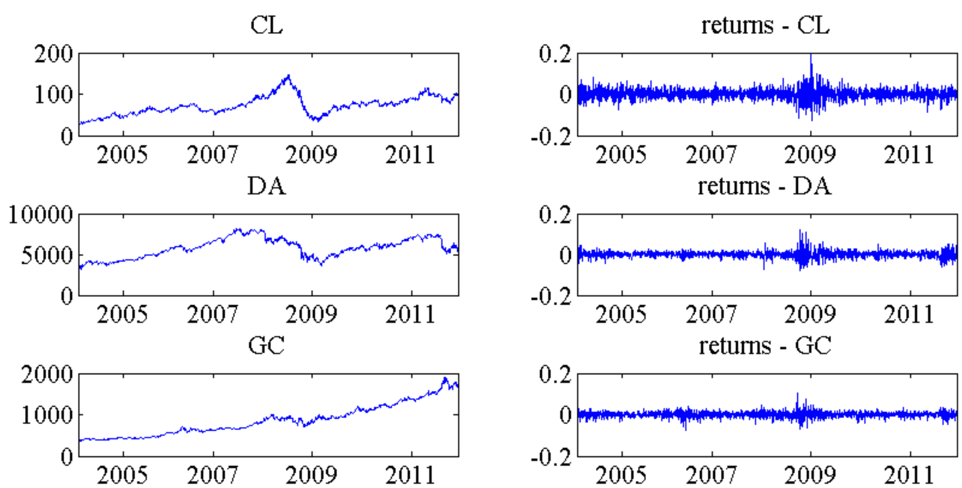
The Figure 3-2 and Figure 3-3 show us prices and corresponding returns for CC and OC data respectively. As in the case of HF data, the significant decrease in the value can be observed for CL and DA during 2008 and beginning of 2009. Price of gold follows its increasing trend throughout the whole dataset. Return series of all assets indicates period of particularly high volatility during second half of 2008 and first half of 2009.

Figure 3-2: daily closing prices and close-close returns



Source: Author's computation

Figure 3-3: open-close prices and open-close returns



Source: Author's computation

Chapter 4

Empirical Findings

In this chapter we present forecasting results obtained from models defined in Chapter 2. To study effects of financial crisis, the models are estimated on two sub-samples, representing period before crisis and during crisis, and full sample, covering period July 8, 2003 to November 29, 2011. The sub-samples are obtained by dividing whole dataset into two equal parts. Period "before crisis" represents 965 days (July 8, 2003 - February 8, 2008). "During crisis" sample starts on February 11, 2008 with last observation on November 29, 2011. Each period is further divided into in-sample and out-of-sample part. In-sample period lasts 713 days for all sub-samples. On the other hand out-of-sample period lasts 252 days which represents one year for before crisis and during crisis sub-sample. In case of full dataset, duration of out-of-sample period is 1 217 days.

The rolling window estimator of length 713 days is used for analysis. We describe rolling window estimation using data from before crisis period. First estimate is obtained using 713 daily observations with first observation at day one (period 1^{st} to 713^{rd} day). Using parameters from the fitted model, we calculate one-step ahead forecast for the day 714. Second estimate and forecast are obtained similarly, but the estimation window is moved by one day to 2^{nd} - 714^{th} observation. The one-step ahead forecast for the day 715 is calculated using estimates from second period of rolling window estimation. Same procedure is applied for 3^{rd} to 251^{st} estimation. Finally we are left with 252 fitted models and 252 forecasts.

Models used in our work are divided into two groups. First group includes models that use returns as input parameters. These models are DCC-GARCH and EWMA with decay factor set to RiskMetrics standards (in the remaining part of the work the model is denoted as RiskMetrics). In second group, four models that use variance-covariance matrix as input parameter are used. First one, HAR model, estimates each element of VCV matrix separately. Second model, Cholesky-HAR, is HAR applied on the elements of Cholesky decomposed VCV matrix. Positive definiteness of forecasts in case of Cholesky HAR is guaranteed by following transformation:

1. original VCV matrix is Cholesky decomposed, $VCV = UU'$, where U is lower triangular matrix
2. standard HAR is applied on elements of U
3. estimates from second step are used for one-step ahead forecast
4. "reverse" Cholesky decomposition on matrix of forecasts, U^F, is applied. The outcome is VCV matrix of forecasts $U^F U^{F'} = VCV^F$

Last two models are WAR(1) and diagonal WAR(1). The diagonal WAR(1) model is restricted version of WAR(1), where the matrix of autoregressive parameters M is diagonal. Data processing, models estimation, forecasting and evaluation of forecasts were done in MATLAB, Version 7.12.0.0635 (R2011a). Estimation of DCC-GARCH was carried out using UCSD GARCH Toolbox[5], RiskMetrics and Multivariate Realized Kernels estimates were obtained using Oxford MFE Toolbox[6]. Both toolboxes are developed by Kevin Sheppard. WAR(1) and diagonal WAR(1) models are estimated using codes written by Matteo Bonato for article Risk spillovers in international equity portfolios by *Bonato et al. (2012)*. Parameter estimates from individual models are presented in the Appendix C.

In sections 4.1 - 4.3 we present results of forecast evaluations for before crisis, during crisis and full sample period respectively.

---

[5] http://www.kevinsheppard.com/wiki/UCSD_GARCH
[6] http://www.kevinsheppard.com/wiki/MFE_Toolbox

4.1. Before crisis

This section summarizes forecasts evaluation within before crisis period. The in-sample period of before crisis sub-sample starts on July 8, 2003 and lasts until February 6, 2007. On the other hand the out-of-sample period starts on February 7, 2007 and lasts until February 8, 2008, which covers 252 days. Four methods are used for evaluation of covariance forecasts. First one Root Mean Square Forecasting Error is followed by Mean-Variance optimization and Global Minimum Variance Portfolio and at the end of the section Value-at-Risk estimates are presented.

4.1.1. Root Mean Square Forecasting Error

In Table 4-1 values of RMSFE for individual models are presented

Table 4-1: before crisis RMSFE (values time 10^{-4})

RMSFE	RCOV			MRK			RCOV	MRK
before crisis	CC	OC	HF	CC	OC	HF		
DCC-GARCH	2.5474	2.4711	1.1905	2.7628	2.6953	1.4066		
RiskMetrics	3.3411	2.9181	1.5386	3.4441	3.0336	1.6219		
HAR							1.1747	1.3632
Cholesky HAR							1.0634	1.2936
WAR(1)							6.1726	6.4110
diagonal WAR(1)							3.5965	3.7032

Note: CC denotes daily returns calculated from daily closing prices, OC denotes daily returns calculated from daily open-close prices and HF denotes daily returns calculated from 5-minute closing prices. Proxy for RMSFE calculation are RCOV and MRK. RCOV and MRK stands for Realized Covariation and Multivariate Realized Kernels respectively.

Source: Author's computation

First part of the table (DCC and RiskMetrics) presents "return as input" models results, second part (HAR - Diagonal WAR) show results of models using VCV matrix as input parameter. RMSFE for DCC-GARCH and RiskMetrics are calculated using both Realized Covariation and Multivariate Realized Kernels covariance matrix estimates as we do not know which one is the true VCV matrix. Moreover models are evaluated separately for daily data returns, CC and OC, and high frequency returns, HF. In case of HAR and WAR models, Realized Covariation estimates are used for RMSFE calculation only when they are also used for model estimation. In case of Multivariate Realized Kernels the same logic is applied.

Concentrating on the first part of the Table 4-1 the best forecasting performance is achieved by DCC-GARCH followed by RiskMetrics both estimated on HF returns. Results are similar using RCOV and also MRK as a proxy for RMSFE calculation. Comparing close-close and open-close returns, the OC returns slightly dominates. Ranking of the models according to returns calculated from different datasets is in line with expected outcome. Once RCOV and MRK, estimated on the restricted dataset is set as a proxy for RMSFE calculation, models using returns from restricted dataset show the best performance. Because dataset of OC returns is much more similar to HF data, their fit compared to CC returns is better. Due to the fact that the overnight trading returns included in close-close data are not included in RCOV and MRK, fit of CC returns is the worst one.

In the second part of the table, models using RCOV estimates of covariance matrix show better results compared to MRK. There is a clear dominance of HAR based model over WAR models for both RCOV and MRK. The best performance among all the models including DCC-GARCH and RiskMetrics is achieved by Heterogeneous Autoregression estimated on the Cholesky decomposed covariance matrix. Poor forecasting power of WAR and diagonal WAR is quite surprising as both models are designed to model covariance matrix directly.

