Three Essays on Risk Modelling and Empirical Asset Pricing

František Čech

Institute of Economic Studies
Faculty of Social Sciences
Charles University

Supervisor:

doc. PhDr. Jozef Baruník, Ph.D.

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, April 1, 2019

Signature

Acknowledgments

I would like to express my gratitude to my supervisor Jozef Baruník. Without his help, constant support, guidance and positive attitude it would be complicated to finish my PhD studies. I especially thank him for the endless patience he has with me.

I am also grateful to my family, especially my parents, for their unconditional support during my studies. Last but not least I am grateful to Martina for her love, patience, support and understanding.

This thesis is part of a project that has received funding from the European Unions Horizon 2020 research and innovation programme under the Marie Skodowska-Curie grant agreement No. 681228 as well as Charles University Grant Agency GAUK No 610317 and 1198214.

Prague, April 2019
František Čech

Abstract

This thesis consists of three papers that focus on risk modelling and empirical asset pricing. In the first paper, we introduce a new model for multivariate volatility modelling and forecasting. By building a system of seemingly unrelated heterogenous autoregressions, we obtain more precise and efficient estimates of the covariance matrices. The complex forecasting exercise carried out on data from the turbulent period of the global financial crisis 2007-2008 demonstrates direct economic benefits of our approach. The second paper moves our research from expected utility to quantile preferences. We concentrate on commonalities in the volatility series that influence the distribution of asset returns. Specifically, we develop a Panel Quantile Regression Model for Returns that can control for otherwise unobserved heterogeneity among financial assets, and allows us to exploit common factors in the panel of volatilities. Results of our empirical application highlights the benefits of our newly proposed model from an economic and statistical point of view. The last paper generalizes our previous results. We show that quantile Euler equation can be transformed into a basic quantile pricing equation and has a stochastic discount factor/pricing kernel representation. We also provide an important link to quantile factor models. The empirical part of this paper demonstrates the validity of our theoretical findings using data of the US and German Treasury futures contracts.

Czech Abstract

Tato práce se skládá ze tří článků, které se zaměřují na modelování rizik a empirické oceňování aktiv. V prvním článků představujeme nový model pro modelování a prognózování vícerozměrné volatility. Budováním systému zdánlivě nesouvisejících heterogenních autoregresí získáme přesnější a účinnější odhady kovariančních matic. Komplexní prognózování dat z turbulentního období globální finanční krize roků 2007-2008 ukazuje přímé ekonomické přínosy našeho přístupu. Druhý článek posouvá náš výzkum z očekávaného užitku na kvantilové preference. Zaměřujeme se na společné rysy řady volatility, které ovlivňují rozdělení výnosů aktiv. Konkrétně jsme vyvinuli Panel Quantile Regression Model for Returns, kterým můžeme kontrolovat jinak nepozorovanou heterogenitu mezi finančními aktivy a umožňuje nám zachytit společné faktory v panelu volatility. Výsledky naší empirické analýzy ukazují výhody našeho nově navrženého modelu jak z ekonomického, tak i statistického hlediska. Poslední článek zobecňuje naše předchozí výsledky. Ukazujeme, že kvantilová Eulerova rovnice může být transformována na základní kvantilovou cenovou rovnici a má reprezentaci stochastického diskontního faktoru / cenového jádra. Poskytujeme také důležité spojení s kvantilovými modely. Empirická část této práce demonstruje platnost našich teoretických poznatků s využitím údajů z termínovaných kontraktů z USA a Německa.

Contents

List of Figures

List of Tables

Introduction

This dissertation consists of three papers that focus on risk modelling and empirical asset pricing. Specifically, the first paper contributes to the literature by providing a method of obtaining more efficient estimates and forecasts of covariance matrices. The second paper identifies common risk factors in panels of volatilities that drives the distribution of asset returns and the third paper introduces basic quantile asset pricing equation with an application to factor pricing. All papers result from the natural collaboration with my supervisor Jozef Barunik who is also co-author of the papers. Therefore, in the rest of the text I stick to "we" when referring to the author. A short summary of the papers follows.

In Chapter 1 - *On the modelling and forecasting multivariate realized volatility: Generalized Heterogeneous Autoregressive (GHAR) model*, we introduce a multivariate extension of the popular Heterogeneous Autoregressive model. This paper is published in Journal of Forecasting (Čech and Baruník, 2017).

Volatility modeling and forecasting are key issues in the area of financial econometrics. In empirical work, researchers and practitioners often study stock market data and find dependencies in the second moment of these data. As shown by Engle (1982), Bollerslev (1986), Nelson (1991) and many others, volatility of financial time series is anything but constant. To deal with this problem, a new family of parametric univariate conditionally heteroscedastic models represented by the Generalized Autoregressive Conditional Heteroscedasticity (GARCH) was developed in the eighties and nineties.

While the search for more accurate volatility models has been the focal point of many researchers, interdependencies among assets and subsequent comovements are of great importance in practice (e.g., asset allocation, portfolio management, risk management, etc.). The natural extension of the family of volatility models is to model the whole covariance structure of the given assets. This gives rise to the development of the multivariate GARCH models. Although the transition from univariate to multivariate GARCH models might seem to be straightforward, it possesses several challenges. Multivariate volatility modeling nowadays offers numerous research opportunities in the form of extension to, or innovation of, current methodologies; as well as developing techniques for solving drawbacks of current approaches (e.g., reduction of dimensionality). Our research contributes to these efforts by introducing a generalization of Heterogeneous Autoregressive (HAR) Model of Corsi (2009).

Increased availability of high-frequency data in the last decade resulted in the development of the new non-parametric approach of treating volatility. In particular, model-free estimator of Realized Volatility in Andersen et al. (2001) makes volatility observable. Theoretical properties of this estimator have been further studied in Andersen et al. (2003) and Barndorff-Nielsen and Shephard (2004a). Barndorff-Nielsen and Shephard (2004a) moreover introduce the concept of Realized Covariation, which is a multivariate extension of Realized Volatility. Market microstructure noise can significantly affect Realized Covariance estimates resulting in not positive semi-definite matrices. A solution to this problem is offered by Barndorff-Nielsen et al. (2011) and their Multivariate Realized Kernels estimator that guarantees the positive semi-definiteness of the covariance matrix.

All the realized measures, univariate or multivariate, are ex-post measures of return (co)variation. These measures need to be further modeled, so they are of some practical use. The research devoted to entire covariance structure modelling is ongoing and growing. Part of the researchers makes use of variants of the Wishart distribution to model the structure of the realized covariances (Gouriéroux et al., 2009; Bonato, 2009; Bonato et al., 2013; Jin and Maheu, 2013). Another stream of researchers decompose realized covariance matrices by matrix exponential-logarithm transformation or Cholesky decomposition and use standard time-series techniques afterwards (Bauer and Vorkink, 2011; Chiriac and Voev, 2011). The advantage of the decomposition approach is a guarantee of the positive-semidefiniteness of the covariance matrix forecasts. Our work contributes to the literature by introducing Generalized Heterogeneous Autoregressive Model (GHAR), a multivariate extension of the popular HAR Model intended for covariance matrix modelling and forecasting.

In our work, we stick to the covariance decomposition stream of literature. Specifically, we model Cholesky factors of the realized covariance matrices as a system of the seemingly unrelated heterogeneous autoregressions. Motivation to build a system of the seemingly unrelated regression (Zellner, 1962) over HAR is the contemporaneous correlation in the residuals of the simple HAR model. The advantage of this approach is that we estimate a multivariate HAR model, which will capture the separate dynamics of the variances and covariances, but also possible common structure. Moreover, it will also yield more efficient estimates - the error terms from simple HAR are heteroscedastic (Corsi et al., 2008), which makes the coefficient estimates less efficient. Furthermore, when there is no information about dependence between equations left in the residuals estimator will converge to a simple Ordinary Least Squares (OLS) estimates, as the diagonal weighting matrix in generalized regression will reduce the estimates to OLS. Therefore, using generalized least squares, we capture dependencies hidden in the residuals delivering more efficient estimates.

In the empirical application, we study portfolios consisting of five, ten and fifteen[1] highly liquid stocks from New York Stock Exchange. We begin our analysis with the one-

[1]Apple Inc. (AAPL), Exxon Mobile Corp. (XOM), Google Inc. (GOOG), Wal-Mart Stores (WMT), Microsoft Corp. (MSFT), General Electric Co (GE), International Business Machines Corp. (IBM), Johnson & Johnson (JNJ), Chevron Corp. (CVX), Procter & Gamble (PG), Pfizer Inc. (PFE), AT&T Inc. (T), Wells Fargo & Co (WFC), JP Morgan Chase & Co (JPM) and Coca-Cola Co. (KO).

step-ahead forecasts for the portfolio consisting of five stocks, whereas we leave portfolios of ten and fifteen stocks and also five- and ten-step ahead forecasts as a robustness check showing that the proposed methodology also works well at larger dimensions and different forecasting horizons. Our dataset consists of tick data and covers the period of Global Financial Crisis, i.e. July, 1 2005 to January, 3 2012 with 1623 trading days. As it is standard in the literature we explicitly exclude weekends and bank holidays (New Year's Day, Independence Day, Thanksgiving Day, Christmas) to ensure sufficient liquidity. From the tick data, we calculate realized covariance matrices using Multivariate Realized Kernels estimator at one-minute frequency, Realized Covariance at one and five-minute frequencies and Sub-Sampled Realized Covariance at five, ten, fifteen and twenty minute frequency. Moreover, we calculate also open-close daily returns. The model estimation and forecasting exercise are carried out using rolling window estimation with a fixed length of 750 days, i.e. three years.

We compare the performance of the GHAR against two covariance based benchmark models (HAR, Vector ARFIMA (Chiriac and Voev, 2011)), and two return based benchmarks (Dynamic Conditional Correlation GARCH (Engle, 2002), RiskMetrics (Longerstaey and Spencer, 1996)) primarily according to economic criteria, i.e. Mean-Variance efficient portfolio of Markowitz (1952) and Global Minimum Variance Portfolio (GMVP) using cumulative and annualized risk. The rationale behind is the importance of well-conditioned and invertible forecasts rather than focusing on unbiasedness, as an unbiased forecast does not necessarily translate into an unbiased inverse (Bauwens et al., 2012). As a robustness check we also provide a ranking of the models based on the Root Mean Squared Error (RMSE) loss functions based on the Frobenius norm and to test the significant differences of competing models, we use the Model Confidence Set (MCS) methodology of Hansen et al. (2011). The MCS procedure sequentially eliminates the worst-performing model from the full set of competing models when the null about the same forecasting performance is rejected.

Overall, the results of our analysis suggest that GHAR provides more precise and more efficient covariance matrix forecasts and they translates to economic gains directly. Specifically, in the one-step-ahead forecasting exercise using portfolio of five stocks, the GHAR shows the best performance according to all economic criteria, i.e. GHAR achieves the best risk-return trade-off in Markowitz optimization, and has the lowest risk according to both cumulative and annualized versions of GMVP. The robustness check, the portfolio of ten/fifteen stocks and five/ten-step-ahead forecasts qualitatively match our previous findings. Moreover, we document the economic benefit of estimating the realized covariance with more efficient multivariate realized kernel and sub-sampled realized covariance estimators using ten to twenty minutes sub-sampling. In the statistical comparison, we obtain a bit mixed results. While in the one-step-ahead forecasts GHAR always belongs to MCS in case of the portfolio of five and ten stocks, it is in MSC only when 5-minutes RCOV is used in case of fifteen stocks portfolio. For the forecasting horizon of five/ten days results do not change substantially. The only notable difference is absence of GHAR in MCS in the case of ten-step ahead forecasts of portfolio consisting of fifteen stocks. We address unambiguous results of the statistical evaluation to a problem of selecting the "correct" proxy for unobservable "true" covariance matrix.

In Chapter 2 - *Measurement of Common Risk Factors in Tails: A Panel Quantile Regression Model for Returns*, we introduce an innovative approach of modelling commonalities in the quantiles of future returns using information from panels of realized measures. The earlier version of the paper was published in Institute of Economic Studies Working Paper series as *IES Working Paper 20/2017* and currently is under review in the *Journal of Financial Econometrics*.

During the last two decades, global financial markets were hit by several crises. The most well-known and important ones are Dot-com Bubble and the Global Financial Crisis of 2007-2012 that includes Icelandic financial crisis and European sovereign debt crisis. The aftermath of these events highlights the necessity of proper risk identification and mitigation. The need for accurate risk measures is important not only from the regulatory point of view to prevent a future crisis but is also crucial for many applications within portfolio and risk management. Recently, the increased availability of high-frequency data resulted in the development of the more accurate volatility estimators commonly referred as Realized Measures. Whether is original Realized Volatility (Andersen et al., 2001), Realized Semivariance (Barndorff-Nielsen et al., 2010), for which the sign of the price change matters, or the adjusted Bi-Power Variation (Andersen et al., 2011) that is robust to jumps in the prices and the certain types of microstructure noise, all these realized measures help us to understand the nature of the data, identify sources and potentially predict the risk.

Although volatility forecasting is essential for many financial applications, it does not help us to specify the conditional distribution of future returns. The classical portfolio theory rather concentrates on the risk-return relationship that has a long history and is well documented. For example in the Capital Asset Pricing Model, the risk of the asset is measured by the covariance between asset return and market return. Market return is just one of the many possible factors affecting an asset's risk. Among other factors, the volatility of the asset plays an essential role in explaining expected returns. The classical asset pricing moreover assumes an economic agent maximizing expected utility. However, the expected utility framework might be too restrictive to describe the real/actual behavior of the economic agents. Recent studies thus assume agents to maximize their quantile utilities, e.g. de Castro and Galvao (2018).

In finance, the Conditional Autoregressive Value-at-Risk (CAViaR) model of Engle and Manganelli (2004) is one of the first examples that focus on the estimation of quantiles of various asset returns, Baur et al. (2012) use quantile autoregressions to study conditional return distributions and Cappiello et al. (2014) detects comovement between random variables with time-varying quantile regression. The work of Žikeš and Baruník (2016), who combine the quantile regression framework (Koenker and Bassett Jr, 1978) with realized volatility, is another important example in this field. In their work, it has been shown that various realized measures are useful in forecasting quantiles of future returns without making assumptions about underlying conditional distributions.

While Žikeš and Baruník (2016) provided an important link between future quantiles of return distribution and its past/ex-ante variation, they concentrate on the univariate time series. Effective risk diversification techniques work not only with the single conditional asset return distribution, but require a deeper knowledge of the dependencies in the joint distributions. In the standard mean-regression framework, Bollerslev et al. (2018) show that realized volatility of the financial time series share many commonalities. In the quantile regression set-up, however, there is no similar study that will try to uncover information captured in the panels of volatility series. To the best of our knowledge, there is no study dealing with estimates of conditional distribution of return series in a multivariate setting that explores ex-post information in volatility.

In this paper, we contribute to the literature by introducing a Panel Quantile Regression Model for Returns - we propose to model the panel of assets returns via its past and/or ex-ante volatility using panel quantile regression techniques. This approach allows us to exploit common factors in volatility series that directly affect quantiles of return series. Moreover, we can control for otherwise unobserved heterogeneity among financial assets. Furthermore, using the fixed effects estimator, we can disentangle overall market risk into the systematic part and idiosyncratic risks. In a sense, we revisit a large literature connecting volatility with the cross-section of returns we model tail events of the conditional distributions via volatility.

In the empirical application, we show that the newly proposed model delivers more accurate estimates than benchmark methods using various data-sets. The gain in accuracy translates into better forecasting performance of Panel Quantile Regression Model for Returns. We test the performance of our model in a portfolio Value-at-Risk forecasting exercise where we concentrate on the statistical and economic evaluation. In the statistical comparison, we distinguish between the absolute and relative performance of the given model. The absolute performance in our work is assessed by the so-called CAViaR test of Berkowitz et al. (2011) and tests whether the model is dynamically correctly specified. For the relative performance, we employ a standard Diebold-Mariano test and we pairwise compare all the competing models. In the economic comparison, we study Global Minimum Value-at-Risk Portfolio(GMVaRP) and the Markowitz like efficient frontiers of the Value-at-Risk Return trade-off. The economic and relative statistical performance is tested against three benchmark models - RiskMetrics (Longerstaey and Spencer, 1996) and two versions of Univariate Quantile Regression Model for Returns (Žikeš and Baruník, 2016).

Our analysis starts with the well-behaved simulated data from Monte-Carlo experiments. Specifically, we simulate 29 continuous price process series using four error distributions - Multivariate normal/fat-tailed Student-t distributions both with the given correlation structure obtained from the stock market data and Univariate normal/fat-tailed Student-t distributions. In total, we run 500 simulations for each error distribution and in each simulation step we use rolling window estimation with a length of 1000 observations. The results of the Monte-Carlo simulation study shows that our model is dynamically well specified and outperforms all the benchmarks in direct statistical comparison when we use more heterogeneous data generated from the univariate error distributions.

Next, we analyze 29 highly liquid stocks[2] from the New York Stock Exchange during

[2]Apple Inc. (AAPL), Amazon.com, Inc. (AMZN), Bank of America Corp (BAC), Comcast Cor-

the period July 1, 2005, to December 31, 2015. In the empirical analysis, we hypothesise that quantile of open-close returns depends on the various ex-post risk measures calculated from the tick data, i.e. Realized Variance, Realized Semivariances and Realized Bi-Power Variation. For the portfolio Value-at-Risk construction we also proxy the covariance structure by the Realized Covariance estimates. Similar to simulation study we use rolling window estimation procedure with the same window length (1000 days) for the estimation and forecasting purposes. In the in-sample analysis, we document unobserved heterogeneity in far quantiles that needs to be controlled. Moreover, all the risk measures show the asymmetric impact on the quantiles of returns, e.g. impact is higher in the bellow median quantiles. In the out-of-sample forecasting exercise, we have found that all the panel quantile regression models are dynamically correctly specified. Importantly, the panel quantile regressions consistently outperform the benchmarks in various quantiles and they are not outperformed by any of the benchmarks. From the economic point of view, newly proposed modeling strategy performs best in all but median quantiles according to GMVaRP criteria and provide us with the best Value-at-Risk Return trade-off.

Our next step was the analysis of the common exogenous risk factor in tails of the returns distribution. We have selected ex-ante measure of the market uncertainty, widely used VIX Index which measures the expectations about the 30-day market volatility, as the exogenous factor. Results of the analysis confirm that VIX carries an important part of the information about risk that is not fully captured by any of the realized measures. Moreover, by controlling for the unobserved heterogeneity and idiosyncratic volatility, VIX proves to be a strong common factor driving the tails of the return distributions. Our findings also hold in the economic comparison where panel quantile regression model with VIX achieves the best performance using both evaluation criteria.

In the last section of the paper, we test the robustness of our previous findings using high dimensional portfolio (496 assets) consisting of the constituents of the S&P 500 index. We found that the VIX Index plays an important role in the high-dimensional application and that anticipation of the future market volatility translates directly to the conditional distribution of future returns.

Overall, the results of our analysis suggest that the Panel Quantile Regression Model for Returns is dynamically correctly specified. Moreover, it dominates benchmark models in the economically important quantiles (5%,10% or 95%) and we find that none of the benchmark models is able to outperform our model consistently. Furthermore, the Panel Quantile Regression Model for Returns provides us with direct economic gains according to both economic evaluation criteria.

---

poration (CMCSA), Cisco Systems, Inc. (CSCO), Chevron Corporation (CVX), Citigroup Inc. (C), Walt Disney Co (DIS), General Electric Company (GE), Home Depot Inc. (HD), International Business Machines Corp. (IBM), Intel Corporation (INTC), Johnson & Johnson (JNJ), JPMorgan Chase & Co. (JPM), The Coca-Cola Co (KO), McDonald's Corporation (MCD), Merck & Co., Inc. (MRK),Microsoft Corporation (MSFT), Oracle Corporation (ORCL), PepsiCo, Inc. (PEP), Pfizer Inc. (PFE), Procter & Gamble Co (PG), QUALCOMM, Inc. (QCOM), Schlumberger Limited. (SLB), AT&T Inc. (T), Verizon Communications Inc. (VZ), Wells Fargo & Co (WFC), Wal-Mart Stores, Inc. (WMT), Exxon Mobil Corporation (XOM).

In Chapter 3 - *Dynamic Quantile Model for Bond Pricing*, we concentrate on the quantile pricing of bond future contracts.

In this work, we study the bond pricing in the tails of the returns distributions. As opposed to classical asset pricing (Sharpe, 1964; Lintner, 1965; Merton, 1973; Ross, 1976), we make a step forward and move from the expected utility set-up to quantile preferences. This transition allows us to study asset pricing given the economic agents differing in their level of risk aversion. In particular, we build on work of de Castro and Galvao (2018) who derive quantile Euler equation using properties of quantile preferences as defined in Manski (1988) and Rostek (2010). We also utilize the advantages of quantile preferences such as robustness to fat tails and the ability to capture heterogeneity through the quantiles. We further extend the results of de Castro and Galvao (2018) into a stochastic discount factor representation of the quantile asset pricing equation and present a link to the factor models.

In the empirical application, we focus on quantile pricing of the two, five, ten and thirty years US and German government bond futures contracts from the Chicago Board of Trade and the EUREX exchanges. The US Treasuries dataset consists of the individual assets tick prices from the period July 1, 2003, to November 30, 2017 during regular trading hours - Sunday to Friday, 5:00 p.m. - 4:00 p.m. Chicago Time. We further consider selected maturities of US forward rates estimates to play an important role in the bond pricing as it is common in the literature, e.g. Cochrane and Piazzesi (2005). These data are obtained from the dataset of Gürkaynak et al. (2007) where the detailed estimation procedure of data creation is described. In the case of German treasury futures, we are working with tick prices from the period October 1, 2005, to November 30, 2017 during standard trading hours - Monday to Friday, 8:00 a.m. - 10:00 p.m. Central European Time. To ensure sufficient liquidity, we explicitly exclude public holidays and days with less than 5 hours of trading. From the raw tick data, we extract 5 minutes prices, and we calculate open-close returns, Realized Volatility (Andersen et al., 2003) and Realized Semi-variance (Barndorff-Nielsen et al., 2010).

For estimation purposes, we adopt recently developed smoothed (Generalized) Method of Moments quantile estimator of de Castro et al. (2018) and the quantile regression of Koenker and Bassett Jr (1978). First, we study single-factor-model with Realized Volatility being the risk factor. In this set-up, we illustrate the proximity of the GMM quantile estimator and the standard quantile regression. Second, relying on the similarity of both methods we study multi-factor-model where we rely solely on the quantile regression approach since the implementation of multiple moment conditions in quantiles is not trivial and is subject to further research. We consider two multi-factor specifications in our work. In the first specification positive and negative Realized Semivariance serve as a risk factor. The second specification is motivated by the Cochrane and Piazzesi (2005) and we consider two to five years forward rates.

Results of our analysis demonstrate a significant influence of the Realized Volatility on the quantiles of the treasury returns. In both US and German treasuries, we obtain qualitatively similar results using both GMM and quantile regression estimators. Quantitatively, however, results differ a bit and we attribute these differences to lower liquidity of the German treasuries. When we concentrate on the comparison of GMM and quantile regression we again obtain qualitatively and also quantitatively similar results. Specifically, the majority of the GMM coefficients estimates lies almost always in the 95% confidence intervals of the quantile regression and vice versa. Interestingly, while in German treasuries quantile regression in almost all quantiles underestimates the influence of the Realized Volatility compared to GMM estimates, in the US Treasuries, GMM and quantile regression estimates intersect frequently, and there is no clear under/overestimation pattern.

In the multi-factor model, our results depend heavily on the factors used for analysis. In case Realized Semivariances are considered being risk factors results of our analysis share many commonalities with single-factor models, e.g. coefficient estimates for quantiles below/above the median have negative/positive signs, the majority of coefficients are statistically significant. Besides similarities, we also document a unique influence of semivariances. In the US Treasuries case, negative semivariance influence lower quantiles relatively more than the upper quantiles while the opposite is true for positive semivariance. Moreover, in the upper quantiles, positive semivariance dominates negative semivariance whereas in the lower quantiles the results are mixed. In contrast, the influence of Realized Semivariances on the quantiles of German government bonds returns is more symmetric and German treasuries look more homogeneous as the coefficient estimates closer to each other.

In the last part of the paper, we study the multi-factor model when forward rates serve as the risk factors. In this part, we concentrate on the US Treasuries only since the data for the German market are not available at the desired (daily) frequency. Our analysis shows that forward rates carry very limited information about bond returns distributions. Specifically, for all the treasuries and all the forward rates, the vast majority of the estimates is statistically insignificant. Hence, the risk-averse investor optimizing quantiles bellow median finds forward rates of limited use since their coefficients are not statistically different from zero. The only exception where a risk-loving investor might consider forward rates to be valid risk factors is the shortest maturity treasury where selected forward rates show partial explanatory power in the upper quantiles of the bond returns distribution.

Chapter 1

On the modelling and forecasting multivariate realized volatility: Generalized Heterogeneous Autoregressive (GHAR) model

Recent multivariate extensions of popular heterogeneous autoregressive model (HAR) for realized volatility leave substantial information unmodelled in residuals. We propose to employ a system of seemingly unrelated regressions to model and forecast realized covariance matrix to capture this information. We find that the newly proposed generalized heterogeneous autoregressive (GHAR) model outperforms competing approaches in terms of economic gains, providing better mean-variance trade-off, while, in terms of statistical precision, GHAR is not substantially dominated by any other model. Our results provide a comprehensive comparison of the performance when realized covariance, sub-sampled realized covariance and multivariate realized kernel estimators are used. We study the contribution of the estimators across different sampling frequencies, and show that the multivariate realized kernel and sub-sampled realized covariance estimators deliver further gains compared to realized covariance estimated on a 5-minutes frequency. In order to show the economic and statistical gains, portfolio of various sizes is used.

1.1 Introduction

The risk of individual financial instruments is crucial for asset pricing, portfolio and risk management. Besides volatility of individual assets, knowledge of covariance structure between assets in portfolio is of great importance. Accurate forecasts of variance-covariance matrices are particularly important in asset allocation and portfolio management.

The traditional approach of obtaining covariance matrix estimates relays on multivariate generalized autoregressive conditional heteroskedasticity (MGARCH) models such as the constant conditional correlation GARCH of Bollerslev (1990), the dynamic conditional correlation GARCH of Engle (2002) or the BEKK of Engle and Kroner (1995) (for a survey of MGARCH models see Bauwens et al. (2006)). These models are popular in the literature although they suffer from curse of dimensionality problem. Increased availability of high-frequency data in the last decade resulted in development of the new non-parametric approach of treating multivariate volatility. A milestone for covariance matrix modelling is the work of Barndorff-Nielsen and Shephard (2004a) where the theory of "realized covariation" is introduced. Realized covariance matrices are expost measures of daily covariation and they need to be further modelled. The research dedicated to modelling the entire covariance matrices is still lively. From the already established methods, let us mention Wishart Autoregression (WAR) of Gouriéroux et al. (2009) with numerous extensions presented in Bonato (2009) and Bonato et al. (2013). A different approach of realized volatility modelling can be found in Bauer and Vorkink (2011), who model realized stock market volatility using matrix-logarithm transformation and primarily concentrate on forecasting performance of the factor model. A more common approach of obtaining positive definite forecasts of covariance matrices is the use of Cholesky decomposition. The use of Cholesky factors, further estimated by Vector Autoregressive Fractionally Integrated Moving Average (VARFIMA), Heterogeneous Autoregression (HAR) or WAR-HAR can be found in the work of Chiriac and Voev (2011). More recently, Amendola and Storti (2015) consider combining predictions from multivariate GARCH models and realized covariance matrices.

In this chapter, we contribute to the literature by proposing a new model for dynamic covariance matrix modelling and forecasting. We model Cholesky factors of the realized covariance matrix as a system of seemingly unrelated heterogeneous autoregressions. The main motivation is that we may expect the residuals from simple HAR model to be contemporaneously correlated and, moreover, heteroscedastic due to well known volatility in the volatility effect (Corsi et al., 2008). Estimating the system of HAR equations using generalized least squares allows us to capture these dependencies. Hence the generalised HAR (GHAR) may provide more precise and more efficient forecasts, which will translate to economic gains directly. On the portfolios of various sizes, we show that GHAR model delivers significant economic gains and, statistically, is not substantially outperformed, when compared to natural benchmark models based on high frequency data (HAR, VARFIMA), as well as daily data (DCC-GARCH, RiskMetrics). In addition, we study the economic benefits of estimating the realized covariance with more efficient sub-sampled realized covariance and multivariate realized kernel estimators.

The rest of the chapter is structured as follows. We provide background for estimation of realized covariation from high frequency data in the next section. The third section describes frameworks for modeling multivariate volatility, and it presents our GHAR model. The fourth section provides description of dataset and research design, including economic as well as statistical evaluation criteria. In the fifth section we discuss out-of-sample forecast evaluation, and sixth section concludes.

1.2 Estimation of covariation from high frequency data

We assume that the q-dimensional efficient price process p_t evolves over time $0 \leq t \leq T$ according to the following dynamics

$$dp_t = \mu_t dt + \Sigma_t dW_t + dJ_t, \tag{1.1}$$

where μ_t is predictable component, Σ_t is real-values $q \times q$ volatility process, $W_1, \ldots, W_q$ is a q-dimensional Brownian motion, and dJ_t is a jump process. A central object of interest is the integrated covariation, which measures the covariance of asset returns over a particular period. Andersen et al. (2003) and Barndorff-Nielsen and Shephard (2004a) suggest estimating the quadratic covariation matrix analogously to the realized variance, by taking the outer product of the observed high-frequency return over the period. This estimation, however, assumes synchronised equidistant data.

In practice, trading is non-synchronous, delivering fresh prices at irregularly spaced times which differ across stocks. In order to estimate the covariance, the data need to be synchronized, meaning that the prices of the q assets need to be collected at the same time stamp. Research of non-synchronous trading has been an active field of financial econometrics in past years: see, for example, Hayashi and Yoshida (2005) and Voev and Lunde (2007). This practical issue induces bias in the estimators and may be partially responsible for the Epps effect (Epps, 1979), a phenomenon of decreasing empirical correlation between the returns of two different stocks with increasing data sampling frequency. Aït-Sahalia et al. (2010) compare various synchronization schemes and find that the estimates do not differ significantly from the estimates using the so called refresh time scheme when dealing with highly liquid assets. The data used further in our study consists of the most liquid U.S. stocks; hence we can restrict ourselves to the refresh time synchronization scheme in our work.

Let $N_{(q)t}$ be the counting process governing the number of observations in the q-th asset up to time t, with times of trades $t_{(q)1}, t_{(q)2}, \ldots$. Following Barndorff-Nielsen et al. (2011), we define the first refresh time as

$$\tau_1 = \max(t_{(1)1}, \ldots, t_{(d)1}), \tag{1.2}$$

for $d = 1, \ldots, q$ assets, and all subsequent refresh times as

$$\tau_{j+1} = \max(t_{(1)N_{(1)\tau_j}+1}, \ldots, t_{(d)N_{(d)\tau_j}+1}), \tag{1.3}$$

with the resulting refresh time sample being of length N. τ_1 is thus the first time that all assets record prices, while τ_2 is the first time that all asset prices are refreshed. In the following analysis, we will always set our clock time to τ_j when using the estimators.

Having synchronized the data, let us denote $\Delta_k p_t = p_{t-1+\tau_k/N} - p_{t-1+\tau_{k-1}/N}$ a discretely sampled vector of k-th intraday log-returns in $[t-1, t]$, with N intraday observations available for each asset q. A simple estimator of realized covariance is then constructed as

$$\widehat{\Sigma}_t^{(RC)} = \sum_{k=1}^{N} (\Delta_k p_t)(\Delta_k p_t)'. \tag{1.4}$$

As shown by Barndorff-Nielsen and Shephard (2004a), realized covariance is a consistent estimator of integrated covariance and is asymptotically mixed normal. However, the estimator is biased and becomes inconsistent in the case that micro-structure noise is present in the data. Sparse sampling is used to mitigate the trade-off between the bias due to noise and variance of the estimator.

To effectively use all available high-frequency data, Zhang et al. (2005) propose to use sub-sampling and averaging for realized variance calculation. In their set-up whole sample is divided into M non-overlapping sub-samples, in each sub-sample realized variance is calculated and averaged across the sub-sampled estimates form the final estimate:

$$\widehat{\Sigma}_t^{(RCSS)} = \frac{1}{M} \sum_{i=1}^{M} \widehat{\Sigma}_{t,i}^{(RC)} \tag{1.5}$$

In addition, the covariance matrix estimated by realized covariance might not necessary be positive semi-definite. To overcome these problems, Barndorff-Nielsen et al. (2011) introduced multivariate realized kernels (MRK) estimator, which guaranties the covariance matrix to be positive semi-definite. Moreover, MRK is more efficient, and it is able to deal with noise. Following Barndorff-Nielsen et al. (2011), the MRK estimator is defined as

$$\widehat{\Sigma}_t^{(MRK)} = \sum_{h=-n}^{n} k\left(\frac{h}{H}\right) \Gamma_h \tag{1.6}$$

where Γ_h stands for h-th realized autocovariance and $k(x)$ is a non-stochastic weight function. In the empirical implementation, we need to choose the kernel function and bandwidth parameter. Following Barndorff-Nielsen et al. (2011), we use a Parzen kernel,[1] which satisfies the smoothness conditions, $K'(0) = K'(1) = 0$, and guarantees $\widehat{\Sigma}_t^{(MRK)}$ to be positive semi-definite. We use the optimal bandwidth derived in Barndorff-Nielsen et al. (2011).

Recently, many new approaches to covariance matrix estimation using high frequency data have emerged in the literature. In addition to estimators used in this study, Realized Co-Range (Bannouh et al., 2009) or Two Scale Realized Covariance (Zhang, 2011) are also becoming increasingly popular. Nowadays, literature also pays attention to disentangling jumps, common jumps and true covariation (see Boudt et al. (2012) or Elst and Veredas (2015)). When the dimension of the problem is high, the estimator of Hautsch et al.

[1]The Parzen kernel function is given by $k(x) = \begin{cases} 1 - 6x^2 + 6x^3 & 0 \le x \le 1/2 \\ 2(1-x)^3 & 1/2 \le x \le 1. \\ 0 & x > 1 \end{cases}$

(2012), which estimates covariance using block-wise Multivariate Realized Kernels, might be of interest.

While the number of recently proposed estimators is growing, we restrict our study on the comparison of the main estimators used in the literature,[2] and focus on the actual estimator of the proposed model.

1.3 Modeling and forecasting multivariate volatility

Modelling and forecasting a conditional covariance matrix of asset returns Σ_t is pivotal to asset allocation, risk management, and option pricing. In order to have a valid multivariate forecasting model, one needs to specify a model that produces symmetric and positive semi-definite covariance matrix predictions. Whereas it is still relatively scarce to use high frequency data in multivariate modelling, the literature dealing with challenging issues is growing quickly. There are three types of approaches proposed recently: modelling the Cholesky factorisation of covariance matrix (Chiriac and Voev, 2011), its matrix-log transformation with the use of latent factors (Bauer and Vorkink, 2011), and direct modelling of the covariance dynamics as a Wishart autoregressive model (Bonato, 2009; Jin and Maheu, 2013).

To ensure positive semi-definiteness of covariance matrix forecasts, we adopt the approach from Chiriac and Voev (2011): we apply the Cholesky decomposition on the covariance matrix. This approach is attractive, as it also helps to reduce the curse of dimensionality, especially in the model structures we are going to use in this study. Following Chiriac and Voev (2011), we model the lower triangular elements of the Cholesky factorization:

$$X_t = vech\left(P_t\right), \tag{1.7}$$

where P_t are Cholesky factors $P_t^{\prime} P_t = \Sigma_t$ and X_t is $m \times 1$ vector, with $m = \frac{q(q+1)}{2}$. Forecasts of the covariance matrix are then obtained by reverse transformation.

1.3.1 Generalized heterogeneous autoregressive (GHAR) model

A simple approximate long-memory model for realized volatility, heterogeneous autoregression (HAR), has been introduced by Corsi (2009). Whereas the approach has been introduced for the univariate volatility modeling, its extension to multivariate volatility has been recently used in the literature (see e.g. Chiriac and Voev (2011) or Bauer and Vorkink (2011)). The original HAR model has an autoregressive structure, and combines volatilities measured at different frequencies (daily, weekly, monthly). Chiriac and Voev (2011) propose a multivariate extension of the HAR to model vector of Cholesky factors X_t, as

$$X_{t+1}^{(1)} = c + \beta^{(1)} X_t^{(1)} + \beta^{(5)} X_t^{(5)} + \beta^{(22)} X_t^{(22)} + \epsilon_t, \quad \epsilon_t \sim i.i.d. \tag{1.8}$$

[2] Realized Covariance sampled at 5 min frequency is the industry standard; Sub-Sampled Realized Covariance enable us to use all data points, resulting in a more efficient estimator, and Multivariate Realized Kernels is able to handle noise and non-synchronous trading

where 1,5, and 22 stands for day, week (5 days) and month (22 days) respectively, c is an $m \times 1$ vector of constants, $\beta^{(\cdot)}$ are scalar parameters, and $X_t^{(\cdot)}$ are averages of lagged daily volatility e.g. $X_t^{(5)} = \frac{1}{5}\sum_{i=0}^{4} X_{t-i}$. To obtain parameter estimates, ordinary least squares (OLS) is used.

One of the disadvantages of this modeling strategy is that we are assuming the same structure for all elements of the Cholesky factors in X_t. Much more importantly, we are leaving a significant amount of information in the error term. One can expect the error term to be heteroscedastic due to volatility of volatility (Corsi et al., 2008) present in the realized measures. More importantly, a common structure of X_t elements may be left unmodelled in residuals. Hence, it may be more natural to estimate the model in Eq. 1.9 as system of equations with some covariance structure of the error terms.

To deal with this problem, we propose to build a system of seemingly unrelated HAR regressions (Zellner, 1962) for all elements of X_t. The advantage of this approach is that we estimate a multivariate HAR model, which will capture the separate dynamics of the variances and covariances, but also possible common structure. Moreover, it will also yield more efficient estimates. As we know, error terms from HAR are heteroscedastic (Corsi et al., 2008), which makes the coefficient estimates less efficient. Moreover, when there is no information about dependence between equations left in the residuals from regression Eq. 1.9, estimator will converge to a simple OLS estimates, as the diagonal weighting matrix in generalized regression will reduce the estimates to OLS. On the other hand, the possible disadvantage is in larger number of parameters to be estimated, which may yield the model unreliable with highly dimensional portfolios.

Let us consider the system of $i = 1, ..., m$ equations, where $m = \frac{q(q+1)}{2}$

$$X_{i,t+1}^{(1)} = \beta_i^{(c)} + \beta_i^{(1)} X_{i,t}^{(1)} + \beta_i^{(5)} X_{i,t}^{(5)} + \beta_i^{(22)} X_{i,t}^{(22)} + \epsilon_{i,t}, \quad \epsilon_{i,t} \sim i.i.d. \tag{1.9}$$

There are m equations representing elements of the Cholesky factors, with T observations. Define the $mT \times 1$ vector of disturbances $\epsilon = (\epsilon_1', \ldots, \epsilon_m')'$, and rewrite the model as

$$\begin{pmatrix} X_{1,t+1}^{(1)} \\ \vdots \\ X_{m,t+1}^{(1)} \end{pmatrix} = \begin{pmatrix} X_{1,t} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & X_{m,t} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix} + \begin{pmatrix} \epsilon_{1,t} \\ \vdots \\ \epsilon_{m,t} \end{pmatrix} \tag{1.10}$$

where $X_{i,t} = \begin{pmatrix} e & X_{i,t}^{(1)} & X_{i,t}^{(5)} & X_{i,t}^{(22)} \end{pmatrix}$ is i-th element of X_t and e being vector of ones, $\beta_i = \begin{pmatrix} \beta_i^{(c)} & \beta_i^{(1)} & \beta_i^{(5)} & \beta_i^{(22)} \end{pmatrix}'$ and $\beta_i^{(c)}$ being estimates of the intercept. It is more convenient to work with this system in the following form:

$$y = Z\beta + \epsilon, \tag{1.11}$$

where $y = \left(X_{1,t+1}^{(1)}, \ldots, X_{m,t+1}^{(1)} \right)'$ and ϵ are of dimension $mT \times 1$, $Z = diag\{X_{1,t}, \ldots, X_{m,t}\}$ is a block diagonal matrix of dimension $mT \times 4m$, and the matrix of parameters $\beta = (\beta_1, \ldots, \beta_m)'$ is of dimension $4m \times 1$.

The disturbances will satisfy strict exogeneity $E[\epsilon|Z] = 0$, but will be correlated across equations, $E[\epsilon_i'\epsilon_j|Z] = \sigma_{ij}I_T$ or

$$\Omega = \begin{pmatrix} \sigma_{11}I_T & \cdots & \sigma_{1m}I_T \\ \vdots & \ddots & \vdots \\ \sigma_{m1}I_T & \cdots & \sigma_{mm}I_T \end{pmatrix} = \Sigma \otimes I_T, \tag{1.12}$$

where $\Sigma = \sigma_{ij}$ for $i, j = 1, ..., m$, $\otimes$ is a Kronecker product and I_T is an identity matrix of dimension $T \times T$. The model parameters are estimated in two step feasible generalized least squares. We run OLS regression in the first step to obtain estimates $\hat{\sigma}_{ij}$ from residuals. In the second step, we run generalized least squares regression using the variance matrix $\widehat{\Omega} = \widehat{\Sigma} \otimes I_T$ as

$$\widehat{\beta} = \left(Z'\widehat{\Omega}^{-1}Z\right)^{-1} Z'\widehat{\Omega}^{-1}y \tag{1.13}$$

The estimator $\widehat{\beta}$ is unbiased, and a consistent estimator of β with asymptotically normal limiting distribution

$$\sqrt{T}\left(\widehat{\beta} - \beta\right) \xrightarrow{d} \mathcal{N}\left(0, \left(\frac{1}{T}Z'\widehat{\Omega}^{-1}Z\right)^{-1}\right) \tag{1.14}$$

While this is a standard estimation technique, we will refrain from discussing any further details about the properties of the generalized least squares estimator.