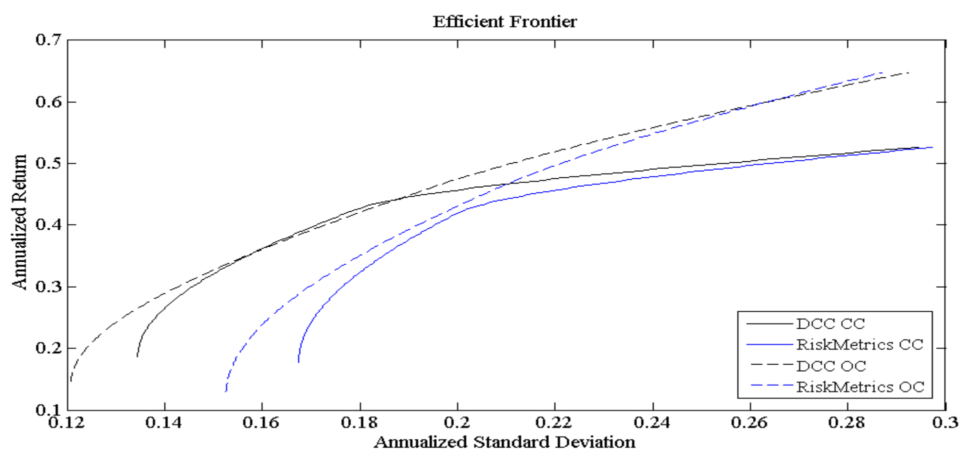
Overall, models using RCOV for estimation purposes and as a proxy for RMSFE calculation left us with better results than MRK alternative. The fit of model with the poorest performance, WAR(1) using MRK estimates, was six times worse than the fit of Cholesky HAR estimated on covariance matrix obtained from RCOV. In addition, WAR models were outperformed by RiskMetrics, which is the simplest model.

4.1.2. Mean-Variance optimization

This section summarizes the risk-return tradeoff of individual models. Results are represented by efficient frontiers shown in following figures. Individual figures are presented for daily and high-frequency datasets. Results are separated due to considerable difference in annualized return level.

Figure 4-1 depicts efficient frontiers of models that use daily data for estimation. As shown, better results are obtained using OC returns. With risk level under 20%, DCC-GARCH outperforms RiskMetrics models. With the risk increase, performance of RiskMetrics estimated on OC returns increases as well. Above 26% risk level, RiskMetrics using OC returns becomes the model with the best risk-return tradeoff.

Figure 4-1: before crisis efficient frontiers - daily data

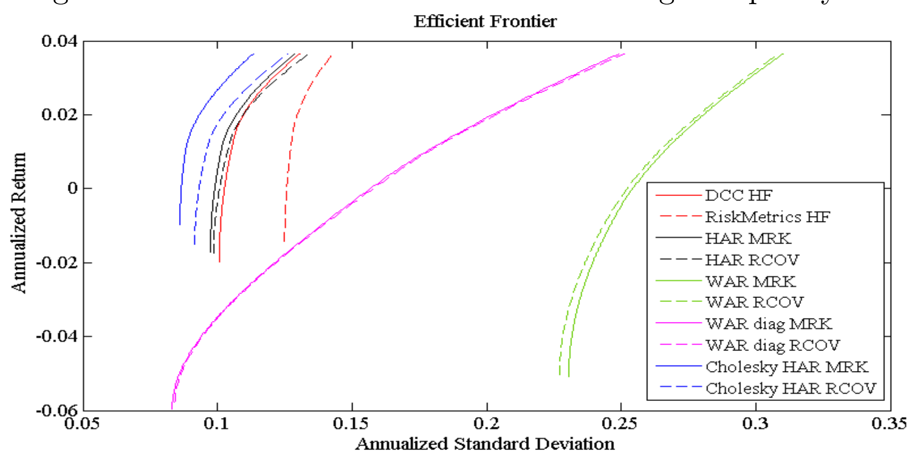


Source: Author's computation

In the Figure 4-2 efficient frontiers of models estimated on high-frequency data are presented. At first sight, bad performance of WAR based models is obvious. For WAR(1) models the lowest risk level achieved is almost twice as high as the maximum risk of simple RiskMetrics model. In case of diagonal WARs the lowest standard deviation is smallest among all the models, however this is accompanied by significantly lower and negative returns. Major drawback of the diagonal WAR models is faster risk increase compared to slower growth of returns. On the other hand HAR based models together with DCC-GARCH and

RiskMetrics show similar paterns of risk-return tradeoff. There is one case when the performance of return based, DCC-GARCH, is better than covariance based HAR model. Cholesky-HAR using MRK estimates is the best mean-variance performance model followed by Cholesky-HAR estimated on RCOV covariances. Multivariate Realized Kernels covariance approximation generally results in better performance of the models.

Figure 4-2: before crisis efficient frontiers - high-frequency data



Source: Author's computation

4.1.3. Global Minimum Variance Portfolio

Forecasting ability of our models from the perspective of finding portfolio with the lowest possible variance is summarized in Table 4-2

Table 4-2: before crisis GMVP (values times 10^{-5})

GMVP before crisis	CC		OC		HF		RCOV		MRK	
	mean	std. dev.	mean	std. dev.	mean	std. dev.	mean	std. dev.	mean	std. dev.
DCC-GARCH	6.41	3.03	4.71	2.95	3.44	1.86				
RiskMetrics	10.0	6.09	8.02	5.88	5.43	4.47				
HAR							3.18	1.70	3.19	1.83
Cholesky HAR							2.86	1.57	2.58	1.50
WAR(1)							20.5	23.6	21.1	28.2

diagonal WAR(1)	2.27	1.97	2.17	2.22

Note: CC denotes daily returns calculated from daily closing prices, OC denotes daily returns calculated from daily open-close prices and HF denotes daily returns calculated from 5-minute closing prices. RCOV and MRK stands for Realized Covariation and Multivariate Realized Kernels respectively.

Source: Author's computation

In Table 4-2 the mean of forecasted portfolio variances with corresponding standard deviations are presented (for the sake of simplicity we use "variance" instead of "mean of variances" in the rest of the section). The DCC-GARCH and RiskMetrics have the lowest variances in case returns calculated from HF dataset are used. Comparing results for HF and CC returns the variance is almost twice as high for CC one. As expected the DCC-GARCH performs better than RiskMetrics except for the case when we directly compare results from CC and HF data. In that particular case the variance is lower for RiskMetrics, however it is compensated by the higher volatility represented by standard deviation.

If we concentrate on models using RCOV and MRK, it is not obvious, which covariance estimates provide us with better results. Variance of HAR and WAR(1) forecasts is lower using RCOV while it is more beneficial for Cholesky-HAR and diagonal WAR(1) to use MRK. Diagonal WAR(1) estimated on MRK covariance is the model with the lowest variance. In contrast, variance of the worst model, WAR(1) using MRK estimates, is almost ten times higher. It is hard to choose the best model according to GMVP criteria because lower variance of diagonal WAR models is compensated with higher volatility. On the other hand lower volatility in case of Cholesky-HAR models is accompanied by slightly higher variance.

Overall, from the GMVP point of view models that use covariance matrices for estimation, except WAR(1), outperform return based models. There is no clear winner between RCOV and MRK. Results of DCC-GARCH estimated on high-frequency returns do not deviate much from covariance based models.

4.1.4. Value-at-Risk

Performance of covariance forecasts according to Value-at-Risk is summarized in Table 4-3. Global minimum variances are used for VaR calculation. Results are presented for both 95% and 99% VaR.

Table 4-3: before crisis Value-at-Risk

Value-at-Risk	CC		OC		HF		RCOV		MRK	
before crisis	N	p=p*	N	p=p*	N	p=p*	N	p=p*	N	p=p*
p*=0.05										
DCC-GARCH	19	accept	19	accept	18	accept				
RiskMetrics	12	accept	13	accept	14	accept				
HAR							23	reject	24	reject
Cholesky HAR							25	reject	28	reject
WAR(1)							0	reject	0	reject
diagonal WAR(1)							24	reject	25	reject
p*=0.01										
DCC-GARCH	8	reject	11	reject	7	accept				
RiskMetrics	4	accept	5	accept	5	accept				
HAR							8	reject	9	reject
Cholesky HAR							13	reject	15	reject
WAR(1)							0	accept	0	accept
diagonal WAR(1)							14	reject	10	reject

Note: "**N**" stands for "number of exceedance", p=p* represents null hypothesis of Proportion of Failures test, CC denotes daily returns calculated from daily closing prices, OC denotes daily returns calculated from daily open-close prices and HF denotes daily returns calculated from 5-minute closing prices. RCOV and MRK stands for Realized Covariation and Multivariate Realized Kernels respectively.

Source: Author's computation

Let us concentrate on 95% VaR (p*=0.05) first. For DCC-GARCH and RiskMetrics evaluation of market risk is appropriate. All tested specifications do not reject null hypothesis of properly specified risk level. On the other hand covariance based models reject null hypothesis thus the market risk is not properly evaluated. For HAR, Cholesky-HAR and diagonal WAR(1) the risk is underestimated while in case of WAR(1) we have overestimated risk.

Underestimation of the risk might lead to potential huge losses. On the other hand opportunity costs are the outcome of risk overestimation.

If we turn to 99% VaR (p*=0.01) situation changes for both model groups. In the first group, estimates of DCC-GARCH using CC and OC returns underestimate risk. In the second group difference is in case of WAR(1). Number of exceedance for that particular model is zero so one has to be really careful interpreting the result. From the definition of VaR, equation (67), we know that VaR depends heavily on variance. For the sample small enough (number of days lower than 611[7] for 99% VaR), with the increasing variance, the number of exceedance is approaching zero. If the variance is high enough we would never reject null hypothesis about properly specified risk level.

Now we turn to description of during crisis sub-period.

4.2. During crisis

In this section, we present results of covariance forecasts evaluation using during crisis period data. The in-sample period of our dataset includes 713 trading days, which starts on February 11, 2008 and lasts until December 6, 2010. The out-of-sample period covers 252 days that lasts from December 7, 2010 until November 29, 2011. Similar to before crisis period, order of evaluation methods is as follows: Root Mean Square Forecasting Error, Mean-Variance optimization, Global Minimum Variance Portfolio and Value-at-Risk.