1.3.2 Competing models

To show the contribution of the GHAR model, we compare the forecasts to several competing alternatives. The first natural choice of benchmark model is a multivariate extension of original HAR. By comparing these two models, we will see the portion of the contribution brought by allowing for correlated residuals in the estimation. Another natural candidate is vector ARFIMA, as Chiriac and Voev (2011) find it to outperform the HAR model slightly, but conclude that HAR performs reasonably well in comparison to VARFIMA. Hence we may have reason to believe that our approach will provide better results than VARFIMA model.

These three main models share the same framework of modeling elements of Cholesky factors from realized covariance matrix. Hence, we also contrast them to two benchmark models, namely popular DCC GARCH [3] of Engle (2002) and risk metrics standard widely used in the business industry. These benchmark models operate on the daily data, so we will have a direct comparison of gains from high-frequency data.

[3] DCC GARCH is an industry standard and we decided to implement it in its original form despite the known problem with consistency of the estimator. For more information about the inconsistency of the DCC see Aielli (2013) an Caporin and McAleer (2013)

HAR

A first, natural competing model to our generalized HAR strategy is multivariate extension of an original HAR, which models vector of Cholesky factors X_t, as

$$X_{t+1}^{(1)} = c + \beta^{(1)} X_t^{(1)} + \beta^{(5)} X_t^{(5)} + \beta^{(22)} X_t^{(22)} + \epsilon_t, \quad \epsilon_t \sim i.i.d. \tag{1.15}$$

where 1,5, and 22 stands for day, week (5 days) and month (22 days) respectively, c is $m \times 1$ vector of constants, $\beta^{()}$ is $m \times 1$ vector of parameters and $X_t^{()}$ are averages of lagged daily volatility e.g. $X_t^{(5)} = \frac{1}{5} \sum_{i=0}^{4} X_{t-i}$. To obtain parameter estimates, OLS is used.

Vector ARFIMA model

A second competing model to the HAR family is the vector autoregressive fractionally integrated moving average (VARFIMA) model of Chiriac and Voev (2011), who use a restricted VARFIMA$(1, d, 1)$ specification to model and forecast dynamics of X_t directly. The authors find that ARFIMA provides a slightly better forecast in comparison to HAR model, which makes it natural candidate to our modeling strategy. We consider the vector ARFIMA model

$$(1 - \phi L)\, D(L)\, [X_t - c] = (1 - \theta L)\, \epsilon_t, \quad \epsilon_t \sim N(0, \Sigma) \tag{1.16}$$

where ϕ and θ are scalars, c is an $m \times 1$ vector of constants and $D(L) = (1 - L)^d I_m$ with a common parameter of fractional integration d for all constituents of X_t. In our case we reject the hypothesis about equality of d; thus we estimated each element of X_t using unique $d_t : D(L) = diag\left\{(1 - L)^{d_1}, \cdots, (1 - L)^{d_m}\right\}$. Hence, we use the model 1 in Chiriac and Voev (2011). In addition, we have experimented with a general VARFIMA(p, d, q), not restricting $p = q = 1$.[4] Comparing the models through information criteria decisively yields VARFIMA$(1, d, 1)$ as the best model; hence we use it as a benchmark to our modeling strategy in the empirical section of the chapter.

RiskMetrics

RiskMetrics of J.P. Morgan Chase, based on an exponentially weighted moving average (EWMA), is a financial industry standard and common benchmark for any volatility model (univariate or multivariate). In our work we use the specification from Longerstaey and Spencer (1996) with decay factor λ set to 0.94. We assume a $q \times 1$ vector of daily returns $r_t = \sum_{k=1}^{n} (\Delta_k p_t)$ for $t = 1, ..., T$ such that $r_t \sim N(\mu_t, \sigma_t^2)$, where μ_t is the conditional mean and σ_t^2 the conditional variance of daily returns. Moreover if we assume $\mu_t = 0$, conditional covariance has the form

$$\sigma_{i,j} = (1 - \lambda) \sum_{t=1}^{T} \lambda^{t-1} r_i r_j. \tag{1.17}$$

[4] These results are available upon request from the authors.

The previous equation can be rewritten in recursive form:

$$\sigma_{i,j,t} = \lambda\sigma_{i,j,t-1} + (1-\lambda) r_{i,t-1} r_{j,t-1} \tag{1.18}$$

where the expression $\sigma_{i,j,t}$ stands for covariance between assets i and j in time t.

DCC-GARCH

The dynamic conditional correlation generalized autoregressive conditional heteroscedasticity (DCC-GARCH) of Engle (2002) is a widely used multivariate GARCH model in practice. It is a generalization of Bollerslev (1990) constant conditional correlation GARCH, with time-varying correlation matrix R. The model is defined as

$$H_t = D_t R_t D_t, \tag{1.19}$$

where D_t is a diagonal matrix of conditional time varying standard deviations, $D_t = diag\left(\sqrt{h_{i,t}}\right)$, and $h_{i,t}$ are univariate GARCH processes, $h_{i,t} = \omega_i + \sum_{p=1}^{P_i} \alpha_{i,p} r^2_{i,t-p} + \sum_{q=1}^{Q_i} \beta_{i,q} h_{i,t-q}$. The dynamics of the correlation matrix are given by transformation:

$$R_t = Q_t^{*-1} Q_t Q_t^{*-1}, \tag{1.20}$$

where $Q_t = \left(1 - \sum_{m=1}^{M} \alpha_m - \sum_{n=1}^{N} \beta_n\right) \bar{Q} + \sum_{m=1}^{M} A_m \left(\epsilon_{t-m} \epsilon^T_{t-m}\right) + \sum_{n=1}^{N} B_n Q_{t-n}$,$\bar{Q}$ is the unconditional covariance matrix of the standardized residuals from the univariate GARCH processes and $Q_t^* = diag\left(\sqrt{q_{ii,t}}\right)$. In our work we use the two-stage estimator presented in Engle (2002) or Engle and Sheppard (2001).

1.4 Data and research design

The dataset consists of tick prices of 15 S&P 500 index constituents with highest liquidity and market capitalization. Final portfolio thus consists[5] of Apple Inc. (AAPL), Exxon Mobile Corp. (XOM), Google Inc. (GOOG), Wal-Mart Stores (WMT), Microsoft Corp. (MSFT), General Electric Co (GE), International Business Machines Corp. (IBM), Johnson & Johnson (JNJ), Chevron Corp. (CVX), Procter & Gamble (PG), Pfizer Inc. (PFE), AT&T Inc. (T), Wells Fargo & Co (WFC), JP Morgan Chase & Co (JPM) and Coca-Cola Co. (KO). We obtain 390, 78, 39, 26 and 19 time-synchronized intraday observations using refresh-time, resulting in 1, 5, 10, 15 and 20 minute intraday returns. Besides 1 to 20 minute returns we also construct open-to-close returns that are used for RiskMetrics and DCC-GARCH models. Moreover, we create sub-portfolios consisting of 5, 10, and 15 assets (assets chosen according to market capitalization). Hence, in total, we study 18 different datasets.

The sample covers the period from July, 1 2005 to January, 3 2012 (1623 trading days), and we consider trades between 9:30 to 16:00 EST time. To ensure sufficient liquidity on

[5] Assets are ordered according to market capitalization.

the market we explicitly exclude weekends and holidays (New Year's Day, Independence Day, Thanksgiving Day, Christmas). For estimation and forecasting purposes we divide our sample into in-sample, spanning from July, 1 2005 to July, 9 2008 and out-of-sample July, 10 2008 to January, 3 2012. For the forecasting, we use rolling window estimation with fixed length of 750 days. Summary statistics of all returns are presented in the Appendix D.

Accuracy of the forecasts is evaluated primarily according to economic criteria. The rationale behind is the importance of well-conditioned and invertible forecasts rather than focusing on unbiasedness, as an unbiased forecast does not necessarily translate into an unbiased inverse (Bauwens et al., 2012). As a robustness check we also provide ranking of the models based on statistical loss functions.

1.4.1 Economic forecasts evaluation

For economic evaluation of volatility forecasts, we use the approach of Markowitz (1952). There are two possibilities of constructing an optimal portfolio. In the first one we specify expected portfolio return and try to find assets weights minimizing the risk. In the second one the expected return of the portfolio is maximized according to a certain risk. Asset weights, $w = (w_1, \dots, w_q)'$, maximizing utility of risk averse investor can be found by solving the following problem:

$$\begin{aligned} \min_{w_{t+1}} \quad & w'_{t+1}\widehat{\Sigma}_{t+1|t}w_{t+1} \\ \text{s.t.} \quad & l'w_{t+1} = 1 \\ & w'_{t+1}\widehat{\mu}_{t+1|t} = \mu_P \end{aligned} \tag{1.21}$$

where w_{t+1} is a $q \times 1$ vector of assets weights, $\widehat{\Sigma}_{t+1|t}$ represents a covariance matrix forecast, l denotes a $q \times 1$ vector of ones, $\widehat{\mu}_{t+1|t}$ is a vector of mean forecasts and μ_P stands for portfolio return. Once the optimization problem is solved for different risk levels we are able to construct efficient frontier. Markowitz-type portfolio relies heavily on mean forecasts. As these forecasts might be noisy, portfolio weights and variance can become notably sensitive to changes in assets mean. To overcome these difficulties we also consider problem of finding the Global Minimum Variance Portfolio (GMVP). Specification of the optimization problem is similar to Markowitz set-up:

$$\begin{aligned} \min_{w_{t+1}} \quad & w'_{t+1}\widehat{\Sigma}_{t+1|t}w_{t+1} \\ \text{s.t.} \quad & l'w_{t+1} = 1 \end{aligned} \tag{1.22}$$

which can be solved analytically[6]

$$w_{t+1}^{GMV} = \frac{\widehat{\Sigma}_{t+1|t}^{-1} l}{l'\widehat{\Sigma}_{t+1|t}^{-1} l}, \tag{1.23}$$

[6] Kempf and Memmel (2006)

with expected return variance being

$$\sigma^{2GMV}_{t+1} = w^{GMV\prime}_{t+1}\widehat{\Sigma}_{t+1|t}w^{GMV}_{t+1} = \frac{1}{l'\widehat{\Sigma}^{-1}_{t+1|t}l}. \tag{1.24}$$

1.4.2 Statistical forecasts evaluation

For statistical evaluation of covariance forecasts, we employ Root Mean Squared Error (RMSE) loss functions based on the Frobenius norm [7]. As a volatility proxy we use Realized Covariance, Sub-Sampled Realized Covariance (RCOV SS) and Multivariate Realized Kernels estimates at given frequencies i.e. to calculate loss function for forecasts based on 5 minutes Realized Covariance we use Realized Covariance estimates based on 5 minutes data as a benchmark. In case of DCC-GARCH and RiskMetrics forecasts we calculate loss functions using all RCOV, RCOV SS and MRK estimates at all frequencies. The measures are calculated for the $t = 1, \ldots, T$ forecasts as

$$e_{t,t+h} = \Sigma_{t+h} - \widehat{\Sigma}_{t+h|t} \tag{1.25}$$

$$\mathcal{L}^{RMSE} = \sqrt{\frac{1}{T-1}\sum_{t=1}^{T}\sum_{i,j}\left|e_{t_{i,j}}\right|^2} \tag{1.26}$$

where $\widehat{\Sigma}_{t+h|t}$ is a covariance matrix forecast and Σ_{t+h} is the volatility proxy.

To test the significant differences of competing models, we use the Model Confidence Set (MCS) methodology of Hansen et al. (2011). Given a set of forecasting models, $\mathcal{M}_0$, we identify the model confidence set $\widehat{\mathcal{M}}^*_{1-\alpha} \subset \mathcal{M}_0$, which is the set of models that contain the "best" forecasting model given a level of confidence α. For a given model $i \in \mathcal{M}_0$, the p-value is the threshold confidence level. Model i belongs to the MCS only if $\widehat{p}_i \geq \alpha$. MCS methodology repeatedly tests the null hypothesis of equal forecasting accuracy

$$H_{0,\mathcal{M}} : E[\mathcal{L}_{i,t} - \mathcal{L}_{j,t}] = 0, \qquad \text{for all } i, j \in \mathcal{M}$$

with $L_{i,t}$ being an appropriate loss function of the i-th model. Starting with the full set of models, $\mathcal{M} = \mathcal{M}_0$, this procedure sequentially eliminates the worst-performing model from $\mathcal{M}$ when the null is rejected. The surviving set of models then belong to the model confidence set $\widehat{\mathcal{M}}^*_{1-\alpha}$. Following Hansen et al. (2011), we implement the MCS using a stationary bootstrap with an average block length of 10 days.[8]

[7] Frobenius norm of $m \times n$ matrix A is defined as $\|A\|_F{}^2 = \sum_{i,j} |a_{i,j}|^2$

[8] We have used different block lengths, including those dependent on the forecasting horizons, to assess the robustness of the results, without any change in the final results. These results are available from the authors upon request.

1.5 Results

For clarity of presentation, we begin with a discussion of the results of one-step-ahead forecasts for the portfolio of five stocks (AAPL, XOM, GOOG, WMT, MSFT), whereas we leave portfolios of ten and fifteen stocks and also five- and ten-step ahead forecasts as a robustness check showing that the methodology also works well at larger dimensions and different forecasting horizons. Focusing on the economic evaluation, we first discuss the results from GMVP,[9] followed by Markowitz approach and statistical evaluation.

We present GMVP comparison through cumulative and annualized risk. In the cumulative approach we use covariance forecasts for daily rebalancing of our portfolio: at each step we calculate optimal asset weights and using these weights we calculate corresponding daily portfolio risk. The results presented in Table 1.1 are sums of portfolio risk $\sigma_{cum.}$ for the whole out-of sample period. Table 1.1 is divided into seven parts according to realized measures and frequencies used for the calculation. For RiskMetrics and DCC-GARCH corresponding $\sigma_{cum.}$ are constant for all frequencies because they are calculated using open-close returns. We present the results of DCC-GARCH and RiskMetrics in all columns of Table 1.1 so we can compare performance of covariance based models estimated on different frequencies with daily data based models.

From the Table 1.1 we can see that the model with the best performance and thus lowest level of risk is GHAR. We can also observe that for various frequencies on which realized measures are calculated DCC-GARCH outperformed covariance based models. However, these results do not indicate superiority of DCC-GARCH compared to covariance based models, but highlight the importance of selecting realized measures properly.

Table 1.1: Cumulative version of GMVP - portfolio of 5 stocks

	MRK	RCOV		Sub-Sampled RCOV			
	1 min	1 min	5 min	5 min	10 min	15 min	20 min
DCC	**30.50**	**30.50**	**30.50**	**30.50**	**30.50**	30.50	30.50
RiskMetrics	40.64	40.64	40.64	40.64	40.64	40.64	40.64
VARFIMA	30.76	34.47	32.44	32.84	31.04	29.86	29.31
GHAR	30.60	34.14	32.22	32.53	30.83	**29.65**	**29.08**
HAR	31.42	34.84	33.05	33.35	31.61	30.50	29.99

Note: Model with the overall best performance is highlighted; for the given frequency model with the lowest risk is presented in bold; values represents percentage level of risk

A disadvantage of model comparison according to cumulative risk is daily rebalancing implying high transaction costs. A more comprehensive method of model comparison is to use annualized portfolio risks. For annualized GMVP calculation we use annualized realized covariance of the whole out–off–sample period calculated as $RCOV_{annualized} =$

[9] With shortselling allowed.

$\frac{\sum_{i=1}^{T} RCOV_i}{\frac{T}{250}}$. In Table 1.2 we present the results for annualized version of GMVP.

Table 1.2: Annualized version of GMVP - portfolio of 5 stocks

	MRK	RCOV		Sub-Sampled RCOV			
	1 min	1 min	5 min	5 min	10 min	15 min	20 min
DCC	17.38	**17.38**	**17.38**	**17.38**	17.38	17.38	17.38
RiskMetrics	23.13	23.13	23.13	23.13	23.13	23.13	23.13
VARFIMA	17.62	19.39	18.44	18.61	17.68	17.04	16.77
GHAR	**17.32**	19.08	18.08	18.27	**17.36**	**16.69**	**16.38**
HAR	18.01	19.61	18.79	18.91	18.01	17.40	17.14

Note: Model with the overall best performance is highlighted; for the given frequency model with the lowest risk is presented in bold; values represents percentage level of risk

Similar to cumulative GMVP, model with the overall lowest achievable risk is GHAR. Remaining results from the Table 1.1 and Table 1.2 partly match the results presented in Chiriac and Voev (2011). The model that scored second is VARFIMA followed by HAR for Sub-Sampled RCOV estimated at 15- and 20- minute frequencies. For the remaining frequencies and realized measures, DCC-GARCH outperform covariance based models. Overall we can say that covariance-based models with proper choice of realized measure outperform return-based models.

To assess the performance of the models not only from the risk minimizing point of view but also return maximization, we present efficient frontiers. In contrast to GMVP we do not allow short selling here.[10] For the calculation of the efficient frontiers we use annualized forecasts of covariance matrices and annualized returns.

Similar to the results from the GMVP evaluation model with the best risk-return tradeoff is the model proposed in this chapter: GHAR. The second-best-performing model is VARFIMA, followed by HAR. From Figure 1.1 we can see that for estimates at 1 minute RCOV and 5 minutes RCOV the score of DCC-GARCH is better than all covariance based models, which is not in line with results presented in Chiriac and Voev (2011) where DCC-GARCH ended in the penultimate position. We can attribute this difference to a different dataset and period that includes financial crisis during which periods of high intraday volatility are observable.

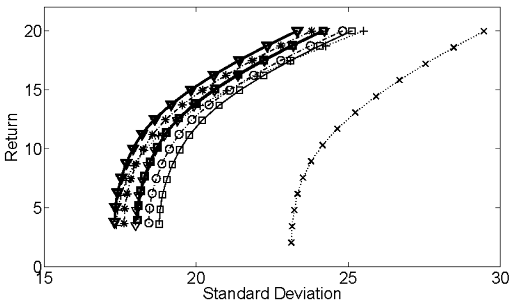
As a robustness check to the economic evaluation, we also provide results from a statistical comparison of forecasting performance of the competing models. In the Table 1.3 a comparison based on the RMSE loss function is presented.

From the RMSE perspective the lowest error is shown by the HAR model, followed by VARFIMA and GHAR. These models always belong to 5% MCS irrespective of the realized measure used for comparison. The worst performance has RiskMetrics, which does not belong to 5% MCS in two cases and it has the highest RMSE in 5 out of 7 cases.

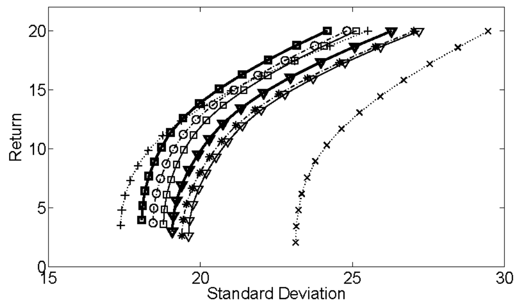
[10] In case the short-selling is allowed the ranking of the models is unchanged only the magnitude differ.

Figure 1.1: Efficient frontiers - portfolio of 5 stocks

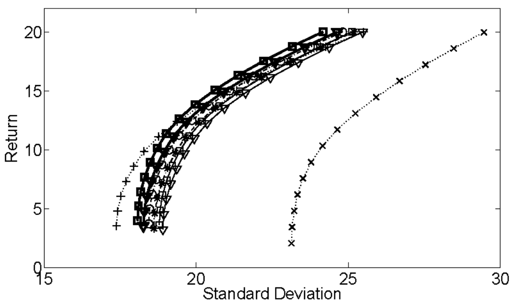
(a) RCOV 5min vs. MRK



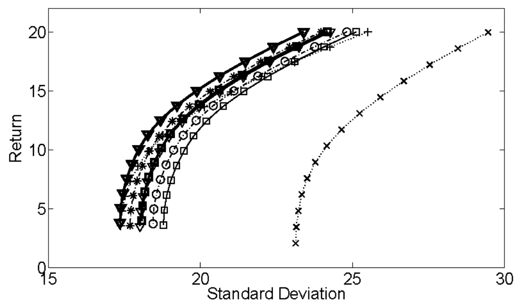
(b) RCOV 5 min vs. RCOV 1 min



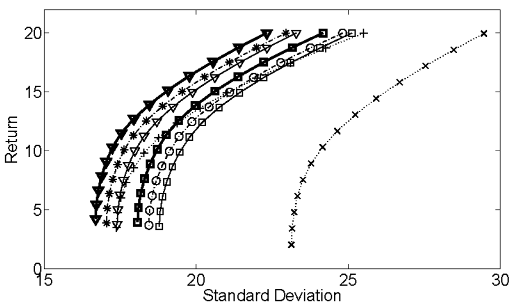
(c) RCOV 5 min vs. RCOV SS 5 min



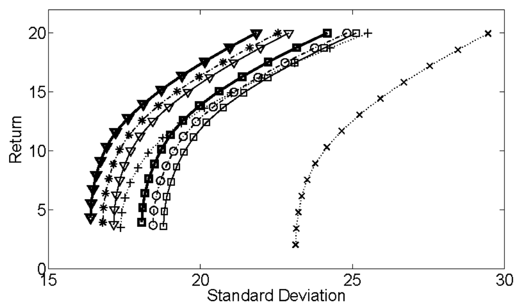
(d) RCOV 5 min vs. RCOV SS 10 min



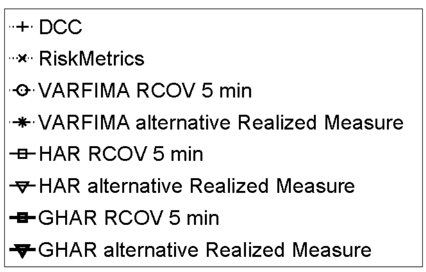
(e) RCOV 5 min vs. RCOV SS 15 min



(f) RCOV 5 min vs. RCOV SS 20 min



(g) legend



Note: Figure displays efficient frontiers of various competing models of portfolio of five stocks based on one-step-ahead forecasts.

Table 1.3: RMSE – portfolio of 5 stocks

	MRK	RCOV		Sub-Sampled RCOV			
	1 min	1 min	5 min	5 min	10 min	15 min	20 min
DCC	1.593	1.730	1.914	1.707	1.547	1.481	1.474
RiskMetrics	1.668	1.728	1.866	1.709	1.646	1.636	1.633
VARFIMA	1.406	1.537	1.682	1.473	1.363	1.331	1.328
GHAR	1.490	1.401	1.740	1.509	1.438	1.430	1.445
HAR	1.190	1.100	1.380	1.162	1.125	1.144	1.158

Note: Values are scaled by 10^{-3}; highlighted cells belongs to 5% MCS

1.5.1 Robustness check

Having discussed the results of one-step-ahead forecasts for portfolio consisting of five stocks, we now turn to evaluation of one-step-ahead forecasts for portfolio consisting of ten (AAPL, XOM, GOOG, WMT, MSFT, GE, IBM, JNJ, CVX, PG), and fifteen (AAPL, XOM, GOOG, WMT, MSFT, GE, IBM, JNJ, CVX, PG, PFE, T, WFC, JPM, KO) stocks and five and ten-step ahead forecasts for portfolios consisting of five, ten and fifteen stocks. We will concentrate on the main differences compared to the smaller portfolio, as we use these results as a robustness check. We also relegate the Tables and Figures to Appendix A: 1 step ahead forecasts, Appendix B: 5 step ahead forecasts and Appendix C: 10 step ahead forecasts.

Portfolio of 10 and 15 stocks

According to GMVP criteria for portfolio consisting of ten stocks, results do not differ from results obtained in portfolio of five stocks. The model with the lowest cumulative and annualized risk is GHAR, estimated on 20-minute Sub-Sampled RCOV. In the case of the portfolio consisting of fifteen stocks, the only difference is that GHAR estimated on MRK covariance matrices outperformed DCC-GARCH.

From the risk-return trade-off point of view there is notable difference for portfolio consisting of ten stocks when the data of higher frequencies (1,5 and 10 minutes) are used. For these frequencies, the model with the best risk-return trade-off is DCC-GARCH. The order of the remaining models is identical to the portfolio of five stocks: GHAR followed by VARFIMA and HAR. If the 15-minute data are used for optimization, GHAR share first place with DCC-GARCH. These two models are closely followed by VARFIMA and HAR. For the 20-minute data ordering of the models is similar to the portfolio consisting of 5 stocks.

Concentrating on statistical evaluation, results of RMSE model comparison for the portfolio consisting of ten stocks are almost identical to results for the portfolio of five stocks, the only difference beeing that RiskMetrics does not belong to 5% MCS in any of the cases. On the other hand, a notable difference occurs in a comparison of the portfolio consisting of fifteen stocks, where GHAR belongs to 5% MCS only in one case (estimated at 5-minute RCOV) and DCC-GARCH and RiskMetrics do not belong to 5% MCS at all. We address unambiguous results of statistical evaluation to problem of selecting the "correct" proxy. These results are also consistent with findings in Kyj et al. (2010), who show that for large portfolios, the pure high frequency based covariance forecasts need to be conditioned in order to achieve the benefits of the high frequency data.

This points us to the result, that unmodelled dependence from HAR and VARFIMA models is increasing with increasing dimension of the portfolio. Hence the GHAR model delivers significant economic gains with increasing dimension of portfolio.

5-step & 10-step ahead forecasts [11]

Extension of forecasting horizon from one to five/ten days does not substantially change the results of our analysis. The only notable difference is absence of GHAR in 5% MCS in the case of ten-step ahead forecasts of portfolio consisting of fifteen stocks. Remaining results supports our previous findings that application of seemingly unrelated regression for HAR estimation delivers significant economic gains regardless the size of the portfolio and/or forecasting horizon.[12]

1.6 Conclusion

In this chapter we propose to employ the seemingly unrelated regression of Zellner (1962) to estimate multivariate extension of the heterogeneous autoregression model in order to improve the variance matrix forecasts. The resulting model, generalized HAR (GHAR), inherit all the favourable properties of HAR, and provides us with more efficient estimator that accounts for otherwise hidden dependencies among variables.

In our setup we closely follow Chiriac and Voev (2011) and model elements of Cholesky decomposed covariance matrices to test the economic and statistical value of the proposed modelling strategy. Moreover, we perform our analysis on portfolios consisting of five, ten and fifteen assets, we include three covariance matrix estimators (realized covariation, sub-sampled realized covariation and multivariate realized kernels), and we obtain covariance matrix estimates using high-frequency data of five different frequencies (1,5,10,15 and 20 minutes). Overall, we test the performance of GHAR estimator on 15 different high-frequency datasets. The resulting forecasts of GHAR prove to perform significantly better than benchmark models according to Global Minimum Variance Portfolio and Mean-Variance evaluation criteria irrespective of frequency or size of the portfolio. Whereas our study focuses on more important economic evaluation of the forecasts, statistical evaluation is used as a robustness check of the results. According to statistical criteria for comparison of models, we find that GHAR is not systematically dominated by any benchmark model, which is supportive result for economic evaluation.

[11] To make the results comparable we scale them according to forecasting horizon

[12] To make the results comparable, we scale them according to the forecasting horizon.

Appendix A: 1 step ahead forecasts

Table 1.4: GMVP - portfolio of 10 stocks

	Cumulative						
	MRK	RCOV		Sub-Sampled RCOV			
	1 min	1 min	5 min	5 min	10 min	15 min	20 min
DCC	**22,14**	**22,14**	**22,14**	**22,14**	**22,14**	22,14	22,14
RiskMetrics	42,15	42,15	42,15	42,15	42,15	42,15	42,15
VARFIMA	23.34	27.70	24.75	25.64	23.82	22.52	21.85
GHAR	22.50	26.71	23.90	24.79	22.98	**21.66**	**20.98**
HAR	24.28	28.30	25.66	26.40	24.63	23.39	22.79
	Annualized						
	MRK	RCOV		Sub-Sampled RCOV			
	1 min	1 min	5 min	5 min	10 min	15 min	20 min
DCC	13.12	**13.12**	**13.12**	**13.12**	13.12	13.12	13.12
RiskMetrics	24.32	24.32	24.32	24.32	24.32	24.32	24.32
VARFIMA	13.74	15.76	14.40	14.84	13.90	13.21	12.88
GHAR	**12.82**	15.00	13.53	14.04	**13.03**	**12.30**	**11.91**
HAR	14.31	16.14	14.96	15.31	14.40	13.74	13.43

Note: Model with the overall best performance is highlighted; for the given frequency model with the lowest risk is presented in bold; values represents percentage level of risk

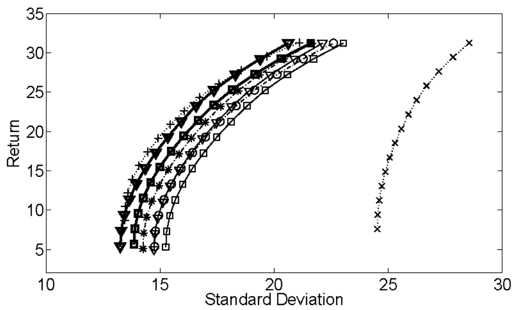
Table 1.5: RMSE – portfolio of 10 stocks

	MRK	RCOV		Sub-Sampled RCOV			
	1 min	1 min	5 min	5 min	10 min	15 min	20 min
DCC	3.242	3.624	3.896	3.600	3.162	3.044	3.085
RiskMetrics	3.808	4.006	4.167	3.949	3.803	3.822	3.846
VARFIMA	2.592	3.028	3.228	2.903	2.551	2.494	2.539
GHAR	3.101	3.109	3.639	3.237	2.988	2.965	3.057
HAR	2.295	2.271	2.837	2.405	2.181	2.213	2.307

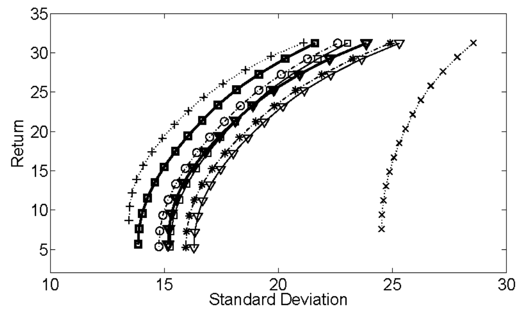
Note: Values are scaled by 10^{-3}; highlighted cells belongs to 5% MCS

Figure 1.2: Efficient frontiers - portfolio of 10 stocks

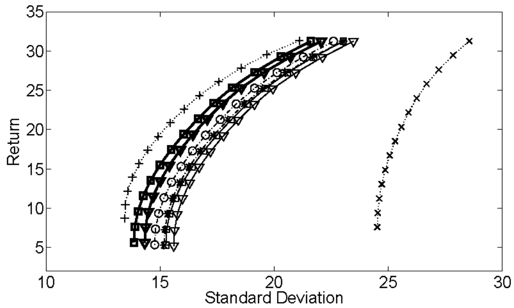
(a) RCOV 5min vs. MRK



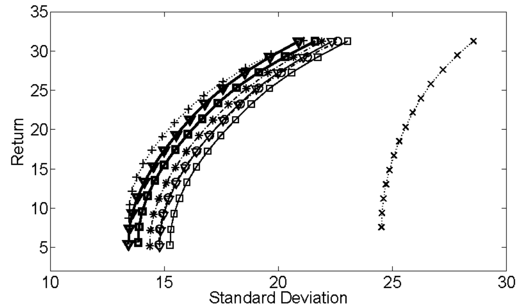
(b) RCOV 5 min vs. RCOV 1 min



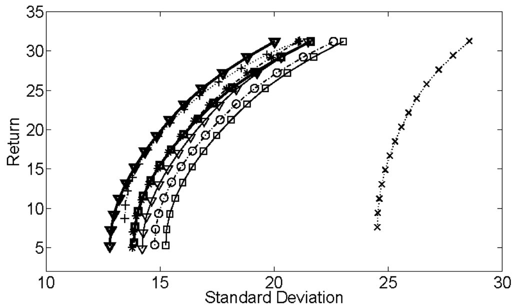
(c) RCOV 5 min vs. RCOV SS 5 min



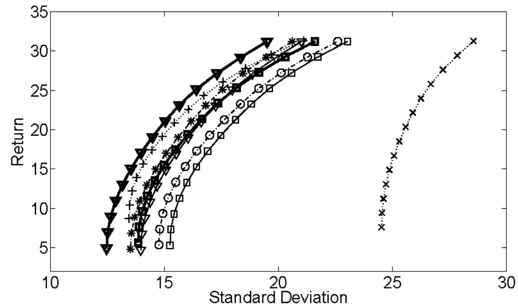
(d) RCOV 5 min vs. RCOV SS 10 min



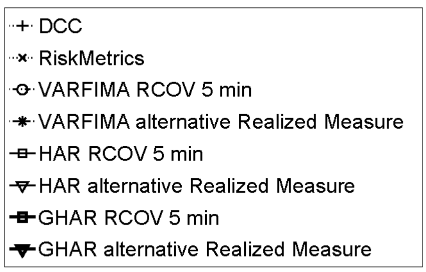
(e) RCOV 5 min vs. RCOV SS 15 min



(f) RCOV 5 min vs. RCOV SS 20 min



(g) legend



Note: Figure displays efficient frontiers of various competing models of portfolio of ten stocks based on one-step-ahead forecasts.

Table 1.6: GMVP - portfolio of 15 stocks

	MRK	RCOV		Sub-Sampled RCOV			
	1 min	1 min	5 min	5 min	10 min	15 min	20 min
Cumulative							
DCC	20.72	**20.72**	**20.72**	**20.72**	**20.72**	20.72	20.72
RiskMetrics	56.67	56.67	56.67	56.67	56.67	56.67	56.67
VARFIMA	21.34	25.63	22.71	23.71	21.91	20.62	19.93
GHAR	**20.37**	24.46	21.75	22.59	20.90	**19.66**	**18.97**
HAR	22.25	26.21	23.52	24.42	22.69	21.47	20.83
Annualized							
DCC	12.64	**12.64**	**12.64**	**12.64**	12.64	12.64	12.64
RiskMetrics	32.19	32.19	32.19	32.19	32.19	32.19	32.19
VARFIMA	12.88	14.80	13.52	13.99	13.06	12.40	12.07
GHAR	**11.64**	13.82	12.39	12.86	**11.91**	**11.22**	**10.83**
HAR	13.43	15.21	14.06	14.45	13.56	12.92	12.62

Note: Model with the overall best performance is highlighted; for the given frequency model with the lowest risk is presented in bold; values represents percentage level of risk

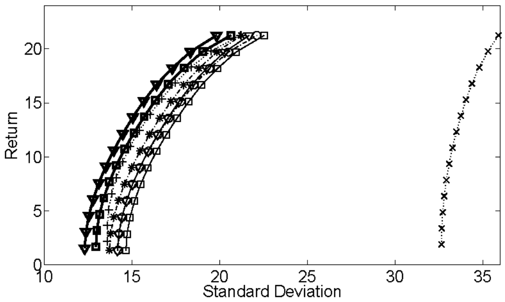
Table 1.7: RMSE – portfolio of 15 stocks

	MRK	RCOV		Sub-Sampled RCOV			
	1 min	1 min	5 min	5 min	10 min	15 min	20 min
DCC	5.323	5.601	6.064	5.793	5.158	5.023	5.058
RiskMetrics	11.905	11.881	12.030	11.902	11.952	12.044	12.030
VARFIMA	4.555	4.809	5.207	4.900	4.374	4.276	4.323
GHAR	5.881	5.352	6.342	5.918	5.565	5.521	5.677
HAR	4.285	3.599	4.832	4.226	3.948	4.005	4.150

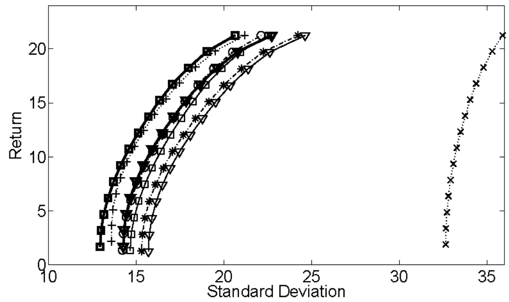
Note: Values are scaled by 10^{-3}; highlighted cells belongs to 5% MCS

Figure 1.3: Efficient frontiers - portfolio of 15 stocks

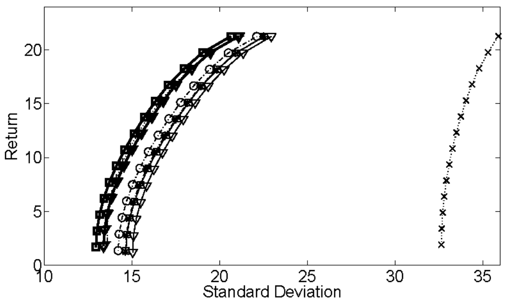
(a) RCOV 5 min vs. MRK



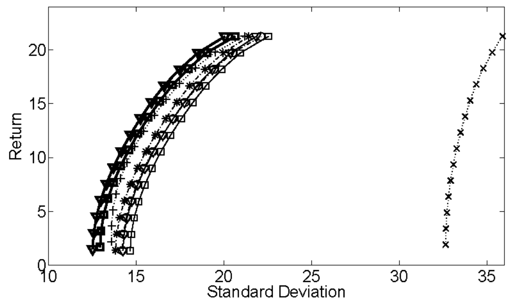
(b) RCOV 5 min vs. RCOV 1 min



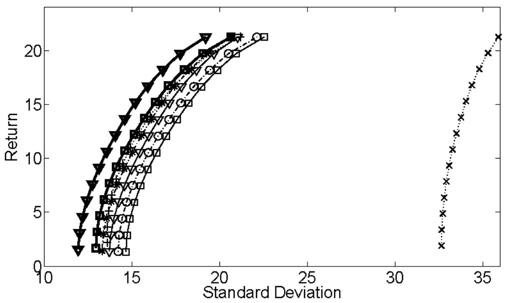
(c) RCOV 5 min vs. RCOV SS 5 min



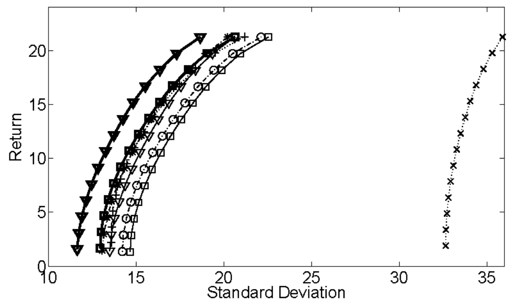
(d) RCOV 5 min vs. RCOV SS 10 min



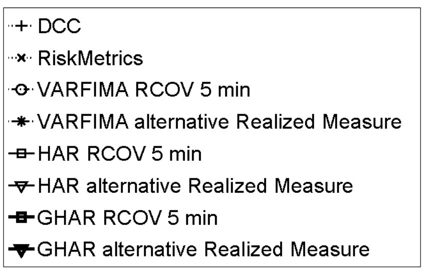
(e) RCOV 5 min vs. RCOV SS 15 min



(f) RCOV 5 min vs. RCOV SS 20 min



(g) legend



Note: Figure displays efficient frontiers of various competing models of portfolio of fifteen stocks based on one-step-ahead forecasts.