4.2.1. Root Mean Square Forecasting Error

Table 4-4 presents fits of the models according to RMSFE criteria for during crisis period.

Table 4-4: during crisis RMSFE (values time 10^{-4})

RMSFE	RCOV			MRK			RCOV	MRK
during crisis	CC	OC	HF	CC	OC	HF		
DCC-GARCH	3.8706	3.6231	2.1759	4.0782	3.8609	2.3849		

[7] author's computation

RiskMetrics	4.4926	4.0484	2.3057	4.6523	4.2534	2.5020		
HAR							2.0548	2.3343
Cholesky HAR							1.9758	2.2609
WAR(1)							11	11
diagonal WAR(1)							4.7459	4.8682

Note: CC denotes daily returns calculated from daily closing prices, OC denotes daily returns calculated from daily open-close prices and HF denotes daily returns calculated from 5-minute closing prices. Proxy for RMSFE calculation are RCOV and MRK. RCOV and MRK stands for Realized Covariation and Multivariate Realized Kernels respectively.

Source: Author's computation

Organization of the Table 4-4 is the same as in before crisis period. Performance of DCC-GARCH and RiskMetrics is the best when HF returns for both proxies are used. Within RCOV and MRK groups open-close returns show the second best performance. The performance of close-close returns is the worst one. RCOV as a proxy for RMSFE calculation ensures generally better fit of the models. The RiskMetrics performance using HF returns within RCOV group is surprisingly better than DCC-GARCH performance using MRK as a proxy.

Covariance based models show better forecasting performance using RCOV. The best fit is presented by Cholesky-HAR followed by HAR model. The diagonal and full WAR show poor performance exactly as in the before crisis period. Order of the models is similar for RCOV and MRK: Cholesky-HAR followed by HAR, diagonal WAR(1) and WAR(1) at the end.

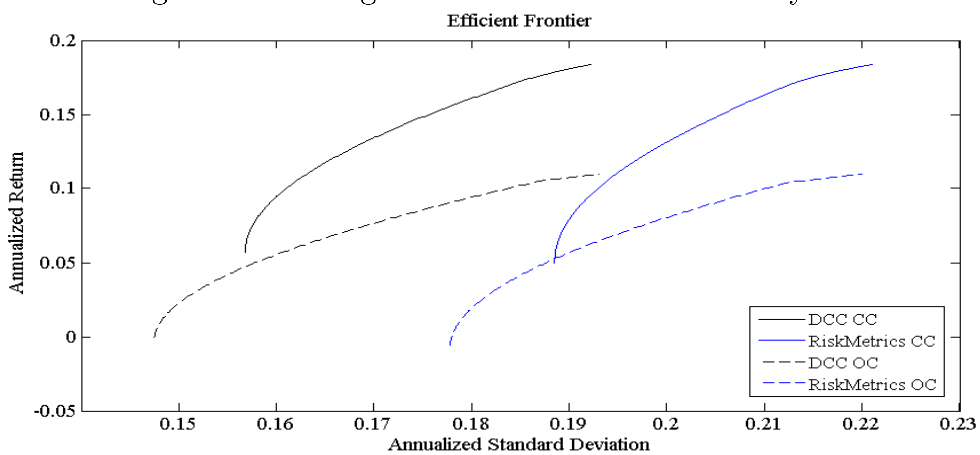
Cholesky-HAR estimated on RCOV has the best fit among all the models used. HAR using RCOV estimates is on the second place. Third best fit achieved by DCC-GARCH using HF returns unexpectedly outperforms HAR models estimated on MRK covariance matrices. Even more surprising is the result of RiskMetrics using RCOV as proxy estimated on HF returns because it outperforms HAR model estimated on MRK. RiskMetrics also outperformed both full and diagonal WAR.

4.2.2. Mean-Variance optimization

Efficient frontiers constructed within the framework of mean-variance optimization using during crisis period forecasts are presented in this section. Similar to before crisis section the results are presented by individual figures for each dataset.

In the Figure 4-3 efficient frontiers of the models using daily data for estimation are presented. Models that use open-close returns are generally less risky but also less profitable. For DCC-GARCH and RiskMetrics better risk-return tradeoff is achieved by the first one.

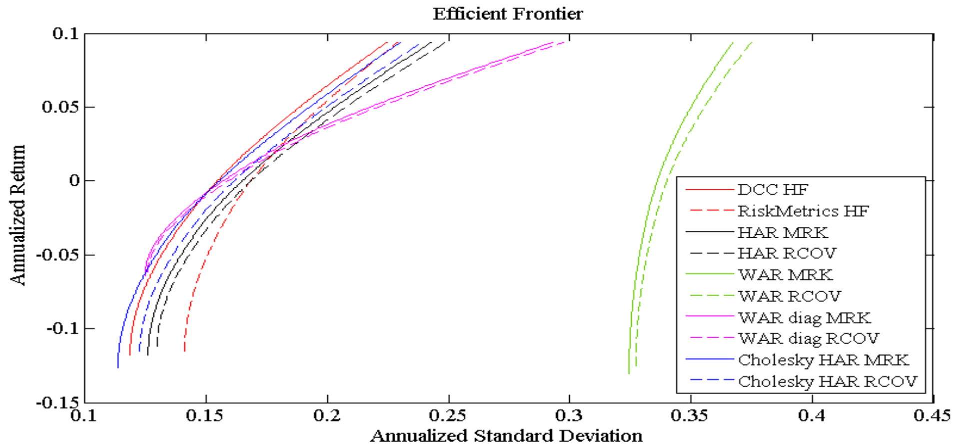
Figure 4-3: during crisis efficient frontiers - daily data



Source: Author's computation

Figure 4-4 displays efficient frontiers of models that use high-frequency data for estimation. Similar to before crisis period, WAR(1) are models with the worst performance. For the rest of the models we concentrate on parts of the efficient frontiers where the returns are positive. Surprisingly, the best mean-variance tradeoff is recorded for DCC-GARCH. Cholesky-HAR using MRK estimates is on the second place. Cholesky-HAR is almost outperformed by RiskMetrics at the higher risk level and that is probably the most surprising result. The score of RiskMetrics for positive returns is also better than HAR and WAR.

Figure 4-4: during crisis efficient frontiers - high-frequency data



Source: Author's computation

4.2.3. Global Minimum Variance Portfolio

Mean and standard deviation of forecasted variances based on financial crisis data are presented in the following table. "Mean of variances" are denoted as "variance" in this section.

Table 4-5: during crisis GMVP (values times 10^{-5})

GMVP during crisis	CC		OC		HF		RCOV		MRK	
	mean	std. dev.	mean	std. dev.	mean	std. dev.	mean	std. dev.	mean	std. dev.
DCC-GARCH	8.78	5.53	7.65	5.59	4.79	3.23				
RiskMetrics	12.9	10.8	11.6	10.1	7.02	6.37				
HAR							6.05	4.60	5.65	4.12
Cholesky HAR							5.31	4.39	4.52	3.50
WAR(1)							42.5	42.1	41.8	43.9
diagonal WAR(1)							5.05	4.17	4.86	4.22

Note: CC denotes daily returns calculated from daily closing prices, OC denotes daily returns calculated from daily open-close prices and HF denotes daily returns calculated from 5-minute closing prices. RCOV and MRK stands for Realized Covariation and Multivariate Realized Kernels respectively.

Source: Author's computation

The lowest variance is achieved by DCC-GARCH followed by RiskMetrics both estimated on HF returns. Moreover, if we take into account all models HF returns based DCC-GARCH achieve the lowest volatility of variance. Variances calculated from open-close and close-close returns are substantially larger for both models. Higher variances are also associated with higher volatility (standard deviation).

If we concentrate on covariance based models, better results are obtained using MRK estimates. The lowest variances are achieved by Cholesky-HAR and diagonal WAR(1) model. As usual, WAR(1) is the model with the worst performance.

Overall, model with the lowest variance is Cholesky-HAR estimated on MRK covariance. Unexpectedly, DCC-GARCH estimated on HF data is the second best one. If the WAR(1) is excluded, covariance based models are better than the return based models.

4.2.4. Value-at-Risk

During crisis market risk evaluation by Value-at-Risk approach is presented in the Table 4-6. Similar to before crisis period, variances for VaR calculations from GMVP approach are used and results for 95% and 99% VaRs are presented.

Table 4-6: during crisis Value-at-Risk

Value-at-Risk	CC		OC		HF		RM		MRK	
during crisis	N	p=p*	N	p=p*	N	p=p*	N	p=p*	N	p=p*
p*=0.05										
DCC-GARCH	15	accept	18	accept	20	reject				
RiskMetrics	12	accept	15	accept	16	accept				
HAR							17	accept	16	accept
Cholesky HAR							18	accept	21	reject
WAR(1)							0	reject	0	reject
diagonal WAR(1)							21	reject	23	reject
p*=0.01										
DCC-GARCH	3	accept	6	accept	8	reject				
RiskMetrics	4	accept	3	accept	6	accept				

HAR	4	accept	3	accept
Cholesky HAR	5	accept	6	accept
WAR(1)	0	accept	0	accept
diagonal WAR(1)	2	accept	1	accept

Note: "**N**" stands for "number of exceedance", p=p* represents null hypothesis of Proportion of Failures test, CC denotes daily returns calculated from daily closing prices, OC denotes daily returns calculated from daily open-close prices and HF denotes daily returns calculated from 5-minute closing prices. RCOV and MRK stands for Realized Covariation and Multivariate Realized Kernels respectively.