Appendix B: 5 step ahead forecasts

Table 1.8: GMVP - portfolio of 5 stocks

	MRK	RCOV		Sub-Sampled RCOV			
Cumulative	1 min	1 min	5 min	5 min	10 min	15 min	20 min
DCC	30.50	**30.50**	**30.50**	**30.50**	**30.50**	30.50	30.50
RiskMetrics	40.61	40.61	40.61	40.61	40.61	40.61	40.61
VARFIMA	30.53	34.06	32.09	32.49	30.78	29.64	29.10
GHAR	**30.49**	33.88	32.07	32.36	30.72	**29.54**	**28.96**
HAR	31.30	34.62	32.86	33.19	31.47	30.38	29.87
Annualized	MRK	RCOV		Sub-Sampled RCOV			
	1 min	1 min	5 min	5 min	10 min	15 min	20 min
DCC	17.38	**17.38**	**17.38**	**17.38**	17.38	17.38	17.38
RiskMetrics	23.17	23.17	23.17	23.17	23.17	23.17	23.17
VARFIMA	17.28	19.02	18.06	18.24	17.35	16.73	16.73
GHAR	**17.18**	18.87	17.93	18.12	**17.23**	**16.57**	**16.57**
HAR	17.85	19.45	18.63	18.75	17.86	17.25	17.25

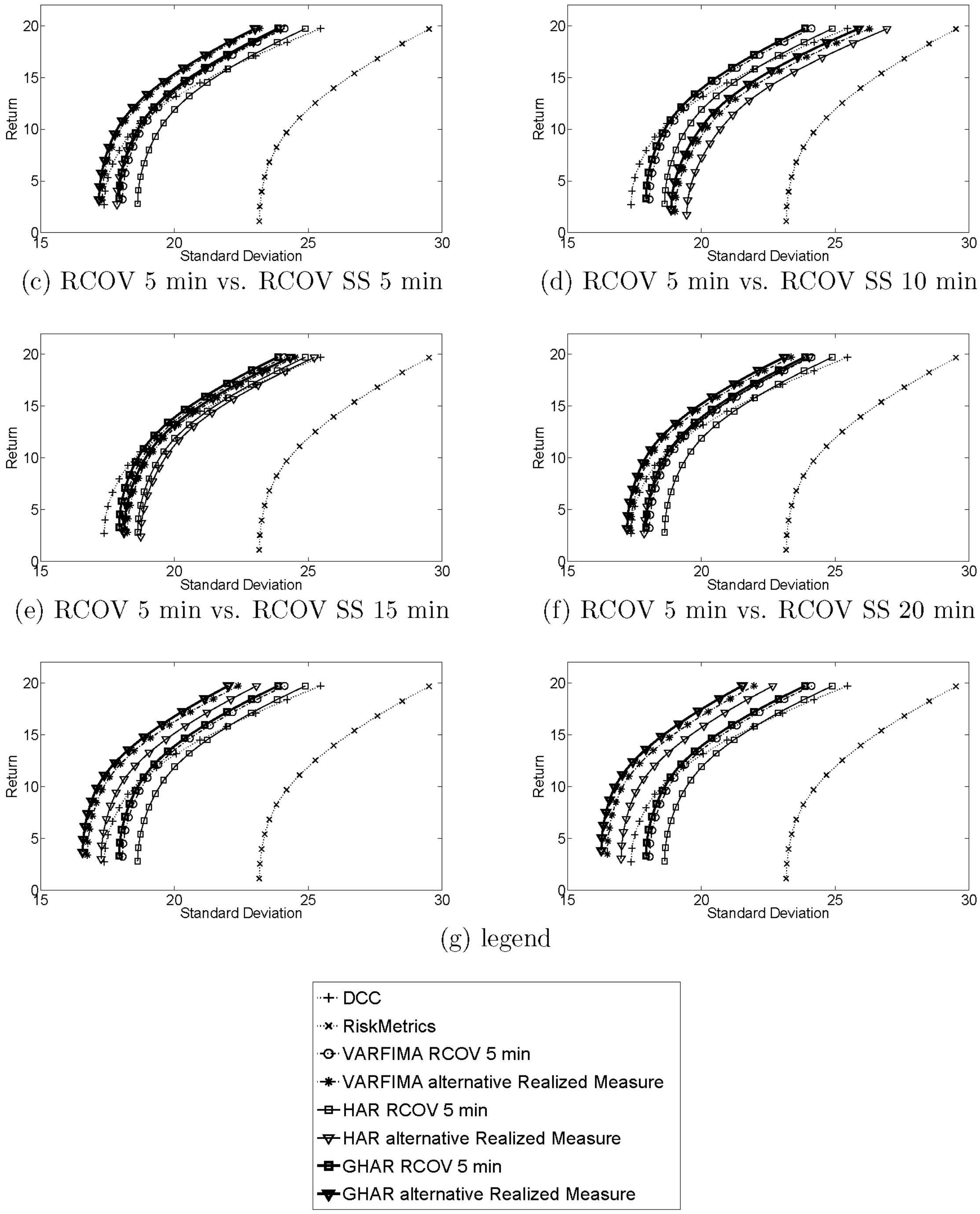
Note: Model with the overall best performance is highlighted; for the given frequency model with the lowest risk is presented in bold; values represents percentage level of risk; values are scaled by forecasting horizon

Table 1.9: RMSE – portfolio of 5 stocks

	MRK	RCOV		Sub-Sampled RCOV			
	1 min	1 min	5 min	5 min	10 min	15 min	20 min
DCC	1.193	1.293	1.376	1.288	1.152	1.081	1.079
RiskMetrics	1.296	1.317	1.330	1.314	1.290	1.288	1.285
VARFIMA	1.043	1.023	1.153	1.055	0.993	0.968	0.978
GHAR	1.261	1.195	1.382	1.273	1.206	1.174	1.189
HAR	1.024	0.980	1.100	1.028	0.968	0.951	0.966

Note: Values are scaled by 10^{-3} and by forecasting horizon; highlighted cells belongs to 5% MCS

Figure 1.4: Efficient frontiers - portfolio of 5 stocks



(a) RCOV 5min vs. MRK

(b) RCOV 5 min vs. RCOV 1 min

(c) RCOV 5 min vs. RCOV SS 5 min

(d) RCOV 5 min vs. RCOV SS 10 min

(e) RCOV 5 min vs. RCOV SS 15 min

(f) RCOV 5 min vs. RCOV SS 20 min

(g) legend

Note: Figure displays efficient frontiers of various competing models of portfolio of five stocks based on five-step-ahead forecasts.

Table 1.10: GMVP - portfolio of 10 stocks

	Cumulative						
	MRK	RCOV		Sub-Sampled RCOV			
	1 min	1 min	5 min	5 min	10 min	15 min	20 min
DCC	**22.10**	**22.10**	**22.10**	**22.10**	**22.10**	22.10	22.10
RiskMetrics	42.12	42.12	42.12	42.12	42.12	42.12	42.12
VARFIMA	23.11	27.25	24.44	25.27	23.55	22.30	21.65
GHAR	22.33	26.45	23.72	24.59	22.80	**21.50**	**20.82**
HAR	24.25	28.14	25.56	26.30	24.57	23.35	22.75
	Annualized						
	MRK	RCOV		Sub-Sampled RCOV			
	1 min	1 min	5 min	5 min	10 min	15 min	20 min
DCC	13.07	**13.07**	**13.07**	**13.07**	13.07	13.07	13.07
RiskMetrics	24.36	24.36	24.36	24.36	24.36	24.36	24.36
VARFIMA	13.38	15.36	14.03	14.44	13.54	12.88	12.54
GHAR	**12.67**	14.80	13.38	13.87	**12.88**	**12.16**	**11.78**
HAR	14.15	15.99	14.81	15.15	14.24	13.59	13.28

Note: Model with the overall best performance is highlighted; for the given frequency model with the lowest risk is presented in bold; values represents percentage level of risk; values are scaled by forecasting horizon

Table 1.11: RMSE – portfolio of 10 stocks

	MRK	RCOV		Sub-Sampled RCOV			
	1 min	1 min	5 min	5 min	10 min	15 min	20 min
DCC	2.487	2.683	2.773	2.690	2.402	2.290	2.309
RiskMetrics	3.232	3.250	3.217	3.222	3.242	3.278	3.267
VARFIMA	1.952	1.966	2.166	2.024	1.867	1.833	1.872
GHAR	2.598	2.480	2.759	2.611	2.481	2.445	2.501
HAR	1.950	1.881	2.103	1.984	1.845	1.826	1.877

Note: Values are scaled by 10^{-3} and by forecasting horizon; highlighted cells belongs to 5% MCS

Table 1.12: GMVP - portfolio of 15 stocks

	Cumulative						
	MRK	RCOV		Sub-Sampled RCOV			
	1 min	1 min	5 min	5 min	10 min	15 min	20 min
DCC	20.70	**20.70**	**20.70**	**20.70**	**20.70**	20.70	20.70
RiskMetrics	56.64	56.64	56.64	56.64	56.64	56.64	56.64
VARFIMA	21.23	25.28	22.52	23.44	21.75	20.52	19.86
GHAR	**20.31**	24.30	21.65	22.45	20.83	**19.62**	**18.92**
HAR	22.31	26.13	23.51	24.40	22.72	21.51	20.89
	Annualized						
	MRK	RCOV		Sub-Sampled RCOV			
	1 min	1 min	5 min	5 min	10 min	15 min	20 min
DCC	12.60	**12.60**	12.60	**12.60**	12.60	12.60	12.60
RiskMetrics	32.25	32.25	32.25	32.25	32.25	32.25	32.25
VARFIMA	12.53	14.43	13.17	13.60	12.72	12.07	11.74
GHAR	**11.53**	13.66	**12.26**	12.70	**11.79**	**11.12**	**10.73**
HAR	13.29	15.07	13.94	14.31	13.42	12.78	12.48

Note: Model with the overall best performance is highlighted; for the given frequency model with the lowest risk is presented in bold; values represents percentage level of risk; values are scaled by forecasting horizon

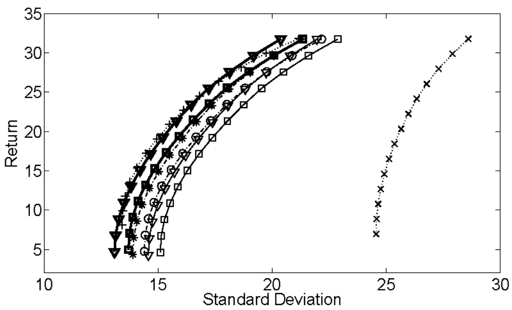
Table 1.13: RMSE – portfolio of 15 stocks

	MRK	RCOV		Sub-Sampled RCOV			
	1 min	1 min	5 min	5 min	10 min	15 min	20 min
DCC	4.110	4.251	4.384	4.329	3.992	3.919	3.949
RiskMetrics	11.404	11.318	11.262	11.260	11.487	11.599	11.573
VARFIMA	3.453	3.223	3.596	3.422	3.239	3.201	3.283
GHAR	4.913	4.490	4.961	4.821	4.644	4.590	4.706
HAR	3.575	3.216	3.644	3.489	3.331	3.314	3.421

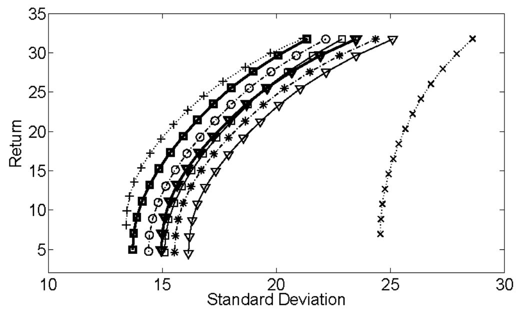
Note: Values are scaled by 10^{-3} and by forecasting horizon; highlighted cells belongs to 5% MCS

Figure 1.5: Efficient frontiers - portfolio of 10 stocks

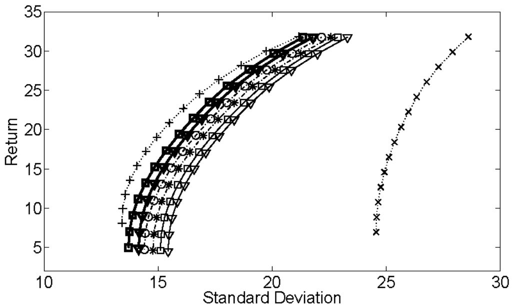
(a) RCOV 5min vs. MRK



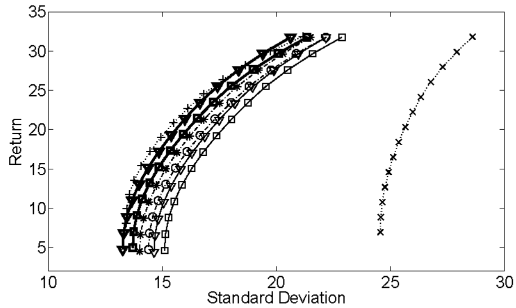
(b) RCOV 5 min vs. RCOV 1 min



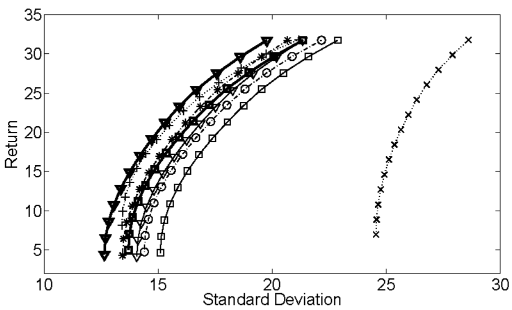
(c) RCOV 5 min vs. RCOV SS 5 min



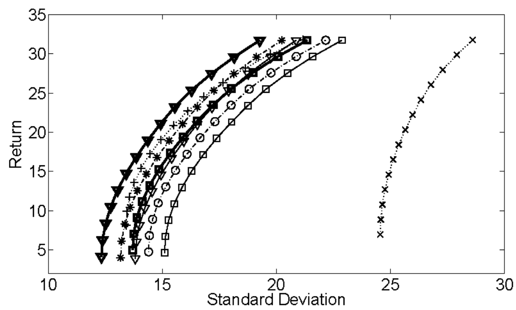
(d) RCOV 5 min vs. RCOV SS 10 min



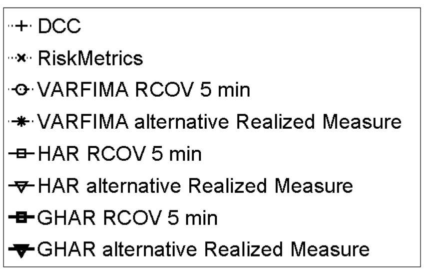
(e) RCOV 5 min vs. RCOV SS 15 min



(f) RCOV 5 min vs. RCOV SS 20 min



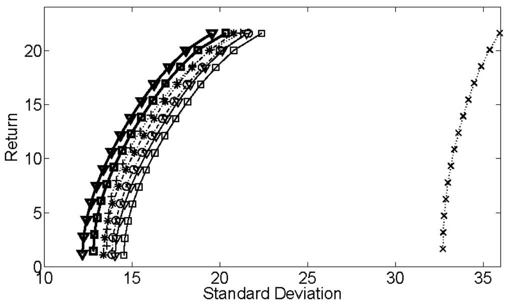
(g) legend



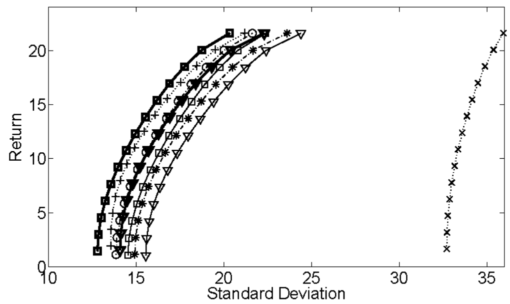
Note: Figure displays efficient frontiers of various competing models of portfolio of ten stocks based on five-step-ahead forecasts.

Figure 1.6: Efficient frontiers - portfolio of 15 stocks

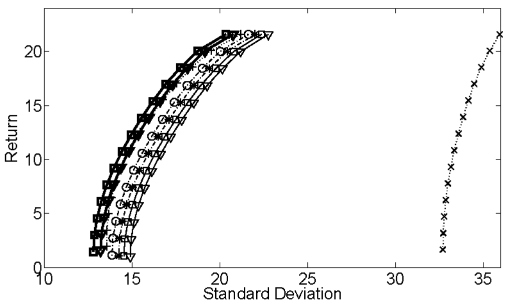
(a) RCOV 5min vs. MRK



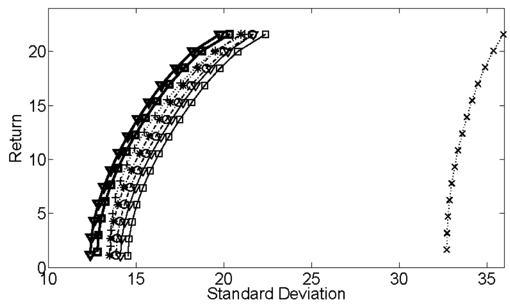
(b) RCOV 5 min vs. RCOV 1 min



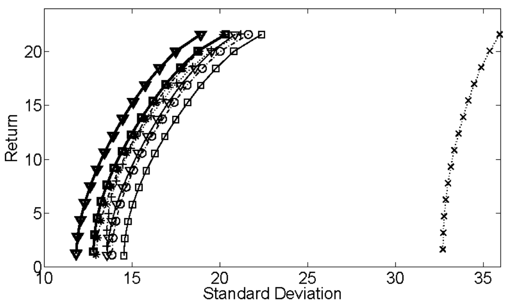
(c) RCOV 5 min vs. RCOV SS 5 min



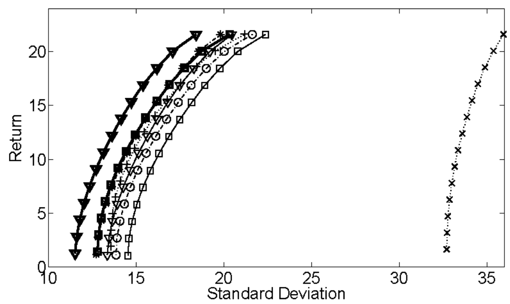
(d) RCOV 5 min vs. RCOV SS 10 min



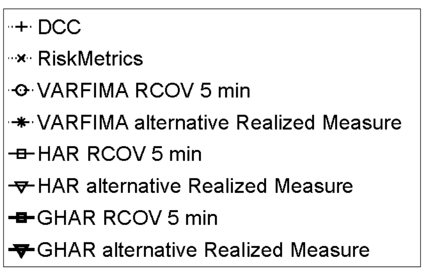
(e) RCOV 5 min vs. RCOV SS 15 min



(f) RCOV 5 min vs. RCOV SS 20 min



(g) legend



Note: Figure displays efficient frontiers of various competing models of portfolio of fifteen stocks based on five-step-ahead forecasts.

Appendix C: 10 step ahead forecasts

Table 1.14: GMVP - portfolio of 5 stocks

	Cumulative						
	MRK	RCOV		Sub-Sampled RCOV			
	1 min	1 min	5 min	5 min	10 min	15 min	20 min
DCC	30.50	**30.50**	**30.50**	**30.50**	**30.50**	30.50	30.50
RiskMetrics	40.58	40.58	40.58	40.58	40.58	40.58	40.58
VARFIMA	30.30	33.75	31.80	32.20	30.53	29.41	28.88
GHAR	**30.35**	33.66	31.94	32.21	30.58	**29.40**	**28.81**
HAR	31.16	34.40	32.68	33.01	31.30	30.24	29.74
	Annualized						
	MRK	RCOV		Sub-Sampled RCOV			
	1 min	1 min	5 min	5 min	10 min	15 min	20 min
DCC	17.39	**17.39**	**17.39**	**17.39**	17.39	17.39	17.39
RiskMetrics	23.22	23.22	23.22	23.22	23.22	23.22	23.22
VARFIMA	17.07	18.82	17.84	18.02	17.15	16.54	16.26
GHAR	**17.11**	18.74	17.86	18.04	**17.15**	**16.49**	**16.16**
HAR	17.72	19.32	18.49	18.62	17.73	17.13	16.86

Note: Model with the overall best performance is highlighted; for the given frequency model with the lowest risk is presented in bold; values represents percentage level of risk; values are scaled by forecasting horizon

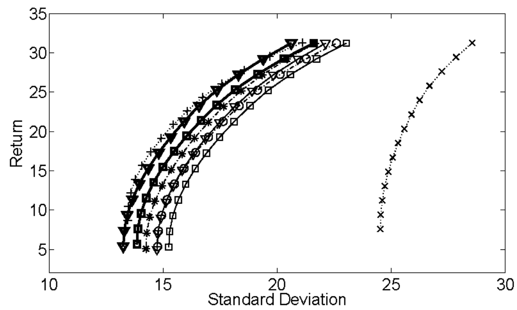
Table 1.15: RMSE – portfolio of 5 stocks

	MRK	RCOV		Sub-Sampled RCOV			
	1 min	1 min	5 min	5 min	10 min	15 min	20 min
DCC	1.208	1.294	1.375	1.291	1.173	1.107	1.101
RiskMetrics	1.389	1.401	1.431	1.404	1.388	1.384	1.380
VARFIMA	1.153	1.147	1.266	1.173	1.106	1.072	1.078
GHAR	1.287	1.256	1.409	1.307	1.237	1.197	1.205
HAR	1.138	1.133	1.242	1.163	1.091	1.058	1.067

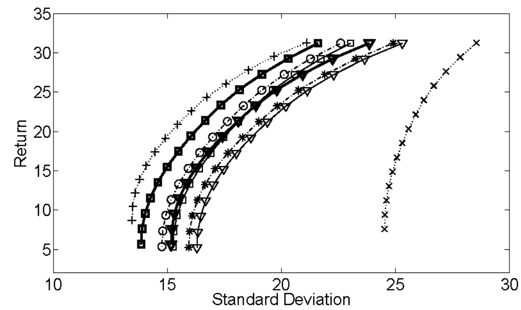
Note: Values are scaled by 10^{-3} and by forecasting horizon; highlighted cells belongs to 5% MCS

Figure 1.7: Efficient frontiers - portfolio of 5 stocks

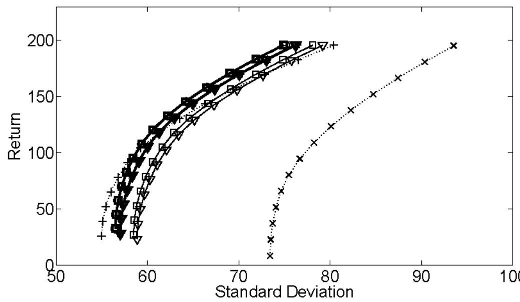
(a) RCOV 5min vs. MRK



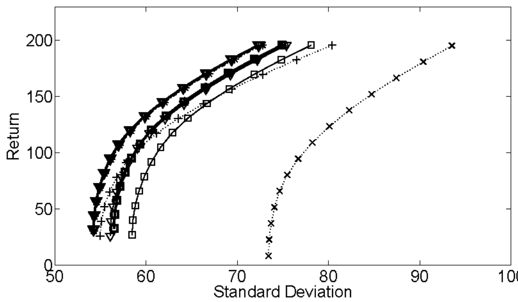
(b) RCOV 5 min vs. RCOV 1 min



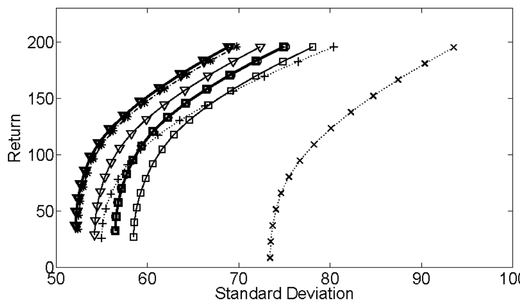
(c) RCOV 5 min vs. RCOV SS 5 min



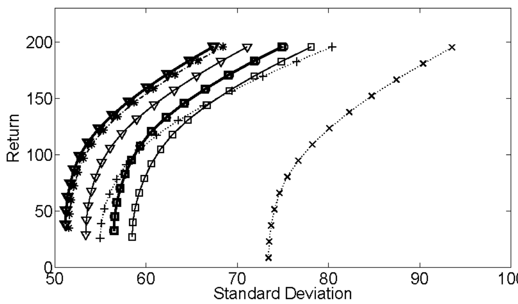
(d) RCOV 5 min vs. RCOV SS 10 min



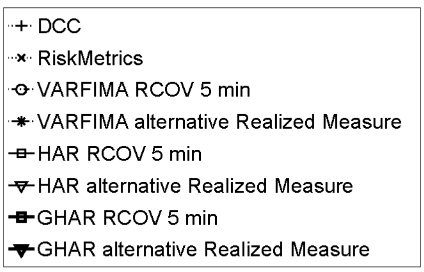
(e) RCOV 5 min vs. RCOV SS 15 min



(f) RCOV 5 min vs. RCOV SS 20 min



(g) legend



Note: Figure displays efficient frontiers of various competing models of portfolio of five stocks based on ten-step-ahead forecasts.

Table 1.16: GMVP - portfolio of 10 stocks

	Cumulative						
	MRK	RCOV		Sub-Sampled RCOV			
	1 min	1 min	5 min	5 min	10 min	15 min	20 min
DCC	**22.05**	**22.05**	**22.05**	**22.05**	**22.05**	22.05	22.05
RiskMetrics	42.08	42.08	42.08	42.08	42.08	42.08	42.08
VARFIMA	22.89	26.91	24.17	24.97	23.31	22.09	21.45
GHAR	22.16	26.23	23.55	24.40	22.61	**21.33**	**20.66**
HAR	24.15	27.94	25.42	26.14	24.45	23.25	22.66
	Annualized						
	MRK	RCOV		Sub-Sampled RCOV			
	1 min	1 min	5 min	5 min	10 min	15 min	20 min
DCC	13.03	**13.03**	**13.03**	**13.03**	13.03	13.03	13.03
RiskMetrics	24.40	24.40	24.40	24.40	24.40	24.40	24.40
VARFIMA	13.16	15.13	13.80	14.20	13.32	12.67	12.33
GHAR	**12.56**	14.67	13.28	13.75	**12.76**	**12.06**	**11.69**
HAR	14.01	15.85	14.67	15.01	14.10	13.45	13.14

Note: Model with the overall best performance is highlighted; for the given frequency model with the lowest risk is presented in bold; values represents percentage level of risk; values are scaled by forecasting horizon

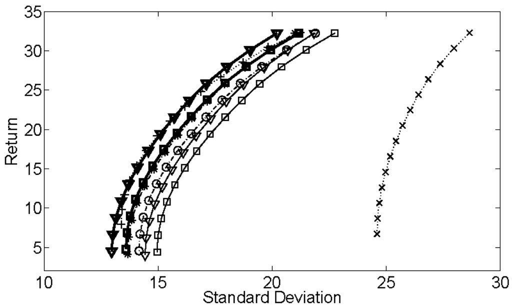
Table 1.17: RMSE – portfolio of 10 stocks

	MRK	RCOV		Sub-Sampled RCOV			
	1 min	1 min	5 min	5 min	10 min	15 min	20 min
DCC	2.437	2.609	2.687	2.610	2.362	2.260	2.271
RiskMetrics	3.445	3.461	3.455	3.448	3.458	3.487	3.481
VARFIMA	2.139	2.165	2.327	2.208	2.057	2.011	2.041
GHAR	2.605	2.514	2.729	2.607	2.494	2.449	2.491
HAR	2.114	2.110	2.276	2.174	2.026	1.986	2.024

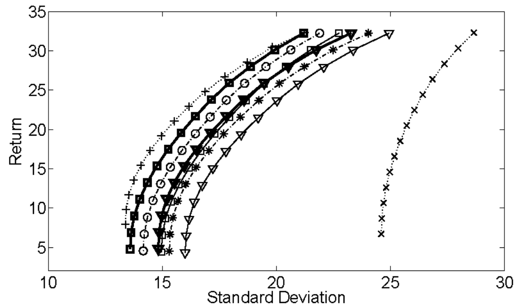
Note: Values are scaled by 10^{-3} and by forecasting horizon; highlighted cells belongs to 5% MCS

Figure 1.8: Efficient frontiers - portfolio of 10 stocks

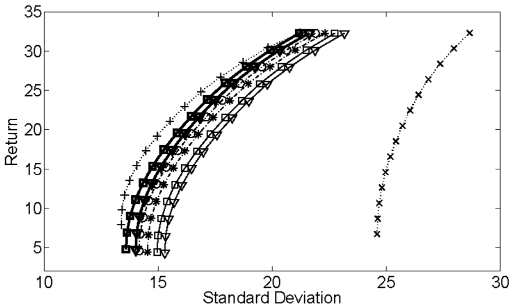
(a) RCOV 5min vs. MRK



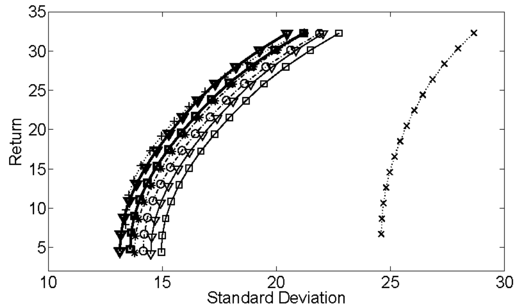
(b) RCOV 5 min vs. RCOV 1 min



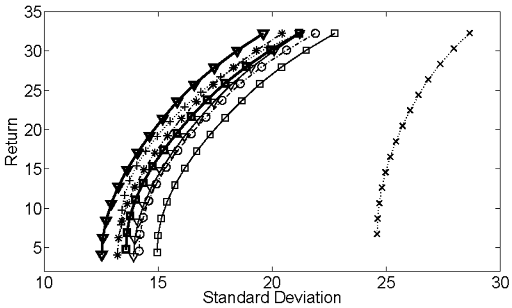
(c) RCOV 5 min vs. RCOV SS 5 min



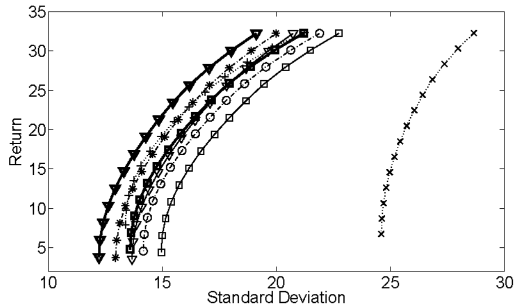
(d) RCOV 5 min vs. RCOV SS 10 min



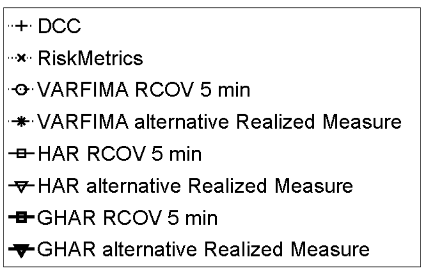
(e) RCOV 5 min vs. RCOV SS 15 min



(f) RCOV 5 min vs. RCOV SS 20 min



(g) legend



Note: Figure displays efficient frontiers of various competing models of portfolio of ten stocks based on ten-step-ahead forecasts.

Table 1.18: GMVP - portfolio of 15 stocks

	MRK	RCOV		Sub-Sampled RCOV			
Cumulative							
	1 min	1 min	5 min	5 min	10 min	15 min	20 min
DCC	20.67	**20.67**	**20.67**	**20.67**	**20.67**	20.67	20.67
RiskMetrics	56.60	56.60	56.60	56.60	56.60	56.60	56.60
VARFIMA	21.08	25.00	22.33	23.21	21.56	20.36	19.72
GHAR	**20.21**	24.13	21.54	22.30	20.72	**19.53**	**18.83**
HAR	22.31	26.00	23.46	24.32	22.68	21.49	20.88
Annualized							
	MRK	RCOV		Sub-Sampled RCOV			
	1 min	1 min	5 min	5 min	10 min	15 min	20 min
DCC	12.56	**12.56**	12.56	**12.56**	12.56	12.56	12.56
RiskMetrics	32.32	32.32	32.32	32.32	32.32	32.32	32.32
VARFIMA	12.32	14.21	12.95	13.38	12.50	11.86	11.53
GHAR	**11.44**	13.55	**12.16**	12.59	**11.70**	**11.04**	**10.66**
HAR	13.19	14.95	13.82	14.19	13.31	12.68	12.37

Note: Model with the overall best performance is highlighted; for the given frequency model with the lowest risk is presented in bold; values represents percentage level of risk; values are scaled by forecasting horizon

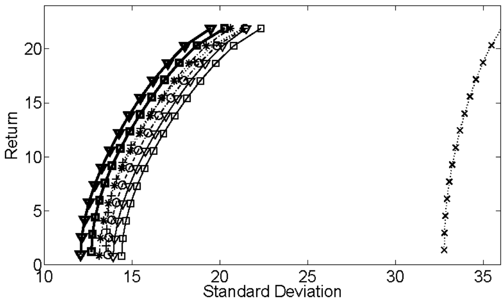
Table 1.19: RMSE – portfolio of 15 stocks

	MRK	RCOV		Sub-Sampled RCOV			
	1 min	1 min	5 min	5 min	10 min	15 min	20 min
DCC	4.141	4.258	4.385	4.323	4.054	3.989	4.010
RiskMetrics	11.806	11.735	11.720	11.719	11.884	11.981	11.961
VARFIMA	3.690	3.542	3.821	3.680	3.496	3.439	3.509
GHAR	4.807	4.514	4.859	4.746	4.571	4.508	4.613
HAR	3.666	3.471	3.767	3.635	3.468	3.424	3.512

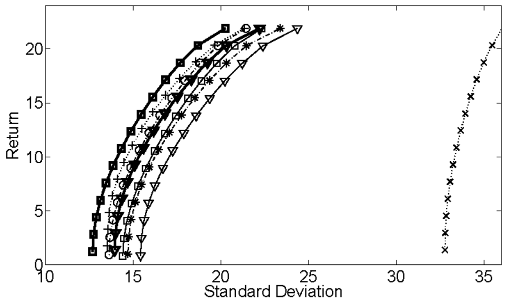
Note: Values are scaled by 10^{-3} and by forecasting horizon; highlighted cells belongs to 5% MCS

Figure 1.9: Efficient frontiers - portfolio of 15 stocks

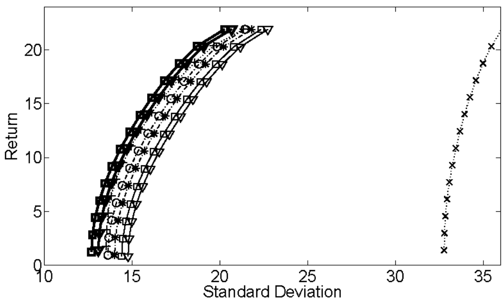
(a) RCOV 5min vs. MRK



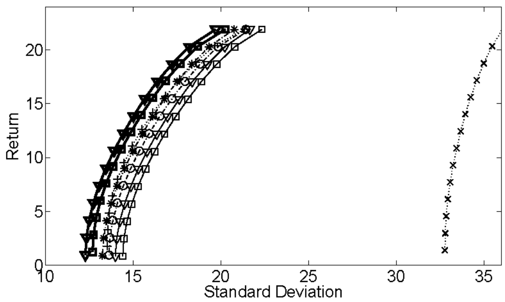
(b) RCOV 5 min vs. RCOV 1 min



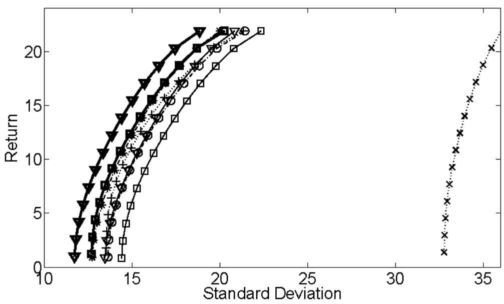
(c) RCOV 5 min vs. RCOV SS 5 min



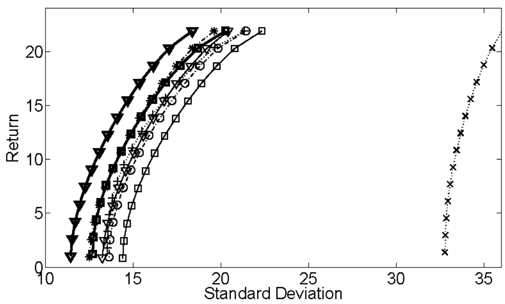
(d) RCOV 5 min vs. RCOV SS 10 min



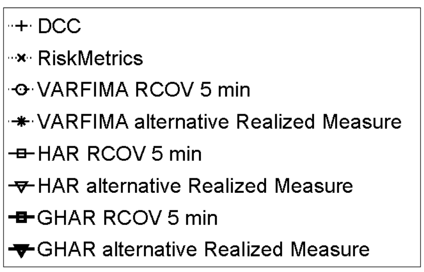
(e) RCOV 5 min vs. RCOV SS 15 min



(f) RCOV 5 min vs. RCOV SS 20 min



(g) legend



Note: Figure displays efficient frontiers of various competing models of portfolio of fifteen stocks based on ten-step-ahead forecasts.

Appendix D

Table 1.20: Descriptive statistics of returns over the period 01.07.2005 – 03.01.2012

	AAPL	CVX	GE	GOOG	IBM	JNJ	JPM	KO	MSFT	PFE	PG	T	WFC	WMT	XOM
	1 min														
Mean	-0.105	0.086	-0.218	-0.204	0.215	-0.035	-0.036	0.011	-0.040	-0.150	0.133	-0.079	-0.049	-0.018	0.099
Max	0.046	0.043	0.050	0.022	0.030	0.044	0.032	0.020	0.019	0.032	0.032	0.039	0.042	0.030	0.040
Min	-0.037	-0.027	-0.032	-0.041	-0.020	-0.033	-0.060	-0.038	-0.025	-0.028	-0.028	-0.038	-0.049	-0.021	-0.034
SD	1.046	0.891	1.088	0.936	0.749	0.559	1.336	0.619	0.833	0.825	0.616	0.834	1.464	0.685	0.827
Skewness	0.037	0.262	0.181	-0.369	0.077	0.394	-0.144	-0.482	-0.063	0.103	-0.066	-0.122	0.068	0.365	-0.188
Kurtosis	37.100	44.353	43.467	34.940	40.526	126.707	41.663	70.490	18.848	30.256	71.747	52.025	40.587	39.587	50.690
	5 min														
Mean	-0.629	0.423	-0.994	-1.037	1.177	-0.132	-0.165	0.113	-0.138	-0.745	0.741	-0.353	-0.252	-0.057	0.526
Max	0.065	0.061	0.052	0.046	0.053	0.032	0.069	0.028	0.030	0.030	0.050	0.034	0.066	0.048	0.053
Min	-0.048	-0.068	-0.046	-0.069	-0.036	-0.040	-0.068	-0.037	-0.028	-0.038	-0.062	-0.073	-0.077	-0.042	-0.059
SD	2.258	1.916	2.280	2.023	1.580	1.174	2.871	1.297	1.756	1.694	1.316	1.779	3.151	1.478	1.779
Skewness	-0.008	-0.062	0.268	-0.509	0.121	-0.127	0.059	-0.300	-0.091	0.109	-0.460	-0.518	0.095	0.432	-0.098
Kurtosis	28.079	37.935	31.858	39.555	37.413	44.133	35.720	33.284	17.395	18.627	86.196	46.840	34.793	42.507	40.512
	10 min														
Mean	-0.960	1.129	-1.740	-1.820	2.856	0.118	-0.611	0.387	0.350	-1.416	1.832	-0.444	-0.827	0.289	1.410
Max	0.050	0.039	0.052	0.043	0.029	0.030	0.067	0.031	0.029	0.029	0.024	0.038	0.069	0.053	0.051
Min	-0.079	-0.034	-0.058	-0.073	-0.036	-0.025	-0.102	-0.040	-0.038	-0.027	-0.031	-0.043	-0.092	-0.035	-0.067
SD	3.150	2.630	3.168	2.776	2.169	1.592	3.937	1.793	2.399	2.297	1.765	2.412	4.372	1.999	2.423
Skewness	-0.301	0.229	0.256	-0.373	-0.161	0.310	0.035	-0.321	-0.018	0.244	-0.042	-0.114	0.082	0.435	-0.098
Kurtosis	24.897	15.055	29.864	25.773	20.025	21.991	31.039	27.228	15.157	12.988	20.916	21.321	30.512	23.805	27.665
	15 min														
Mean	-1.415	2.229	-2.493	-2.598	4.830	0.250	-0.769	0.643	0.765	-1.986	2.917	-0.370	-1.003	0.712	3.193
Max	0.058	0.046	0.071	0.049	0.038	0.025	0.113	0.030	0.032	0.039	0.028	0.046	0.099	0.051	0.047
Min	-0.053	-0.037	-0.070	-0.068	-0.053	-0.024	-0.086	-0.041	-0.042	-0.029	-0.034	-0.053	-0.075	-0.035	-0.037
SD	3.801	3.186	3.877	3.350	2.630	1.946	4.896	2.175	2.921	2.794	2.138	2.951	5.335	2.445	2.925
Skewness	-0.012	0.237	0.161	-0.242	-0.159	0.332	0.314	-0.319	-0.050	0.264	0.086	-0.052	0.421	0.562	0.263
Kurtosis	16.529	15.547	31.899	22.586	21.422	19.781	35.685	22.721	14.785	13.138	21.422	21.883	30.411	20.882	19.090
	20 min														
Mean	-1.950	2.227	-4.353	-3.743	5.912	-0.276	-1.903	0.412	0.445	-2.604	3.494	-0.812	-2.259	0.542	3.371
Max	0.050	0.059	0.062	0.043	0.036	0.034	0.074	0.035	0.034	0.041	0.026	0.053	0.086	0.053	0.069
Min	-0.048	-0.037	-0.068	-0.118	-0.040	-0.021	-0.102	-0.040	-0.038	-0.029	-0.029	-0.049	-0.080	-0.024	-0.067
SD	4.245	3.608	4.350	3.775	2.939	2.159	5.381	2.446	3.259	3.131	2.380	3.300	6.064	2.741	3.325
Skewness	-0.075	0.252	0.130	-0.867	-0.034	0.420	-0.051	-0.148	-0.044	0.256	0.087	-0.086	0.161	0.510	0.177
Kurtosis	13.207	14.377	26.975	42.831	17.258	17.410	23.170	19.581	12.493	11.697	15.784	19.932	25.774	17.183	25.829

Note: The means are scaled by 10^5, the standard deviations are scaled by 10^3

Chapter 2

Measurement of Common Risk Factors in Tails: A Panel Quantile Regression Model for Returns

This chapter investigates how to measure common market risk factors in tails of the return distributions using newly proposed Panel Quantile Regression Model for Returns. By exploring the fact that volatility crosses all quantiles of the return distribution and using a penalized fixed effects estimator, we are able to control for otherwise unobserved heterogeneity among financial assets. Direct benefits of the proposed approach are revealed in a portfolio Value–at–Risk forecasting application, where our modeling strategy performs significantly better than several benchmark models according to both statistical and economic comparison. In particular Panel Quantile Regression Model for Returns consistently outperforms all the competitors in the 5% and 10% quantiles. Sound statistical performance translates directly into economic gains which is demonstrated in the Global Minimum Value–at–Risk Portfolio and Markowitz-like comparison. Overall results of our research are important for correct identification of the sources of systemic risk, and are particularly attractive for high dimensional applications.

2.1 Introduction

Many studies document cross-sectional relations between risk and expected returns, generally measuring a stock's risk as the covariance between its return and some factor. In this laborious search for proper risk factors,[1] volatility still plays a central role in

[1] See for example Harvey et al. (2016); Feng et al. (2019) for recent very complete overviews. This research dates back to French et al. (1987).

explaining expected stock returns for decades. Most recent efforts explore increasingly available datasets, and make measurement of ex-post volatility more precise than ever before. In turn, these measures can be used for more precise identification of market risk. Although predictions about expected returns are essential for understating of classical asset pricing, little is known about potential of the factors to precisely identify extreme tail events of the returns distribution. More importantly, even less is known about commonalities between more assets with this respect. Our research attempts to contribute in this direction.

Asset pricing models explaining risk valuation theoretically assume an economic agent who decides based on the preference about her consumption by maximizing expected utility function. However, these preferences may be too restrictive to deliver satisfactory description of the real behavior of agents. Instead of working with standard expected utilities, recent literature strives to incorporate heterogeneity into dynamic economic models assuming agents maximize their stream of future quantile utilities (Chambers, 2007; Rostek, 2010; de Castro and Galvao, 2018). We contribute to these efforts by developing a Panel Quantile Regression Model for Returns that is able to control for otherwise unobserved heterogeneity among financial assets and allows us to exploit common factors in volatility that directly affect future quantiles of returns. In a sense, we revisit a large literature connecting volatility with the cross-section of returns, as by construction, we model tail events of the conditional distributions via volatility.

Since the seminal work of Koenker and Bassett Jr (1978), quantile regression models have been increasingly used in many disciplines. In finance, Engle and Manganelli (2004) were among the first to use quantile regression to develop the Conditional Autoregressive Value–at–Risk (CAViaR) model and capture conditional quantiles of the asset returns. Baur et al. (2012) use quantile autoregressions to study conditional return distributions, Cappiello et al. (2014) detects comovement between random variables with time-varying quantile regression. Žikeš and Baruník (2016) show that various volatility measures are useful in forecasting quantiles of future returns without making assumptions about underlying conditional distributions. The resulting semi-parametric modeling strategy captures conditional quantiles of financial returns well in a flexible framework of quantile regression. Moving towards a multivariate framework, and concentrating on interrelations between quantiles of more assets, White et al. (2015) pioneer the extension. Different stream of multivariate quantile regression based literature concentrates on the analysis using factors (Gonzalo et al., 2017; Ando and Bai, 2018).[2] From a theoretical point of view, Giovannetti (2013) derives an asset pricing model in which equity premium is no longer based on the covariance between return and consumption. Instead, Giovannetti (2013) argue that under optimism, higher volatility can be connected to high chance of high returns leading to increased prices, hence decreasing expected returns, and vice

[2] Panel quantiles methods are useful in the other areas of economics besides finance. They are mostly applied in the labour economics (Billger and Lamarche (2015), Dahl et al. (2013), Toomet (2011)), banking and economic policy analysis (Covas et al. (2014), Klomp and De Haan (2012)), economics of education (Lamarche (2008), Lamarche (2011)), energy and environmental economics (You et al. (2015), Zhang et al. (2015)) or international trade (Dufrenot et al. (2010), Foster-McGregor et al. (2014), Powell and Wagner (2014)).

versa under pessimism. Based on Choquet utility functions, Bassett et al. (2004) show that pessimistic optimization may be formulated as a linear quantile regression problem, and can lead to optimal portfolio allocation.

In this respect, work by Žikeš and Baruník (2016) is important as it provides link between future quantiles of return distribution and its past variation. As the financial sector is highly connected and the co-movements in asset prices are common, there is a need for proper identification of dependencies in joint distributions. In the classical mean-regression framework, Bollerslev et al. (2018) showed that realized volatility of financial time series shares many commonalities. In the quantile regression set-up, however, there is no similar study that attempts to uncover information captured in the panels of volatility series. Moreover, to the best of our knowledge there is no study estimating the conditional distribution of returns in a multivariate setting that explores ex-post information in the volatility.

In this chapter, we contribute to the literature by introducing a Panel Quantile Regression Model for Returns that allows to measure common risk factors in tails of the return distributions. Our model utilize all the advantages offered by panel quantile regression and financial market datasets. In particular, we are able to control for otherwise unobserved heterogeneity among financial assets and reveal common factors in volatility that have direct influence on the future quantiles of returns. To the best of our knowledge this is one of the first applications of the panel quantile regression using a dataset where the time dimension T is much greater than cross-sectional dimension N, i.e. $T >> N$. As a result we are able to obtain estimates of quantile specific individual fixed effects that represents the idiosyncratic part of market risk.

In an empirical application, we hypothesize that newly proposed model will deliver more accurate estimates compared to currently established methods. These estimates moreover translates into better forecasting performance of the Panel Quantile Regression Model for Returns. In addition, using a penalized fixed effect estimator we will be able to disentangle overall market risk into systematic and idiosyncratic parts. Actual performance of our model is tested in a portfolio Value–at–Risk forecasting exercise. Before the analysis of the empirical dataset (29 highly liquid stocks from the New York Stock Exchange), we run a small Monte-Carlo experiment that enable us to study well-behaved data. For the robustness reasons we evaluate forecasts from both a statistical and economic perspective. In the statistical comparison we furthermore distinguish between absolute and relative performance of the given model.

Results of our analysis suggest that the Panel Quantile Regression Model for Returns is dynamically correctly specified. Moreover it dominates the benchmark models in the economically important quantiles (5%,10% or 95%). Overall we find that according to statistical comparison none of the benchmark models is able to outperform our model consistently. Furthermore the model we introduce in this chapter provide us with direct economic gains according to both economic evaluation criteria.

2.2 Risk Measurement using High Frequency Data

We naturally begin with the definition of risk measures used in the study. Let's assume that the efficient logarithmic price process $p_{i,t}$ of ith asset evolves over time $0 \leq t \leq T$ according to the following dynamics

$$dp_{i,t} = \mu_{i,t}dt + \sigma_{i,t}dW_{i,t} + dJ_{i,t}, \tag{2.1}$$

where $\mu_{i,t}$ is a predictable component, $\sigma_{i,t}$ is a cadlag process, $W_{i,t}$ is a standard Brownian motion, and $J_{i,t}$ is a jump process.

The volatility of the logarithmic price process can be measured by quadratic return variation which can be decomposed into integrated variance (IV) of the price process and the jump variation (JV):

$$QV_{i,t} = \underbrace{\int_{t-1}^{t} \sigma_{i,s}^2 ds}_{IV_{i,t}} + \underbrace{\sum_{l=1}^{N_{i,t}} \kappa_{i,t,l}^2}_{JV_{i,t}}, \tag{2.2}$$

where $N_{i,t}$ is total number of jumps during day t and $\sum_{l=1}^{N_t} \kappa_{i,t,l}^2$ represents magnitude of the jumps. As shown by Andersen et al. (2003) Realized Variance estimator can be simply constructed by squaring intraday returns:

$$\widehat{RV}_{i,t} = \sum_{k=1}^{N} \left(\Delta_k p_{i,t}\right)^2, \tag{2.3}$$

where $\Delta_k p_{i,t} = p_{i,t-1+\nu_k/N} - p_{i,t-1+\nu_{k-1}/N}$ is a discretely sampled vector of k-th intraday log-returns of ith asset in $[t-1, t]$, with N intraday observations. Realized Variance estimator moreover converges uniformly in probability to $QV_{i,t}$ as the sampling frequency goes to infinity

$$\widehat{RV}_{i,t} \xrightarrow[N\to\infty]{p} \int_{t-1}^{t} \sigma_{i,s}^2 ds + \sum_{l=1}^{N_t} \kappa_{i,t,l}^2$$

Building on the concept of Realized Variance Barndorff-Nielsen and Shephard (2004b) and Barndorff-Nielsen and Shephard (2006) introduced the bipower variation estimator that is robust to jumps and thus able to consistently estimate $IV_{i,t}$. Furthermore, Andersen et al. (2011) adjust original estimator, which helps render it robust to certain types of microstructure noise:

$$\widehat{IV}_{i,t}^{BPV} = \mu_1^{-2} \left(\frac{N}{N-2}\right) \sum_{k=3}^{N} |\Delta_{k-2} p_{i,t}||\Delta_k p_{i,t}|,$$

where $\mu_\alpha = E(|Z^\alpha|)$, and $Z \sim N(0,1)$. Having an estimator of $IV_{i,t}$ in hand, jump variation can be consistently estimated[3] as a difference between Realized Variance and

[3] Asymptotic behaviour and further details of the estimator can be found in Barndorff-Nielsen and Shephard (2006).

the bipower variation:

$$\left(\widehat{RV}_{i,t} - \widehat{IV}_{i,t}^{BPV}\right) \xrightarrow[N\to\infty]{p} \sum_{l=1}^{N_t} \kappa_{i,t,l}^2.$$

For many financial applications not only the magnitude of the variation, but also its sign is important. Therefore Barndorff-Nielsen et al. (2010) introduce an innovative approach for measuring negative and positive variation in data called Realized Semivariance. They showed that Realized Variance can be decomposed to realized downside semivariance ($RS_{i,t}^-$) and realized upside semivariance ($RS_{i,t}^+$):

$$RV_{i,t} = RS_{i,t}^+ + RS_{i,t}^-,$$

where $RS_{i,t}^+$ and $RS_{i,t}^-$ are defined as follows,

$$\widehat{RS}_{i,t}^+ = \sum_{k=1}^{N} (\Delta_k p_{i,t})^2 I(\Delta_k p_{i,t} > 0) \xrightarrow{p} \frac{1}{2} IV_{i,t} + \sum_{l=1}^{N_t} \kappa_{i,t,l}^2 I(\kappa_{i,t,l} > 0) \tag{2.4}$$

$$\widehat{RS}_{i,t}^- = \sum_{k=1}^{N} (\Delta_k p_{i,t})^2 I(\Delta_k p_{i,t} < 0) \xrightarrow{p} \frac{1}{2} IV_{i,t} + \sum_{l=1}^{N_t} \kappa_{i,t,l}^2 I(\kappa_{i,t,l} < 0). \tag{2.5}$$

Consequently, the negative and positive semivariance provides information about variation associated with movements in the tails of the underlying variable. Similar to Patton and Sheppard (2015) and Bollerslev et al. (2017), we use negative semivariance as a proxy to the bad state of the returns, and positive semivariance as an empirical proxy of the good state of the underlying variable.

Since correlation is inevitably important in portfolio applications, and we use it later in our portfolio Value–at–Risk application, we also define Realized Covariance estimator (Barndorff-Nielsen and Shephard, 2004a) as

$$\widehat{\Sigma}_t = \sum_{k=1}^{N} (\mathbf{\Delta_k p_t})(\mathbf{\Delta_k p_t})',$$

where $\mathbf{\Delta_k p_t} = (\Delta_k p_{1,t}, ..., \Delta_k p_{q,t})'$ is vector containing log-returns of q individual assets.

2.3 Panel Quantile Regression Model for Returns

Having briefly described realized measures that we need for model construction, we now propose simple linear models for cross-section of quantiles of future returns. We base our model in a recent theoretical endeavor to move from expected values to quantiles, thereby understanding heterogeneity in asset prices. Based on the risk preferences of quantile maximizers defined by Manski (1988); Rostek (2010) and de Castro and Galvao (2018) develop a dynamic model of rational behavior under uncertainty, in which an agent maximizes streams of future quantile utilities. This is in sharp contrast to the mainstream literature that assumes the decision making process to be driven by maximization of the expected utility instead. In the spirit of Bassett et al. (2004), general version of our model can be viewed as linear asset pricing equation

$$Q_{r_{i,t+1}}\left(\tau | v_{i,t}, \alpha_i(\tau)\right) = \underbrace{\alpha_i(\tau)}_{\text{Unobserved Heterogeneity}} + \underbrace{v_{i,t}^{\top}\beta(\tau)}_{\text{Idiosyncratic Risk}} + \underbrace{F_t^{\top}\gamma(\tau)}_{\text{Common Factors}}, \quad \tau \in (0,1), \tag{2.6}$$

where $r_{i,t+1} = p_{i,t+1} - p_{i,t}$ are logarithmic daily returns, $\alpha_i(\tau)$ represents individual fixed effects that accounts for unobserved heterogeneity, $v_{i,t}$ are measures of quadratic variation as defined in previous section and accounts for the firm specific (idiosyncratic) risk and F_t represents exogenous common factors. This model enables us to study influence of the various sources of risk on the specific quantiles of the future returns. Further, in case of $\gamma(\tau) \neq 0$ for a given τ, the model allows to capture the common risk factors in the tails.

While Equation 2.6 accommodates many possible specifications, we are interested in the set of following models. In the first set of model specifications, quantiles of the return series depends on the unobserved heterogeneity and idiosyncratic risk measured by one of the realized measures:

- *PQR-RV* with Realized Volatility defined as

$$Q_{r_{i,t+1}}\left(\tau | RV_{i,t}^{1/2}, \alpha_i(\tau)\right) = \alpha_i(\tau) + \beta_{RV^{1/2}}(\tau) RV_{i,t}^{1/2}, \tag{2.7}$$

- *PQR-RSV* with Realized Semivariance defined as

$$\begin{aligned} Q_{r_{i,t+1}}\left(\tau | RS_{i,t}^{+\,1/2}, RS_{i,t}^{-\,1/2}, \alpha_i(\tau)\right) =& \alpha_i(\tau) + \beta_{RS^{+1/2}}(\tau) RS_{i,t}^{+\,1/2} + \\ &+ \beta_{RS^{-1/2}}(\tau) RS_{i,t}^{-\,1/2}, \end{aligned} \tag{2.8}$$

- *PQR-BPV* with Realized Bi-Power Variation defined as

$$\begin{aligned} Q_{r_{i,t+1}}\left(\tau | BPV_{i,t}^{1/2}, JV_{i,t}^{1/2}, \alpha_i(\tau)\right) =& \alpha_i(\tau) + \beta_{BPV^{1/2}}(\tau) BPV_{i,t}^{1/2} + \\ &+ \beta_{JV^{1/2}}(\tau) JV_{i,t}^{1/2}. \end{aligned} \tag{2.9}$$

In the second set of model specifications, we study role of the ex-ante measure of market volatility, i.e. VIX index, that we consider to be a good proxy for the common exogenous factor. These specifications will measure the direct influence of the common market factor once we control for the asset specific volatility as well as unobserved heterogeneity:

- *PQR-RV-VIX* defined as

$$Q_{r_{i,t+1}}\left(\tau | RV_{i,t}^{1/2}, \alpha_i(\tau)\right) = \alpha_i(\tau) + \beta_{RV^{1/2}}(\tau) RV_{i,t}^{1/2} + \gamma_{VIX}(\tau) VIX_t, \tag{2.10}$$

- *PQR-RSV-VIX* defined as

$$\begin{aligned} Q_{r_{i,t+1}}\left(\tau|RS_{i,t}^{+\,1/2}, RS_{i,t}^{-\,1/2}, \alpha_i(\tau)\right) =&\alpha_i(\tau) + \beta_{RS^{+1/2}}(\tau)RS_{i,t}^{+\,1/2} + \\ &+ \beta_{RS^{-1/2}}(\tau)RS_{i,t}^{-\,1/2} + \gamma_{VIX}(\tau)VIX_t, \end{aligned} \tag{2.11}$$

- *PQR-BPV-VIX* defined as

$$\begin{aligned} Q_{r_{i,t+1}}\left(\tau|BPV_{i,t}^{1/2}, JV_{i,t}^{1/2}, \alpha_i(\tau)\right) =&\alpha_i(\tau) + \beta_{BPV^{1/2}}(\tau)BPV_{i,t}^{1/2} + \\ &+ \beta_{JV^{1/2}}(\tau)JV_{i,t}^{1/2} + \gamma_{VIX}(\tau)VIX_t. \end{aligned} \tag{2.12}$$

Details of this specifications are described in the Section Results: Common Risk Factors in Tails. Generally, Equation 2.6 can be easily extend by another exogenous variables such as factors used in Fama and French (1993), as already attempted by Galvao et al. (2017), however, this is beyond the scope of this chapter.

2.3.1 Estimation

In our work, we concentrate on commonalities in the quantiles of several return series. To obtain parameters estimates of the general model defined in Equation 2.6 we use panel quantile regression as introduced in Koenker (2004). In the seminal work, Roger Koenker proposed a penalized fixed effects estimator as a general way of estimating quantile regression models in the panel data framework. Recently Lamarche (2010) studied penalized quantile regression estimator, and Galvao (2011) introduced a fixed effects model for dynamic panels. Galvao and Montes-Rojas (2010) moreover shown that bias in dynamic panels can be reduced using a penalty term. Further, Canay (2011) introduced a simple two-step approach to the estimation of panel quantile regression and showed consistency and asymptotic normality of the proposed estimator. Other influential works developing theory of panel quantile methods are Harding and Lamarche (2009), Galvao and Montes-Rojas (2015), Galvao and Wang (2015), Galvao and Kato (2015), Graham et al. (2015), Harding and Lamarche (2014) or Kato et al. (2012).

Although literature devoted to panel quantile estimators is growing and many interesting alternatives have been introduced, we use original penalized fixed effects estimator. The advantage of this approach is the ability to account, and control for, unobserved heterogeneity among financial assets, which will yield more precise quantile specific estimates. As a consequence these estimates will translate into better forecasting performance directly. Moreover one can use this approach to obtain precise estimates of the Value–at–Risk (VaR) which is commonly used financial industry risk measure. In the VaR application panel data will utilize all the favorable properties of the standard time series. In addition, the cross-sectional dimension will help us to account for common shocks among the assets.

To obtain parameter estimates we solve following optimization problem

$$\min_{\alpha(\tau),\beta(\tau),\gamma(\tau)} \sum_{t=1}^{n}\sum_{i=1}^{t_i} \rho_\tau(r_{i,t+1} - \alpha_i(\tau) - v_{i,t}^\top \beta(\tau) - F_t^\top \gamma(\tau)) + \lambda \sum_{i=1}^{n} |\alpha_i(\tau)|, \tag{2.13}$$

where $\rho_\tau(u) = u\,(\tau - I(u(<0)))$ is the quantile loss function (Koenker and Bassett Jr, 1978) and $\sum_{i=1}^{n}|\alpha_i|$ is l_1 penalty that controls variability introduced by the large number of estimated parameters. The general form of our model consider penalty term λ from range $\langle 0, \infty)$. In case $\lambda = 0$, we obtain full set of asset specific fixed effects, for $\lambda > 0$ fixed effects of some assets shrink toward zero and as $\lambda \to \infty$ we have model without fixed effects. One might consider shrinking some of the fixed effects toward zero in a high cross-sectional dimension problems thus keeping the number of estimated parameters reasonable.

In the empirical application we choose the penalty term by minimizing Bayesian Information Criterion (BIC) as proposed in Galvao and Montes-Rojas (2010)

$$BIC(p_\lambda) = log\hat{\sigma}_\lambda + 1/2\,(NT)^{-1}\,p_\lambda logNT,$$

where $\hat{\sigma}_\lambda = NT^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\rho_\tau\left(r_{i,t+1} - \hat{\alpha}_i(\tau,\lambda) - v_{i,t}^\top\hat{\beta}(\tau,\lambda) - F_t^\top\hat{\gamma}(\tau,\lambda)\right)$, and p_λ is a measure of the effective dimension of the fitted model with penalty parameter λ. In the p_λ calculation we consider both Method 1 and 2, where

- Method 1: p_λ is the dimension of the set $\{\beta \cup \gamma \cup \{i||\alpha_i|> \kappa\}\}$
- Method 2: $p_\lambda = \sum_{i=1}^{N}\sum_{t=1}^{T} I\left[|\hat{\xi}_{i,t+1}(\tau,\lambda)|< \kappa\right]$, where $\hat{\xi}_{i,t+1}(\tau,\lambda) = r_{i,t+1} - \hat{\alpha}_i(\tau,\lambda) - v_{i,t}^\top\hat{\beta}(\tau,\lambda) - F_t^\top\hat{\gamma}(\tau,\lambda)$ is the τ-quantile residual sequence for a given λ.

We have set tolerance parameter κ to various values ranging from 10^{-2} to 10^{-7} and it turns out to be the crucial part of the analysis. Unlike the results of Galvao and Montes-Rojas (2010) where model selection was not affected by the values of κ, in our empirical application, the optimal lambda differ substantially for different λ. Moreover, results from Method 1 and 2 sometimes contradict each other, i.e. for $\kappa = 10^{-4}$ Method 1 suggests to shrink all the fixed effects while Method 2 suggest zero penalization in the 75% quantile. Furthermore, using both methods for p_λ calculation, the differences in the BICs of unpenalized model and the model with "optimaly" selected λ are very small, i.e no greater than 4×10^{-3} or 0.4%.

Since the results of penalty selection are inconclusive and the increasing time dimension reduces the usefulness of the shrinkage method (Galvao and Montes-Rojas, 2010), we concentrate in our further analysis on the model without penalty, i.e. $\lambda = 0$, and we apply standard pure fixed effects model. This approach allows us to obtain estimates of all individual quantile specific fixed effects, i.e. account for unonserved heterogenity among assets. As a robustness check[4] we also carried out the analysis with values of λ

[4] Results are presented in Appendix.

from range (0; 1) as in Damette and Delacote (2012) and Covas et al. (2014) and with $\lambda = 1$ as in Koenker (2004), Bache et al. (2008), Matano and Naticchioni (2011), Lee et al. (2012) and You et al. (2015). Overall we find that the choice of λ does not affect precision of our analysis. We address this finding to the structure and characteristics of the dataset (high time dimension T compared to low cross-section dimension N).

In the Equation 2.13 we further consider individual fixed effects to have distributional effects and we concentrate on each quantile separately rather than solving optimization problem through several quantiles simultaneously. In contrast, Koenker (2004) and vast majority of the theoretical and applied works consider α_i to have a pure location shift effect on the conditional quantiles. This restriction is a consequence of the structure of the usual panel-datasets where cross-sectional dimension is much larger than time dimension[5]. This problem is not so severe in our application since majority of assets have a long history and thus consist of thousands of observations. Moreover analysis of the specific quantiles is essential for many financial applications including popular Value–at–Risk in which we are most often interested in finding 1-day 5% VaR or 10-day 1% VaR, as historically recommended by Basel Committee on Banking Supervision.

2.4 Competing Models and Evaluation

In the previous section we introduce Panel Quantile Regression Model for Returns which will be used in the applied part of the chapter to analyze simulated and empirical data. In this section, we describe alternative approaches that can be viewed as the direct competitors to our model. Benchmarks in our work includes popular and widely used RiskMetrics model that is the industry standard for the risk evaluation in high-dimensional problems and two applications of the Univariate Quantile Regression Model for Returns.

2.4.1 RiskMetrics

Based on Exponentially Weighted Moving Average, J.P. Morgan Chase in 1996 introduced new methodology for accessing the financial risk called RiskMetrics. It is considered to be the baseline benchmark model for numerous fiancial applications. For our benchmark purposes, we adopt the specification in its original form as defined in Longerstaey and Spencer (1996) with decay factor, λ set to 0.94. We assume a $q \times 1$ vector of daily returns $r_t = \sum_{k=1}^{n} (\Delta_k p_t)$ for $t = 1, ..., T$ such that $r_t \sim N(\mu_t, \sigma_t^2)$, where μ_t is conditional mean and σ_t^2 is conditional variance of daily returns. We also assume that $\mu_t = 0$ and therefore conditional covariance has the form

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

where $\sigma_{i,j,t}$ denotes covariance between assets i and j at time t.

[5]As detailed in Koenker (2004) it is not advisable to estimate τ-specific α_i in problems with small/medium T.

2.4.2 Univariate Quantile Regression Model for Returns

As already mentioned Žikeš and Baruník (2016) introduced an elegant framework for modelling and obtaining forecasts of the conditional quantiles of future returns in a univariate setting. They proposed to model quantiles of return series according to:

$$Q_{r_{i,t+1}}(\tau|v_{i,t}, z_t) = \alpha_i(\tau) + v_{i,t}^{\top}\beta_i(\tau) + z_t^{\top}\gamma(\tau), \tag{2.14}$$

where $r_{i,t+1} = p_{i,t+1} - p_{i,t}$ is return series of ith asset, $v_{i,t} = \left(\widehat{QV}_{i,t}^{1/2}, \widehat{QV}_{i,t-1}^{1/2}, \ldots, \widehat{IV}_{i,t}^{1/2}, \widehat{IV}_{i,t-1}^{1/2}, \ldots, \widehat{JV}_{i,t}^{1/2}, \widehat{JV}_{i,t-1}^{1/2}, \ldots\right)$ are components of quadratic variation and z_t is vector of vector of weakly exogenous variables. Estimates of asset i quantile specific β from Equation 2.14 are obtained by minimizing following objective function:

$$\min_{\alpha_i(\tau),\beta_i(\tau)} \frac{1}{n}\sum_{t=1}^{n} \rho_\tau\left(r_{i,t+1} - \alpha_i(\tau) - v_{i,t}^{\top}\beta_i(\tau) - z_t^{\top}\gamma(\tau)\right), \tag{2.15}$$

where $\rho_\tau(u) = u(\tau - I(u < 0)))$ is the quantile loss function defined in Koenker and Bassett Jr (1978). The application of the model in a multivariate setting is further described in the following section.

2.4.3 Forecasting Exercise and Forecast Evaluation

In order to evaluate the performance of the newly proposed Panel Quantile Regression Model for Returns we conduct forecasting exercise in which we study portfolio Value–at–Risk from a statistical and economic point of view. We decided to concentrate on both statistical and economic evaluation in order to get a complete picture of behavior of the new model. Moreover concentrating on statistical evaluation only might get us into trouble because good statistical performance might not necessarily translate into economic gains. Therefore to make our results robust we apply two statistical and two economic evaluation criteria.

In the statistical comparison we focus on the absolute and relative performance of the considered models in an equally weighted portfolio set-up. By focusing on an equally weighted portfolio, we refrain from specifying complicated weighting schemes which might affect the overall performance.

In the economic comparison, we study the efficient frontier of the Value–at–Risk - return trade-off and also Global Minimum Value–at–Risk Portfolio (GMVaRP). As both approaches by definition tries to find optimal weights of the assets we are not using equally weighted portfolio here anymore.

Portfolio Value–at–Risk

Value–at–Risk is an elegant way of quantifying the risk of an investment. Its simplicity makes it popular in the financial industry because it provides us with single number that represents the potential loss we can incur, at a certain probability level during pre-defined period of time. Using VaR as the only risk measure however has some limitations. There are well known problems of VaR generally not being a coherent risk measure because it violates the subadditivity criteria (Artzner et al., 1999). However, Daníelsson et al. (2013) show that under reasonable assumptions VaR might be subadditive. In this chapter we decided to use a VaR framework because forecasts we obtain from the Panel Quantile Regression Model for Returns are by definition semi-parametric VaRs[6]. Moreover we are not trying to introduce new measures of financial risk, rather we want to show accuracy of the model we proposed in the standard set–up.

Having briefly discussed our motivation to concentrate on the VaR in our analysis we now turn to Value–at–Risk framework itself. Generally there are two main approaches of calculating VaR: (semi)parametric estimation vs. historical simulation. In our work we will concentrate on the parametric approach because it directly enables us to compare forecasts from several benchmark models.

The original parametric VaR calculation was introduced by J.P.Morgan. In their set-up, VaR is derived from the quantile of a standard normal distribution,

$$VaR_i = \gamma_\tau \sigma_i, \tag{2.16}$$

where γ_τ is the τ quantile of the standard normal distribution and σ_i is the volatility of the asset i. If we would like to study VaR of the portfolio instead of the individual assets, σ_i is replaced by the portfolio volatility σ_P. Under the assumption of the multivariate normality σ_P is calculated as

$$\sigma_P = \sqrt{w^\top \Sigma w},$$

where Σ is the covariance matrix and w is the vector of asset weights. We can therefore calculate percentage Value–at–Risk ($\%VaR$) of the given portfolio as a

$$\%VaR_P = \sqrt{\gamma_\tau^2 w^\top \Sigma w}. \tag{2.17}$$

We can rewrite Equation 2.17 in terms of VaRs of the individual assets as

$$\%VaR_P = \sqrt{(w^\top \odot \%VaR^\top)\Omega(w \odot \%VaR)}, \tag{2.18}$$

where $\%VaR$ is a vector of individual percentage VaR estimates, Ω stands for correlation matrix and $\odot$ is the Hadamar product. Alternatively we can also write it as

$$\%VaR_P = \sqrt{\sum_{i=1}^{N} (w_i \%VaR_i)^2 + 2\sum_{i=1}^{N}\sum_{j=i+1}^{N} w_i w_j \%VaR_i \%VaR_j \rho_{i,j}}$$

where w_i is the weight of asset i, $\%VaR_i$ is the percentage VaR of the i^{th} asset and $\rho_{i,j}$ represents correlation between asset i and j.

In the forecasting exercise we will study portfolio Value-at-Risk performance of the 4 benchmark model specifications:

[6] According to Jorion (2007) p.17 "Value–at–Risk describes the *quantile* of the projected distribution of gains an losses over the targeted horizon." Since the VaR is a quantile of returns, and we model quantiles of returns directly by panel quantile regression, we therefore obtain semi-parametric VaR estimates.

- RiskMetrics,
- Panel Quantile Regression(PQR) Model for Returns,
- Univariate Quantile Regression(UQR) Model for Returns,
- portfolio version of Univariate Quantile Regression(Portfolio UQR) Model for Returns.

For calculation of portfolio VaR using the *RiskMetrics* approach, we directly apply Equation 2.17 where Σ is the covariance matrix obtained from RiskMetrics and γ_τ is a cut–off point of standard normal distribution at a given quantile τ.

In case of *PQR* and *UQR*, forecasts of quantiles of return series are considered to be a semi-parametric percentage VaR. The correlation matrix Ω is obtained from the Realized Covariance matrix estimate, Σ, as

$$\Omega = (diag(\Sigma))^{-1/2}\,\Sigma\,(diag(\Sigma))^{-1/2}$$

and therefore Equation 2.18 can be used for VaR calculation.

In contrast to previous approaches, *Portfolio UQR* is calculated in a different fashion. We first create portfolio returns and portfolio volatility series using individual returns and correlation structure obtained from Realized Covariance matrix, Σ, as

$$r_{t,P} = w^\top r_t$$

and

$$\sigma_{t,P} = \sqrt{w^\top \Sigma_t w},$$

where $r_{t,P}$ and $\sigma_{t,P}$ is portfolio return and portfolio volatility at time t respectively and r_t is vector of individual returns at time t. Series $r_{t,P}$ and $\sigma_{t,P}$ are further modeled using Univariate Quantile Regression Model for Returns and the forecasts of the quantiles of the portfolio return series are considered to be semi-parametric percentage portfolio VaR.

Statistical Evaluation

In the statistical comparison, we study absolute performance which tells us whether a model is dynamically correctly specified, i.e. we study goodness-of-fit, and relative performance in which we compare models against each other. For the absolute performance evaluation we use modified version of the Dynamic Quantile test (Engle and Manganelli, 2004), referred to as the CAViaR test by Berkowitz et al. (2011). In their work, Berkowitz et al. (2011) define a "hit" variable in a way that

$$hit_{t+1} = \begin{cases} 1 & if \quad r_{t+1} \leq Q_{r_{t+1}}(\tau) \\ 0 & otherwise \end{cases}$$

i.e. hit_{t+1} is a binary variable taking values 1 if conditional quantile is violated and 0 otherwise. Hit series of a dynamically correctly specified series should be i.i.d Bernoulli distributed with parameter τ

$$hit_{t+1} \sim iid(\tau, \tau(1-\tau)).$$

By construction, *hit* is a binary variable, therefore Berkowitz et al. (2011) propose to test the hypothesis of correct dynamic specification using following logistic regression

$$hit_t = c + \sum_{d=1}^{n} \beta_{1d} hit_{t-l} + \sum_{d=1}^{n} \beta_{2d} Q_{r_{t-d+1}}(\tau) + u_t$$

where u_t is assumed to have a logistic distribution. We use a likelihood ratio test to verify null hypothesis that β's are equal to zero and $\mathbb{P}(hit_t = 1) = \dfrac{e^c}{1+e^c} = \tau$. Exact finite sample critical values for the likelihood ratio test are obtained from Monte Carlo simulation as suggested by Berkowitz et al. (2011).

Relative performance of benchmark models is tested using expected tick loss for pair-wise model comparison (Giacomini and Komunjer, 2005; Clements et al., 2008). The loss function is defined as

$$\mathcal{L}_{\tau,m} = E\left(\left(\tau - I\left(e^m_{t+1} < 0\right) e^m_{t+1}\right)\right),$$

where $I(\cdot)$ is indicator function, $e^m_{t+1} = r_{t+1} - Q^m_{r_{t+1}}(\tau)$ and $Q^m_{r_{t+1}}(\tau)$ is the m'th model quantile forecast. Forecasting accuracy of two models is assessed using Diebold and Mariano (1995) test. Null hypothesis of the test that expected losses of two models are equal i.e. $H_0 : \mathcal{L}_{\tau,1} = \mathcal{L}_{\tau,2}$ is tested against general alternative.

Economic Evaluation

In the economic evaluation, we study portfolio Value-at-Risk forecasts in the modifed Markowitz (1952) approach. From the original work of Markowitz (1952) it differs in a way that we concentrate on the relationship of the return and Value-at-Risk compared to original risk–return trade-off [7]. To overcome the difficulties of specifying a proper model for returns and covariance/correlation matrices we decide to use their ex-post realizations i.e. for day T we use returns realized in day T, realized covariance/correlation matrix in day T and forecasts of univariate VaR for day T.

In general, the efficient frontier of the optimal portfolio can be constructed in two equivalent ways:

1. Expected portfolio return is maximized for various levels of portfolio Value-at-Risk
2. Portfolio Value-at-Risk is minimized for various levels of expected portfolio return

In both approaches asset weights, $w = (w_1, \ldots, w_q)'$, maximizing utility of risk averse investor can be found by solving following problem:

$$\min_{w_{t+1}} \quad w'_{t+1} \widehat{\Xi}_{t+1|t} w_{t+1} \tag{2.19}$$

$$\text{s.t.} \quad l' w_{t+1} = 1$$

[7] Note that if we assume that quantiles of returns are standard normally distributed and we use standard cut–off points, i.e. -1.645 for the 5% quantile, both approaches are equivalent

$$w'_{t+1} \geq 0^8$$
$$w'_{t+1}\widehat{\mu}_{t+1} = \mu_P$$

where w_{t+1} is $n \times 1$ vector of assets weights, l denotes a $n \times 1$ vector of ones, $\widehat{\mu}_{t+1}$ is a vector of ex-post returns, μ_P stands for portfolio return and $\widehat{\Xi}_{t+1|t} = diag\left(\widehat{\%VaR_{t+1|t}}\right) * \widehat{\Omega}_{t+1} * diag\left(\widehat{\%VaR_{t+1|t}}\right)$ represents a correlated Value-at-Risk covariance matrix where $\widehat{\%VaR_{t+1|t}}$ is $n \times 1$ vector of univariate %VaR forecast and $\widehat{\Omega}_{t+1}$ is correlation matrix obtained from realized covariance matrix estimate. Once we solve the optimization problem for different levels of risk, we construct efficient frontier. In the Markowitz-type portfolio optimization exercise we do not allow short-selling in order to meet restrictions imposed mainly by regulators on certain types of investors (pension funds etc.).

The second economic evaluation criteria used in our study is the Global Minimum Value-at-Risk Portfolio. The basic problem of GMVaRP is similar to Markowitz, there are only two differences in the set-up. The first one is the existence of the closed-form solution. As a consequence we are not restricting asset weights because the global minimum of the optimization problem might require negative weights of some assets. The second difference is the absence of a targeted portfolio return. Therefore in some cases we might get negative portfolio return for the asset weights minimizing the overall risk of the portfolio. GMVaRP optimization problem can be written as

$$\underset{w_{t+1}}{min} \quad w'_{t+1}\widehat{\Xi}_{t+1|t}w_{t+1} \tag{2.20}$$

$$\text{s.t.} \quad l'w_{t+1} = 1.$$

In the Kempf and Memmel (2006) paper, it was shown that the analytical solution of the problem is

$$w_{t+1}^{GMVaR} = \frac{\widehat{\Xi}_{t+1|t}^{-1}l}{l'\widehat{\Xi}_{t+1|t}^{-1}l}, \tag{2.21}$$

and portfolio Value-at-Risk corresponding to calculated asset weights is finally obtained as

$$\%VaR_{t+1}^{GMVaR} = w_{t+1}^{GMVaR\prime}\widehat{\Xi}_{t+1|t}w_{t+1}^{GMVaR}.$$

2.5 Simulation Study

Before we analyze the empirical data we would like to show performance of the newly proposed model in a controlled environment. Our aim is to show how various error distributions used for continuous price process simulation affect the performance of the Panel Quantile Regression for Returns model.

[8] We do not allow short-selling in this set-up.

As it is common in the literature, let's assume that the price processes follow jump diffusion processes with stochastic volatility:

$$\begin{aligned} dp_t &= \left(\mu - \frac{\sigma_t^2}{2}\right) dt + \sigma_t dW_{1t} + c_t dN_t \\ d\sigma_t^2 &= \kappa\left(\alpha - \sigma_t^2\right) dt + \gamma\sigma_t dW_{2t}, \end{aligned} \tag{2.22}$$

where W_1 and W_2 are Brownian motions, $c_t dN_t$ is a compound Poisson process with random jump size distributed as $N(0, \sigma_J)$ and $\sigma_J = 0.01$. Parameters in Equation 2.22 are set to the values which are reasonable for a stock price, i.e. $\alpha = 0.04$, $\kappa = 5$, $\gamma = 0.5$ as in Zhang et al. (2005) and $\mu = 0$ because we assume that returns are zero-mean. The volatility parameters satisfy Feller's condition $2\kappa\alpha \geq \gamma^2$, which keeps the volatility process away from the zero boundary. Moreover we assume that W_1 comes from one of the following distributions with Σ being Realized Covariance matrix obtained from the empirical data:

- Multivariate normal distribution, $N(0, \Sigma)$.
- Multivariate Student-t distribution with 9 degrees of freedom, $t_9(0, \Sigma)$.
- Univariate normal distribution, $N(0, 1)$.
- Univariate Student-t distribution with 9 degrees of freedom, $t_9(0, 1)$.

To work with a similar environment as the empirical data, each simulation step consists of 29 time series containing of 7 hours of 1 minutes intra-day prices for 2613 days. From the intra-day prices we calculate daily returns and all the realized measures. In case of multivariate normal and multivariate Student-t distribution we use empirical estimate of Realized Covariance matrix for given day as the input for multivariate random number generation. For each error distribution we run 500 simulations. In each simulation step we use same estimation procedure as in case of empirical data - rolling window of length 1000.

2.5.1 In-Sample Fit

We start with a description of the results with data generated from Multivariate Normal Distribution i.e. $N(0, \Sigma)$. Table 2.1 shows detailed estimation results for 5%, 10%, 90% and 95% quantiles that are most important from an economic point of view for all three model specifications. To get a better view of quantile dynamics we also report lower and upper quartile together with median. The results of the other distributions are presented in the Appendix - Table 2.11, 2.12, 2.13 and we comment here only main differences from Multivariate Normal Distribution.

Table 2.1 reveals significant estimates (except median) for PQR-RV model, with parameter values increasing in quantiles. The median coefficient is zero as a consequence of setting μ in Equation 2.22. Similar to the PQR-RV model, all but median quantiles are statistically significant also for the second model, PQR-RSV. We can notice differences in smaller magnitudes of coefficients in comparison to PQR-RV. Since both positive and negative semivariance should carry equal information in Multivariate Normal distribution, we expect equal coefficients. Finally, PQR-BPV model shows insignificant estimates for jump component, while coefficients for the volatility component are equal to PQR-RV model. This again is consistent with our expectation, as simulated jump variation in the simulations is too small. We conclude with observation that results for all three models are symmetric, as expected.

Tables 2.11, 2.12, and 2.13 reveal similar patterns. Heavy tails introduced to the data with Student-t distribution cause higher coefficients on both tails.

Table 2.1: Multivariate Normal Distribution – Mean of coefficients estimates from Monte-Carlo simulations

τ	5%	10%	25%	50%	75%	90%	95%
				PQR-RV			
$\hat{\beta}_{RV^{1/2}}$	-1.54	-1.15	-0.57	0	0.56	1.14	1.55
	(-18.9)	(-18.37)	(-12.51)	(-0.06)	(11.85)	(17.72)	(17.97)
				PQR-RSV			
$\hat{\beta}_{RS^{+1/2}}$	-1.12	-0.83	-0.43	-0.02	0.36	0.76	1.04
	(-2.17)	(-2.35)	(-2.23)	(-0.18)	(1.96)	(2.29)	(2.07)
$\hat{\beta}_{RS^{-1/2}}$	-1.06	-0.79	-0.38	0.02	0.44	0.86	1.15
	(-2.06)	(-2.21)	(-1.95)	(0.15)	(2.42)	(2.63)	(2.32)
				PQR-BPV			
$\hat{\beta}_{BPV^{1/2}}$	-1.55	-1.15	-0.57	0	0.57	1.15	1.55
	(-18.9)	(-18.46)	(-12.49)	(-0.06)	(11.83)	(17.81)	(17.87)
$\hat{\beta}_{JV^{1/2}}$	0.06	0.04	0.02	0	-0.03	-0.05	-0.06
	(0.49)	(0.56)	(0.45)	(0.01)	(-0.47)	(-0.63)	(-0.52)

Note: Table displays mean of coefficient estimates with corresponding t-statistics in parentheses. Individual fixed effects $\alpha_i(\tau)$ are not reported for brevity.

2.5.2 Out-of-Sample Performance

In the out-of-sample forecasting exercise we start with comparison of absolute performance represented by various measures of unconditional coverage ($\widehat{\tau}_{avg}$, $\widehat{\tau}_{max}$, $\widehat{\tau}_{min}$, $\widehat{\tau}_{avg-dev}$) and dynamic quantile CAViaR test ($\widehat{DQ}_{violations}$) followed by pair-wise relative comparison according to Diebol-Mariano test (DM). For the unconditional coverage we report average unconditional coverage ($\widehat{\tau}_{avg}$) from the Monte-Carlo simulation which indicates how close our model was to theoretical quantile hit rate (i.e. for 5% quantile we expect unconditional coverage to be somewhere around 5%), maximum and minimum unconditional coverage ($\widehat{\tau}_{max}$, $\widehat{\tau}_{min}$) which show the range of possible movements of unconditional coverage rate and average deviation from the theoretical quantile hit rate ($\widehat{\tau}_{avg-dev}$) that shows on average how close our estimates were to theoretical values. Results of Diebold-Mariano test shows us percentage values when the benchmark model was outperformed by its competitors.