Source: Author's computation

On the 5% probability level the covariance forecasts correctly evaluate the risk in case of 8 models out of 14. There is one risk underestimation in case of return based model. By one VaR exceedance, the null hypothesis of correctly specified risk level is rejected for DCC-GARCH estimated on HF returns. Covariance based models underestimate risk in 3 cases. Namely, these models are diagonal WARs using both RCOV and MRK estimates and Cholesky HAR estimated on MRK covariances. The risk is overestimated in case of WAR(1).

We accept that the risk is correctly specified for all but one model at 1% probability level. Similar to 95% VaR, rejection occurs due to one additional exceedance in case of DCC-GARCH estimated on HF returns. As in case of before crisis period we have to be careful about not rejecting results for WAR(1) specifications. Zero number of VaR exceptions might not be result of correctly specified risk level but possibly overestimated portfolio variance.

Having described all the forecast evaluation methods within during crisis period we turn to full dataset results.

4.3. Full sample

Finally, we present performance of covariance forecasts for the whole dataset. In-sample period of the whole dataset is identical to the in-sample before crisis period. The out-of-sample period covers 1 217 trading days (approximately 4.5 years) and starts on February 7, 2007 with the last observation on November 29, 2011. Structure of the section is the same as for before and during crisis periods. Firstly, we present results of Root Mean Square Forecasting Error, followed by Mean-Variance optimization and Global Minimum Variance Portfolio and we present results for Value-at-Risk approach in the end.

4.3.1. Root Mean Square Forecasting Error

Statistical evaluation of one-step ahead covariance forecasts represented by Root Mean Square Forecasting Error is shown in the Table 4-7. The structure of the Table is the same as in both before and during crisis sections.

Table 4-7: full sample RMSFE (values time 10^{-4})

RMSFE	RCOV			MRK			RCOV	MRK
full sample	CC	OC	HF	CC	OC	HF		
DCC-GARCH	5.0974	4.1015	2.4199	5.3267	4.4451	2.8335		
RiskMetrics	6.5171	5.1420	2.9864	6.6526	5.3339	3.2490		
HAR							2.2476	2.7336
Cholesky HAR							2.1168	2.5907
WAR(1)							12	12
diagonal WAR(1)							4.7654	5.0420

Note: CC denotes daily returns calculated from daily closing prices, OC denotes daily returns calculated from daily open-close prices and HF denotes daily returns calculated from 5-minute closing prices. Proxy for RMSFE calculation are RCOV and MRK. RCOV and MRK stands for Realized Covariation and Multivariate Realized Kernels respectively.

Source: Author's computation

Best performance in return based model group is achieved by DCC-GARCH that uses high-frequency returns for estimation and RCOV as a proxy for RMSFE calculation. Moreover, DCC-GARCH outperforms RiskMetrics within each the return and the proxy group. Ranking of the forecasts according to return groups, starting with the best result, is following: high-frequency, open-close and close-close returns. According to proxy selection, better results are obtained for RCOV.

Turning our attention to covariance based models, RCOV estimates show better performance. The best results are achieved by Cholesky-HAR and HAR.

The HAR based models outperform WAR one for the both covariance specifications. WAR(1) is model with the worst fit.

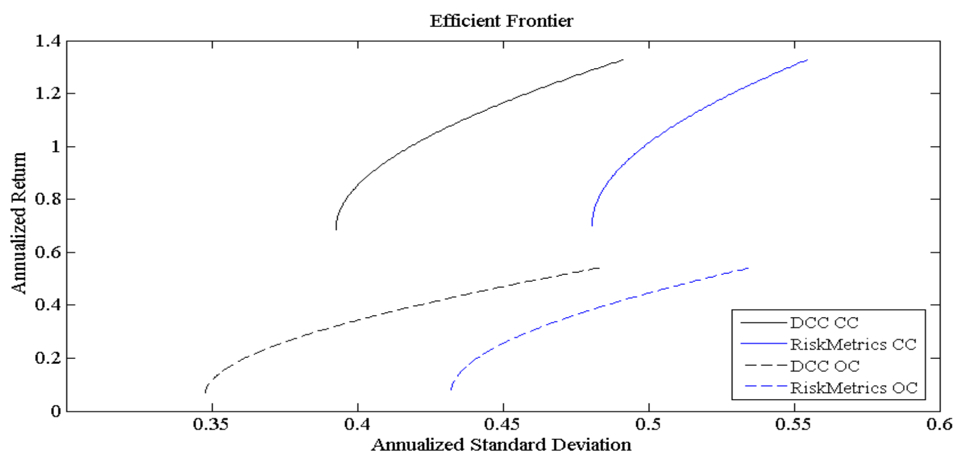
Overall, models with the best fit are Cholesky-HAR and HAR using Realized Covariation estimates. DCC-GARCH estimated on HF returns with Realized Covariation set as a proxy has the third best fit. WAR(1) models are always outperformed. On the other hand, diagonal WAR(1) perform better than two DCC-GARCH and four RiskMetrics specifications.

4.3.2. Mean-Variance optimization

The mean-variance tradeoff of covariance forecasts calculated at full dataset is presented in this section. The section is divided similarly to before and during crisis period into models using daily and high-frequency data.

Efficient frontiers displayed in Figure 4-5 are constructed on the basis of daily data covariance forecasts. Open-close return based models are less risky but also less profitable. The difference in achievable returns is more than 80% comparing OC and CC returns. A better risk-return tradeoff is recorded in the case of DCC-GARCH.

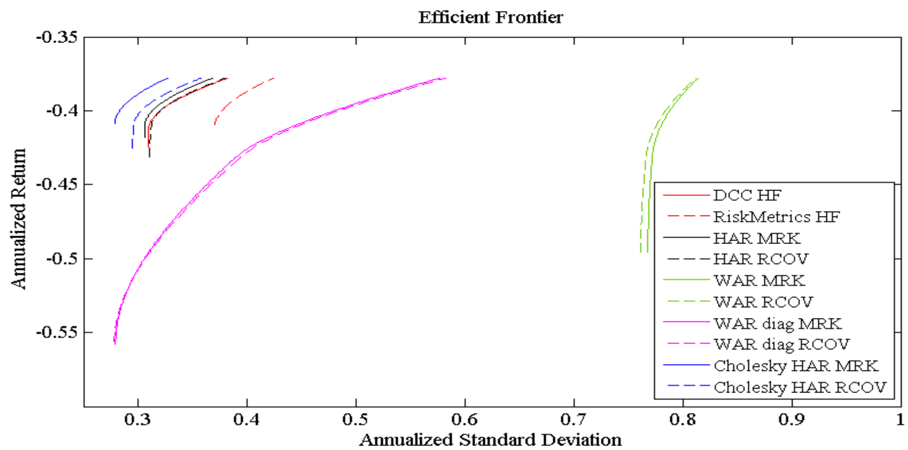
Figure 4-5: full-sample efficient frontiers - daily data



Source: Author's computation

For high-frequency data the efficient frontiers are presented in Figure 4-6. If we look at it we can see negative annualized returns as a consequence of huge losses experienced during financial crisis. WAR(1) and diagonal WAR(1) specifications are substantially outperformed by all remaining models. If we do not take WARs into account the rank of the models is in line with what we have expected. The best risk-return tradeoff is achieved by Cholesky-HAR and HAR models. DCC-GARCH performance is almost identical to HAR model using RCOV and the worst mean-variance tradeoff belongs to RiskMetrics.

Figure 4-6: full-sample efficient frontiers - high-frequency data



Source: Author's computation

4.3.3. Global Minimum Variance Portfolio

Average values of one-step ahead global minimum variance portfolio forecasts with corresponding standard deviations are presented in the Table 4-8. We keep notation from before and during crisis period so by "variance" we denote "mean of variances" in the rest of the section.

Table 4-8: full sample GMVP (values times 10^{-5})

GMVP	CC		OC		HF		RCOV		MRK	
full sample	mean	std. dev.	mean	std. dev.	mean	std. dev.	mean	std. dev.	mean	std. dev.
DCC-GARCH	11.1	11.0	8.68	7.91	6.85	6.66				
RiskMetrics	17.5	22.1	13.9	15.3	10.1	11.9				
HAR							6.95	7.23	6.68	6.93
Cholesky HAR							6.25	6.57	5.54	5.87
WAR(1)							46.9	54.2	47.7	58.0

diagonal WAR(1)	5.25	5.20	5.11	5.37

Note: CC denotes daily returns calculated from daily closing prices, OC denotes daily returns calculated from daily open-close prices and HF denotes daily returns calculated from 5-minute closing prices. RCOV and MRK stands for Realized Covariation and Multivariate Realized Kernels respectively.

Source: Author's computation

It is obvious from the Table 4-8 that the full sample, including forecasts during financial crisis (second half of 2008 and beginning of 2009), is the most volatile period. For almost all the models the standard deviation reaches values higher than variance.