In the $PanelA.1$ and $PanelA.2$ of the Table 2.2 we present absolute performance of the PQR and benchmark models respectively. Overall we can say that all the models are dynamically correctly specified in the majority of simulation trials for all the quantiles but median. Models with the lowest average deviation from studied quantile τ are all PQR specifications and UQR for all but median quantile. In case of median Portfolio UQR is the winner. Similarly to in-sample fit we obtain qualitatively identical results when we study data simulated from Multivariate Student-t distribution. When we switch to univariate error distributions the situation changes and the Portfolio UQR seems to be the model with lowest average deviation and the lowest number of dynamically not correctly specified models. However, we must stress that for all but median quantile, all the results are close to each other which indicates that none of the models are systematically misspecified.

A more interesting comparison comes from $PanelB.1$ and $PanelB.2$ of Table 2.2 where PQR models are compared to benchmarks directly. All the PQR variants outperform significantly Portfolio UQR in all studied quantiles and RiskMetrics in all quantiles but median. Median RiskMetrics performance is overall the best which we attribute to the fact that median cut-off point for VaR calculation is zero and by construction series we simulated are supposed to be zero mean. When we concentrate on the comparison of the PQR to UQR situation is identical in both tails - UQR outperforms slightly all PQR specifications. We address this result to the nature of simulated data - data generating process is driven by generated random numbers and contains just little heterogeneity that could possibly translate into the gains using PQR. Median performance however is better for PQR which is result of the averaging in the PQR median calculation. Moreover as the number of estimated parameters is significantly lower in case of PQR compared to UQR, median forecasts are less noisy which translates to better median PQR performance directly. Qualitatively similar results are obtained also for Multivariate Student-t distribution. If we turn to the comparison with univariate distributions, PQR outperform UQR significantly in all studied quantiles. The source of this interesting fact lies in the degree of heterogeneity present in the data. The only source of heterogeneity in our simulated data is the random number generation process. In case of univariate distributions each generated time series has errors that are independent from remaining time series. However, in the multivariate distributions all error terms are affected by each other because we assume some correlation/covariance structure. As a result multivariate random numbers are less heterogeneous compared to univariate one.

Generally, results obtained from the Monte-Carlo simulations helps us justify the use of panel quantile regressions for modelling quantiles of future returns. Our main results are that whatever error distribution for simulation we use, PQR models are specified well dynamically and they dominate RiskMetrics and Portfolio UQR benchmark models. When we use univariate error distributions for random data generation PQR also outperform UQR. In case multivariate error distributions are used, PQR is slightly outperformed by UQR because the simulated data are less heterogeneous compared to univariate error distributions. We also show the importance of covariance structure in the comparison of the results of multivariate and univariate distributions.

Table 2.2: Models performance using data simulated from Multivariate Normal Distribution

		PQR-RV					PQR-RSV					PQR-BPV				
Panel A.1		5%	10%	50%	90%	95%	5%	10%	50%	90%	95%	5%	10%	50%	90%	95%
	$\widehat{DQ}_{violations}$	6.2	4.8	12.2	2.8	7.8	6.8	5	13	2.8	7.8	6.2	5.2	12.4	2.6	7
	$\widehat{\tau}_{avg}$	5.0	10.1	51.4	90.1	95.1	5.0	10.1	51.4	90.1	95.0	5.1	10.1	51.4	90.1	95.0
	$\widehat{\tau}_{max}$	6.5	11.8	55.4	91.6	96.4	6.5	11.7	55.6	91.7	96.4	6.3	11.7	55.4	91.7	96.2
	$\widehat{\tau}_{min}$	3.6	8.4	47.5	88.5	93.7	3.5	8.3	47.4	88.5	93.7	3.5	8.5	47.4	88.5	93.6
	$\widehat{\tau}_{avg-dev}$	0.0	0.1	1.4	0.1	0.1	0.0	0.1	1.4	0.1	0.0	0.1	0.1	1.4	0.1	0.0

		RiskMetrics					UQR					Portfolio UQR				
Panel A.2		5%	10%	50%	90%	95%	5%	10%	50%	90%	95%	5%	10%	50%	90%	95%
	$\widehat{DQ}_{violations}$	9.2	14	6.6	18.8	7.4	6.8	5.2	14.2	3	8.2	7	4.8	3	4.2	7
	$\widehat{\tau}_{avg}$	5.2	9.3	50.5	91.1	95.0	5.0	10.1	51.5	90.1	95.1	4.8	9.6	49.9	90.4	95.2
	$\widehat{\tau}_{max}$	7.1	11.2	54.1	92.7	96.4	6.3	11.7	55.6	91.8	96.4	6.1	11.5	52.5	91.8	96.3
	$\widehat{\tau}_{min}$	3.3	7.1	46.5	89.3	93.6	3.5	8.4	47.7	88.5	93.5	3.7	8.1	47.4	88.5	94.0
	$\widehat{\tau}_{avg-dev}$	0.2	-0.7	0.5	1.1	0.0	0.0	0.1	1.5	0.1	0.1	-0.2	-0.4	-0.1	0.4	0.2

		benchmark														
		RiskMetrics					UQR					Portfolio UQR				
Panel B.1		5%	10%	50%	90%	95%	5%	10%	50%	90%	95%	5%	10%	50%	90%	95%
PQR-RV	*DM*	58.8	62.6	0.2	65.6	54.8	3.4	2.6	9.4	4.4	4.2	47.4	43.6	20.4	44.8	48.6
PQR-RSV	*DM*	58.4	62.2	0.2	64.4	54	2.8	2	9.4	2.6	3	46.6	42.6	20.4	44.2	47.2
PQR-BPV	*DM*	58	61.8	0.2	63.8	54	1.6	1	8.8	1.6	1.2	44.2	40.4	20.6	42.4	44.8

		PQR-RV					PQR-RSV					PQR-BPV				
Panel B.2		5%	10%	50%	90%	95%	5%	10%	50%	90%	95%	5%	10%	50%	90%	95%
RiskMetrics	*DM*	0.2	0	13.8	0	0	0.2	0	14.2	0	0	0.2	0	13	0	0
UQR	*DM*	6.8	8.8	2.6	9.2	9	12	13.6	2.8	15.2	12.2	8.8	11.2	3	12.4	12
Portfolio UQR	*DM*	0.2	1	0.6	0	0.6	0.2	1.6	0.6	0	0.6	0.2	1	0.6	0	0.8

Note: Table displays absolute and relative performance of PQR models for returns with RV, RSV and BPV as regressors and benchmark models. All values are in %.

Panel A.1 reports absolute performance of PQR models, *Panel A.2* reports absolute performance of benchmark models. For each model and quantile τ, percentage of violations of the CAViaR test for correct dynamic specification ($\widehat{DQ}_{violations}$), average unconditional coverage ($\widehat{\tau}_{avg}$),maximum unconditional coverage ($\widehat{\tau}_{max}$), minimum unconditional coverage ($\widehat{\tau}_{min}$) and average deviation of unconditional coverage from given quantile τ ($\widehat{\tau}_{avg-dev}$)

Panel B.1 and *Panel B.2* report relative performance of Panel Quantile Regression Models for Returns in comparison to benchmark models and relative performance of benchmark models in comparison to Panel Quantile Regression Models for Returns respectively. For each specification and quantile τ we report percentage of statistically better performance according to Diebold-Mariano(DM) test at 5% significance level.

2.6 Results: the Role of Unobserved Heterogeneity in Tails

Confident about the performance of our modeling strategy in a controlled environment, we turn to applications of the proposed models on empirical data. First, we describe the in-sample fit of the PQR-RV, PQR-RSV and PQR-BPV model specifications. Second, we present results for our out-of-sample Value–at–Risk forecasting exercise. Third, we complement our statistical evaluation by computing a simple portfolio allocation exercise where we study Global Minimum Value-at-Risk Portfolios and Markowitz like relationships between Value-at-Risk and return of the portfolio.

Our empirical application is carried out using 29 U.S. stocks[9] that are traded at New York Stock Exchange. These stocks have been chosen according to market capitalization and their liquidity. Sample we study spans from July 1, 2005 to December 31, 2015 and we consider trades executed within U.S. business hours (9:30 – 16:00 EST). In order to ensure sufficient liquidity and eliminate possible bias we explicitly exclude weekends and bank holidays (Christmas, New Year's Day, Thanksgiving Day, Independence Day). In total, our final dataset consists of 2613 trading days. Basic descriptive statistic of the data can be found in Table 2.10 in Appendix.

For estimation and forecasting purposes we use a rolling window with fixed length of 1000 observations,[10] hence our model is always calibrated on a 4 year history. Our analysis is restricted to 5 minutes intraday log-returns that are used for computation of the daily returns and realized measures.

All the results presented in this section were obtained using pure fixed effects panel quantile regression, i.e. penalty parameter λ set to 0. In the Appendix[11] we present also estimation results when $\lambda = 1$ which serves as a robustness check.

2.6.1 In-Sample Fit

Estimation results are detailed in the Table 2.3. In addition, to get a better view of the dynamics, we show results of the PQR-RV, PQR-RSV and PQR-BPV also graphically in the Figures 2.1, 2.2 and 2.3 respectively.

[9] Apple Inc. (AAPL), Amazon.com, Inc. (AMZN), Bank of America Corp (BAC), Comcast Corporation (CMCSA), Cisco Systems, Inc. (CSCO), Chevron Corporation (CVX), Citigroup Inc. (C), Walt Disney Co (DIS), General Electric Company (GE), Home Depot Inc. (HD), International Business Machines Corp. (IBM), Intel Corporation (INTC), Johnson & Johnson (JNJ), JPMorgan Chase & Co. (JPM), The Coca-Cola Co (KO), McDonald's Corporation (MCD), Merck & Co., Inc. (MRK),Microsoft Corporation (MSFT), Oracle Corporation (ORCL), PepsiCo, Inc. (PEP), Pfizer Inc. (PFE), Procter & Gamble Co (PG), QUALCOMM, Inc. (QCOM), Schlumberger Limited. (SLB), AT&T Inc. (T), Verizon Communications Inc. (VZ), Wells Fargo & Co (WFC), Wal-Mart Stores, Inc. (WMT), Exxon Mobil Corporation (XOM).

[10] We have tried different length of rolling window with the qualitative results of our analysis remaining unchanged. These results are available from authors upon request.

[11] Table 2.17, Figure 2.11, Figure 2.12 and Figure 2.13

Table 2.3: Coefficient estimates of Panel Quantile Regressions

τ	5%	10%	25%	50%	75%	90%	95%
				PQR-RV			
$\hat{\beta}_{RV^{1/2}}$	-1.5	-1.16	-0.6	-0.01	0.56	1.11	1.42
	(-23.5)	(-20.62)	(-15.65)	(-0.2)	(20.37)	(24.84)	(20.7)
				PQR-RSV			
$\hat{\beta}_{RS^{+1/2}}$	-0.97	-0.75	-0.44	-0.16	0.18	0.41	0.53
	(-12.74)	(-11.98)	(-8.31)	(-2.73)	(2.69)	(4.55)	(4.51)
$\hat{\beta}_{RS^{-1/2}}$	-1.18	-0.9	-0.41	0.14	0.62	1.14	1.49
	(-11.72)	(-14.05)	(-9.9)	(2.7)	(9.17)	(13.66)	(10.39)
				PQR-BPV			
$\hat{\beta}_{BPV^{1/2}}$	-1.55	-1.18	-0.62	0	0.59	1.15	1.44
	(-19.5)	(-18.15)	(-16.27)	(-0.13)	(23.84)	(23.22)	(25.72)
$\hat{\beta}_{JV^{1/2}}$	-0.25	-0.21	-0.14	-0.03	0.06	0.21	0.44
	(-3.24)	(-3.54)	(-3.39)	(-0.58)	(1.11)	(1.9)	(2.56)

Note: Table displays coefficient estimates with bootstraped t-statistics in parentheses. Individual fixed effects $\alpha_i(\tau)$ are not reported for brevity - they are available from authors upon request.

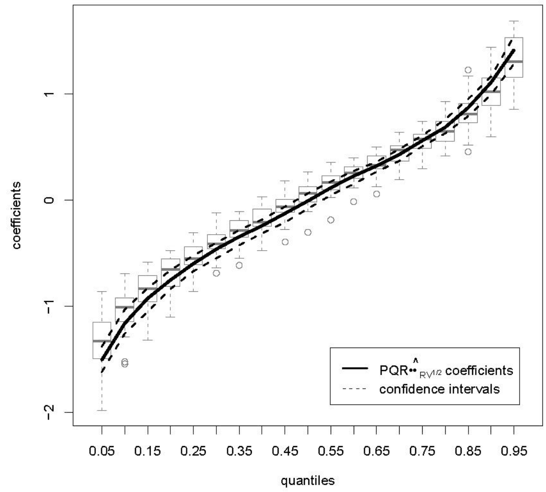
Table 2.3 reveals that parameters of the first model specification (PQR-RV) where lagged volatility is used to explain conditional quantiles of returns are significantly different from zero for all quantiles except median. Moreover, signs of the estimated parameters correspond to our expectations – coefficients at lower (upper) quantiles are negative (positive). Note that these values can be interpreted directly as semi-parametric estimates of Value–at–Risk. Our model hence shows that standard VaR from RiskMetrics in which quantiles of standard normal distribution are rescaled by volatility overestimates both left as well as right tails (corresponding values for the 5% and 95% quantiles of standard normal distribution are -1.645 and 1.645 respectively). Furthermore, insignificant parameter estimate at median confirms the hypothesis about the randomness/unpredictability of the short-term returns.

In the Table 2.3, we can also see that absolute values of parameter estimates are not symmetric around median which highlight the relative importance of the realized volatility on the estimation of the lower quantiles of returns. We arrive to a similar conclusion also when looking at the Figure 2.1 that compares and displays PQR-RV estimates together with their corresponding 95% confidence intervals and individual UQR-RV parameter estimates plotted in boxplots. Importantly, Figure 2.1 shows that once we control for unobserved heterogeneity by the PQR-RV, past volatility has a larger influence on both the lower and the upper quantiles of returns than the majority of individual UQR-RV.

This is highlighted in far quantiles, e.g. coefficient of PQR-RV in 5% quantile is -1.5 whereas median of individual UQR-RV coefficient is -1.33 (mean -1.36) or 95% quantile PQR-RV coefficient is 1.42 and median of individual UQR-RV is only 1.30 (mean 1.31).

This finding constitutes an important empirical result, as we document unobserved heterogeneity in far quantiles that needs to be controlled.

Figure 2.1: PQR-RV parameter estimates



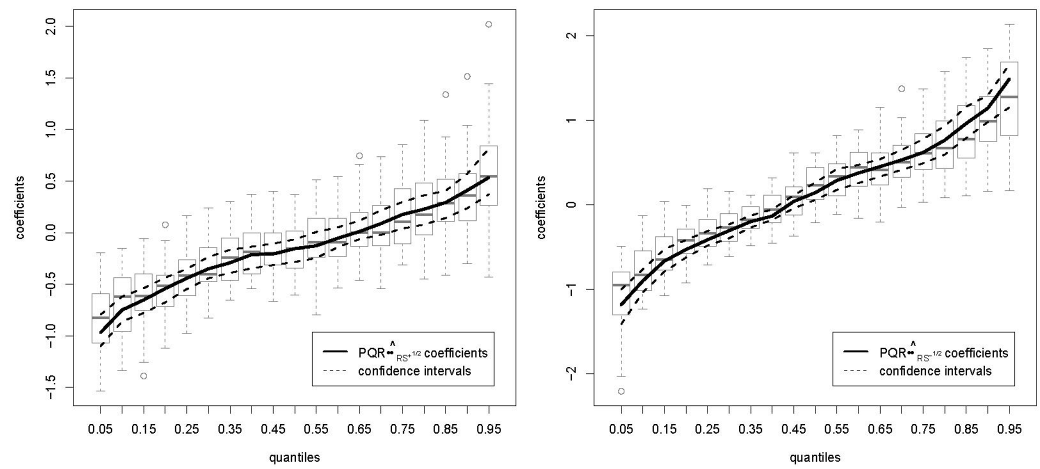
Note: Parameter estimates with corresponding 95% confidence intervals from the PQR-RV specification are plotted by solid and dashed lines respectively. Individual UQR-RV estimates are plotted in boxplots.

Coefficients from the second model specification (PQR-RSV), where Realized Variance is decomposed into realized downside (RS^-) and upside (RS^+) semivariance are significantly different from zero for all considered quantiles. The magnitude of the coefficients driving impact of both variables is highest at far quantiles showing strongest impact of both negative, and positive semivariance on tails of the returns distributions. However, influence of RS^- is far more important in the upper quantiles where it dominates RS^+. On the contrary, in the lower quantiles, values of parameters are close to each other and therefore we cannot draw the similar conclusion as in upper quantiles. Median performance is bit different from PQR-RV case.

We can see that coefficients for both RS^- and RS^+ are statistically significant and in the case that magnitude of RS^- and RS^+ is equal, they sum to -0.02 which translates into loss in 50% of cases. However as theory and stylized facts about financial time series suggest influence of negative returns and subsequently negative semivariances should be greater than the effect of positive ones. Therefore one can not draw straightforward conclusions about the sign and magnitude of median return.

Figure 2.2: PQR-RSV parameter estimates

(a) $\hat{\beta}_{RS^{+1/2}}$ (b) $\hat{\beta}_{RS^{-1/2}}$



Note: For both realized upside and downside semivariance parameters estimates with corresponding 95% confidence intervals are plotted by solid and dashed lines respectively. Individual UQR-RSV estimates are plotted in boxplots.

The careful reader might also notice that median coefficient of $\hat{\beta}_{RS^{+1/2}}$ is negative and opposite is true for $\hat{\beta}_{RS^{-1/2}}$. Explanation of this feature rely on short and long term mean-reversion nature of the returns and the fact that we are using lagged values of realized semivariances as regressors. If we put it together negative return at day $t-1$ will cause that $RS^{-}_{t-1} > RS^{+}_{t-1}$ and prediction of the median quantile for day t will be positive because $\hat{\beta}_{RS^{-1/2}}$ is positive and vice versa for positive return and subsequent $RS^{-}_{t-1} < RS^{+}_{t-1}$. This behaviour leads to mean-reversion. Results of our analysis are also supported by the Figure 2.2.

Similar to the PQR-RV specification, we can see in Figure 2.2 that controlling for unobserved heterogeneity among financial assets is important because the influence of both downside and upside semivariance is greater in the lower quantiles than in individual UQR-RSV. For example in 5% quantile coefficients obtained by PQR-RSV are -0.97 and -1.18 for RS^{+} and RS^{-} respectively, however median values of individual UQR-RSV are -0.82 (mean -0.84) for RS^{+} and -0.95 (mean -1.1) for RS^{-}. Moreover, in the upper quantiles of negative semivariance (Figure 2.2b) PQR-RSV coefficients differs substantially from individual UQR-RSV (95% quantile $\hat{\beta}_{RS^{-1/2}}$ coefficient of 1.49 vs. individual UQR-RSV median/mean coefficient of 1.28/1.27), however, the opposite is true for RS^{+} (95% quantile $\hat{\beta}_{RS^{+1/2}}$ coefficient of 0.54 vs. individual UQR-RSV median/mean coefficient of 0.55/0.55). These findings support our previous conclusion that RS^{-} influences future upper quantiles of returns more than RS^{+}.

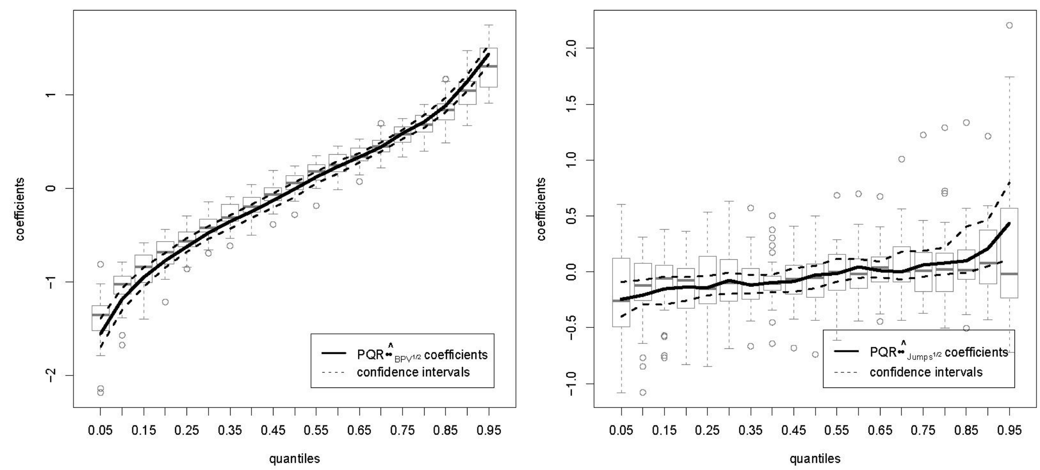
Finally, Table 2.3 reveals interesting results about parameter estimates of the third model specification (PQR-BPV), where the Bi-Power Variation and Jump Component are used to drive the return quantiles. We can infer that jumps have significant impact on both far upper and lower quantiles of future returns. To be precise, magnitude of the jump coefficient $\hat{\beta}_{JV^{1/2}}$ is highest for 95% quantile with the value of 0.44. For the remaining above median quantiles, jumps are not statistically significant and therefore the influence of the Quadratic Variation reduces to Integrated Variance represented by Bi-Power Variation. We observe the opposite situation for the below median quantiles where $\hat{\beta}_{JV^{1/2}}$ coefficients are always significant.

Figure 2.3 helps us to confirm results of our previous analysis graphically. If we compare Figure 2.3a to Figure 2.1 we get an almost identical picture. Moreover, in Figure 2.3b, we can see that from the 45% to 85% quantiles confidence intervals of the jump component are getting wider and include zero. Once we combine these two findings, we can state that for these quantiles Quadratic Variation reduces to Integrated Variance. In contrast none of the confidence intervals of the 5% to 40% quantiles contain zero which highlights the relative importance of the jump component in the modelling lower future quantiles of returns.

Overall, results of the in-sample analysis show asymmetric impact of the regressors on the quantiles of future returns. This impact is higher in the below median quantiles. We have also found evidence for positive/negative news asymmetry. This asymmetry is the highest in the 95% quantile (0.53 coefficient of RS^{+} vs. 1.49 of RS^{-}) while 5% quantile shows only little asymmetry (-0.97 in case of RS^{+} vs. -1.18 for RS^{-}). In addition we show importance of jumps for below median and far above median quantiles. Importantly, we document unobserved heterogeneity in far quantiles. We have also tested all three models (PQR-RV, PQR-RSV, PQR-BPV) for correct dynamic specification and we have found that none of them is systematically misspecified.

Figure 2.3: PQR-BPV parameter estimates

(a) $\hat{\beta}_{BPV^{1/2}}$ (b) $\hat{\beta}_{JV^{1/2}}$



Note: For both realized bi-power variation and jump component parameters estimates with corresponding 95% confidence intervals are plotted by solid and dashed lines respectively. Individual UQR-BPV estimates are plotted in boxplots

2.6.2 Out-of-Sample Performance

We now turn to results of our out-of-sample forecasting exercises. Similar to our simulation study, we are analyzing the Value–at–Risk performance of an equally weighted portfolio of the 29 stocks described earlier. Results of our analysis are presented in the following way: first we comment on the absolute performance of the PQR models; second the absolute performance of the benchmark models is discussed; third we concentrate on the most interesting relative performance comparison of the PQR models with respect to the benchmark models. All results are summarized in the Table 2.4.

The unconditional coverage, $\widehat{\tau}$, shown in *Panel A.1* and *Panel A.2* of Table 2.4 reveals that almost all models underestimate risk. Specifically, values of unconditional coverage are higher than corresponding quantiles τ, with few exceptions. Median quantiles, as well as the 5% quantile of Portfolio UQR and 90% quantile of the PQR-RSV overestimate risk. We must also stress here that the deviation from nominal quantile rates is generally lower than 1%, and we can not reject hypothesis of correct unconditional coverage.

If we turn to median performance, we can see that all the models overestimate risk. Moreover we can see that deviations from the nominal quantiles are higher compared to off-median quantiles. We address this finding to the nature of financial time series especially to stylized fact about the unpredictability of the returns. More importantly, this result corresponds to our motivation of explaining quantiles of the cross-section of market returns instead of expected value. This is in line with our previous result that median estimates are not statistically significant.

If we concentrate on the correct dynamic specification of the models represented by CAViaR test (i.e. the second and third line of the *Panel A.1* and *A.2*), we see that all the models in all quantiles are dynamically correctly specified except for the median of RiskMetrics. In this case we strongly reject null hypothesis of proper dynamic specification given p-value<0.01. We attribute the poor median RiskMetrics performance to the construction of Equation 2.17 where the cut-off point at 50% quantile, $\gamma_{50\%}$, is 0[12].

Relative performance of the PQR models is summarized in the *Panel B*[13]. Results of our analysis indicate good relative performance of PQR models. All three Panel Quantile model specifications (PQR-RV, PQR-RSV and PQR-BPV) significantly outperform RiskMetrics in all studied quantiles. Moreover, all PQR specifications consistently outperform Portfolio UQR in upper quantiles and UQR in several quantiles i.e. PQR-RV outperform individual UQR estimates in 10% quantile, however performance of PQR-RSV is significantly better in the 95% quantile and PQR-BPV delivers significantly more accurate forecasts than individual UQR in 5% and 10% quantiles. If we concentrate on the full pair-wise comparison, the most important is the performance of the UQR as the main competitor of the PQR specifications. In all of the studied quantiles, UQR is not able to outperform any of the PQR specification. This fact highlights the importance to control for unobserved heterogeneity among the assets. Moving from comparison of PQR and UQR models to the relative performance of the Portfolio UQR, we can see that

[12] The median of standard normal distribution is 0.

[13] For brevity we report in Table 2.4 only pair-wise comparison against benchmark models, full matrix of pairwise comparison is available from authors upon request.

it outperforms RiskMetrics only at 5% and 10% quantiles. In contrast UQR, similar to PQR, outperform RiskMetrics in all studied quantiles. These results reveal the importance of the asset specific contribution to overall future portfolio risk as the approach of firstly aggregating data and secondly modeling them is not able to capture dynamics creating variation in the distribution of future portfolio returns.

Table 2.4: Out-of-sample performance of various specifications of Panel Quantile Regression Model for Returns

		PQR-RV					PQR-RSV					PQR-BPV				
Panel A.1	τ	5%	10%	50%	90%	95%	5%	10%	50%	90%	95%	5%	10%	50%	90%	95%
	$\widehat{\tau}$	0.060	0.108	0.465	0.901	0.959	0.059	0.107	0.465	0.899	0.960	0.058	0.107	0.465	0.902	0.960
	$\widehat{DQ}$	8.917	3.373	10.157	6.939	5.686	8.180	3.339	10.129	1.476	9.152	7.956	4.625	10.157	5.298	6.210
	p-val	0.178	0.761	0.118	0.326	0.459	0.225	0.765	0.119	0.961	0.165	0.241	0.593	0.118	0.506	0.400
		RiskMetrics					UQR					Portfolio UQR				
Panel A.2	τ	5%	10%	50%	90%	95%	5%	10%	50%	90%	95%	5%	10%	50%	90%	95%
	$\widehat{\tau}$	0.061	0.094	0.451	0.919	0.958	0.061	0.107	0.467	0.902	0.960	0.043	0.099	0.491	0.909	0.955
	$\widehat{DQ}$	9.652	3.096	20.600	9.452	10.899	8.323	3.041	9.067	7.174	6.796	9.426	5.988	3.273	4.507	3.238
	p-val	0.140	0.797	0.002	0.150	0.092	0.215	0.804	0.170	0.305	0.340	0.151	0.425	0.774	0.608	0.778

		benchmark														
		RiskMetrics					UQR					Portfolio UQR				
Panel B	τ	5%	10%	50%	90%	95%	5%	10%	50%	90%	95%	5%	10%	50%	90%	95%
PQR-RV	DM	**-2.430**	**-2.259**	**-3.347**	**-2.127**	**-1.935**	0.125	**-1.734**	1.350	0.801	-0.362	-0.733	-1.590	-0.310	**-2.053**	**-2.260**
	p-val	0.008	0.012	0.000	0.017	0.027	0.550	0.041	0.911	0.788	0.359	0.232	0.056	0.378	0.020	0.012
PQR-RSV	DM	**-2.368**	**-2.249**	**-3.561**	**-2.367**	**-2.242**	1.244	-1.569	-0.438	-1.268	**-2.023**	-0.558	-1.580	-0.758	**-2.921**	**-3.099**
	p-val	0.009	0.012	0.000	0.009	0.012	0.893	0.058	0.331	0.102	0.022	0.289	0.057	0.224	0.002	0.001
PQR-BPV	DM	**-2.540**	**-2.422**	**-3.304**	**-2.055**	**-1.851**	**-1.796**	**-1.887**	1.424	0.839	0.703	-1.191	**-1.887**	-0.294	**-1.978**	**-1.705**
	p-val	0.006	0.008	0.000	0.020	0.032	0.036	0.030	0.923	0.799	0.759	0.117	0.030	0.384	0.024	0.044

Note: Table displays absolute and relative performance of PQR models for returns with RV, RSV and BPV as regressors and benchmark models.

Panel A.1 reports absolute performance of PQR models, *Panel A.2* reports absolute performance of benchmark models. For each model and quantile τ, unconditional coverage ($\widehat{\tau}$), the value of the CAViaR test for correct dynamic specification ($\widehat{DQ}$) with corresponding Monte Carlo based p-value. Not correctly dynamically specified models are underlined.

Panel B reports relative performance of Panel Quantile Regression Models for Returns. For each specification and quantile τ we report Diebold-Mariano test statistics for pairwise comparison with benchmark models($\widehat{DM}$) with corresponding p-value. Significantly more accurate forecasts with respect to benchmark models at the 5% significance level are in bold. Full matrix of pairwise comparison is available from authors upon request

2.6.3 Economic Evaluation

We would like to see if statistical gains also translate to economic value. We concentrate on the comparison of 3 models – PQR-RV, UQR and RiskMetrics, and refrain from presenting results for PQR-RSV and PQR-BPV for the sense of brevity. The construction of Portfolio UQR rules out economic evaluation in our set-up because asset weights will be set before applying the quantile regression, and therefore results will be driven by covariance structure only.

We start description of the results by Global Minimum Value-at-Risk Portfolio followed by Markowitz like optimization where we show Value-at-Risk – Return relationship. In both approaches, we use annualized portfolio returns [14] and annualized portfolio Value-at-Risks [15] of the whole out–of–sample period. In the GMVaRP comparison, we focus on both left and right tail together with median because we do not set any constraints regarding asset weights - according to Equation 2.21 GMVaRP has a closed form solution. On the contrary, the Markowitz like optimization is purely numeric and does not offer a closed form solution. Therefore we restrict our analysis on long only positions. As a result we concentrate on the left tail of the return distribution only which shows potential loss for the investor.

Results of the GMVaRP analysis are displayed in Table 2.5. The PQR-RV model performs best in all quantiles except for the median, where UQR has the lowest VaR. RiskMetrics ended last, and we must note that for the median quantile we were not able to calculate value of GMVaRP due to the problem of singularity of the correlated Value-at-Risk matrix.[16]

Table 2.5: Global Minimum Value-at-Risk Portfolio

τ	5%	10%	50%	90%	95%
PQR-RV	**11.76**	**8.69**	0.02	**9.46**	**12.37**
UQR	11.85	8.79	**0.01**	9.52	12.43
RiskMetrics	12.77	9.95	NaN	9.95	12.77

Note: Table displays absolute percentage values of Global Minimum Value-at-Risk Portfolio for given quantile τ. Best model for given quantile is reported in bold.

Efficient frontiers of Value-at-Risk – return trade-off are plotted in Figure 2.4a for 5% and Figure 2.4b for 10% quantile. In both quantiles the model with the best performance is PQR-RV. Similarly to GMVaRP analysis second best performance is achieved by UQR and the model with the worst VaR–return trade–off is RiskMetrics. In Figure 2.4b we can also see that benefits from using PQR are greater for lower values of Value-at-Risk.

[14] $\left(\prod_{t=1}^{T}(1+r_t)\right)^{\frac{250}{T}}$

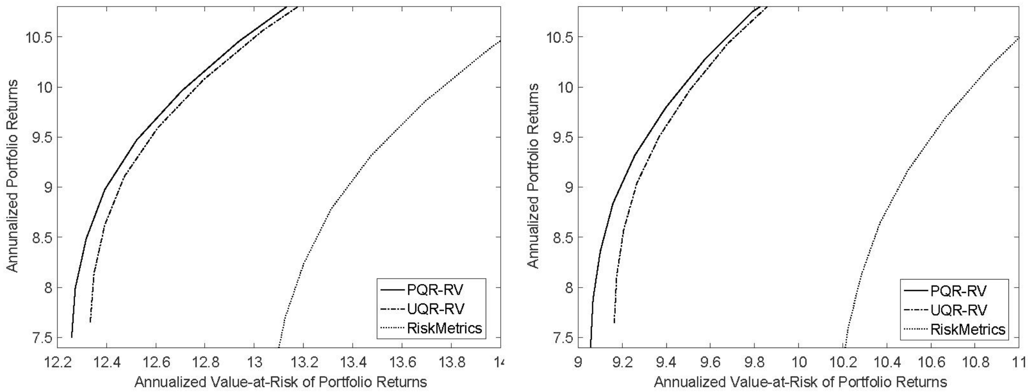
[15] $\sqrt{250}\frac{\sum_{t=1}^{T}\%VaR_t}{T}$

[16] If we set cut-off point in Equation 2.17 to zero we get a singular matrix of zeros that is not invertible.

Overall we can say that Panel Quantile Regression Model for Returns generates better economic performance than the remaining benchmark models.

Figure 2.4: Value-at-Risk – Return efficient frontiers

(a) 5%VaR (b) 10%VaR



Note: Percentage values of portfolio VaR and returns are displayed.

2.7 Results: Common Risk Factors in Tails

Further, we select VIX Index as exogenous factor that have a potential to drive the tails of the return distributions. VIX is often used as the measure of the ex-ante/anticipated uncertainty and it well complements the realized volatility used in our previous analysis. It is also referred to as the "fear gauge" since it measure the expectations about the 30-day volatility using the weighted and aggregated prices of the call and put options with various strike prices on the S&P 500 Index. The value of the index is calculated according to VIX methodology[17] as

$$\sigma^2_{VIX} = \frac{2}{T} \sum_i \frac{\Delta K_i}{K_i^2} e^{RT} Q(K_i) - \frac{1}{T} \left(\frac{F}{K_0} - 1 \right)^2, \tag{2.23}$$

where T is time to expiration; F is the forward index level derived from index option prices; K_0 is the first strike below the forward index level F; K_i is the strike price of ith out-of-the-money option (call if $K_i > K_0$, put if $K_i < K_0$); ΔK_i is the interval between strike prices; R is the risk-free interest rate to expiration; and $Q(K_i)$ is the midpoint of the bid-ask spread for each option with strike K_i.

The value of the Volatility Index represents the annual percentage volatility and it is reported by CBOE as

$$VIX = 100\sigma_{VIX}. \tag{2.24}$$

The daily counterpart of the annual option implied volatility measure is constructed by dividing index by $\sqrt{250}$. We further divide daily VIX by 100 to scale it to units of Realized Volatility, i.e. $VIX_{daily} = \frac{VIX_{annual}}{\sqrt{250}} \frac{1}{100}$. The historical data can be freely downloaded from the Federal Reserve Bank at Saint Louis.[18]. Moreover, since the index was launched at 1993 it effectively covers sample period previously used in our empirical analysis.

In the empirical application, we estimate panel quantile regression models containing VIX Index as defined in Equation 2.10, 2.11 and 2.12. In the optimization we have set penalty λ to zero since the minimization of the Bayesian Information Criteria produce almost identical result as in the models without VIX. The results of the in-sample fit are summarized in the Table 2.6 and Figure 2.5, 2.6 and 2.7

[17] Full details of the VIX calculation can be found at `http://www.cboe.com/micro/vix/vixwhite.pdf`

[18] `https://fred.stlouisfed.org/series/VIXCLS`

Table 2.6: Coefficient estimates of Panel Quantile Regressions

τ	5%	10%	25%	50%	75%	90%	95%
				PQR-RV-VIX			
$\hat{\beta}_{RV^{1/2}}$	-1.33	-0.99	-0.53	-0.04	0.4	0.81	1.12
	(-10.87)	(-9.65)	(-8.77)	(-0.83)	(11)	(7.76)	(9.09)
$\hat{\gamma}_{VIX}$	-0.32	-0.29	-0.12	0.05	0.26	0.44	0.44
	(-4.39)	(-4.49)	(-3.45)	(1.9)	(9.58)	(6.3)	(5.38)
				PQR-RSV-VIX			
$\hat{\beta}_{RS^{+1/2}}$	-0.78	-0.65	-0.4	-0.18	0.08	0.28	0.52
	(-9.46)	(-5.91)	(-6.12)	(-3.07)	(1.22)	(3.5)	(4.45)
$\hat{\beta}_{RS^{-1/2}}$	-1.09	-0.75	-0.35	0.11	0.49	0.89	1.1
	(-7.76)	(-10.54)	(-6.9)	(2)	(6.32)	(5.63)	(5.83)
$\hat{\gamma}_{VIX}$	-0.33	-0.28	-0.12	0.06	0.26	0.41	0.42
	(-4.63)	(-4.58)	(-3.35)	(1.99)	(9.28)	(5.73)	(4.7)
				PQR-BPV-VIX			
$\hat{\beta}_{BPV^{1/2}}$	-1.35	-0.99	-0.54	-0.03	0.42	0.82	1.14
	(-9.88)	(-9.5)	(-7.94)	(-0.65)	(9.42)	(7.68)	(10.34)
$\hat{\beta}_{JV^{1/2}}$	-0.21	-0.14	-0.14	-0.04	0.07	0.17	0.44
	(-2.76)	(-2.43)	(-2.79)	(-0.77)	(1.03)	(1.39)	(2.58)
$\hat{\gamma}_{VIX}$	-0.33	-0.3	-0.12	0.04	0.26	0.45	0.44
	(-4.1)	(-4.9)	(-3.57)	(1.87)	(8.35)	(6.87)	(5.7)

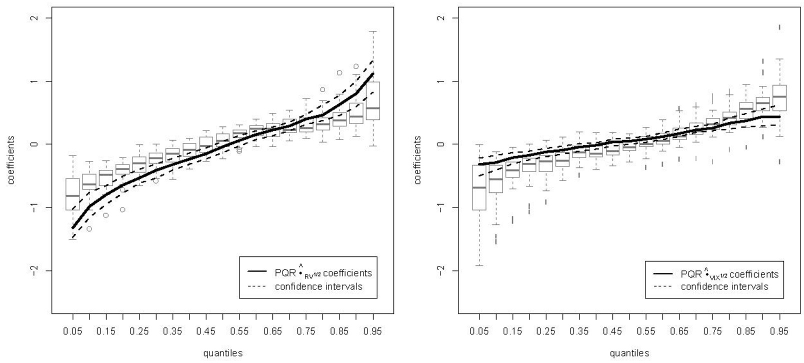
Note: Table displays coefficient estimates with bootstraped t-statistics in parentheses. Individual fixed effects $\alpha_i(\tau)$ are not reported for brevity - they are available from authors upon request.

Table 2.6 documents stability of the relative influence of the VIX on the quantiles of future returns - the VIX coefficient estimates are of the same magnitude for all realized measures model specifications. Although market participants perceive VIX as the fear-index, our analysis reveals higher relative influence in the upper quantiles compared to lower ones, e.g. in *RV + VIX* model specification 0.44 coefficient estimate of 95% quantile vs. -0.32 of 5% quantile. Moreover, when we compare Table 2.6 to Table 2.3 (PQR models with and without VIX) we can see that VIX index reduce relative influence of the realized measures more in the upper than in lower quantiles. In the *"RV + VIX"* and *"BPV + VIX"* specifications the coefficients are reduced by 0.17 and 0.20 respectively in 5% quantile while in the 95% quantile the reduction is 0.30 in both specifications. In the *"RSV + VIX"* specification the total reduction in the coefficients is higher than in the previous two cases (0.28 and 0.40 in 5% and 95% quantiles respectively) and the influence reduction of the positive semivariance is higher than that of negative semivariance in the 5% quantile and vice-versa in 95% quantile where positive semivariance is almost not reduced and the influence of the negative semivariance is lowered by 0.39. Figure 2.5, 2.6 and 2.7 support our findings also graphically - in all three figures, the patterns of the $\hat{\gamma}_{VIX}$ coefficient estimates are almost identical.

Overall, we conclude that VIX carries an important part of the information about risk that is not fully captured by any of the realized measures, and the expectations about the future risk affects higher quantiles more than the lower quantiles. Controlling for the unobserved heterogeneity and idiosyncratic volatility, VIX proves to be a strong common factor driving the tails of the return distributions.

Figure 2.5: PQR-RV-VIX parameter estimates

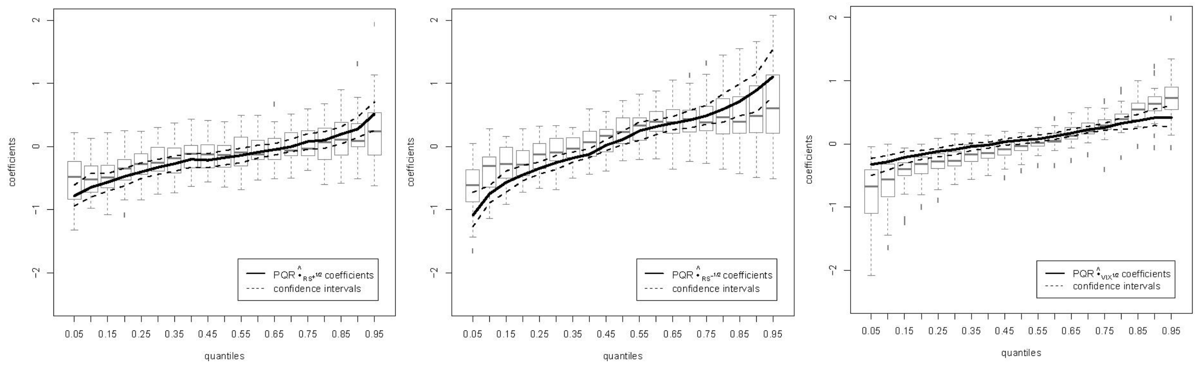
(a) $\hat{\beta}_{RV^{1/2}}$ (b) $\hat{\gamma}_{VIX}$



Note: For both realized volatility and VIX index parameters estimates with corresponding 95% confidence intervals are plotted by solid and dashed lines respectively. Individual UQR-RV-VIX estimates are plotted in boxplots.

Figure 2.6: PQR-RSV-VIX parameter estimates

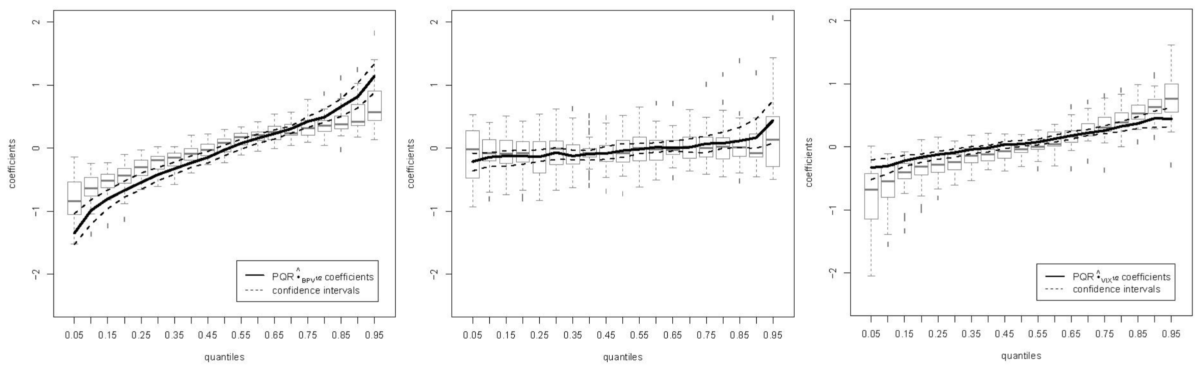
(a) $\hat{\beta}_{RS^{+1/2}}$ (b) $\hat{\beta}_{RS^{-1/2}}$ (c) $\hat{\gamma}_{VIX}$



Note: For all realized upside semivariance, downside semivariance and VIX index parameters estimates with corresponding 95% confidence intervals are plotted by solid and dashed lines respectively. Individual UQR- RSV-VIX estimates are plotted in boxplots.

Figure 2.7: PQR-BPV-VIX parameter estimates

(a) $\hat{\beta}_{BPV^{1/2}}$ (b) $\hat{\beta}_{JV^{1/2}}$ (c) $\hat{\gamma}_{VIX}$



Note: For all realized bi-power variation, jump component and VIX index parameters estimates with corresponding 95% confidence intervals are plotted by solid and dashed lines respectively. Individual UQR-BPV-VIX estimates are plotted in boxplots.

The out-of-sample performance of the models containing VIX index is summarized in Table 2.7. In the absolute performance, *Panel A.1* and *Panel A.2* of the Table 2.7, there is only one change compared to models without VIX - the 5% quantile of the PRQ-RV-VIX specification is not dynamically correctly specified. All the remaining results of absolute performance qualitatively match the results of the models without VIX index presented in Table 2.7, i.e. unconditional coverage $\hat{\tau}$ is close to the nominal quantile rates and all the models in all quantiles are dynamically correctly specified except for the median of RiskMetrics.

In the relative performance comparison presented in *Panel B* of the Table 2.7, all the PQR model specifications significantly outperforms RiskMetrics in all studied quantiles and Portfolio UQR in the upper quatiles similarly to models without VIX (see *Panel B*, Table 2.4). We can also see that PRQ-RSV-VIX specification outperforms UQR in the 95% quantile and PRQ-BPV-VIX outperform Portfolio UQR in the 10% quantile. In contrast to results without VIX, UQR significantly outperforms both PRQ-RV-VIX and PRQ-BPV-VIX specifications in the median quantile.

Table 2.7: Out-of-sample performance of various specifications of Panel Quantile Regression Model for Returns with VIX Index

		PQR-RV-VIX					PQR-RSV-VIX					PQR-BPV-VIX				
Panel A.1		5%	10%	50%	90%	95%	5%	10%	50%	90%	95%	5%	10%	50%	90%	95%
	$\widehat{\tau}$	0.060	0.107	0.462	0.896	0.956	0.060	0.107	0.464	0.895	0.956	0.059	0.106	0.462	0.898	0.957
	$\widehat{DQ}$	12.928	3.081	11.825	3.475	5.230	12.309	3.081	11.425	2.060	5.230	11.030	2.470	11.825	3.973	5.881
	p-val	0.044	0.799	0.066	0.747	0.515	0.055	0.799	0.076	0.914	0.515	0.087	0.872	0.066	0.680	0.437
		RiskMetrics					UQR					Portfolio UQR				
Panel A.2		5%	10%	50%	90%	95%	5%	10%	50%	90%	95%	5%	10%	50%	90%	95%
	$\widehat{\tau}$	0.061	0.094	0.451	0.919	0.958	0.043	0.099	0.491	0.909	0.955	0.061	0.094	0.451	0.919	0.958
	$\widehat{DQ}$	9.652	3.096	20.600	9.452	10.899	8.323	3.041	9.067	7.174	6.796	9.426	5.988	3.273	4.507	3.238
	p-val	0.140	0.797	0.002	0.150	0.092	0.215	0.804	0.170	0.305	0.340	0.151	0.425	0.774	0.608	0.778
		benchmark														
		RiskMetrics					UQR					Portfolio UQR				
Panel B		5%	10%	50%	90%	95%	5%	10%	50%	90%	95%	5%	10%	50%	90%	95%
PQR-RV-VIX	DM	**-2.320**	**-2.339**	**-2.452**	**-2.825**	**-2.370**	1.460	-0.463	*1.871*	**-1.981**	-1.426	-0.310	-1.624	0.051	**-4.574**	**-3.045**
	p-val	0.010	0.010	0.007	0.002	0.009	0.928	0.322	0.969	0.024	0.077	0.378	0.052	0.520	0.000	0.001
PQR-RSV-VIX	DM	**-2.278**	**-2.319**	**-2.254**	**-2.942**	**-2.630**	*1.838*	-0.356	1.335	**-2.194**	**-2.192**	-0.207	-1.568	-0.039	**-4.676**	**-3.503**
	p-val	0.011	0.010	0.012	0.002	0.004	0.967	0.361	0.909	0.014	0.014	0.418	0.058	0.484	0.000	0.000
PQR-BPV-VIX	DM	**-2.440**	**-2.464**	**-2.433**	**-2.755**	**-2.348**	0.056	-0.890	*1.888*	**-1.786**	-1.291	-0.703	**-1.919**	0.058	**-4.357**	**-2.860**
	p-val	0.007	0.007	0.007	0.003	0.009	0.522	0.187	0.971	0.037	0.098	0.241	0.028	0.523	0.000	0.002

Note: Table displays absolute and relative performance of PQR models for returns with RV, RSV and BPV as regressors combined with VIX and benchmark models.

Panel A.1 reports absolute performance of PQR models, *Panel A.2* reports absolute performance of benchmark models. For each model and quantile τ, unconditional coverage ($\widehat{\tau}$) and the value of the CAViaR test for correct dynamic specification ($\widehat{DQ}$) with corresponding Monte Carlo based p-value. Not correctly dynamically specified models are underlined.

Panel B reports relative performance of Panel Quantile Regression Models for Returns. For each specification and quantile τ we report Diebold-Mariano test statistics for pairwise comparison with benchmark models($\widehat{DM}$) with corresponding p-value. Significantly more/less accurate forecasts with respect to benchmark models at the 5% significance level are in bold/italic. Full matrix of pairwise comparison is available from authors upon request

2.7.1 Economic Evaluation

Further, we report the results from economic evaluation of the models. Table 2.8 and Figure 2.8 reveals qualitatively similar patterns to models without VIX (Table 2.5 and Figure 2.4). In the Global Minimum Value-at-Risk Portfolio application panel quantile regression PQR-RV-VIX specification achieve the lowest values of VaR in all studied quantiles. It also provides us with the best *Value-at-Risk –return* trade–off in the Markowitz like set-up since the efficient frontiers of PQR-RV-VIX model are to the right of the competitors.

In the comparison of models with and without exogenous factor we can see direct economic benefits of using the VIX index. In the GMVaRP comparison (Table 2.5 vs. Table 2.8) the reductions in the Value-at-Risks are 0.34, 0.173, 0.014, 0.422 and 0.609 percentage points or 2.89, 1.99, 68.54, 4.46, 4.92 percent for 5%, 10%, 50%, 90% and 95% quantiles respectively. In the visual comparison we can see that efficient frontiers of model with VIX shifts to the right thus having superior *VaR-return* trade-off. Generally we document direct economic benefits of the ex-ante volatility measure.

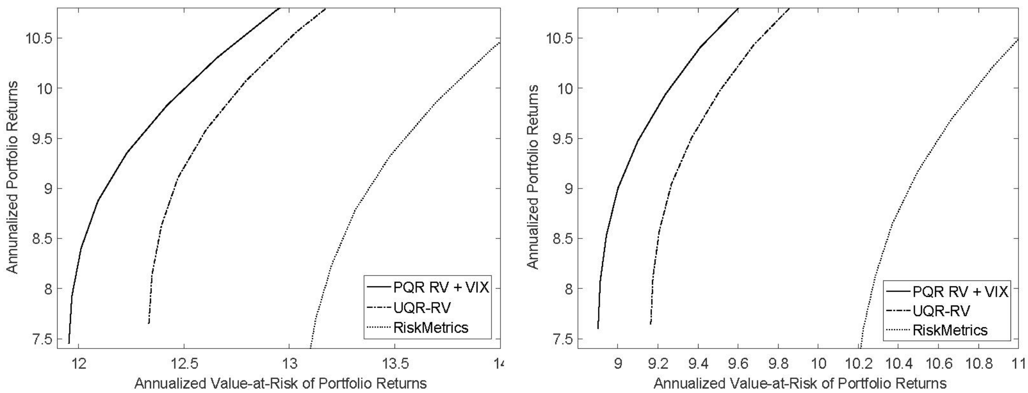
Table 2.8: Global Minimum Value-at-Risk Portfolio - PQR with VIX

τ	5%	10%	50%	90%	95%
PQR-RV-VIX	**11.42**	**8.51**	**0.007**	**9.04**	**11.76**
UQR-RV	11.85	8.79	0.011	9.52	12.43
RiskMetrics	12.77	9.95	NaN	9.95	12.77

Note: Table displays absolute percentage values of Global Minimum Value-at-Risk Portfolio for given quantile τ. Best model for given quantile is reported in bold.

Figure 2.8: Value-at-Risk – Return efficient frontiers - PQR with VIX

(a) $5\% VaR$ (b) $10\% VaR$



Note: Percentage values of portfolio VaR and returns are displayed.

2.7.2 Robustness Check: Portfolio of S&P 500 Constituents

To see how robust the findings are on the portfolio with large number of stocks, we apply our methodology on the constituents of the S&P 500 index. Since the firms included in S&P 500 vary substantially over time, we include in our analysis firms that

- were included in the index at least once during period July 1, 2005 to December 31, 2015
- have full history for the period July 1, 2005 to December 31, 2015
- were liquid enough, i.e. there were at least five active trading hours during a day.

Similar to previous analysis, we consider trades executed within U.S. business hours and we explicitly exclude weekends and bank holidays. In total, our dataset consists of 496 firms over 2613 trading days.

In our analysis we concentrate on two Panel Quantile Regression specification: *PQR-RV* and *PQR-RV-VIX*. In both specification we estimates models with $\lambda = 0$ since the results of minimization of the Bayesian Information Criteria were again inconclusive[19]. Results of our analysis are summarized in the Table 2.9, Figure 2.9 and Figure 2.10a and they reveal interesting finding about the role of the ex-ante volatility. In the PQR-RV specification the coefficient estimates $\hat{\beta}_{RV^{1/2}}$ are almost identical to one obtained using portfolio of 29 stocks (Table 2.3). In contrast, in the PQR-RV-VIX case, both $\hat{\beta}_{RV^{1/2}}$ and $\hat{\gamma}_{VIX}$ differ substantially for 496 and 29 stocks. Specifically, the relative influence of the realized volatility is lower as we move from 29 to 496 stocks in the bellow median quantiles, e.g. 5% quantile coefficient estimate of 29 stocks $\hat{\beta}_{RV^{1/2}} = -1.33$ vs. 496 stocks of $\hat{\beta}_{RV^{1/2}} = -1.17$, while it remains at the same level in the above median quantiles. In contrast, the $\hat{\gamma}_{VIX}$ coefficient of the large portfolio almost doubled in all bellow median quantiles and raised by fifty percent in the above median quantiles. Moreover, relative influence of the VIX index become almost symmetric, while in the 29-stocks portfolio the upper quantiles were influenced more. Overall, results of our analysis suggest that anticipation of the future market volatility translates directly to the conditional distribution of future returns of the firms listed in the New York Stock Exchange.

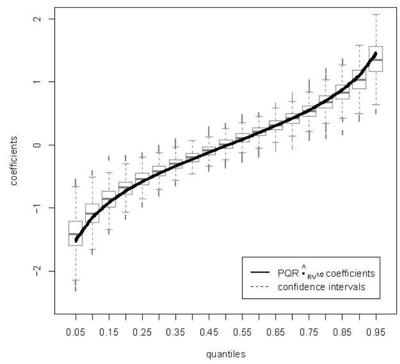
[19] The differences in the BICs of unpenalized model and the model with “optimaly” selected λ are higher than that we obtain in analysis of 29 stocks, however, they are still rather small i.e no greater than 6×10^{-3} or 0.6%.

Table 2.9: Coefficient estimates of Panel Quantile Regressions

τ	5%	10%	25%	50%	75%	90%	95%
				PQR-RV			
$\hat{\beta}_{RV^{1/2}}$	-1.51	-1.15	-0.58	-0.01	0.55	1.1	1.45
	(-96.03)	(-94.56)	(-78.79)	(-1.73)	(60.04)	(90.56)	(88.03)
				PQR-RV-VIX			
$\hat{\beta}_{RV^{1/2}}$	-1.17	-0.88	-0.46	-0.04	0.38	0.81	1.11
	(-42.81)	(-43.89)	(-38.05)	(-4.5)	(37.12)	(39.23)	(40.67)
$\hat{\gamma}_{VIX}$	-0.67	-0.54	-0.26	0.06	0.36	0.57	0.63
	(-21.62)	(-21.9)	(-20.25)	(7.23)	(27.85)	(26.06)	(25.04)

Note: Table displays coefficient estimates with bootstraped t-statistics in parentheses. Individual fixed effects $\alpha_i(\tau)$ are not reported for brevity - they are available from authors upon request.

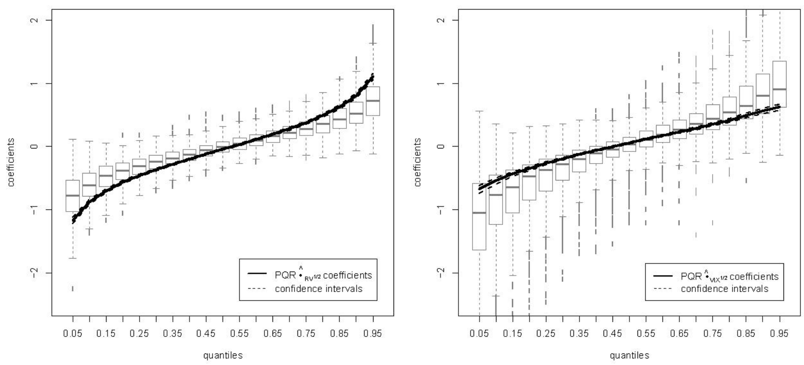
Figure 2.9: PQR-RV parameter estimates



Note: Parameter estimates with corresponding 95% confidence intervals from the PQR-RV specification are plotted by solid and dashed lines respectively. Individual UQR-RV estimates are plotted in boxplots.

Figure 2.10: PQR-RV-VIX parameter estimates

(a) $\hat{\beta}_{RV^{1/2}}$ (b) $\hat{\gamma}_{VIX}$



Note: For both realized volatility and VIX index parameters estimates with corresponding 95% confidence intervals are plotted by solid and dashed lines respectively. Individual UQR-RV-VIX estimates are plotted in boxplots.

2.8 Conclusion

In this chapter, we employ a series of panel quantile regressions together with non-parametric measures of quadratic return variation and ex-ante measure of market volatility as the common factor, to model conditional quantiles of financial assets return series. For estimation purposes we use penalized fixed effects estimator as introduced in Koenker (2004). Resulting Panel Quantile Regression Model for Returns inherit all favorable properties offered by panel data and quantile regression. A key attraction of the proposed methodology is the ability to control for otherwise unobserved heterogeneity among financial assets so it is possible to disentangle overall market risk into its systemic and idiosyncratic parts. Another attraction is the dimensionality reduction because the number of estimated parameters is always less than or equal to $k+n$, where k is the number of regressors and n number of assets. Last but not least, to the best of our knowledge, this is one of the first applications of the panel quantile regression using dataset where $T >> N$. As a result we are able to obtain estimates of quantile specific individual fixed effects that accounts for unobserved heterogeneity and represents the idiosyncratic part of market risk. Moreover, these estimates translate into better forecasting performance compared to traditional benchmarks. Overall, we test accuracy and performance of the Panel Quantile Regression Model for Returns in a simple portfolio Value-at-Risk forecasting exercise using simulated and empirical data. The Monte-Carlo experiment shows that our model is dynamically well specified. Moreover when we use heterogeneous data it is able to outperform benchmark models in direct statistical comparison. In our empirical application, the in-sample model fit highlights the importance of different components of quadratic variation and ex-ante volatility measure for various quantiles of return series. Out-of-sample statistical comparison shows superiority of the new approach. Better statistical performance moreover translates directly into economic gains as shown by Global

Minimum Value-at-Risk Portfolio set-up and efficient frontiers of the Value–at–Risk - Return trade-off.

Our results make the model attractive not only from an academic but also from the practitioners point of view. In particular it is highly attractive for portfolio and risk managers, because of its ability to handle high-dimensional problems. More importantly, it can be easily used to obtain Value-at-Risk measures of portfolios consisting of a high number of assets.

2.9 Appendix

Table 2.10: Descriptive statistics of daily returns: 1.7.2005 - 31.12.2015

	Mean	Max	Min	St. Dev.	Skewness	Kurtosis
AAPL	-0.05	10.62	-12.29	1.72	-0.14	7.09
AMZN	0.09	12.32	-12.96	1.95	0.33	8.27
BAC	-0.17	19.09	-25.09	2.77	-0.20	20.64
CMCSA	0.03	12.77	-13.63	1.57	-0.33	12.09
CSCO	-0.02	7.26	-8.69	1.35	-0.14	7.34
CVX	0.02	11.01	-10.50	1.31	-0.08	11.29
C	-0.27	19.92	-40.33	2.93	-2.48	38.66
DIS	0.06	11.03	-10.29	1.36	0.34	11.11
GE	-0.03	10.96	-10.52	1.51	-0.36	14.16
HD	0.03	11.03	-7.68	1.47	0.62	9.40
IBM	0.05	6.19	-5.93	1.06	-0.10	7.36
INTC	0.01	9.20	-9.43	1.42	0.13	7.41
JNJ	0.01	11.19	-7.77	0.85	0.75	21.90
JPM	0.01	13.85	-19.75	2.08	0.15	16.17
KO	0.02	7.14	-7.37	0.93	-0.08	11.52
MCD	0.03	8.76	-7.53	1.02	0.17	9.26
MRK	0.00	9.75	-8.09	1.29	-0.08	9.72
MSFT	0.02	9.96	-7.01	1.28	0.06	7.88
ORCL	0.04	7.56	-8.90	1.36	-0.09	6.85
PEP	0.04	8.44	-6.27	0.90	0.32	10.24
PFE	-0.03	6.49	-7.46	1.14	-0.07	7.02
PG	0.05	7.07	-5.62	0.86	0.00	9.50
QCOM	-0.01	9.04	-8.15	1.45	-0.10	6.31
SLB	0.00	11.34	-15.62	1.85	-0.33	9.57
T	-0.01	11.42	-6.56	1.11	0.50	13.58
VZ	0.01	8.62	-7.72	1.12	0.40	10.41
WFC	0.00	18.29	-18.73	2.23	0.45	18.50
WMT	0.00	7.71	-10.60	0.97	-0.08	14.66
XOM	0.03	8.90	-11.76	1.22	-0.11	13.33

Note: Values for Mean, Max, Min and St. Dev are displayed in %.

Table 2.11: Univariate Normal Distribution – Mean of coefficients estimates from Monte-Carlo simulations

τ	5%	10%	25%	50%	75%	90%	95%
				PQR-RV			
$\hat{\beta}_{RV^{1/2}}$	-1.56	-1.22	-0.64	-0.01	0.62	1.2	1.54
	(-44.19)	(-43.75)	(-30.85)	(-0.6)	(30.86)	(44.63)	(46.36)
				PQR-RSV			
$\hat{\beta}_{RS^{+1/2}}$	-1.12	-0.87	-0.46	-0.01	0.45	0.84	1.07
	(-5.38)	(-6.06)	(-5.34)	(-0.12)	(5.55)	(5.66)	(5.28)
$\hat{\beta}_{RS^{-1/2}}$	-1.09	-0.86	-0.46	-0.01	0.45	0.86	1.11
	(-5.29)	(-6.07)	(-5.44)	(-0.18)	(5.58)	(5.78)	(5.5)
				PQR-BPV			
$\hat{\beta}_{BPV^{1/2}}$	-1.58	-1.25	-0.67	-0.01	0.65	1.23	1.56
	(-45.75)	(-45.44)	(-31.71)	(-0.6)	(32.26)	(46.06)	(47.31)
$\hat{\beta}_{JV^{1/2}}$	0.08	0.06	0.03	0	-0.03	-0.06	-0.08
	(1.1)	(1.14)	(0.74)	(-0.03)	(-0.83)	(-1.1)	(-1.05)

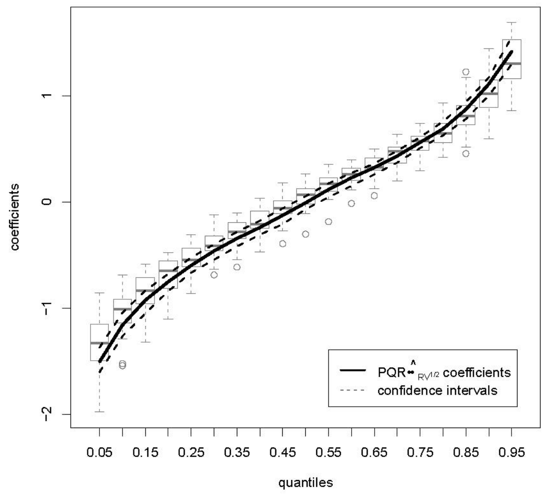
Note: Table displays mean of coefficient estimates with corresponding t-statistics in parentheses. Individual fixed effects $\alpha_i(\tau)$ are not reported for brevity.

Table 2.12: Multivariate Student-t Distribution – Mean of coefficients estimates from Monte-Carlo simulations

τ	5%	10%	25%	50%	75%	90%	95%
				PQR-RV			
$\hat{\beta}_{RV^{1/2}}$	-1.56	-1.16	-0.58	-0.01	0.57	1.15	1.55
	(-18.43)	(-19.14)	(-13.3)	(-0.17)	(12.9)	(18.73)	(19.43)
				PQR-RSV			
$\hat{\beta}_{RS^{+1/2}}$	-1.12	-0.83	-0.42	-0.01	0.39	0.82	1.09
	(-2.55)	(-2.66)	(-2.2)	(-0.12)	(2.05)	(2.55)	(2.43)
$\hat{\beta}_{RS^{-1/2}}$	-1.09	-0.82	-0.4	0	0.41	0.81	1.1
	(-2.48)	(-2.62)	(-2.13)	(0.04)	(2.15)	(2.52)	(2.42)
				PQR-BPV			
$\hat{\beta}_{BPV^{1/2}}$	-1.62	-1.21	-0.6	-0.01	0.59	1.19	1.6
	(-18.27)	(-19)	(-13.14)	(-0.17)	(12.64)	(18.41)	(19.05)
$\hat{\beta}_{JV^{1/2}}$	-0.06	-0.04	-0.02	0	0.02	0.05	0.06
	(-0.45)	(-0.45)	(-0.29)	(0.04)	(0.29)	(0.47)	(0.45)

Note: Table displays mean of coefficient estimates with corresponding t-statistics in parentheses. Individual fixed effects $\alpha_i(\tau)$ are not reported for brevity.

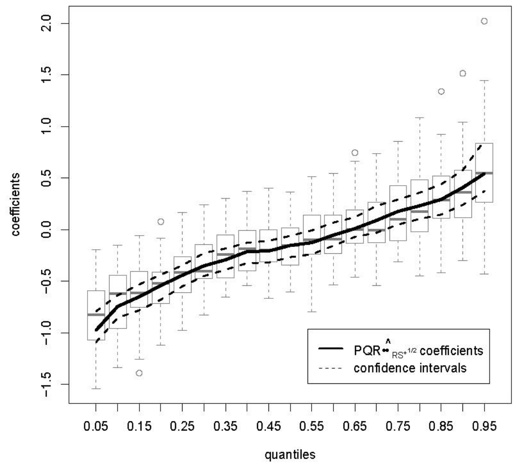
Figure 2.11: PQR-RV parameter estimates: $\lambda = 1$



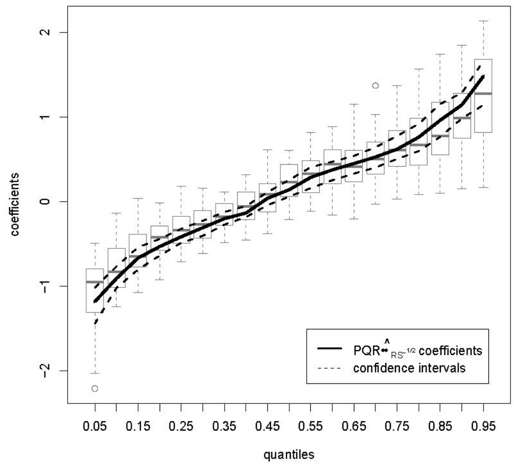
Note: Parameter estimates with corresponding 95% confidence intervals from the PQR-RV specification are plotted by solid and dashed lines respectively. Individual UQR-RV estimates are plotted in boxplots.

Figure 2.12: PQR-RSV parameter estimates: $\lambda = 1$

(a) $\hat{\beta}_{RS^{+1/2}}$



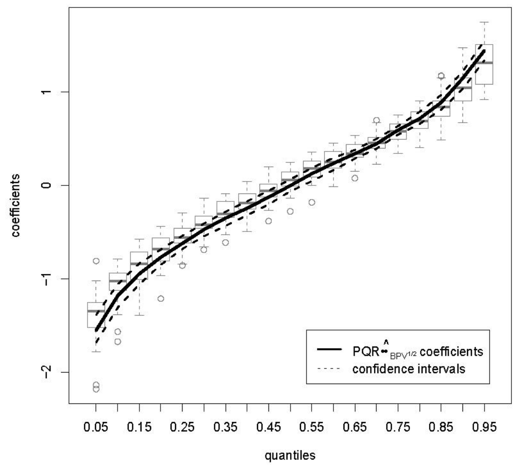
(b) $\hat{\beta}_{RS^{-1/2}}$



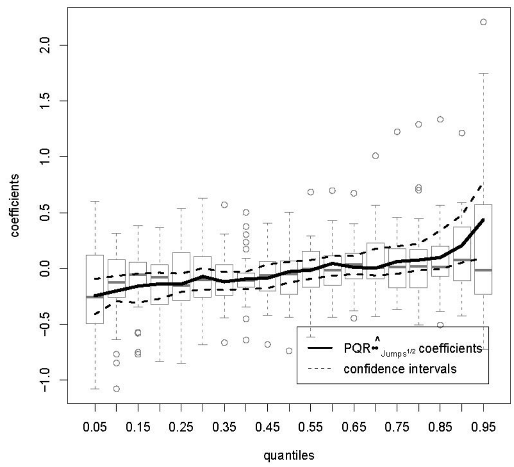
Note: For both realized upside and downside semivariance parameters estimates with corresponding 95% confidence intervals are plotted by solid and dashed lines respectively. Individual UQR-RSV estimates are plotted in boxplots.

Figure 2.13: PQR-BPV parameter estimates: $\lambda = 1$

(a) $\hat{\beta}_{BPV^{1/2}}$



(b) $\hat{\beta}_{JV^{1/2}}$



Note: For both realized bi-power variation and jump component parameters estimates with corresponding 95% confidence intervals are plotted by solid and dashed lines respectively. Individual UQR-BPV estimates are plotted in boxplots

Table 2.13: Univariate Student-t Distribution – Mean of coefficients estimates from Monte-Carlo simulations

τ	5%	10%	25%	50%	75%	90%	95%
				PQR-RV			
$\hat{\beta}_{RV^{1/2}}$	-1.56	-1.23	-0.65	-0.01	0.63	1.21	1.54
	(-52.18)	(-52.37)	(-36.12)	(-0.53)	(35.09)	(55.55)	(52.99)
				PQR-RSV			
$\hat{\beta}_{RS^{+1/2}}$	-1.12	-0.89	-0.47	-0.01	0.44	0.85	1.08
	(-6.47)	(-6.84)	(-5.75)	(-0.25)	(5.48)	(6.39)	(6.13)
$\hat{\beta}_{RS^{-1/2}}$	-1.09	-0.86	-0.45	0	0.46	0.87	1.12
	(-6.23)	(-6.57)	(-5.65)	(0.01)	(5.84)	(6.55)	(6.39)
				PQR-BPV			
$\hat{\beta}_{BPV^{1/2}}$	-1.64	-1.29	-0.69	-0.01	0.67	1.27	1.62
	(-51.41)	(-52.71)	(-36.1)	(-0.52)	(35.3)	(55.28)	(52.55)
$\hat{\beta}_{JV^{1/2}}$	-0.02	-0.01	0	0	0	0.01	0.02
	(-0.3)	(-0.21)	(-0.11)	(-0.02)	(0.01)	(0.16)	(0.24)

Note: Table displays mean of coefficient estimates with corresponding t-statistics in parentheses. Individual fixed effects $\alpha_i(\tau)$ are not reported for brevity.

Table 2.14: Models performance using data simulated from Multivariate Student-t Distribution

Panel A.1		PQR-RV					PQR-RSV					PQR-BPV				
		5%	10%	50%	90%	95%	5%	10%	50%	90%	95%	5%	10%	50%	90%	95%
	$\widehat{DQ}_{violations}$	6.4	5.8	11	5.6	7.8	6	6.2	11.4	5.2	7	6.8	5.8	10.6	5.2	7.6
	$\widehat{\tau}_{avg}$	5.1	10.2	51.3	90.0	95.0	5.1	10.2	51.4	90.0	95.0	5.1	10.2	51.4	90.0	95.0
	$\widehat{\tau}_{max}$	6.5	12.4	55.3	91.7	96.3	6.5	12.4	55.1	91.7	96.2	6.5	12.5	55.3	91.8	96.3
	$\widehat{\tau}_{min}$	3.8	7.9	48.1	87.8	93.7	3.7	7.9	48.2	88.0	93.8	3.8	7.9	48.2	88.1	93.7
	$\widehat{\tau}_{avg-dev}$	0.1	0.2	1.3	0.0	0.0	0.1	0.2	1.4	0.0	0.0	0.1	0.2	1.4	-0.1	0.0

Panel A.2		RiskMetrics					UQR					Portfolio UQR				
		5%	10%	50%	90%	95%	5%	10%	50%	90%	95%	5%	10%	50%	90%	95%
	$\widehat{DQ}_{violations}$	5.6	12.4	5.6	18.6	8.2	6.4	5	13.4	6.2	7.6	5.8	6.2	4.2	6.8	6.6
	$\widehat{\tau}_{avg}$	5.2	9.2	50.4	91.0	94.9	5.0	10.2	51.5	90.0	95.0	4.8	9.6	50.0	90.4	95.2
	$\widehat{\tau}_{max}$	6.8	11.7	54.0	92.8	96.3	6.4	12.3	55.2	91.7	96.5	6.1	11.2	53.1	92.2	96.3
	$\widehat{\tau}_{min}$	4.0	7.2	46.2	89.1	93.4	3.7	7.9	48.0	87.9	93.9	3.7	7.8	46.8	88.1	93.7
	$\widehat{\tau}_{avg-dev}$	0.2	-0.8	0.4	1.0	-0.1	0.0	0.2	1.5	0.0	0.0	-0.2	-0.4	0.0	0.4	0.2

		benchmark														
Panel B.1		RiskMetrics					UQR					Portfolio UQR				
		5%	10%	50%	90%	95%	5%	10%	50%	90%	95%	5%	10%	50%	90%	95%
PQR-RV	*DM*	54	58.2	0.2	62.6	57.8	2.4	2.4	12	3.6	3.6	43	40	18.4	46.4	49.2
PQR-RSV	*DM*	54.2	57.4	0.2	62.2	57.4	1.4	1.2	9.6	1.4	3	41	39	19.2	45	48.2
PQR-BPV	*DM*	53	57.6	0.2	62.4	57.4	1.6	2	10	3	5.2	42.4	38.6	18.8	44.8	45.4

Panel B.2		PQR-RV					PQR-RSV					PQR-BPV				
		5%	10%	50%	90%	95%	5%	10%	50%	90%	95%	5%	10%	50%	90%	95%
RiskMetrics	*DM*	0	0	14.4	0	0	0	0	14.8	0	0	0	0	14.6	0	0
UQR	*DM*	9	10.6	2.8	9.8	10.6	10.4	11	2.8	11.2	9.6	10.4	11.6	2.8	8.2	7.8
Portfolio UQR	*DM*	0.8	0.4	1	0.6	0.4	1	0.4	1.4	0.6	0.4	0.8	0.6	1.2	0.8	0.2

Note: Table displays absolute and relative performance of PQR models for returns with RV, RSV and BPV as regressors and benchmark models.

Panel A.1 reports absolute performance of PQR models, *Panel A.2* reports absolute performance of benchmark models. For each model and quantile τ, percentage of violations of the CAViaR test for correct dynamic specification ($\widehat{DQ}_{violations}$), average unconditional coverage ($\widehat{\tau}_{avg}$),maximum unconditional coverage ($\widehat{\tau}_{max}$), minimum unconditional coverage ($\widehat{\tau}_{min}$) and average deviation of unconditional coverage from given quantile τ ($\widehat{\tau}_{avg-dev}$)

Panel B.1 and *Panel B.2* report relative performance of Panel Quantile Regression Models for Returns in comparison to benchmark models and relative performance of benchmark models in comparison to Panel Quantile Regression Models for Returns respectively. For each specification and quantile τ we report percentage of statistically better performance according to Diebold-Mariano(DM) test at 5% significance level.

Table 2.15: Models performance using data simulated from Univariate Normal Distributions

		PQR-RV					PQR-RSV					PQR-BPV				
Panel A.1		5%	10%	50%	90%	95%	5%	10%	50%	90%	95%	5%	10%	50%	90%	95%
	$\widehat{DQ}_{violations}$	8.2	7.4	26	4.4	8.6	9	8.4	26.6	4.6	7.4	9.4	9.6	26.4	4.2	6.8
	$\widehat{\tau}_{avg}$	5.3	10.6	52.5	90.3	95.2	5.4	10.6	52.5	90.3	95.2	5.4	10.7	52.5	90.3	95.2
	$\widehat{\tau}_{max}$	7.1	12.7	55.6	92.9	96.9	7.1	12.6	55.7	92.8	96.8	7.2	12.8	55.7	92.7	96.8
	$\widehat{\tau}_{min}$	3.8	8.8	49.3	88.1	93.5	3.9	8.8	49.3	88.1	93.4	3.8	8.6	49.3	88.1	93.5
	$\widehat{\tau}_{avg-dev}$	0.3	0.6	2.5	0.3	0.2	0.4	0.6	2.5	0.3	0.2	0.4	0.7	2.5	0.3	0.2
		RiskMetrics					UQR					Portfolio UQR				
Panel A.2		5%	10%	50%	90%	95%	5%	10%	50%	90%	95%	5%	10%	50%	90%	95%
	$\widehat{DQ}_{violations}$	18.8	13.4	10	5.4	9.8	7.4	8.2	47.6	4.2	7.4	5.4	4.4	3	3	4.4
	$\widehat{\tau}_{avg}$	5.8	11.0	51.2	90.2	94.9	5.3	10.6	53.3	90.3	95.2	5.1	10.1	50.0	89.9	94.9
	$\widehat{\tau}_{max}$	7.3	12.7	54.1	92.1	96.2	7.1	12.6	56.4	92.9	96.9	6.3	11.6	52.3	91.7	96.1
	$\widehat{\tau}_{min}$	4.3	9.1	47.8	88.0	93.4	3.8	8.7	50.2	88.3	93.4	4.1	8.4	47.7	88.2	93.7
	$\widehat{\tau}_{avg-dev}$	0.8	1.0	1.2	0.2	-0.1	0.3	0.6	3.3	0.3	0.2	0.1	0.1	0.0	-0.1	-0.1
							benchmark									
		RiskMetrics					UQR					Portfolio UQR				
Panel B.1		5%	10%	50%	90%	95%	5%	10%	50%	90%	95%	5%	10%	50%	90%	95%
PQR-RV	*DM*	62.6	56.6	0	57	60.6	9	6	72.4	6.6	7.4	16.4	17.4	4.6	17.4	16.4
PQR-RSV	*DM*	63.2	57.6	0	56.8	61.8	12	7	71.8	9.2	12.4	18.4	18.8	4.6	17.2	17.4
PQR-BPV	*DM*	65.4	58.8	0	58.6	62.2	12.8	11.2	71.4	12.4	14.4	19.2	18.6	4.6	19.4	16.4
		PQR-RV					PQR-RSV					PQR-BPV				
Panel B.2		5%	10%	50%	90%	95%	5%	10%	50%	90%	95%	5%	10%	50%	90%	95%
RiskMetrics	*DM*	0	0	44.2	0	0	0	0	44.2	0	0	0	0	44.4	0.2	0
UQR	*DM*	2.8	2.8	0	2.8	2.8	2.2	2	0	2.2	1.8	2.4	1.8	0	3	1.2
Portfolio UQR	*DM*	0.2	0.4	2.4	0.6	1	0.2	0.4	2.4	0.6	0.8	0.2	0.2	2.4	1.2	0.4

Note: Table displays absolute and relative performance of PQR models for returns with RV, RSV and BPV as regressors and benchmark models.

Panel A.1 reports absolute performance of PQR models, *Panel A.2* reports absolute performance of benchmark models. For each model and quantile τ, percentage of violations of the CAViaR test for correct dynamic specification ($\widehat{DQ}_{violations}$), average unconditional coverage ($\widehat{\tau}_{avg}$),maximum unconditional coverage ($\widehat{\tau}_{max}$), minimum unconditional coverage ($\widehat{\tau}_{min}$) and average deviation of unconditional coverage from given quantile τ ($\widehat{\tau}_{avg-dev}$)

Panel B.1 and *Panel B.2* report relative performance of Panel Quantile Regression Models for Returns in comparison to benchmark models and relative performance of benchmark models in comparison to Panel Quantile Regression Models for Returns respectively. For each specification and quantile τ we report percentage of statistically better performance according to Diebold-Mariano(DM) test at 5% significance level.

Table 2.16: Models performance using data simulated from Univariate Student-t Distributions

	PQR-RV					PQR-RSV					PQR-BPV				
Panel A.1	5%	10%	50%	90%	95%	5%	10%	50%	90%	95%	5%	10%	50%	90%	95%
$\widehat{DQ}_{violations}$	7.2	8.0	23.4	4.0	6.4	7.0	8.2	23.2	4.0	6.6	7.6	8.2	23.6	4.2	6.8
$\widehat{\tau}_{avg}$	5.3	10.5	52.3	90.2	95.2	5.3	10.5	52.3	90.2	95.2	**5.3**	10.5	52.4	90.2	95.1
$\widehat{\tau}_{max}$	7.0	12.8	56.0	92.0	96.6	7.0	12.8	56.0	91.9	96.6	7.0	12.7	56.1	91.9	96.6
$\widehat{\tau}_{min}$	3.9	8.3	47.7	87.8	93.6	3.8	8.2	47.7	87.8	93.6	4.0	8.1	47.8	87.8	93.4
$\widehat{\tau}_{avg-dev}$	0.3	0.5	2.3	0.2	0.2	0.3	0.5	2.3	0.2	0.2	0.3	0.5	2.4	0.2	0.1

	RiskMetrics					UQR					Portfolio UQR				
Panel A.2	5%	10%	50%	90%	95%	5%	10%	50%	90%	95%	5%	10%	50%	90%	95%
$\widehat{DQ}_{violations}$	14.8	10.0	7.6	5.6	6.4	6.8	7.4	46.4	5.0	7.0	3.4	4.2	2.6	3.8	3.2
$\widehat{\tau}_{avg}$	5.8	10.9	51.2	90.1	94.8	5.3	10.5	53.2	90.2	95.2	5.1	10.1	50.0	89.9	94.9
$\widehat{\tau}_{max}$	7.3	12.8	55.0	91.7	96.3	7.0	12.6	56.9	91.9	96.5	6.8	12.5	52.2	91.2	96.0
$\widehat{\tau}_{min}$	4.3	8.9	46.8	88.1	93.4	3.9	8.3	48.5	87.9	93.5	4.0	8.2	47.4	88.0	93.7
$\widehat{\tau}_{avg-dev}$	0.8	0.9	1.2	0.1	-0.2	0.3	0.5	3.2	0.2	0.2	0.1	0.1	0.0	-0.1	-0.1

		benchmark														
		RiskMetrics					UQR					Portfolio UQR				
Panel B.1		5%	10%	50%	90%	95%	5%	10%	50%	90%	95%	5%	10%	50%	90%	95%
PQR-RV	*DM*	66.4	57.8	0.4	57.4	59.4	6.4	6.4	67.4	9.2	7.8	18.6	17.0	7.6	23.2	20.4
PQR-RSV	*DM*	65.8	58.2	0.4	58.0	59.6	6.8	7.8	68.6	10.2	8.2	17.2	17.4	7.4	23.2	21.6
PQR-BPV	*DM*	67.0	59.6	0.4	59.4	61.2	9.0	6.8	66.8	10.6	10.8	20.6	16.6	7.6	22.4	21.6

		PQR-RV					PQR-RSV					PQR-BPV				
Panel B.2		5%	10%	50%	90%	95%	5%	10%	50%	90%	95%	5%	10%	50%	90%	95%
RiskMetrics	*DM*	0.0	0.0	39.6	0.0	0.0	0.0	0.0	39.8	0.0	0.0	0.0	0.0	39.6	0.0	0.0
UQR	*DM*	4.0	4.0	0.0	4.0	4.4	4.6	3.2	0.2	3.8	3.6	3.4	2.4	0.0	2.6	3.0
Portfolio UQR	*DM*	0.2	0.0	2.6	0.0	0.4	0.2	0.0	2.8	0.0	0.2	0.2	0.0	2.8	0.0	0.4

Note: Table displays absolute and relative performance of PQR models for returns with RV, RSV and BPV as regressors and benchmark models.