When we compare DCC-GARCH and RiskMetrics, the lowest variance can be obtained estimating DCC-GARCH on HF returns. DCC-GARCH is also the model for which standard deviation does not exceed the value of variance. The best performance of models is reached by HF based returns followed by OC one. From the GMVP optimization point of view, the CC alternative is the worst one.

Looking at the results of covariance based models, overall better performance is achieved by MRK estimates. Diagonal WAR(1) estimated on the MRK covariance has the lowest variance followed by diagonal WAR(1) using RCOV estimates. The third and the fourth places are taken by Cholesky-HAR models using MRK and RCOV estimates respectively. Variances of the full WARs are the highest for all estimated models in the thesis.

The overall performance is better for the covariance based models. From the return based group the only DCC-GARCH specification using high-frequency returns is able to compete.

4.3.4. Value-at-Risk

Finally, Value-at-Risk results are presented in the Table 4-9. Again, variances from GMVP are used for VaR calculations and we present results for both 95% and 99% VaR.

Table 4-9: full sample Value-at-Risk

Value-at-Risk	CC		OC		HF		RM		MRK	
full sample	N	p=p*	N	p=p*	N	p=p*	N	p=p*	N	p=p*
p*=0.05										
DCC-GARCH	75	accept	90	reject	89	reject				
RiskMetrics	56	accept	64	accept	62	accept				
HAR							87	reject	86	reject
Cholesky HAR							93	reject	107	reject
WAR(1)							0	reject	0	reject
diagonal WAR(1)							124	reject	127	reject
p*=0.01										
DCC-GARCH	23	reject	30	reject	29	reject				
RiskMetrics	17	accept	15	accept	22	accept				
HAR							26	reject	27	reject
Cholesky HAR							33	reject	41	reject
WAR(1)							0	reject	0	reject
diagonal WAR(1)							48	reject	36	reject

Note: "**N**" stands for "number of exceedance", p=p* represents null hypothesis of Proportion of Failures test, CC denotes daily returns calculated from daily closing prices, OC denotes daily returns calculated from daily open-close prices and HF denotes daily returns calculated from 5-minute closing prices. RCOV and MRK stands for Realized Covariation and Multivariate Realized Kernels respectively.

Source: Author's computation

We reject the null hypothesis of properly evaluated risk for all covariance based models at 5% probability level. In case of HAR, Cholesky-HAR and diagonal WAR(1) the risk is underestimated. On the other hand, WAR(1) overestimates market risk. For the return based DCC-GARCH model we reject that the risk is set properly for all but one specification. Both rejected specifications, HF returns and OC returns, ends up with underestimated risk. RiskMetrics is the only model for which the risk level was set correctly.

On 1% probability level, the risk is correctly specified only for RiskMetrics. For all but one remaining model, WAR(1), the market risk is underestimated.

Due to the length of the forecasting period the zero number of exceedance for WAR(1) does not end up in acceptance of properly specified risk.

Chapter 5

Discussion of results

In this chapter we summarize and discuss our empirical findings. We start with combining results for individual models from different time periods and we try to asses general performance of the models. To be more specific we try to answer the following questions. Which model provide us with appropriate forecasts? Do we gain some advantages using more sophisticated models compared to simple ones? What kind of data are to be used in order to minimize the risk of the portfolio? At the end of the chapter we add some comments and remarks on the data and portfolio selections.

5.1. Overall performance

In this section we summarize the performance of estimated models. At the beginning we focus on return based, next on evaluation of covariance based and at the end we comment on performance of models used to obtain covariance estimates from high-frequency data.

Description of return based models starts with RiskMetrics. Forecasting performance of the RiskMetrics is the most stable one. From the Value-at-Risk perspective it is the only model with correctly specified risk level within all examined periods. Results of remaining evaluation methods show similar patterns for all periods, although they are not the best ones. The division into sub samples does not affect performance of RiskMetrics much, making it applicable not only during the stable but also the volatile times.

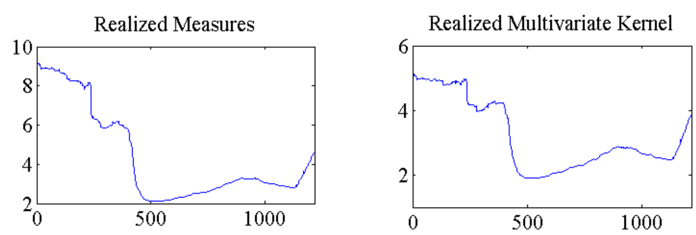
Second representative of the return based models, Dynamic Conditional Correlation GARCH, shows similar patterns for all evaluating methods except Value-at-Risk. From the RMSFE, GMVP and Mean-variance optimization point of view, DCC-GARCH substantially outperforms RiskMetrics and both WAR models during all the periods. Value-at-Risk performance of DCC-GARCH can be characterized as time dependant. Use of shorter time horizons leads to better performance. In the short sample, financial crisis does not affect results much, while in the long one, crisis might be the reason of worse performance.

Now we turn to covariance based models. The first one is Heterogeneous Autoregression. Performance of HAR is very similar for the during crisis and the full sample period. According to Value-at-Risk, model shows best performance in during crisis period. The risk is specified correctly for both 95% and 99% VaRs. In case of before crisis and full sample period risk is underestimated. According to remaining forecasts evaluation methods, HAR is the model that outperformed almost all the other models.

Cholesky-HAR is the absolute winner if we take into account RMSFE, GMVP and Mean-variance optimization criteria and shows the best performance in all the time periods. From the Value-at-Risk point of view, similar to HAR, the risk is correctly specified for during crisis period and underestimated in case of before crisis and full sample period.

Wishart Autoregressive model and diagonal Wishart Autoregressive model are the models with the worst forecasting performance. Although for diagonal WAR the lowest variance among all the models is achieved, the results are not conclusive. Estimated degrees of freedom fall below minimum level where no density function is specified for the covariance distribution.

Figure 5-1: degrees of freedom for Wishart Autoregression



Source: Author's computation

Degrees of freedom for WAR and diagonal WAR are the same across different periods thus results of the longest period are presented. Before crisis period is the only one with sufficient degrees of freedom. Generally, WAR models show a bad performance independent on the time-period.

The last part of the section is dedicated to Realized Covariation and Multivariate Realized Kernels comparison. Differences in the performance of both methods are minor. According to results of the RMSFE, GMVP and Value-at-Risk comparisons both methods show similar performance. From the Mean-variance optimization point of view slightly better performance is achieved by Multivariate Realized Kernels. If both methods are compared across different time periods, results indicates that the performance of covariance estimates is not affected by financial crisis.

Results of our analysis partially correspond to results of *Voev (2009)* and *Chiriac & Voev (2011)* where the Cholesky-HAR shows good forecasting performance. On the other hand, DCC-GARCH was outperformed by diagonal and full WAR which is not in line with our results. In the work of *Bonato et al. (2009)* where a set of different WAR specifications and the DCC-GARCH are estimated, diagonal WAR outperforms the DCC-GARCH while score of full WAR is the worse. Possible sources of differences in the results are the estimation time periods and the assets chosen for the purpose of analysis. In the above mentioned works, the period up to 2008 is considered for analysis while in our work financial crisis 2008/2009 is analyzed. Assets used in *Voev (2009)* and *Chiriac & Voev (2011)* include six S&P 500 constituents. Two currencies and two bonds are used in *Bonato et al. (2009)*. Within both asset groups similar characteristics (mean, standard deviation ...) are observed for all assets while in our work data are more volatile.

5.2. Simple or sophisticated model?

In an ideal world, the more sophisticated model we use, the better performance of the forecasts we get. However, situation in real world is more complicated and the previous statement might not be necessarily true. Easy interpretation and implementation with low time and computing demands speak in favour of simple models. On the other hand, more sophisticated models based on advanced economic and mathematical theory perform well during simulation studies. However, software implementation, difficult economic interpretation of the estimated parameters, high time and technology requirements are their major disadvantages.

Simple models presented in our work are RiskMetrics, HAR and Cholesky-HAR. Except Cholesky-HAR, where the economic interpretation of the coefficient is ruled out by Cholesky decomposition, all above mentioned advantages can be found in the group. The major advantage is duration of the estimation and forecasting procedure. All results are obtained within a minute.

DCC-GARCH and both WAR specifications belong to sophisticated models group. The main disadvantage in case of DCC-GARCH and full WAR is their time-consumption. The rolling window estimation for period of 713 days estimated for 252 consecutive days can take more than half an hour. Diagonal WAR, restricted and simplified version of full WAR, reduce time necessary for estimation to the level of simple models. Another disadvantage of these models is their software implementation. To our best knowledge there is no software with directly implemented WAR models.

Besides covariance forecasting models, Realized Covariation and Multivariate Realized Kernels were used in the thesis. Realized Covariation can be characterized as easy to implement technique with straightforward interpretation of the estimation procedure, although theory behind it requires deep mathematic knowledge. In contrast, implementation, interpretation of estimation procedure as well as theory of Multivariate Realized Kernels is rather complicated.

Final choice of preferred methods for obtaining covariance forecasts is complicated. It always depends on needs, requirements and limitations of individual investors.

5.3. Daily or High-frequency data?

The choice between daily and high-frequency data might be extremely difficult. The main advantage of daily data is that they are freely available and the major drawback is that the information about prices is limited and not suitable for intraday trading. On the other hand, high-frequency data provide us with more information and also the intraday trading is not a problem. Using HF data for covariance forecasting is problematic when individual portfolio assets are traded during not fully overlapping hours. By synchronization of the dataset considerable amount of information might be lost resulting in poor performance of forecasts compared to daily data.