Panel A.1 reports absolute performance of PQR models, *Panel A.2* reports absolute performance of benchmark models. For each model and quantile τ, percentage of violations of the CAViaR test for correct dynamic specification ($\widehat{DQ}_{violations}$), average unconditional coverage ($\widehat{\tau}_{avg}$),maximum unconditional coverage ($\widehat{\tau}_{max}$), minimum unconditional coverage ($\widehat{\tau}_{min}$) and average deviation of unconditional coverage from given quantile τ ($\widehat{\tau}_{avg-dev}$)

Panel B.1 and *Panel B.2* report relative performance of Panel Quantile Regression Models for Returns in comparison to benchmark models and relative performance of benchmark models in comparison to Panel Quantile Regression Models for Returns respectively. For each specification and quantile τ we report percentage of statistically better performance according to Diebold-Mariano(DM) test at 5% significance level.

Table 2.17: *Coefficient estimates of Panel Quantile Regressions:* $\lambda = 1$

τ	5%	10%	25%	50%	75%	90%	95%
				PQR-RV			
const	0	0	0	0	0	0	0
	(-1.47)	(-1.27)	(-0.61)	(-0.58)	(0.9)	(1.64)	(2.08)
$\hat{\beta}_{RV^{1/2}}$	-1.51	-1.16	-0.6	-0.01	0.56	1.11	1.42
	(-24.24)	(-21.41)	(-16.36)	(-0.24)	(20.15)	(24.62)	(21.11)
				PQR-RSV			
const	0	0	0	0	0	0	0
	(-1.6)	(-1.21)	(-0.63)	(-0.44)	(0.97)	(2.07)	(2.41)
$\hat{\beta}_{RS^{+1/2}}$	-0.97	-0.74	-0.44	-0.15	0.18	0.41	0.54
	(-12.92)	(-13.02)	(-8.54)	(-2.9)	(2.82)	(4.39)	(4.3)
$\hat{\beta}_{RS^{-1/2}}$	-1.18	-0.91	-0.41	0.14	0.62	1.15	1.49
	(-11.14)	(-14.29)	(-10.12)	(2.78)	(9.23)	(13.91)	(10.06)
				PQR-BPV			
const	0	0	0	0	0	0	0
	(-1.4)	(-1.36)	(-0.6)	(-0.67)	(0.77)	(1.86)	(2.67)
$\hat{\beta}_{BPV^{1/2}}$	-1.55	-1.18	-0.62	0	0.59	1.15	1.44
	(-20.25)	(-17.58)	(-16.49)	(-0.15)	(24.16)	(22.79)	(25.91)
$\hat{\beta}_{JV^{1/2}}$	-0.24	-0.2	-0.14	-0.03	0.06	0.21	0.44
	(-3.2)	(-3.41)	(-3.41)	(-0.62)	(1.03)	(1.88)	(2.73)

Note: Table displays coefficient estimates with bootstraped t-statistics in parentheses. Individual fixed effects $\alpha_i(\tau)$ are not reported for brevity - they are available from authors upon request.

Chapter 3

Dynamic Quantile Model for Bond Pricing

This chapter introduces a dynamic quantile model for bond pricing with an agent who values securities by maximizing the quantile level of her utility function. The transition from traditional to quantile preferences allows us to study the pricing of the term structure of interest rates by economic agents differing in their levels of risk aversion. Moreover, the framework is robust to fat tails commonly observed in the empirical data. In the application, we focus on the quantile pricing of the two, five, ten and thirty years US and German government bonds. For the analysis, we use flexible quantile regression framework which is applied over highly liquid bond futures contract from the Chicago Board of Trade and EUREX exchanges.

3.1 Introduction

> *Asset pricing theory all stems from one simple concept... : price equals expected discounted payoff. The rest is elaboration, special cases, and a closet full of tricks that make the central equation useful for one or another application*
>
> —Cochrane (2009)

Asset pricing theory has a long history in economic and finance literature. Whether it is work of Markowitz (1952), who set the foundation of the modern portfolio theory, Sharpe (1964) and Lintner (1965) and their Capital Asset Pricing Model further developed in Merton (1973), the Arbitrage Pricing Theory of Ross (1976) or the factor pricing represented by Fama and French (1993) they all work within expected utility framework. In our work, we move from traditional expected utility asset pricing to a more general quantile preferences. The advantages of quantile preferences are for example: robustness to fat tails and the ability to capture heterogeneity through the quantiles. This framework also allows us to study asset pricing given economic agents who differ in their levels of risk aversion.

This chapter introduces the dynamic quantile factor assets pricing model. We build on the work of de Castro and Galvao (2018) who derive the quantile Euler equation, and we modify it to a *basic quantile pricing equation*. In our empirical application, we focus on the quantile pricing of the two, five, ten and thirty years US and German government bonds. For the analysis, we use a flexible quantile regression framework which is applied over highly liquid bond futures contract from the Chicago Board of Trade and the EUREX exchanges. In particular, we employ the recently developed smoothed GMM quantile estimator of de Castro et al. (2018) and the quantile regression of Koenker and Bassett Jr (1978).

Results of our analysis demonstrate a significant influence of different Realized Measures on the quantiles of the treasury returns. In contrast, forward rates as used in Cochrane and Piazzesi (2005) to forecast short term returns are of limited use in quantile asset pricing.

3.2 Asset pricing

The basic idea of classical asset pricing is that price equals discounted payoff (Cochrane, 2009). Formally we can summarize this relationship by the equation

$$p_t = E(m_{t+1} * x_{t+1}), \tag{3.1}$$

where p_t is the asset price, x_{t+1} is a payoff of the asset and m_{t+1} is a stochastic discount factor. Often the (3.1) is referred to as the *basic pricing equation* and it applies to all kind of assets (stocks, bonds, options, etc.) Moreover, it can be seen as a generalization of the popular consumption-based asset pricing or the factor pricing (e.g., Capital Asset Pricing Model).

The logic behind (3.1) can be traced back to the concept of utility, applied to the agent who is deciding how much to consume now, and how much to save for the next period. Formally we can describe this problem by the agents utility function as

$$U(c_t, c_{t+1}) = u(c_t) + \beta E_t[u(c_{t+1})], \tag{3.2}$$

where c_t is consumption at a time t, $u(\cdot)$ is period utility function which is increasing and concave in consumption, $E_t[\cdot]$ are the expectations made at time t, and β represents discount factor. Since our agent wants to maximize her utility from the consumption, we solve (3.2) subject to her budget constraints

$$\max_{\xi} \quad u(c_t) + E_t \beta u(c_{t+1}) \tag{3.3}$$

$$\text{s.t.} \quad c_t = y_t - p_t \xi$$

$$c_{t+1} = y_{t+1} + x_{t+1} \xi,$$

where y_t is the initial income/wealth, p_t is the price of the asset, ξ is the amount of asset agent is willing to buy and x_{t+1} is the payoff of the asset at a time $t+1$. Often it is assumed that $x_{t+1} = p_{t+1} + d_{t+1}$, i.e., payoff equals price of the asset plus dividend. First order conditions of the agent are obtained by substituting budget constraints into the utility function and differentiating it w.r.t. ξ:

$$0 = -p_t u'(y_t - p_t\xi) + E_t[\beta u'(y_{t+1} + x_{t+1}\xi)x_{t+1}], \tag{3.4}$$

which results in

$$p_t u'(c_t) = E_t[\beta u'(c_{t+1})x_{t+1}], \tag{3.5}$$

where $u'(\cdot)$ is interpreted as marginal utility. Equation 3.5 tells us that a decrease in the utility from the purchase of ξ units of the asset at time t must be compensated by the increase in expected utility from the payoff at time $t+1$. After rearranging we obtain the basic(fundamental) pricing equation

$$p_t = E_t\left[\beta \frac{u'(c_{t+1})}{u'(c_t)} x_{t+1}\right]. \tag{3.6}$$

Substituting $m_{t+1} = \beta\frac{u'(c_{t+1})}{u'(c_t)}$ give us exactly (3.1) and the term m_{t+1} is called *stochastic discount factor*, *marginal rate of substitution* or *pricing kernel*. In practice, it is more convenient to work with returns instead of prices. Defining asset return as

$$1 + r_{t+1} = \frac{p_{t+1} + d_{t+1}}{p_t} \tag{3.7}$$

and substituting it into (3.6) we arrive at convenient form of the Euler equation

$$1 = E_t\left[\beta \frac{u'(c_{t+1})}{u'(c_t)} (1 + r_{t+1})\right]. \tag{3.8}$$

Moreover, substituting $m_{t+1} = \beta\frac{u'(c_{t+1})}{u'(c_t)}$, using a definition of covariance $E(mx) = E(m)E(x) + cov(mx)$ and after rearranging (3.8) becomes

$$E[1 + r_{t+1}] = \frac{1}{E_t[m_{t+1}]} - \frac{cov(m_{t+1}(1 + r_{t+1}))}{E_t[m_{t+1}]}. \tag{3.9}$$

Further more defining $\alpha = \frac{1}{E_t[m_{t+1}]}$, $\beta = \frac{cov(m_{t+1}(1+r_{t+1}))}{var[m_{t+1}]}$ and $\lambda = -\frac{var(m_{t+1})}{E_t[m_{t+1}]}$ we arrive to beta representation for asset i

$$E[1 + r^i_{t+1}] = \alpha + \beta_{i,m_{t+1}}\lambda_{m_{t+1}}. \tag{3.10}$$

If we now assume that $m_{t+1} = \beta\frac{u'(c_{t+1})}{u'(c_t)}$ can be approximated by linear function $a + b_1 f^1_{t+1} + b_2 f^2_{t+1}$... we get factor pricing model

$$E[1 + r_{t+1}] = \alpha + \beta\lambda. \tag{3.11}$$

3.2.1 From Classical to Quantile Asset Pricing

The previous section describes classical asset pricing in an expected utility framework. To be more precise von-Neumann-Moregenster's expected utility is applied.

Von-Neumann-Moregenster's theorem states that under completeness, transitivity, continuity, and independence there exists preferences $\succeq$, if and only if there exists utility function $u'(\cdot)$ such that

$$X \succeq Y \Leftrightarrow E[u'(X)] \geq E[u'(Y)]. \tag{3.12}$$

Using quantile preferences, a similar result was obtain by Manski (1988) and later by Rostek (2010)

$$X \succeq Y \Leftrightarrow Q_\tau[u'(X)] \geq Q_\tau[u'(Y)], \tag{3.13}$$

where $Q_\tau[\cdot]$ is the τ-quantile function. The importance of quantile preferences also lies in linking risk-aversion to quantiles. Manski (1988) showed that agents maximizing different quantiles have different risk aversion, i.e., agent maximizing lower quantile is more risk-averse than agent maximizing higher quantile.

Analogously to transition from von-Neuman Moregenster's preferences to quantile preferences, de Castro and Galvao (2018) show that we can move from the Euler equation in expectations to quantile Euler equation. In particular we can rewrite (3.4) as

$$0 = -p_t u'(y_t - p_t\xi) + \beta_\tau Q_\tau[u'(y_{t+1} + x_{t+1}\xi)x_{t+1}|\Omega_t], \tag{3.14}$$

which after substituting asset returns from (3.7) results in

$$1 = Q_\tau\left[\beta\frac{u'(c_{t+1})}{u'(c_t)}(1 + r_{t+1})|\Omega_t\right], \tag{3.15}$$

where Ω_t denotes information set.

The previous equation can be estimated using Smoothed (G)MM Quantile estimator of de Castro et al. (2018). In their paper de Castro et al. (2018) estimate the elasticity of intertemporal substitution of the consumption-based model of the agent with isoelastic utility,

$$u(c_t) = \frac{1}{1-\gamma}c_t^{1-\gamma}. \tag{3.16}$$

with corresponding marginal utilities

$$\frac{u'(c_{t+1})}{u'(c_t)} = \left(\frac{c_{t+1}}{c_t}\right)^{-\gamma}. \tag{3.17}$$

Combining (3.15) and (3.17) they arrive to Euler equation in form

$$1 = Q_\tau\left[\beta\left(\frac{c_{t+1}}{c_t}\right)^{-\gamma}(1 + r_{t+1})|\Omega_t\right]. \tag{3.18}$$

Using $Q_\tau(\mathcal{F}) = Q_{1-\tau}(-\mathcal{F})$ and $ln(Q_\tau(\mathcal{F})) = Q_\tau(ln(\mathcal{F}))$(i.e., log-lineariztion), where $\mathcal{F}$ is a general random variable, the equation they estimate is

$$0 = Q_\tau \left[ln\left(\frac{c_{t+1}}{c_t}\right) - \gamma^{-1}ln(\beta) - \gamma^{-1}ln(1+r_{t+1})|\Omega_t \right]. \tag{3.19}$$

Here we show that from (3.14) we can arrive at the quantile version of the fundamental pricing equation. Rearranging (3.14) we get

$$p_t u'(c_t) = Q_\tau[\beta_\tau u'(c_{t+1})x_{t+1}|\Omega_t]. \tag{3.20}$$

Using $m_{t+1} = \beta\frac{u'(c_{t+1})}{u'(c_t)}$ finally we have

$$p_t = Q_\tau[m_{t+1}x_{t+1}|\Omega_t], \tag{3.21}$$

or in terms of returns

$$1 = Q_\tau[m_{t+1}(1+r_{t+1})|\Omega_t]. \tag{3.22}$$

By log-linearizing we get

$$0 = Q_\tau\left[ln(m_{t+1}) + ln(1+r_{t+1})|\Omega_t\right], \tag{3.23}$$

which is a generalized version of the (3.18). Indeed if we substitute $m_{t+1} = \beta\left(\frac{c_{t+1}}{c_t}\right)^{-\gamma}$ into (3.23) and use $Q_\tau(\mathcal{F}) = Q_{1-\tau}(-\mathcal{F})$ we get (3.19).

Importance of the log-linearization of the quantile fundamental pricing equation is that now we can create a family of quantile factor pricing models. Analogously to factor models in expectation, we can define

$$ln(m_{t+1}) = a + \mathbf{bf}, \quad or \quad m_{t+1} = e^{a+\mathbf{bf}}, \tag{3.24}$$

where $\mathbf{f}$ is vector of factors f with corresponding vector of parameters $\mathbf{b}$. The final form of quantile factor pricing equation is

$$0 = Q_\tau\left[ln(1+r_{t+1}) + a + \mathbf{bf}|\Omega_t\right]. \tag{3.25}$$

We can also log-linearize the quantile pricing equation by dropping conditioning of the whole formula on information set, Ω_t, and instead condition only returns. As a result, we rewrite (3.23) in the form of a standard quantile regression of Koenker and Bassett Jr (1978)

$$Q_\tau\left[ln(1+r_{t+1})|\Omega_t\right] = ln(m_{t+1}), \tag{3.26}$$

which after substitution for $ln(m_{t+1})$ and approximating $ln(1+r_{t+1})$ by r_{t+1} we get

$$Q_\tau\left[r_{t+1}|\Omega_t\right] = a + \mathbf{bf}. \tag{3.27}$$

The equation in a similar form was studied by Žikeš and Baruník (2016) who showed that quantiles of future returns depend on their past quadratic variation and weakly exogenous variables. Formally

$$Q_\tau(r_{t+1}|\Omega_t) = \alpha(\tau) + \beta_v(\tau)'v_t + \beta_z(\tau)'z_t, \tag{3.28}$$

where $v_t = \left(\widehat{QV}_t^{1/2}, \widehat{QV}_{t-1}^{1/2}, \ldots, \widehat{IV}_t^{1/2}, \widehat{IV}_{t-1}^{1/2}, \ldots, \widehat{JV}_t^{1/2}, \widehat{JV}_{t-1}^{1/2}, \ldots\right)$ are components of quadratic variation and z_t is vector of weakly exogenous variables. Estimates of quantile specific β from (3.28) are obtained by minimizing the following objective function:

$$\min_{\alpha(\tau),\beta(\tau)} \frac{1}{n} \sum_{t=1}^{n} \rho_\tau \left(r_{t+1} - \alpha(\tau) - \beta_v(\tau)' v_t - \beta_z(\tau)' z_t\right), \tag{3.29}$$

where $\rho_\tau(u) = u(\tau - I(u < 0)))$ is the quantile loss function defined in Koenker and Bassett Jr (1978).

An application of (3.23) and (3.27) can be also illustrated within a simple CAPM example. If we set $ln(m_{t+1}) = a + \beta ln(1 + r_{t+1}^M)$ where r_{t+1}^M is market return and using the approximation $ln(1 + x) = x$ we get

$$\begin{aligned} 0 &= Q_\tau \left[r_{t+1} + a + \beta r_{t+1}^M | \Omega_t\right] \\ Q_\tau \left[r_{t+1} | \Omega_t\right] &= a + \beta r_{t+1}^M \end{aligned} \tag{3.30}$$

which are quantile and approximated quantile version of CAPM respectively.

3.3 Smoothed (Generalized) Method of Moments Quantile Estimator

We estimate the quantile pricing equation using the Smoothed (Generalized)Method of Moments Quantile Estimator introduced in de Castro et al. (2018) [1]. Smoothed MM and GMM estimator differ in identification restriction - the former can be applied to exactly identified models while the later is used in case of over-identification.

It was shown in de Castro et al. (2018) that the nonlinear quantile model

$$Q_\tau \left[\Lambda \left(Y_i, X_i, \beta_{0\tau}\right) | Z_i\right] = 0, \tag{3.31}$$

where Y_i is an endogenous variable vector, Z_i is the full instrument vector, X_i is subset of Z_i, and $\Lambda(\cdot)$ is the residual function. The above can be rewritten in terms of conditional quantile moments as

$$0 = E\left[\mathbf{1}\left\{\Lambda\left(Y_i, X_i, \beta_{0\tau}\right) \leq 0\right\} - \tau | Z_i\right], \tag{3.32}$$

where $\mathbf{1}(\cdot)$ is indicator function and estimates of $\beta_{0\tau}$ can be obtain by unconditional moments

$$0 = E\left\{Z_i\left[\mathbf{1}\left\{\Lambda\left(Y_i, X_i, \beta_{0\tau}\right) \leq 0\right\} - \tau\right]\right\}. \tag{3.33}$$

de Castro et al. (2018) also argue that the discontinuity of the indicator function $\mathbf{1}(\cdot)$ makes the GMM minimization problem computationally difficult. To overcome this issue they suggest to smooth the indicator function so the smoothed population moment

[1] For estimation, we use R code of M. Kaplan available at his web-page: `https://faculty.missouri.edu/~kaplandm/code/dCGKL_2018_code.zip`

conditions are

$$\begin{aligned} g_{ni}(\beta,\tau) &\equiv g_n(Y_i, X_i, Z_i, \beta, \tau) \equiv Z_i \left[\tilde{I}\left(-\Lambda\left(Y_i, X_i, \beta\right)/h_n\right) - \tau \right] \\ \tilde{M}_n(\beta,\tau) &\equiv \tfrac{1}{n}\sum_{i=1}^{n} g_{ni}(\beta,\tau), \end{aligned} \tag{3.34}$$

where h_n is a bandwidth and $\tilde{I}(\cdot)$ is smoothed version of indicator function $\mathbf{1}(\cdot \geq 0)$.

In case of exact identification, the method of moments estimator solves the smoothed sample moment conditions

$$\hat{M}_n\left(\hat{\beta}_{MM}, \tau\right) = 0 \tag{3.35}$$

If the model is over-identified, however, there is no guarantee that the solution of method of moments estimator is global, and not the local minimum. Thus for the over-identified model, the generalized method of moments estimator is used

$$\hat{\beta}_{GMM} = \underset{\beta\in B}{arg\,min} \left[\sum_{i=1}^{n} g_{ni}(\beta,\tau)\right]' \hat{W} \left[\sum_{i=1}^{n} g_{ni}(\beta,\tau)\right] = \underset{\beta\in B}{arg\,min}\, \hat{M}_n(\beta,\tau)' \hat{W} \hat{M}_n(\beta,\tau), \tag{3.36}$$

where $\hat{W}$ is a symmetric, positive-definite weighting matrix. de Castro et al. (2018) proposed to use an estimator of the inverse long-run variance of the sample moments $\hat{W}^* = \bar{\Omega}^{-1} \xrightarrow{p} \Omega^{-1}$ with $\bar{\Omega}$ depending on initial estimates from the method of moments estimator. The final estimator

$$\hat{\beta}_{2-step\,GMM} = \underset{\beta\in B}{arg\,min}\, \hat{M}_n(\beta,\tau)' \bar{\Omega}^{-1} \hat{M}_n(\beta,\tau) \tag{3.37}$$

uses a simulated annealing algorithm for finding global minimum.

In our work, we concentrate on exactly identified models due to the relative time-consuming estimation of the GMM model via simulated annealing algorithm.

3.4 Empirical Application

In our empirical application, we study the pricing of US and German government bonds. In particular, we are concentrating on factor pricing of futures contracts using both the Method of Moments estimator, and its approximation by standard quantile regression.

The majority of the literature devoted to bond pricing concentrates on monthly, quarterly or yearly data (e.g., Cochrane and Piazzesi (2005)). Our approach, however, concentrates on high-frequency data which also affect factors included in pricing kernel we use in our analysis. Žikeš and Baruník (2016) identified realized volatility to play an important role in the pricing of future return quantiles of S&P 500 and WTI Crude oil future contracts. Not to fully deviate from bond pricing literature we also study the influence of forward rates as in Fama and Bliss (1987) or Cochrane and Piazzesi (2005).

3.4.1 Data

Tick prices for both US and German government bonds were obtained from Tick Data[2]. For both countries two, five, ten and thirty years treasury futures are considered. Throughout the text, we use following abbreviations

Table 3.1: Treasuries abbreviations

	Name	Maturity	Abbreviation
US Treasury		2-years	TU
		5-years	FV
		10-years	TY
		30-years	US
EURO treasury	Euro - Schatz	2-years	BZ
	Euro - Bobl	5-years	BL
	Euro - Bund	10-years	BN
	Euro - Buxl	30-years	BX

US treasuries futures considered in our work are traded at CME Globex platform under Chicago Board of Trade rules. We consider trades from the period July 1, 2003 until November 30, 2017 during regular trading hours - Sunday to Friday, 5:00 p.m. - 4:00 p.m. Chicago Time. [3]. Euro bond futures are traded at EUREX exchange and we study period October 1, 2005 until November 30, 2017 during regular trading hours - Monday to Friday, 8:00 a.m. - 10:00 p.m.. To ensure sufficient liquidity, we explicitly exclude public holidays (Christmas, Thanksgiving, Independence Day, etc.) and days with less than 5 hours of trading. From the raw tick data, we extract 5 minutes prices using the last-tick method. For the analysis we calculate open-close returns, Realized Volatility (Andersen et al., 2003) and Realized Semi-variance (Barndorff-Nielsen et al., 2010).

Daily US forward rates are taken from the dataset of Gürkaynak et al. (2007)[4]. In particular we are working with two, three, four and five year forwards as in Cochrane and Piazzesi (2005). We synchronize the forward rate data with the treasury futures data resulting in 22 days without no forward rate available. These missing data are replaced using linear interpolation.

[2] `https://www.tickdata.com/`

[3] Monday to Thursday 4:00 p.m. - 5:00 p.m. Chicago Time is the maintenance period

[4] `https://www.federalreserve.gov/pubs/feds/2006/200628/200628abs.html`

3.4.2 Model Specification

We now describe the structure of quantile factor models used in the empirical application. We begin by computationally easier, at least from the quantile method of moments perspective, single factor model. We then follow by approximated multi-factor models since the implementation of multiple moment conditions in quantiles is not trivial and is subject to further research.

The single factor quantile pricing model in our work summarizes classical risk-return relationship. In particular we concentrate on relationship between returns and corresponding realized volatility so we rewrite (3.25) as

$$0 = Q_\tau \left[ln(1 + r_{t+1}) + a + \beta \widehat{RV}_{t+1}^{1/2} | \Omega_t \right], \tag{3.38}$$

where $\widehat{RV}_{t+1}^{1/2} = \sqrt{\sum_{k=1}^{N} (\Delta_k p_{i,t})^2}$ denotes realized volatility computed from a discretely sampled vector of k-th intraday log-returns in $[t-1, t]$ with N intraday observations, i.e. $\Delta_k p_t = p_{t-1+\nu_k/N} - p_{t-1+\nu_{k-1}/N}$. Since we have two parameters α and β to be estimated we use two instruments to have exactly identified model which allows us to use the smoothed method of moments estimator. The instruments we consider for estimation are lagged values of returns and realized volatility. Since the extension of single factor model into the multi-factor framework is computationally challenging, we illustrate the proximity of results of the method of moments estimation and standard quantile regression. We further build on these results and illustrate the estimation of multi-factor models in quantile regression framework.

Using theoretical results from Section 3.2.1 we rewrite single-factor (3.38) in the form of the classic quantile regression as

$$Q_\tau \left[r_{t+1} | \Omega_t \right] = a + \beta \widehat{RV}_t^{1/2}. \tag{3.39}$$

Extension to multi-factor is straightforward and for the realized semivariances we estimate

$$Q_\tau \left[r_{t+1} | \Omega_t \right] = a + \beta_{RS^+} \widehat{RS_t^+}^{1/2} + \beta_{RS^-} \widehat{RS_t^-}^{1/2}, \tag{3.40}$$

where RS_t^+ and RS_t^- are realized positive and negative semivariance respectively. Formally we define semivariances as $RV_t = RS_t^+ + RS_t^-$ with $\widehat{RS}_t^+ = \sum_{k=1}^{N} (\Delta_k p_t)^2 I(\Delta_k p_t > 0) \xrightarrow{p} \frac{1}{2} IV_t + \sum_{l=1}^{N_t} \kappa_{t,l}^2 I(\kappa_{t,l} > 0)$ and $\widehat{RS}_t^- = \sum_{k=1}^{N} (\Delta_k p_t)^2 I(\Delta_k p_t < 0) \xrightarrow{p} \frac{1}{2} IV_t + \sum_{l=1}^{N_t} \kappa_{t,l}^2 I(\kappa_{t,l} < 0)$ where IV_t stands for integrated variance, $N_{i,t}$ is total number of jumps during day t and $\sum_{l=1}^{N_t} \kappa_{i,t,l}^2$ represents magnitude of the jumps.

The final set of multi-factor models is based on forward rates which are frequently used in bond pricing literature. The general form of the forward rates factor model is

$$Q_\tau \left[r_{t+1} | \Omega_t \right] = a + \boldsymbol{\beta}_{\mathbf{FW}}{}' \mathbf{FW_t}, \tag{3.41}$$

where $\boldsymbol{\beta}_{\mathbf{FW}}$ is $p \times 1$ parameter vector and $\mathbf{FW_t}$ is $p \times T$ matrix of p forward rates of length T.

The estimation of all classical quantile regression single and multi-factor models (3.39, 3.40, 3.41) is carried in R using package *quantreg* [5].

3.5 Results

The presentation of results starts with a single-factor model for both US and German treasuries. Next, the multi-factor model based on Realized Semivariances is described. Finally the results of US forward rates multi-factor model are presented.

3.5.1 Single Factor Model

Results of Method of Moments estimator and Quantile Regression for the US treasuries are presented in left and right panels of Table 3.2 respectively. We also show the visual comparison for individual treasuries in Figure 3.1.

Concentrating on results of Method of Moments estimator displayed in the MM part of the Table 3.2 we can see that all but median coefficients are statistically significant which is in line with the theory of unpredictability of returns. Second, for all the assets, signs of the $\hat{\beta}_{RV^{1/2}}$ coefficients are consistent with classical risk-return trade-off (e.g., Value-at-Risk). In particular, we observe negative coefficients for quantiles below the median and positive coefficients for quantiles above the median. Financially speaking we can say that a risk-averse investor who optimizes the 5% quantile of asset TU, is not going to face loss higher than $-1.76 \times {RV_t}^{1/2}$. Similarly, a risk-taking investor optimizing at the 95% quantile of the same asset, will not have profit higher than $2.12 \times {RV_t}^{1/2}$. Third, there is an asymmetric influence of Realized Volatility on the quantiles of returns. The relative influence is stronger in upper quantiles than in lower quantiles, e.g., the absolute value of 1.76 of 5% quantile compared to 2.12 of 95% quantile. Fourth, unconditionally the least volatile asset TU [6] has the highest conditional relative risk-return trade-off, i.e., the absolute values of TU $\hat{\beta}_{RV^{1/2}}$ coefficients are the highest among US Treasuries.

After the description of MM results, we now turn to estimates obtained by Quantile Regression. Generally, results are very similar to the one obtained by Method of Moments estimator. In particular we can see in the QR part of the Table 3.2 that coefficient estimates are significant for all but median quantile, signs of the coefficients are as expected (negative sign for below median quantile, positive sign for above median quantile) and the least volatile asset has the highest conditional relative risk-return trade-off in the extreme quantiles (e.g. 5%, 10%, 90% and 95%). There is also asymmetry in the influence of Realized Volatility on the quantiles of returns. However, in contrast to the MM estimates for all assets except TU, there is a higher relative influence in lower quantiles than in higher quantiles.

We arrive to similar conclusions as presented in previous paragraphs also by visual inspection of Figure 3.1. In particular, all figures of slope coefficients are upward-sloping, and there is asymmetric influence of the Realized Volatility in the lower and upper

[5] `https://cran.r-project.org/web/packages/quantreg/index.html`

[6] The unconditional standard deviation of TU returns is the lowest among US treasuries.

Table 3.2: US Treasuries - Method of Moments and Quantile Regression estimates

	0.05	0.1	0.25	0.5	0.75	0.9	0.95	0.05	0.1	0.25	0.5	0.75	0.9	0.95
	MM							QR						
2-years maturity (TU)														
const	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)
$\hat{\beta}_{RV^{1/2}}$	-1.76	-1.37	-0.75	0.05	1.01	1.69	2.12	-1.62	-1.25	-0.52	0.08	0.83	1.41	1.79
	(0.12)	(0.16)	(0.09)	(0.09)	(0.07)	(0.14)	(0.16)	(0.13)	(0.1)	(0.07)	(0.07)	(0.07)	(0.08)	(0.14)
5-years maturity (FV)														
const	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)
$\hat{\beta}_{RV^{1/2}}$	-1.42	-1.13	-0.66	0.02	0.85	1.4	1.65	-1.56	-1.16	-0.52	0.03	0.63	1.21	1.39
	(0.11)	(0.1)	(0.08)	(0.09)	(0.06)	(0.11)	(0.11)	(0.13)	(0.08)	(0.07)	(0.07)	(0.07)	(0.1)	(0.1)
10-years maturity (TY)														
const	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)
$\hat{\beta}_{RV^{1/2}}$	-1.36	-1.13	-0.65	-0.13	0.86	1.47	1.59	-1.47	-1.24	-0.65	-0.01	0.64	1.15	1.42
	(0.1)	(0.08)	(0.08)	(0.13)	(0.1)	(0.14)	(0.13)	(0.14)	(0.08)	(0.08)	(0.08)	(0.08)	(0.1)	(0.11)
30-years maturity (US)														
const	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)
$\hat{\beta}_{RV^{1/2}}$	-1.58	-1.24	-0.87	-0.1	0.69	1.4	1.77	-1.61	-1.25	-0.88	-0.12	0.57	1.13	1.38
	(0.14)	(0.13)	(0.1)	(0.13)	(0.12)	(0.18)	(0.21)	(0.14)	(0.1)	(0.08)	(0.09)	(0.07)	(0.12)	(0.14)

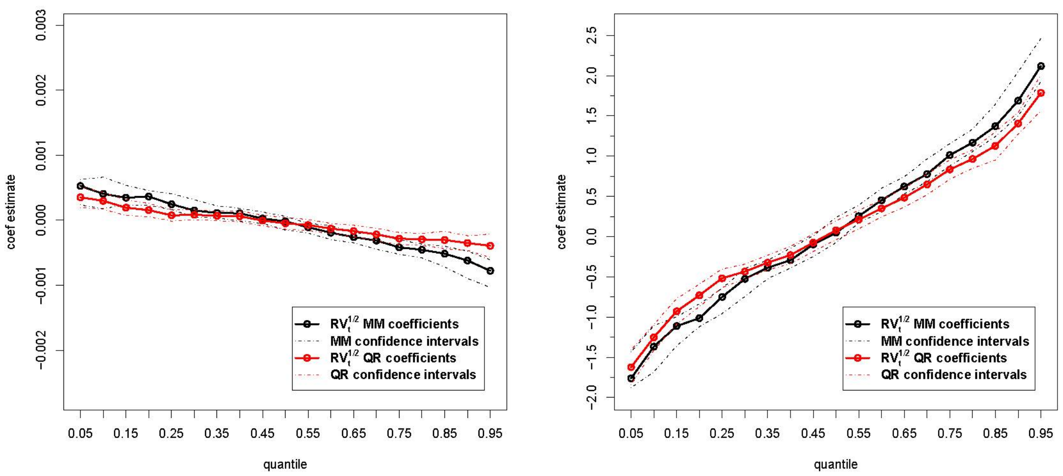
Note: Table displays coefficient estimates with corresponding standard errors in parentheses.

quantiles of returns. In these figures, it is also visible that quantile regression estimates (in red) underestimate the relative influence of RV in the upper quantiles compared to MM estimates. Moreover, the majority of quantiles of the Method of Moments estimates lies in the 95% confidence intervals of the Quantile Regression and vice-versa. This pattern is clearly visible in lower quantiles.

Figure 3.1: US Treasuries estimates - MM and QR

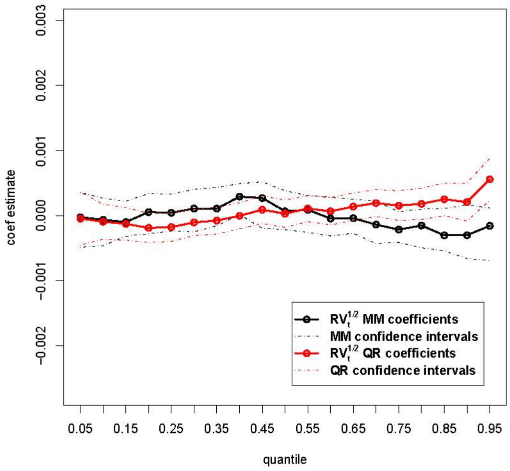
2-years maturity

(a) *TU intercept* (b) *TU slope*

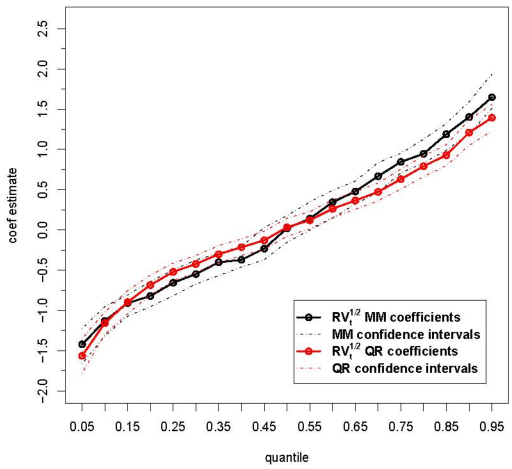


5-years maturity

(c) *FV intercept*

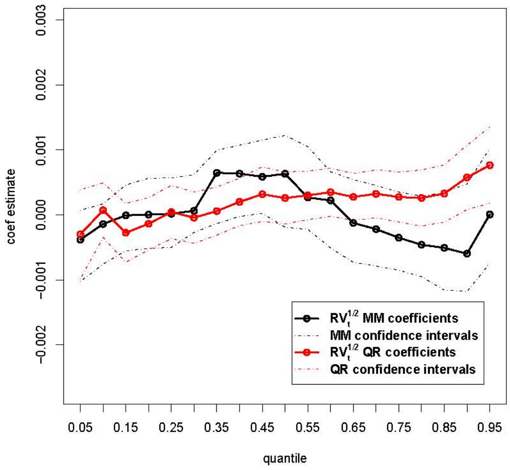


(d) *FV slope*

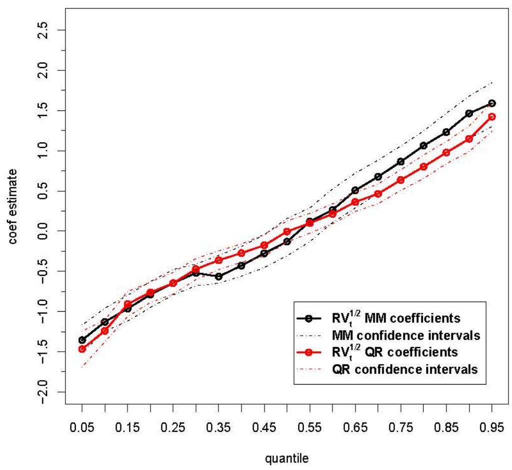


10-years maturity

(e) *TY intercept*

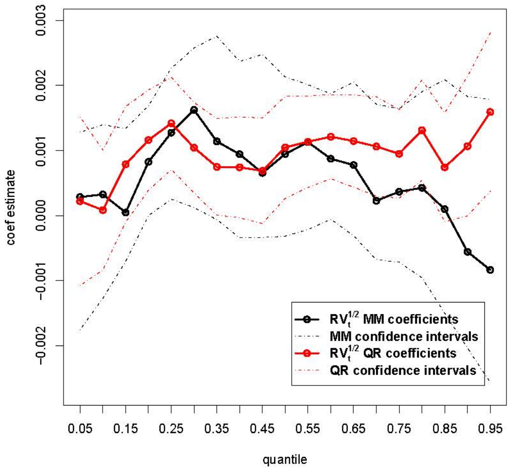


(f) *TY slope*

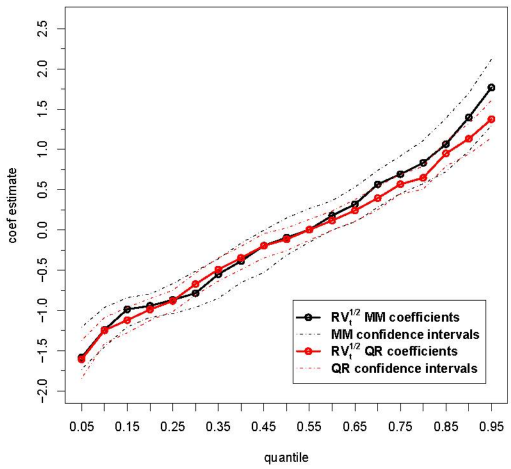


30-years maturity

(g) *US intercept*



(h) *US slope*



Note: Parameter estimates (solid lines) with corresponding 95% confidence intervals (dashed lines) from the MM and QR specification are plotted in black and red respectively.

Having described results of the US Treasuries analysis, we now turn to the case of German government bonds. In both Method of Moments and Quantile Regression estimators, we can spot similarities with US Treasuries results. First, all but some median coefficients are statistically significant. These significant median estimates are Method of Moments coefficient estimate of asset BZ (coefficient of 0.14 with standard error 0.05), Quantile Regression coefficient estimates of assets BN (coefficient of -0.19 with standard error 0.07) and BX (coefficient of -0.24 with standard error 0.07). These median anomalies deserve further research because they are in slight dispute with the efficient market hypothesis which states that returns should be unpredictable since prices follow random-walk. Second, $\hat{\beta}_{RV^{1/2}}$ coefficients have expected signs, i.e., quantiles below the median are negative, above the median are positive. Third, quantiles of returns are influenced asymmetrically by Realized Volatility. In the MM case, upper quantiles of short-term securities (maturity two and five years) are influenced more than lower quantiles (absolute value of 1.78 and 1.56 of 5% quantile compared to 1.87 and 1.74 of 95% quantile of assets BZ and BL) while situation in long-term securities (maturity ten and thirty years) is opposite (absolute value of 1.69 and 1.51 of 5% quantile compared to 1.59 and 1.37 of 95% quantile of assets BN and BX). In contrast, the relative influence of Realized Volatility in the Quantile Regression estimates is always greater in the lower quantiles. Fourth, the shortest maturity bond (BZ) is unconditionally the least volatile asset, however, conditionally it has the highest relative risk-return trade-off.