In our work 56.2% of the 5-minute closing prices were thrown away during the process of synchronization. For individual assets the level of preserved observations is 34.1% in case of crude oil, 36.8% in case of gold and 83.5% in case of DAX index. The impact of such a huge reduction of the dataset is mostly visible in the return level. Cumulative daily returns for individual assets for the out-of sample period of full sample dataset are presented in the Table 5-1

Table 5-1: cumulative daily returns

	close-close	open-close	5-minute
crude oil	5.92 %	5.39 %	-84.39 %
DAX index	-19.75 %	-15.82 %	-46.95 %
gold	18.36 %	10.95 %	-37.79 %

Source: Author's computation

We can observe huge difference in return level between daily and high-frequency data. There is also a difference between close-close and open-close returns but it is a minor one. Comparing close-close and 5-minute returns decline is more than 90, 25 and 50 percent for crude oil, DAX and gold respectively. From our point of view, possible source of a huge drop are the trading hours for particular commodities. Although crude oil and gold are traded during DAX (European) trading hours, the primary market for these assets is US market thus the most of the trades are executed within US trading sessions. Once we do not include information from the prime trading hours higher volatility and miscalculation of returns occur.

5.4. Concluding remarks

Our analysis shows that the performance of models highly depends on datasets and also on chosen assets. Here we present comments on the portfolio selection. Assets included in the portfolio have to be chosen according to certain criteria. If the daily data are used for optimization, the most important thing we have to care for is similarity of the assets. The more similar assets are used, chance to obtain better results increases. By similarity of assets, the statistical properties like mean and standard deviation are meant. On the other hand, if assets from different risk levels are used (variances of the assets are significantly different), asset weights of global minimum variance portfolio are highest (more than 50 %) for the least risky one. It might happen that the entire portfolio consists of only one asset in an extreme case.

By using high-frequency data, besides similarity of assets, we have to add one more constraint. In order not to throw away significant amount of data by synchronization procedure, trading hours of all assets have to be (almost) the same.

Chapter 6

Conclusion

In this thesis we examine dynamic relationships between crude oil, DAX index and gold futures by covariance modeling during the period 2003 - 2011. Our main contribution comes from the use of European asset, represented by DAX index, together with the worldwide heavily traded assets (crude oil and gold) during times of financial crisis for estimation and direct comparison of performance of daily and high-frequency data.

In the first part of the thesis theoretical background of the models used for the analysis is presented. We start with RiskMetrics and DCC-GARCH description, the models that require returns for calculation. After that, Realized Covariation and Multivariate Realized Kernels methods of obtaining covariance estimates from the high-frequency data are described. Then we follow with the description of HAR, Cholesky-HAR, WAR(1) and the diagonal WAR(1), the models that use covariance matrices for estimation. We end the theoretical part with the definition of variance-covariance matrices forecasts evaluation methods.

The next part of the thesis starts with description of the data. Firstly, data synchronization procedure is described in detail. Secondly, the main characteristics of daily and high-frequency datasets follows. Thirdly, we conclude with our empirical findings which is the most import part of the thesis.

The empirical part of the thesis is divided into the before crisis, during crisis and full sample periods where performance of the models is summarized. From the perspective of risk-minimizing investor who seeks for least volatile portfolio it is optimal to use Cholesky-HAR model with MRK covariance estimates. Moreover, stable performance of the model across all examined periods is an advantage in volatile times similar to financial crisis. On the other hand, both the WAR specifications are not able to deal with periods of particularly high volatility and are outperformed by models that use daily data. According to Value-at-Risk perspective which is used in banking regulation the most appropriate models are RiskMetrics and DCC-GARCH.

Finally, the most interesting findings arise from risk-return tradeoff results. In that particular case, daily data based models substantially outperform high-frequency covariance based models for all the time periods. This fact is clearly visible in full sample case when risk level is similar for both model groups, however returns are positive for daily data while in case of high-frequency data returns become negative. In our opinion it is a consequence of data synchronization procedure thanks to which more than a half of original data do not appear in final dataset.

Bibliography

ANDERSEN, T.G., BOLLERSLEV, T. & DIEBOLD, F.X. (2005): "Roughing it up: Including Jump Components in the Measurement, Modeling and Forecasting of Return Volatility." *NBER Working Paper Series*, (11775).

ANDERSEN, T.G., BOLLERSLEV, T., DIEBOLD, F.X. & LABYS, P. (2001): "The Distribution of Realized Exchange Rate Volatility." *Journal of the American Statistical Association*.

ANDERSEN, T.G., BOLLERSLEV, T., DIEBOLD, F.X. & LABYS, P. (2003) "Modeling and Forecasting Realized Volatility." *Econometrica*, ,Vol. 71(2), pp.579-625.

ANDERSEN, T.G., BOLLERSLEV, T., CHRISTOFFERSEN, P. & DIEBOLD, F.X. (2007): "Practical Volatility and Correlation Modeling for Financial Market Risk Management." In Carey, M. & Stulz, R.M. *The Risks of Financial Institutions*. University of Chicago Press. pp.513-548.

BACK, K.(1991): "Asset pricing for general processes." *Journal of Mathematical Economics*, pp.371-395.

BARNDORFF-NIELSEN, O.E., HANSEN, P.R., LUNDE, A. & 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), pp.149-169.

BARNDORFF-NIELSEN, O.E. & SHEPHARD, N. (2004): "Econometric Analysis of Relaized Covariation: High Frequency Based Covariance, Regression, and Correlation in Financial Economics." *Econometrica*, 72(3), pp.885-925.

BAUWENS, L., LAURENT, S. & ROMBOUST, F.(2006): "Multivariate GARCH Models: A Survey." *Journal of Applied Econometrics*, 21, pp.79-109.

BOLLERSLEV, T. (1986): "Generalized Autoregressive Conditional Heteroscedasticity." *Journal of Econometrics*, 31, pp.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), pp.498-505.

BONATO, M. (2009): "Estimating the degrees of freedom of the Realized Volatility Wishart Autoregressive model." *Working paper*.

BONATO, M., CAPORIN, M. & RANALDO, A. (2009): "Forecasting realized (co)variances with a block structure Wishart autoregressive model." *Working paper*.

BONATO, M., CAPORIN, M. & RANALDO, A. (2012): "Risk spillovers in international equity portfolios." *Swiss National Bank Working Papers, 2012*

BROWN, A. (2008): "Private Profits and Socialized Risk - Countepoint: Capital Inadequacy." *Global Association of Risk Professional*, (42), pp.19-26.

CORSI, F. (2009): "A Simple Approximate Long-Memory Model of Realized Volatility." *Journal of Financial Econometrics*, 7(2), p.174–196.

DELBAEN, F. & SCHACHERMAYER, W. (1994): "A general version of the fundamental theorem." In *Mathematische Annalen*. Springer-Verlag. pp.463-520.

ENGLE, R.F. (1982): "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation." *Econometrica*, 50(4), pp.987-1007.

ENGLE, R.F., (2002): "Dynamic Conditional Correlation - A Simple Class of Multivariate GARCH Models." *Forthcoming Journal of Business and Economic Statistics 2002*.

ENGLE, R.F. & KRONER, K.F. (1995): "Multivariate Simultaneous Generalized Arch." *Econometric Theory*, 11(1), pp.122-150.

ENGLE, R.F. & SHEPPARD, K. (2001): "Theoretical and Empirical Properties of Dynamic Conditional Correlation Multivariate GARCH." *NBER Working Paper Series*, p.Working Paper 8554.

GOURIEROUX, C., JASIAK, J. & SUFANA, R. (2009): "The Wishart Autoregressive process of multivariate stochastic volatility." *Journal of Econometrics*, 150, pp.167-181.

HAUTSCH, N. (2011): "*Econometrics of Financial High-Frequency Data.*" Springer.

CHIRIAC, R. & VOEV, V. (2011): "Modelling and Forecasting Multivariate Reralized Volatility." *Journal of Applied Econometrics*, pp.922-947.

J.P.MORGAN & REUTERS (1996): "*RiskMetricsTM - Technical Document.*" 4th ed.

JORION, P. (2007): "*Value at Risk: The New Benchmark for Managing Financial Risk.*" 3rd ed. McGraw-Hill.

KEMPF, A. & MEMMEL, C. (2006): "Estimating the Global Minimum Variance Portfolio." *Schmalenbach Business Review*, 58, pp.332-348.

KUPIEC, P.H. (1995): "Techniques for Verifying the Accuracy of Risk Measurement Models." *The Journal of Derivatives*, pp.73-84.

MARKOWITZ, H. (1952): "Portfolio Selection." *The Journal of Finance*, 7(1), pp.77-91.

MERTON, R.C. (1972): "An Analytic Derivation of the Efficient Portfolio Frontier." *The Journal of Financial and Quantitative Analysis*, 7(4), pp.1851-1872.

MÜLLER, U.A. et al. (1993): "Fractals and Intrinsic Time – A Challenge to Econometricians." In *International Conference of the Applied Econometrics Association (AEA), Real Time Econometrics — Submonthly Time Series.* Ascona (Switzerland), 1993.

PROTTER, P.E. (2004): "*Stochastic Integration and Differential Equations.*" Second Edition ed. Springer.

VOEV, V. (2007): "Dynamic Modelling of Large Dimensional Covariance Matrices." *CoFE Discussion Paper 07-01.*

VOEV, V. (2009): "On the Economic Evaluation of Volatility Forecasts." *CREATES Research Paper 2009-56.*

WATKINS, D.S.(2002): "*Fundamentals of Matrix Computations*." 2nd ed. New York: John Wiley & Sons, Inc..

Appendix A

Table A-1: Non-rejection region for Proportion of Failures Value-at-Risk test

	Non rejection region for x, n=252 days	Non rejection region for x, n=1 217 days
p*=0.05	$6<x<20$	$46<x<77$
p*=0.01	$x<8$	$4<x<23$

Source: Author's computation

Appendix B

Table B-1: Root Mean Square Forecasting Error results (values time 10^{-4})

RMSFE	RCOV			MRK			RCOV	MRK
	CC	OC	HF	CC	OC	HF		
before crisis								
DCC-GARCH	2.5474	2.4711	1.1905	2.7628	2.6953	1.4066		
RiskMetrics	3.3411	2.9181	1.5386	3.4441	3.0336	1.6219		
HAR							1.1747	1.3632
Cholesky HAR							1.0634	1.2936
WAR(1)							6.1726	6.4110
diagonal WAR(1)							3.5965	3.7032
during crisis								
DCC-GARCH	3.8706	3.6231	2.1759	4.0782	3.8609	2.3849		
RiskMetrics	4.4926	4.0484	2.3057	4.6523	4.2534	2.5020		
HAR							2.0548	2.3343
Cholesky HAR							1.9758	2.2609
WAR(1)							11	11
diagonal WAR(1)							4.7459	4.8682
full sample								
DCC-GARCH	5.0974	4.1015	2.4199	5.3267	4.4451	2.8335		
RiskMetrics	6.5171	5.1420	2.9864	6.6526	5.3339	3.2490		
HAR							2.2476	2.7336
Cholesky HAR							2.1168	2.5907
WAR(1)							12	12
diagonal WAR(1)							4.7654	5.0420

Note: CC denotes daily returns calculated from daily closing prices, OC denotes daily returns calculated from daily open-close prices and HF denotes daily returns calculated from 5-minute closing prices. Proxy for RMSFE calculation are RCOV and MRK. RCOV and MRK stands for Realized Covariation and Multivariate Realized Kernels respectively.

Source: Author's computation

Table B-2: Global Minimum Variance Portfolio results (values time 10^{-5})

GMVP	CC		OC		HF		RCOV		MRK	
	mean	std. dev.	mean	std. dev.	mean	std. dev.	mean	std. dev.	mean	std. dev.
before crisis										
DCC-GARCH	6.41	3.03	4.71	2.95	3.44	1.86				
RiskMetrics	10.0	6.09	8.02	5.88	5.43	4.47				
HAR							3.18	1.70	3.19	1.83
Cholesky HAR							2.86	1.57	2.58	1.50
WAR(1)							20.5	23.6	21.1	28.2
diagonal WAR(1)							2.27	1.97	2.17	2.22
crisis										
DCC-GARCH	8.78	5.53	7.65	5.59	4.79	3.23				
RiskMetrics	12.9	10.8	11.6	10.1	7.02	6.37				
HAR							6.05	4.60	5.65	4.12
Cholesky HAR							5.31	4.39	4.52	3.50
WAR(1)							42.5	42.1	41.8	43.9
diagonal WAR(1)							5.05	4.17	4.86	4.22
full sample										
DCC-GARCH	11.1	11.0	8.68	7.91	6.85	6.66				
RiskMetrics	17.5	22.1	13.9	15.3	10.1	11.9				
HAR							6.95	7.23	6.68	6.93
Cholesky HAR							6.25	6.57	5.54	5.87
WAR(1)							46.9	54.2	47.7	58.0
diagonal WAR(1)							5.25	5.20	5.11	5.37

Note: CC denotes daily returns calculated from daily closing prices, OC denotes daily returns calculated from daily open-close prices and HF denotes daily returns calculated from 5-minute closing prices. RCOV and MRK stands for Realized Covariation and Multivariate Realized Kernels respectively.

Source: Author's computation

Table B-3: Value-at-Risk - 95%

$VaR_{95\%}$	CC		OC		HF		RCOV		MRK	
	N	p=p*	N	p=p*	N	p=p*	N	p=p*	N	p=p*
before crisis										
DCC-GARCH	19	accept	19	accept	18	accept				
RiskMetrics	12	accept	13	accept	14	accept				
HAR							23	reject	24	reject
Cholesky HAR							25	reject	28	reject
WAR(1)							0	reject	0	reject
diagonal WAR(1)							24	reject	25	reject
crisis										
DCC-GARCH	15	accept	18	accept	20	reject				
RiskMetrics	12	accept	15	accept	16	accept				
HAR							17	accept	16	accept
Cholesky HAR							18	accept	21	reject
WAR(1)							0	reject	0	reject
diagonal WAR(1)							21	reject	23	reject
full sample										
DCC-GARCH	75	accept	90	reject	89	reject				
RiskMetrics	56	accept	64	accept	62	accept				
HAR							87	reject	86	reject
Cholesky HAR							93	reject	107	reject
WAR(1)							0	reject	0	reject
diagonal WAR(1)							124	reject	127	reject

Note: "**N**" stands for "number of exceedance", p=p* represents null hypothesis of Proportion of Failures test, CC denotes daily returns calculated from daily closing prices, OC denotes daily returns calculated from daily open-close prices and HF denotes daily returns calculated from 5-minute closing prices. RCOV and MRK stands for Realized Covariation and Multivariate Realized Kernels respectively.

Source: Author's computation

Table B-4: Value-at-Risk - 99%

$VaR_{99\%}$	CC		OC		HF		RCOV		MRK	
	N	p=p*	N	p=p*	N	p=p*	N	p=p*	N	p=p*
before crisis										
DCC-GARCH	8	reject	11	reject	7	accept				
RiskMetrics	4	accept	5	accept	5	accept				
HAR							8	reject	9	reject
Cholesky HAR							13	reject	15	reject
WAR(1)							0	accept	0	accept
diagonal WAR(1)							14	reject	10	reject
crisis										
DCC-GARCH	3	accept	6	accept	8	reject				
RiskMetrics	4	accept	3	accept	6	accept				
HAR							4	accept	3	accept
Cholesky HAR							5	accept	6	accept
WAR(1)							0	accept	0	accept
diagonal WAR(1)							2	accept	1	accept
full sample										
DCC-GARCH	23	reject	30	reject	29	reject				
RiskMetrics	17	accept	15	accept	22	accept				
HAR							26	reject	27	reject
Cholesky HAR							33	reject	41	reject
WAR(1)							0	reject	0	reject
diagonal WAR(1)							48	reject	36	reject

Note: "**N**" stands for "number of exceedance", p=p* represents null hypothesis of Proportion of Failures test, CC denotes daily returns calculated from daily closing prices, OC denotes daily returns calculated from daily open-close prices and HF denotes daily returns calculated from 5-minute closing prices. RCOV and MRK stands for Realized Covariation and Multivariate Realized Kernels respectively.

Source: Author's computation

Appendix C

Table C-1: Heterogeneous Autoregression -parameter estimates

	RCOV					MRK				
	c	$\beta^{(d)}$	$\beta^{(w)}$	$\beta^{(m)}$	R^2	c	$\beta^{(d)}$	$\beta^{(w)}$	$\beta^{(m)}$	R^2
before crisis										
crude	4.47e-05 ***	0.021	0.432 ***	0.319 ***	0.176	5.35e-05 ***	-0.038	0.425 ***	0.294 ***	0.104
crude-DAX	-9.04e-08	0.176 ***	0.250 ***	0.351 ***	0.163	1.63e-07	0.129 ***	0.227 ***	0.405 ***	0.131
crude-gold	1.93e-06 **	0.024	0.249 ***	0.599 ***	0.254	4.66e-06 ***	0.054	0.218 ***	0.522 ***	0.162
DAX	1.19e-05 **	0.683 ***	-0.096 *	0.238 ***	0.462	1.29e-05 ***	0.641 ***	-0.107 **	0.257 ***	0.407
DAX-gold	8.52e-07	0.284 ***	0.100	0.254 **	0.138	1.52e-06	0.312 ***	0.092	0.257 **	0.159
gold	7.99e-06 **	0.103 ***	0.429 ***	0.368 ***	0.401	1.04e-05 **	0.082 **	0.348 ***	0.435 ***	0.282
crisis										
crude	1.54e-05	0.132 ***	0.386 ***	0.445 ***	0.661	1.71e-05	0.151 ***	0.330 ***	0.475 ***	0.616
crude-DAX	1.08e-05 **	0.337 ***	0.281 ***	0.275 ***	0.479	1.42e-05 **	0.156 ***	0.321 ***	0.400 ***	0.361
crude-gold	6.67e-06 **	0.213 ***	0.450 ***	0.232 ***	0.472	1.01e-05 **	0.190 ***	0.294 ***	0.375 ***	0.318
DAX	2.69e-05 **	0.459 ***	0.167 ***	0.257 ***	0.501	2.70e-05 **	0.419 ***	0.126 **	0.335 ***	0.452
DAX-gold	2.86e-06	0.206 ***	0.512 ***	0.091	0.366	4.22e-06	0.153 ***	0.489 ***	0.116	0.268
gold	7.73e-06	0.250 ***	0.260 ***	0.427 ***	0.543	8.14e-06	0.173 ***	0.096	0.661 ***	0.459

	full sample									
crude	1.18e-05 *	0.113 ***	0.397 ***	0.451 ***	0.654	1.26e-05 *	0.113 ***	0.355 ***	0.487 ***	0.601
crude-DAX	3.78e-06 *	0.329 ***	0.286 ***	0.309 ***	0.556	5.16e-06 *	0.156 ***	0.321 ***	0.434 ***	0.433
crude-gold	3.27e-06 **	0.192 ***	0.450 ***	0.278 ***	0.530	5.81e-06 ***	0.171 ***	0.294 ***	0.416 ***	0.355
DAX	1.54e-05 ***	0.486 ***	0.140 ***	0.270 ***	0.532	1.50e-05 ***	0.443 ***	0.104 **	0.348 ***	0.489
DAX-gold	1.66e-06 *	0.222 ***	0.454 ***	0.128 ***	0.341	2.62e-06 *	0.187 ***	0.415 ***	0.149 ***	0.249
gold	6.59e-06 **	0.224 ***	0.291 ***	0.421 ***	0.538	7.40e-06 **	0.153 ***	0.163 ***	0.608 ***	0.439

Note: ***, ** and * denote significance at the 1%, 5% and 10% level respectively

Source: Author's computation

Table C-2: Cholesky Heterogeneous Autoregression -parameter estimates

	RCOV					MRK				
	c	$\beta^{(d)}$	$\beta^{(w)}$	$\beta^{(m)}$	R^2	c	$\beta^{(d)}$	$\beta^{(w)}$	$\beta^{(m)}$	R^2
before crisis										
crude	0.0024 ***	0.065 *	0.472 ***	0.285 ***	0.259	0.0033 ***	0.064	0.450 ***	0.280 ***	0.150
crude-DAX	-8.50e-07	0.247 ***	0.325 ***	0.239 **	0.266	2.34e-05	0.150 ***	0.320 ***	0.328 ***	0.194
crude-gold	1.17e-04 *	-0.003	0.200 **	0.695 ***	0.286	2.77e-04 **	-0.032	0.324 ***	0.547 ***	0.206
DAX	9.32e-04 ***	0.501 ***	0.180 ***	0.191 ***	0.477	0.0011 ***	0.350 ***	0.252 ***	0.229 ***	0.346
DAX-gold	3.61e-05	0.032	0.108	0.531 ***	0.074	8.15e-05	0.051	0.080	0.495 ***	0.059
gold	6.02e-04 **	0.153 ***	0.410 ***	0.363 ***	0.483	6.75e-04 **	0.110 ***	0.343 ***	0.452 ***	0.378
crisis										
crude	6.41e-04	0.229 ***	0.408 ***	0.328 ***	0.717	7.24e-04	0.150 ***	0.415 ***	0.393 ***	0.653
crude-DAX	3.04e-04 **	0.353 ***	0.354 ***	0.234 ***	0.656	4.41e-04 **	0.181 ***	0.378 ***	0.369 ***	0.515
crude-gold	3.49e-04 ***	0.228 ***	0.527 ***	0.142 **	0.525	5.54e-04 ***	0.182 ***	0.350 ***	0.322 ***	0.330
DAX	6.32e-04 **	0.522 ***	0.260 ***	0.164 ***	0.737	6.35e-04 **	0.443 ***	0.252 ***	0.246 ***	0.672
DAX-gold	4.80e-05	0.205 ***	0.401 ***	0.234 ***	0.350	5.51e-05	0.082 **	0.301 ***	0.339 ***	0.141
gold	4.07e-04 *	0.284 ***	0.339 ***	0.330 ***	0.659	4.37e-04 *	0.182 ***	0.328 ***	0.433 ***	0.568

full sample										
crude	6.41e-04 **	0.175 ***	0.438 ***	0.347 ***	0.679	7.06e-04 **	0.099 ***	0.437 ***	0.416 ***	0.608
crude-DAX	1.01e-04 *	0.335 ***	0.361 ***	0.268 ***	0.722	1.55e-04 *	0.177 ***	0.377 ***	0.398 ***	0.582
crude-gold	1.64e-04 ***	0.154 ***	0.515 ***	0.260 ***	0.569	3.30e-04 ***	0.102 ***	0.365 ***	0.416 ***	0.354
DAX	4.88e-04 ***	0.516 ***	0.243 ***	0.190 ***	0.733	4.83e-04 ***	0.414 ***	0.258 ***	0.272 ***	0.668
DAX-gold	4.18e-05	0.117 ***	0.325 ***	0.355 ***	0.223	6.48e-05	0.064 **	0.207 ***	0.412 ***	0.098
gold	4.50e-04 ***	0.237 ***	0.366 ***	0.344 ***	0.614	4.84e-04 ***	0.156 ***	0.332 ***	0.447 ***	0.518

Note: ***, ** and * denote significance at the 1%, 5% and 10% level respectively

Source: Author's computation

Table C-3: DCC-GARCH parameter estimates

	GARCH parameters									DCC parameters						Loglikelihood		
	ω			α			β			A			B					
	CC	OC	HF	CC	OC	HF	CC	OC	HF	CC	OC	HF	CC	OC	HF	CC	OC	HF
before crisis																		
crude	3.97e-06	1.59e-05	4.60e-06	0.030	0.043	0.036	0.960	0.916	0.937									
	(1.41e-11)	(1.01e-10)	(1.38e-11)	(2.54e-04)	(2.34e-05)	(3.83e-04)	(5.05e-04)	(1.27e-02)	(1.36e-02)									
DAX	4.65e-06	2.81e-06	2.80e-06	0.106	0.094	0.093	0.859	0.867	0.868	0.008	0.011	0.014	0.986	0.985	0.976	8.41 e+03	8.83 e+03	9.46 e+03
	(3.66e-12)	(5.64e-12)	(5.67e-12)	(9.74e-04)	(2.42e-02)	(2.40e-02)	(1.12e-02)	(5.47e-02)	(5.62e-02)	(7.94e-06)	(1.33e-05)	(3.28e-05)	(2.45e-05)	(4.77e-05)	(2.00e-04)			
gold	1.18e-06	9.34e-07	6.90e-07	0.035	0.047	0.032	0.958	0.947	0.960									
	(4.92e-13)	(4.20e-13)	(1.84e-13)	(6.90e-05)	(1.14e-04)	(7.60e-05)	(6.26e-05)	(1.30e-04)	(1.11e-04)									
crisis																		
crude	8.73e-06	8.44e-06	2.29e-06	0.059	0.057	0.043	0.928	0.930	0.950									
	(2.43e-11)	(1.87e-11)	(2.18e-12)	(2.83e-04)	(2.31e-04)	(8.88e-05)	(4.03e-04)	(3.36e-04)	(1.17e-04)									
DAX	2.74e-06	2.66e-06	2.47e-06	0.094	0.084	0.094	0.900	0.905	0.897	0.043	0.051	0.046	0.930	0.913	0.921	7.88 e+03	8.09 e+03	8.72 e+03
	(1.80e-12)	(2.15e-12)	(1.70e-12)	(3.36e-04)	(5.01e-04)	(5.96e-04)	(2.97e-04)	(5.82e-04)	(6.28e-04)	(5.02e-05)	(9.18e-05)	(7.62e-05)	(1.22e-04)	(2.59e-04)	(3.04e-04)			
gold	1.82e-06	1.71e-06	9.42e-07	0.063	0.067	0.048	0.929	0.926	0.944									
	(8.66e-13)	(8.40e-13)	(2.24e-13)	(2.74e-04)	(2.65e-04)	(1.08e-04)	(1.79e-04)	(1.81e-04)	(1.06e-04)									

									full sample									
crude	5.86e-06	7.80e-05	1.99e-06	0.045	0.050	0.040	0.944	0.934	0.952									
	(6.68e-12)	(9.73e-12)	(7.87e-13)	(9.59e-05)	(1.13e-04)	(6.51e-05)	(1.50e-04)	(2.17e-04)	(9.87e-05)									
DAX	2.63e-06	1.44e-06	1.53e-06	0.093	0.085	0.090	0.895	0.906	0.901	0.030	0.035	0.034	0.953	0.940	0.950	6.19 e+03	6.53 e+03	6.33 e+03
	(8.17e-13)	(3.97e-13)	(4.12e-13)	(2.56e-04)	(4.15e-04)	(4.50e-04)	(2.48e-04)	(4.51e-04)	(4.88e-04)	(2.14e-05)	(3.53e-05)	(4.45e-05)	(6.96e-05)	(1.70e-04)	(1.46e-04)			
gold	1.55e-06	1.37e-06	8.15e-07	0.048	0.059	0.041	0.944	0.934	0.951									
	(3.57e-13)	(3.40e-13)	(9.95e-14)	(7.53e-05)	(9.98e-05)	(4.53e-05)	(6.02e-05)	(8.97e-05)	(5.71e-05)									

Note: Standard errors of parameter estimates are presented in parentheses, CC denotes daily returns calculated from daily closing prices, OC denotes daily returns calculated from daily open-close prices and HF denotes daily returns calculated from 5-minute closing prices

Source: Author's computation