Table 3.3: EU Treasuries - Method of Moments and Quantile Regression estimates

	0.05	0.1	0.25	0.5	0.75	0.9	0.95	0.05	0.1	0.25	0.5	0.75	0.9	0.95
	MM							QR						
2-years maturity (BZ)														
const	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)
$\hat{\beta}_{RV^{1/2}}$	-1.78	-1.38	-0.74	0.14	0.99	1.54	1.87	-1.95	-1.55	-0.71	0.09	0.82	1.43	1.78
	(0.12)	(0.1)	(0.08)	(0.05)	(0.11)	(0.14)	(0.11)	(0.15)	(0.1)	(0.07)	(0.06)	(0.08)	(0.1)	(0.12)
5-years maturity (BL)														
const	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)
$\hat{\beta}_{RV^{1/2}}$	-1.56	-1.26	-0.65	0.07	0.93	1.48	1.74	-1.78	-1.44	-0.68	0.01	0.73	1.37	1.64
	(0.07)	(0.09)	(0.1)	(0.07)	(0.1)	(0.12)	(0.08)	(0.11)	(0.11)	(0.08)	(0.07)	(0.08)	(0.1)	(0.1)
10-years maturity (BN)														
const	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)
$\hat{\beta}_{RV^{1/2}}$	-1.69	-1.28	-0.66	-0.05	0.79	1.41	1.59	-1.7	-1.44	-0.87	-0.19	0.48	1.31	1.5
	(0.13)	(0.12)	(0.11)	(0.11)	(0.15)	(0.17)	(0.14)	(0.11)	(0.12)	(0.09)	(0.07)	(0.1)	(0.12)	(0.12)
30-years maturity (BX)														
const	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)
$\hat{\beta}_{RV^{1/2}}$	-1.51	-1.33	-0.63	-0.16	0.61	1.3	1.37	-1.74	-1.42	-0.83	-0.24	0.47	1.12	1.39
	(0.12)	(0.09)	(0.08)	(0.1)	(0.12)	(0.12)	(0.1)	(0.1)	(0.11)	(0.09)	(0.07)	(0.09)	(0.1)	(0.14)

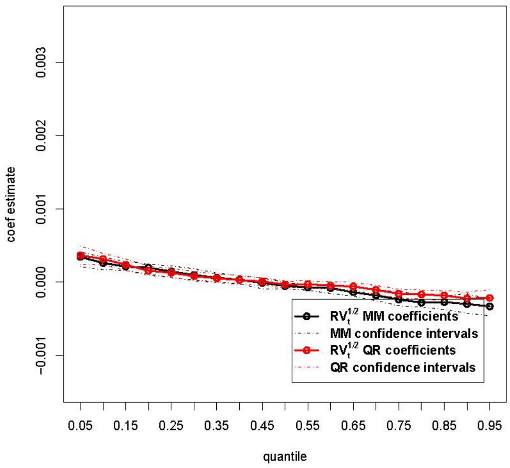
Note: Table displays coefficient estimates with corresponding standard errors in parentheses.

Visual inspection of the results presented in Figure 3.2 reveals similarity but also an interesting difference from the US Treasuries. The similarity is in the width of the confidence intervals of the slope estimates. Specifically, Method of Moments estimates lies almost always in the confidence intervals of Quantile Regression estimates and vice-versa. The difference from the US Treasuries is that Quantile Regression $\hat{\beta}_{RV^{1/2}}$ estimates for almost all the quantiles of all the assets relatively underestimate the influence of the Realized Volatility compared to Method of Moments Estimates. In contrast, MM and QR estimates of US Treasuries intersects frequently, and there is no clear under/overestimation of any estimator.

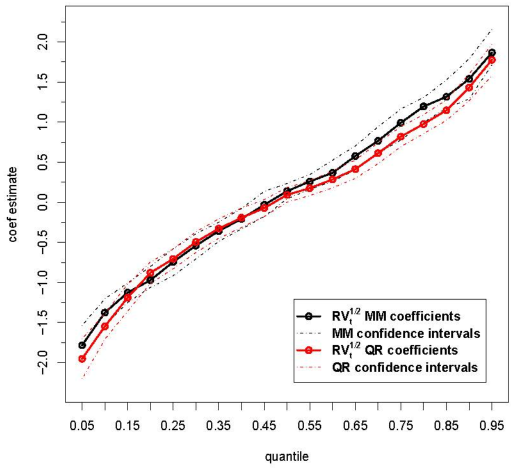
Figure 3.2: German Treasuries estimates - MM and QR

2-years maturity

(a) *BZ intercept*

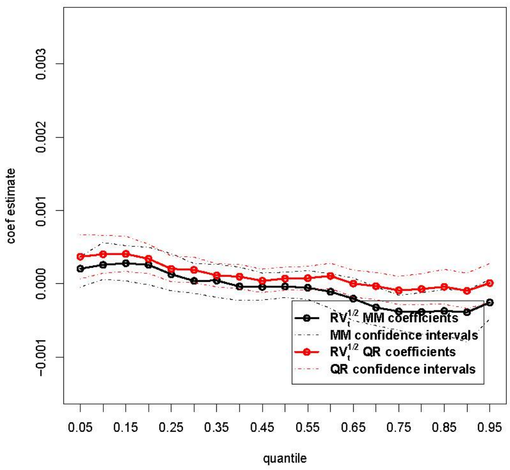


(b) *BZ slope*

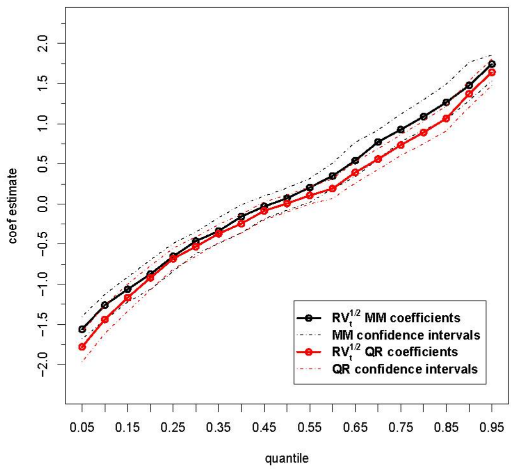


5-years maturity

(c) *BL intercept*

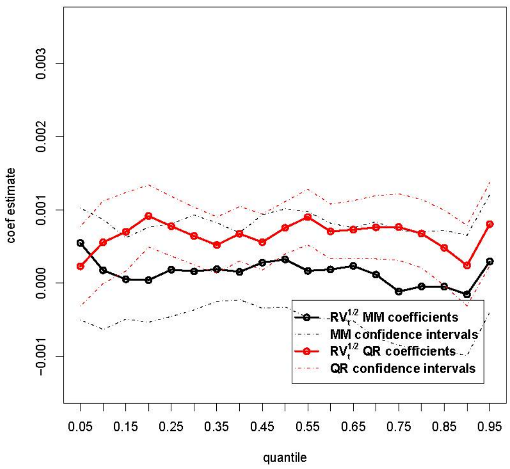


(d) *BL slope*

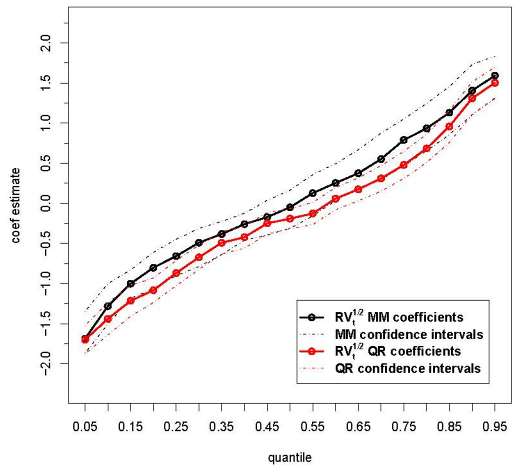


10-years maturity

(e) *BN intercept*

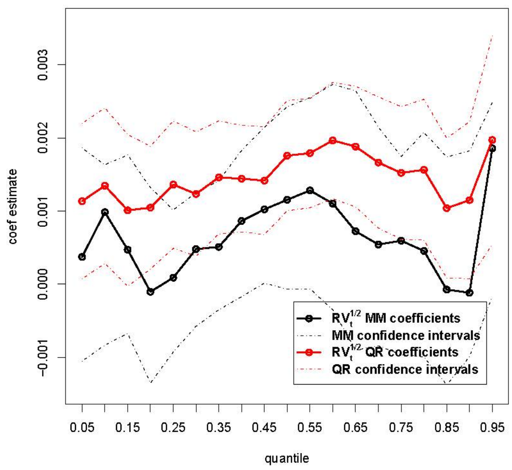


(f) *BN slope*

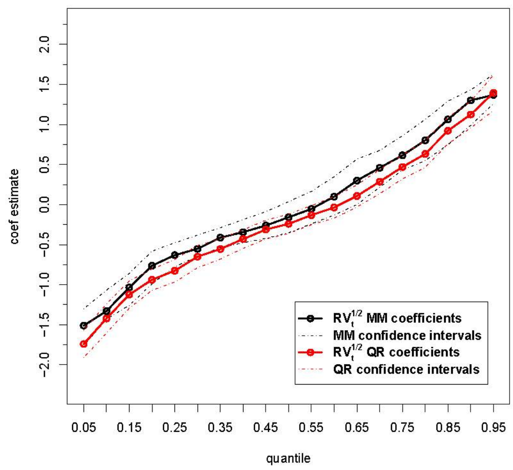


30-years maturity

(g) *BX intercept*



(h) *BX slope*



Note: Parameter estimates (solid lines) with corresponding 95% confidence intervals (dashed lines) from the MM and QR specification are plotted in black and red respectively.

3.5.2 Multi Factor Model

After the description of the single-factor models results, we now turn to multi-factor models. We present two-factor models of the Realized Semivariances for both US and German Treasuries and the forward rates factor model for US Treasuries. Specifically, we study four-factor forward rate model similar to one in Cochrane and Piazzesi (2005). As previously mentioned in the Section 3.4.2 we will concentrate on Quantile Regression estimation only.

Realized Semivariance

Results of the analysis when the Realized Semivariances are used as factors share many commonalities with single-factor models. For both US and German Treasuries, coefficient estimates of $\hat{\beta}_{RS^{+1/2}}$ and $\hat{\beta}_{RS^{-1/2}}$ have negative signs for quantiles below the median, positive for quantiles above the median. Moreover, figures of $\hat{\beta}_{RS^{+1/2}}$ and $\hat{\beta}_{RS^{-1/2}}$ are upward sloping. Furthermore, all the coefficients are statistically significant for all but median quantiles in US Treasuries, while ten and thirty years German treasuries have statistically significant median coefficients of the $\hat{\beta}_{RS^{+1/2}}$. In addition, the shortest maturity treasuries TU and BZ have the highest conditional influence of the Realized Semivariances on the quantiles of returns.

Besides commonalities, there are also features specific to Realized Semivariances. In the US Treasuries case, negative semivariance influence lower quantiles relatively more than the upper quantiles (-1.24, -1.13 or -1.61 coefficients of 5% quantiles vs. 0.34, 0.88 or 0.51 of 95% quantiles of TU, TY, and US respectively) while the opposite is true for positive semivariance. We can also see that for upper quantiles RS^{+} clearly dominates RS^{-} whereas in the lower quantiles the results are mixed. To illustrate, let us consider FV in which RS^{+} has higher influence than RS^{-} in 5 and 10% quantiles while the opposite is true for TY and finally in case of TU, 5% quantile is dominated by RS^{-} whereas 10% quantile is dominated by RS^{+}. Turning to German Treasuries, we comment the main differences from US Treasuries. In the German government bonds, we observe more symmetric influence of Realized Semivariances on the quantiles of returns, particularly in the shortest and the longest maturity bonds. Moreover, German treasuries look more homogeneous as the estimates of all the bonds for given quantiles are closer to each other. Furthermore, the relative influence of the negative semivariance is dominant in upper quantiles (exact opposite to US Treasuries) while in the lower quantiles neither of the semivariances plays dominant role similarly to US Government bonds. Asymmetric influence documented in this multi-factor specification has many potential sources. Besides difference of the Realized Volatility and Realized Semivariance, we attribute it also to the institutional differences of the US and EU markets, e.g. different trading hours (US market 23 hours a day vs 14 hours EU market), investors‘ perception of US Treasury futures being well established investment instruments (history of 30-years US Tresury futures date back to October 1982 while 30-year German Treasury futures were introduced in September 2005), etc.. Last but not least, liquidity connected to different trading hours also plays a role.

The important difference is also in the median estimates. While in the US Treasuries majority of estimates are either zero or close to zero with high standard errors this is not the case of the German Treasuries. We can see in the EU part of the Table 3.4 that none of the median coefficient (except constant terms) are zero, moreover, some of the estimates are statistically significant as we mentioned at the beginning of the section.

Table 3.4: US & EU Treasuries - Quantile Regression Realized Semivariances estimates

	0.05	0.1	0.25	0.5	0.75	0.9	0.95	0.05	0.1	0.25	0.5	0.75	0.9	0.95
	USA							EU						
	2-years maturity (TU)							*2-years maturity (BZ)*						
const	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)
$\hat{\beta}_{RS^{+1/2}}$	-1.08	-1.15	-0.38	0.12	0.86	1.39	2.06	-1.49	-1.44	-0.64	-0.17	0.19	0.68	1.04
	(0.25)	(0.18)	(0.14)	(0.16)	(0.14)	(0.18)	(0.22)	(0.35)	(0.19)	(0.18)	(0.12)	(0.2)	(0.23)	(0.26)
$\hat{\beta}_{RS^{-1/2}}$	-1.24	-0.67	-0.39	0	0.34	0.71	0.34	-1.3	-0.74	-0.37	0.3	0.97	1.3	1.5
	(0.17)	(0.21)	(0.13)	(0.15)	(0.12)	(0.09)	(0.25)	(0.35)	(0.24)	(0.16)	(0.14)	(0.19)	(0.22)	(0.32)
	5-years maturity (FV)							*5-years maturity (BL)*						
const	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)
$\hat{\beta}_{RS^{+1/2}}$	-0.49	-0.39	-0.15	0	0.27	0.42	0.55	-1	-1.05	-0.76	-0.17	0.34	0.66	0.79
	(0.02)	(0.08)	(0.05)	(0.04)	(0.05)	(0.08)	(0.09)	(0.21)	(0.17)	(0.17)	(0.14)	(0.16)	(0.19)	(0.18)
$\hat{\beta}_{RS^{-1/2}}$	-0.36	-0.22	-0.14	0	0.14	0.36	0.51	-1.59	-1.02	-0.3	0.19	0.72	1.31	1.55
	(0.07)	(0.08)	(0.05)	(0.04)	(0.05)	(0.08)	(0.11)	(0.24)	(0.22)	(0.15)	(0.14)	(0.18)	(0.22)	(0.2)
	10-years maturity (TY)							*10-years maturity (BN)*						
const	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)
$\hat{\beta}_{RS^{+1/2}}$	-0.98	-0.8	-0.59	-0.05	0.56	1.18	1.17	-1.16	-1.15	-0.93	-0.36	-0.16	0.42	0.52
	(0.23)	(0.17)	(0.15)	(0.15)	(0.14)	(0.17)	(0.2)	(0.19)	(0.2)	(0.16)	(0.14)	(0.16)	(0.19)	(0.23)
$\hat{\beta}_{RS^{-1/2}}$	-1.13	-0.99	-0.3	0.06	0.38	0.49	0.88	-1.3	-0.87	-0.36	0.12	0.86	1.44	1.63
	(0.27)	(0.17)	(0.16)	(0.16)	(0.13)	(0.14)	(0.2)	(0.24)	(0.21)	(0.17)	(0.16)	(0.18)	(0.21)	(0.16)
	30-years maturity (US)							*30-years maturity (BX)*						
const	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)
$\hat{\beta}_{RS^{+1/2}}$	-0.77	-0.75	-0.61	-0.11	0.43	1.2	1.58	-1.33	-1.17	-0.83	-0.45	0.09	0.6	1.09
	(0.27)	(0.16)	(0.14)	(0.16)	(0.14)	(0.19)	(0.19)	(0.22)	(0.21)	(0.16)	(0.16)	(0.18)	(0.2)	(0.27)
$\hat{\beta}_{RS^{-1/2}}$	-1.61	-1.07	-0.67	-0.03	0.44	0.48	0.51	-1.15	-0.85	-0.41	0.13	0.56	1.04	0.9
	(0.25)	(0.2)	(0.16)	(0.17)	(0.13)	(0.15)	(0.08)	(0.22)	(0.22)	(0.17)	(0.16)	(0.18)	(0.19)	(0.29)

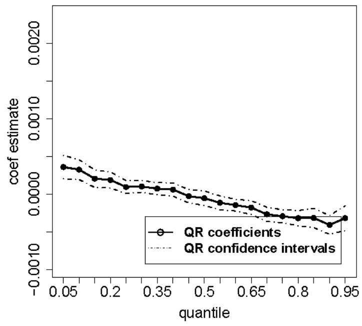
Note: Table displays coefficient estimates with corresponding standard errors in parentheses.

Results of our analysis are also confirmed by visual inspection of Figure 3.3 and Figure 3.4. In all of the RS^+ and RS^- subfigures, we can see the upward sloping pattern. In the US Treasuries, the dominance of negative semivariance in lower quantiles and positive semivariance in the upper quantiles is well documented in Figure 3.3k and Figure 3.3l. The opposite is true for EU Treasuries where comparison of Figure 3.4b and Figure 3.4c suggest that the below median quantiles are more influenced by positive semivariance and the above median quantiles by the negative one. When we compare total dispersion of the RS^+ and RS^- the US Treasuries seems to be more heterogeneous (figures 3.3b, 3.3c vs 3.3e, 3.3f) than all the RS^+ and RS^- pairs of the German Treasuries. As a possible source of this heterogeneity, we identify the overall higher dispersion of the average liquidity of the US Treasuries. While the average daily volume of EU bond futures is around 340, 400 and 704 thousand for 2,5 and 10 years maturity respectively, the US volumes are 190, 500, 980 thousand. However, the liquidity of the 30-years bonds is a bit puzzling for us - 13 thousands of EU bond vs 270 thousand in the US case.

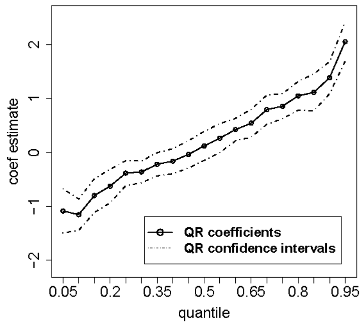
Figure 3.3: US Treasury estimates - QR RSV

2-years maturity

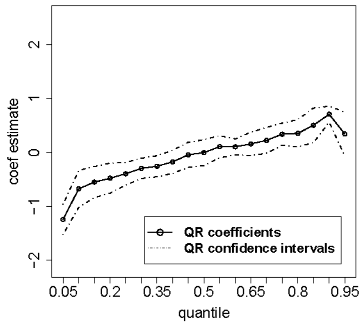
(a) *TU intercept*



(b) *TU RS*$^{+}$

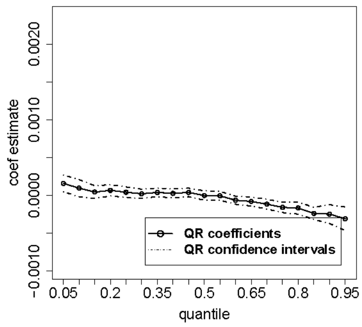


(c) *TU RS*$^{-}$

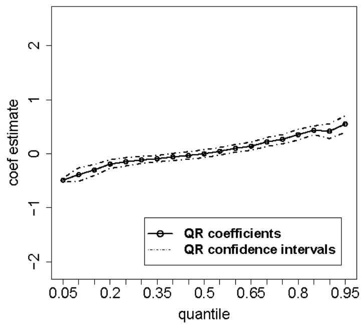


5-years maturity

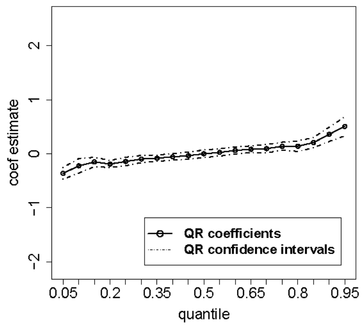
(d) *FV intercept*



(e) *FV RS*$^{+}$

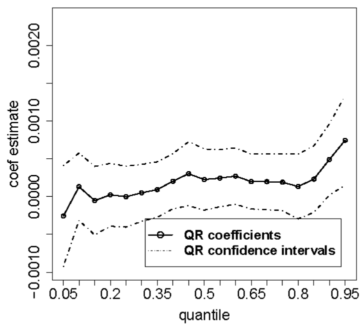


(f) *FV RS*$^{-}$

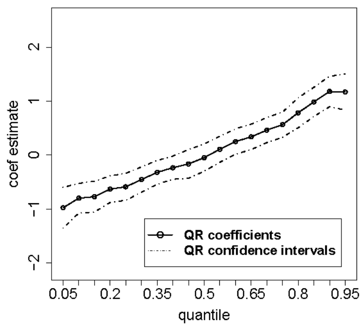


10-years maturity

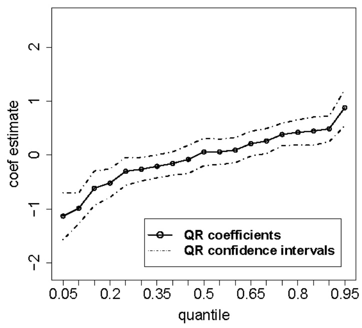
(g) *TY intercept*



(h) *TY RS*$^{+}$

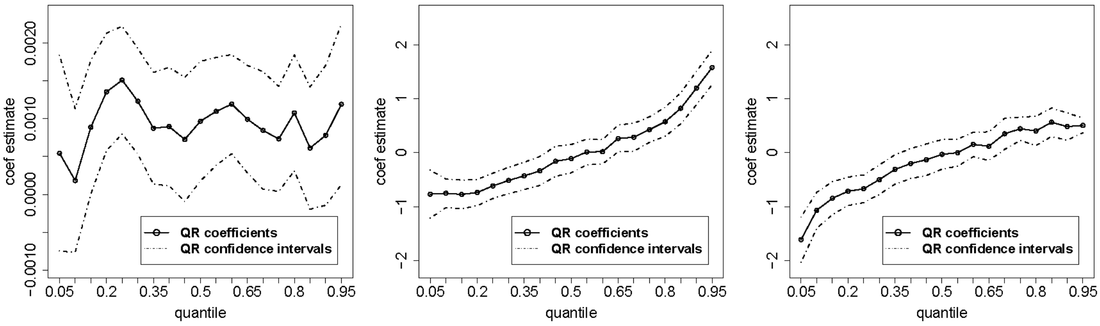


(i) *TY RS*$^{-}$



30-years maturity

(j) $US\,intercept$ (k) $US\,RS^{+}$ (l) $US\,RS^{-}$

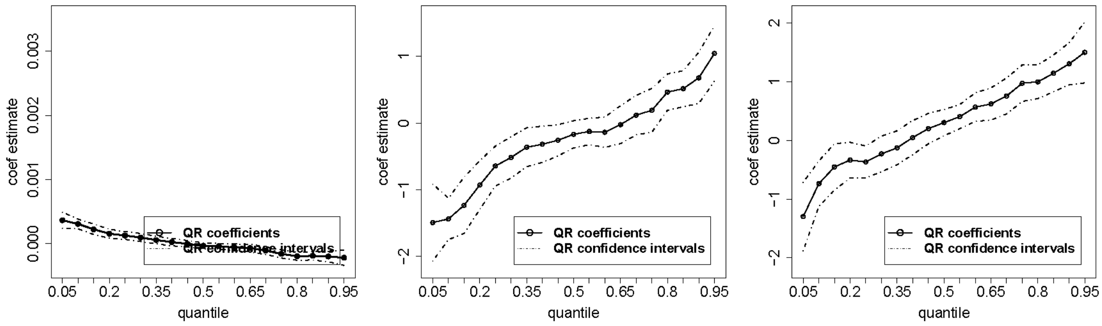


Note: Parameter estimates with corresponding 95% confidence intervals are plotted in solid and dashed lines respectively.

Figure 3.4: EU Treasury estimates - QR RSV

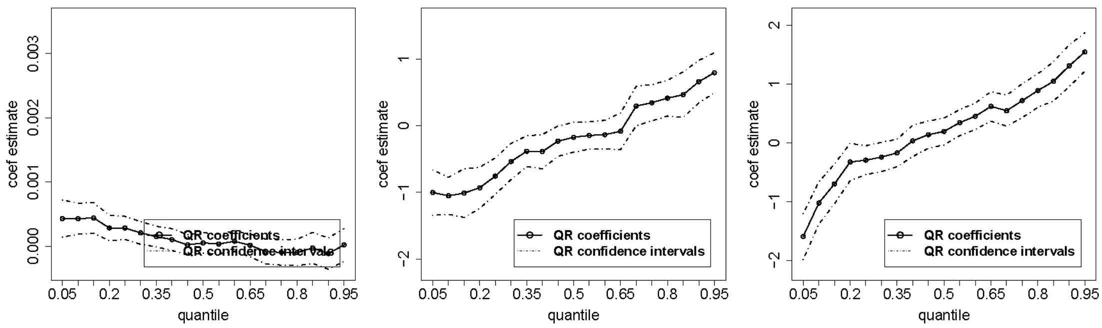
2-years maturity

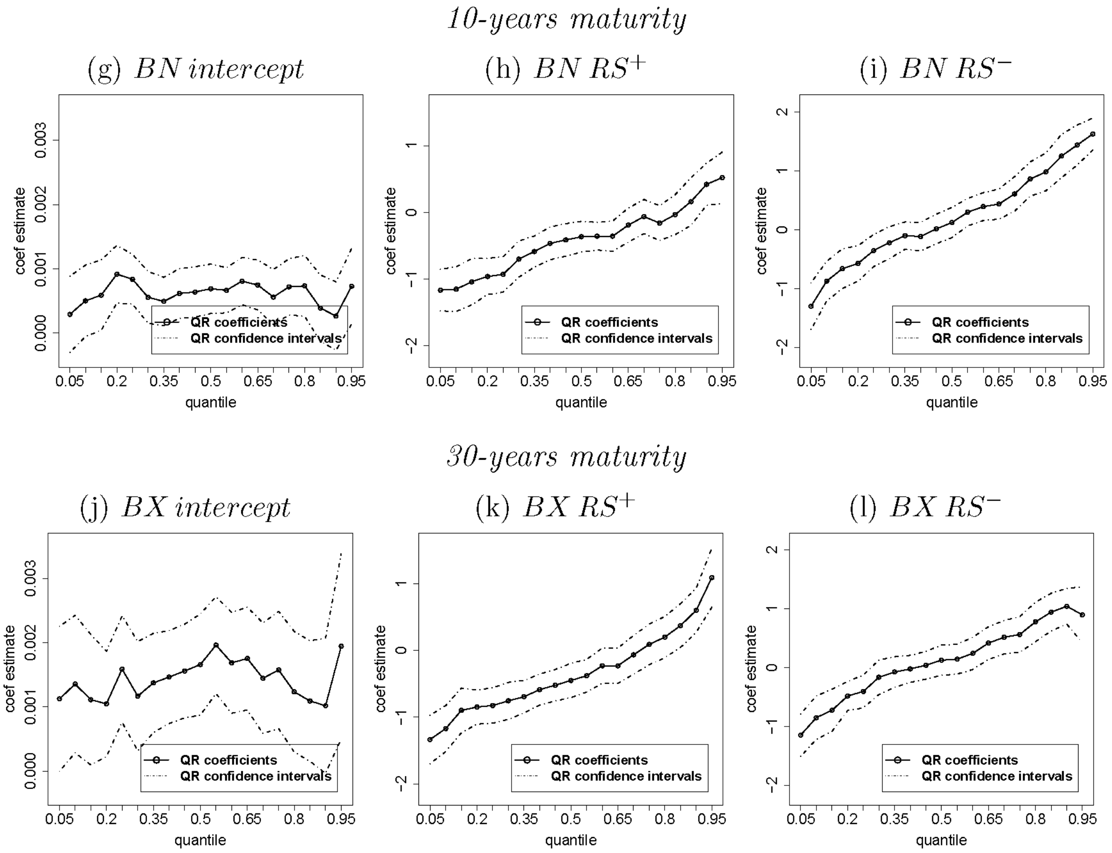
(a) $BZ\,intercept$ (b) $BZ\,RS^{+}$ (c) $BZ\,RS^{-}$



5-years maturity

(d) $BL\,intercept$ (e) $BL\,RS^{+}$ (f) $BL\,RS^{-}$





Note: Parameter estimates with corresponding 95% confidence intervals are plotted in solid and dashed lines respectively.

US Treasuries - Forward rates

We finalize results of our analysis by four-factor forward rates model applied on the US Treasuries. We are working with two, three, four and five years forwards and we adopt naming from the original paper of Gürkaynak et al. (2007) e.g., the name of two-years forward is SVENF02, etc..

At first sight, results presented in Table 3.5 differ substantially from all the results presented so far. The biggest difference is the statistical significance of the coefficient estimates.For all the treasuries and all the forward rates, the vast majority of the estimates is statistically insignificant. On one side, for the risk-averse investor who optimizes lower quantiles of the returns distribution, forward rates are of limited use since their coefficients are not statistically different from zero. On the other side, a risk-taking investor might find forward rates useful when concentrating on the upper quantiles of the shortest maturity treasury TU. For the rest of the treasuries, forward rates do not seem to play a very important role in determining the behavior of the quantiles of the returns. This is in contrast with Cochrane and Piazzesi (2005) who use monthly data and show that linear combination of the short-term forward rates can be useful in forecasting excess returns of the short-term bonds. We address this difference to the structure of the dataset. The theory suggests that lower frequency returns (e.g., monthly, quarterly) exhibit certain degree of predictability(e.g. Cochrane and Piazzesi (2005),Fama and French (1993)), however, returns at higher frequencies (e.g., daily) should not be forecastable as suggested by Efficient Market Hypothesis (Fama, 1970) what median results of our analysis confirm.

Table 3.5: US treasuries - QR forward rates

	0.05	0.1	0.25	0.5	0.75	0.9	0.95	0.05	0.1	0.25	0.5	0.75	0.9	0.95
	2-years maturity (TU)							*5-years maturity (FV)*						
const	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)
SVENF02	-0.02	0.03	0.02	0.01	0.06	0.18	0.29	0.15	0.29	0.09	0.13	0.21	0.17	0.39
	(0.08)	(0.07)	(0.04)	(0.04)	(0.04)	(0.08)	(0.09)	(0.25)	(0.22)	(0.12)	(0.11)	(0.12)	(0.21)	(0.26)
SVENF03	0.07	-0.07	-0.08	-0.05	-0.23	-0.63	-1	-0.27	-0.8	-0.19	-0.49	-0.82	-0.76	-1.49
	(0.28)	(0.24)	(0.14)	(0.12)	(0.14)	(0.26)	(0.33)	(0.85)	(0.74)	(0.42)	(0.38)	(0.42)	(0.73)	(0.89)
SVENF04	-0.1	0.04	0.07	0.06	0.28	0.75	1.18	0.37	0.89	0.13	0.54	0.92	0.87	1.65
	(0.35)	(0.29)	(0.16)	(0.14)	(0.17)	(0.31)	(0.4)	(1.03)	(0.88)	(0.51)	(0.46)	(0.51)	(0.89)	(1.08)
SVENF05	0.01	-0.03	-0.03	-0.02	-0.09	-0.26	-0.42	-0.32	-0.43	-0.06	-0.19	-0.28	-0.22	-0.47
	(0.14)	(0.12)	(0.07)	(0.06)	(0.07)	(0.13)	(0.16)	(0.42)	(0.36)	(0.21)	(0.19)	(0.21)	(0.36)	(0.45)
	10-years maturity (TY)							*30-years maturity (US)*						
const	0	0	0	0	0	0	0	-0.01	-0.01	0	0	0	0.01	0.01
	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)
SVENF02	0.2	0.53	0.19	0.21	0.19	0.05	0.18	0.48	0.9	0.31	0.21	-0.18	0.01	0.11
	(0.39)	(0.32)	(0.19)	(0.18)	(0.2)	(0.32)	(0.34)	(0.67)	(0.49)	(0.29)	(0.33)	(0.31)	(0.42)	(0.56)
SVENF03	-0.39	-1.64	-0.51	-0.73	-0.63	-0.23	-0.64	-0.86	-2.82	-0.95	-0.68	0.66	-0.32	-0.29
	(1.36)	(1.1)	(0.66)	(0.64)	(0.71)	(1.1)	(1.2)	(2.4)	(1.71)	(1.05)	(1.15)	(1.08)	(1.51)	(1.99)
SVENF04	0.74	2.09	0.61	0.8	0.46	-0.08	0.24	1.13	3.64	1.33	0.71	-1.18	0.03	-0.61
	(1.63)	(1.33)	(0.8)	(0.78)	(0.85)	(1.33)	(1.46)	(2.97)	(2.1)	(1.31)	(1.41)	(1.32)	(1.88)	(2.48)
SVENF05	-0.61	-1.02	-0.32	-0.29	0.02	0.31	0.29	-0.72	-1.7	-0.71	-0.25	0.69	0.28	0.75
	(0.66)	(0.55)	(0.33)	(0.32)	(0.35)	(0.55)	(0.6)	(1.24)	(0.88)	(0.55)	(0.58)	(0.55)	(0.79)	(1.04)

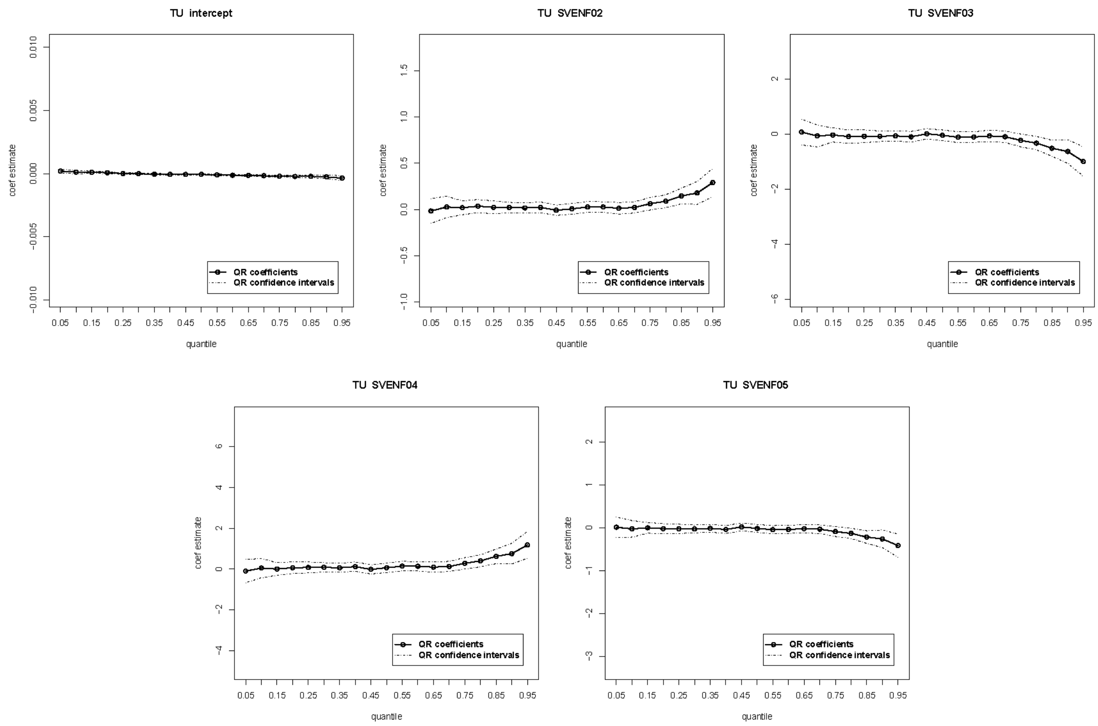
Note: Table displays coefficient estimates with corresponding standard errors in parentheses.

The visual inspection of the results presented in Figure 3.5 confirm the results displayed in the Table 3.5. In almost all but intercept figures, parameters estimates are close to zero. Moreover, confidence intervals are rather wide, so we can not say that parameter estimates are not zero. We can also see that coefficient estimates lack the dynamic and are stable across quantiles (e.g. Figure 3.5b, subplot $FV\,SVENF05$). This suggests that forward rates might not drive the dynamics of the quantiles of the US treasury returns as it was the case in the single-factor Realized volatility or multi-factor Realized Semivariance models.

Figure 3.5: US Treasuries - QR forward rates estimates

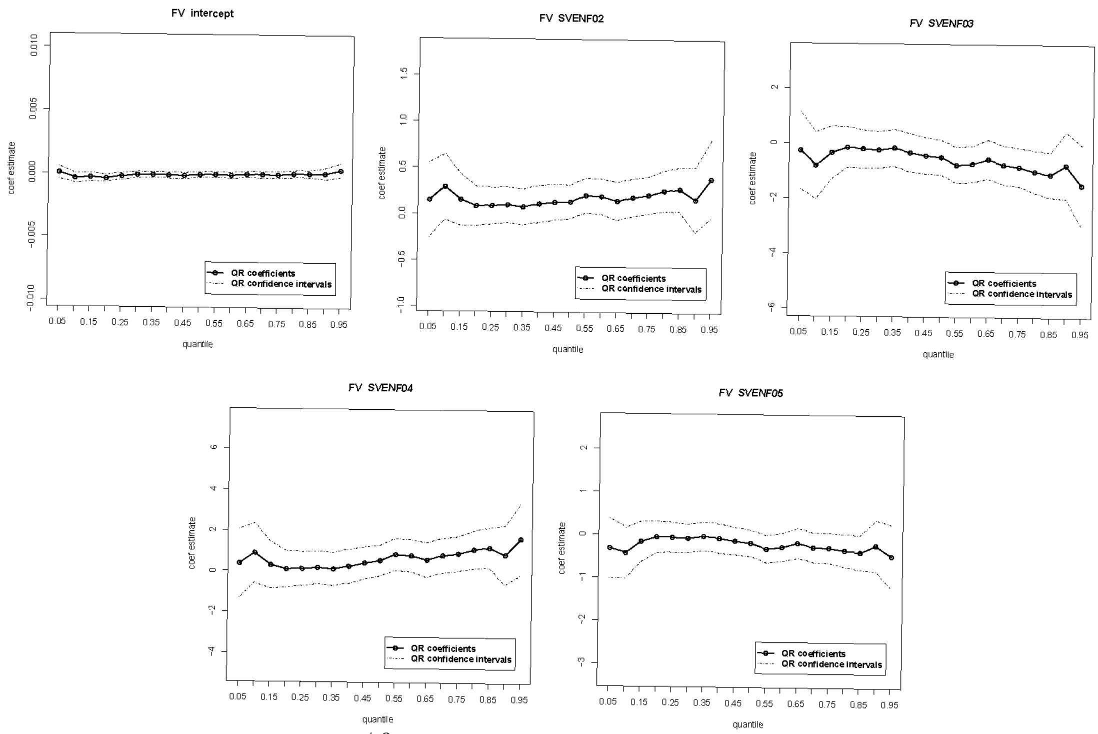
2-years maturity estimates

(a) TU



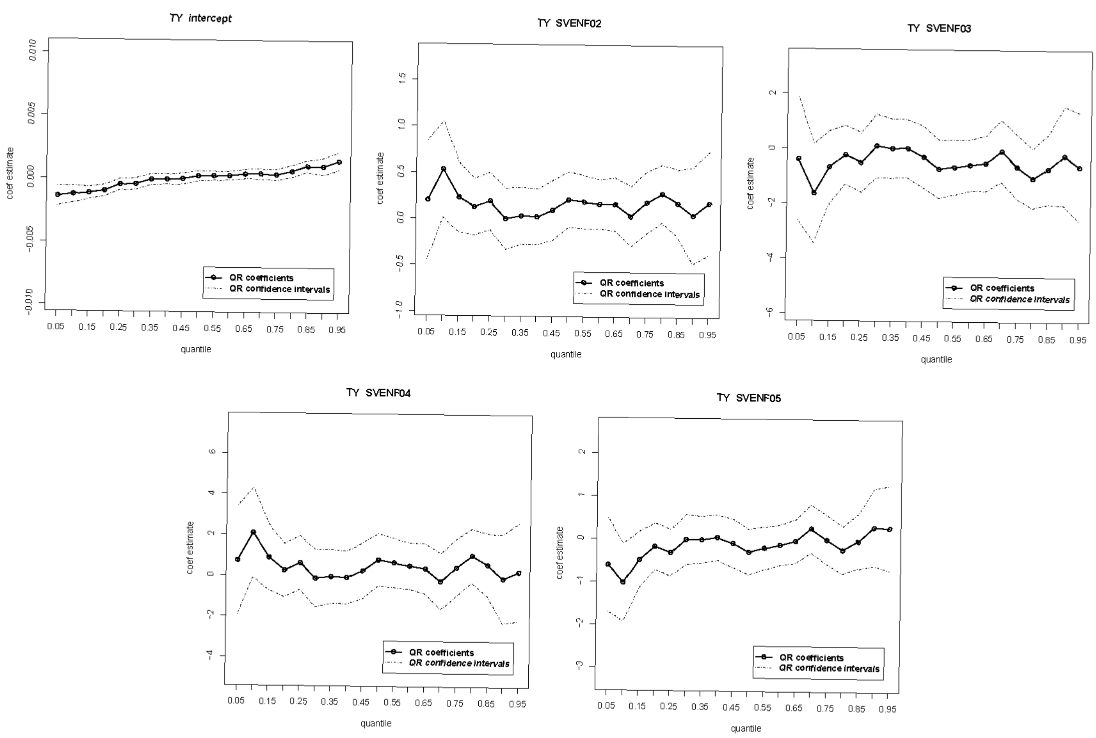
5-years maturity estimates

(b) FV



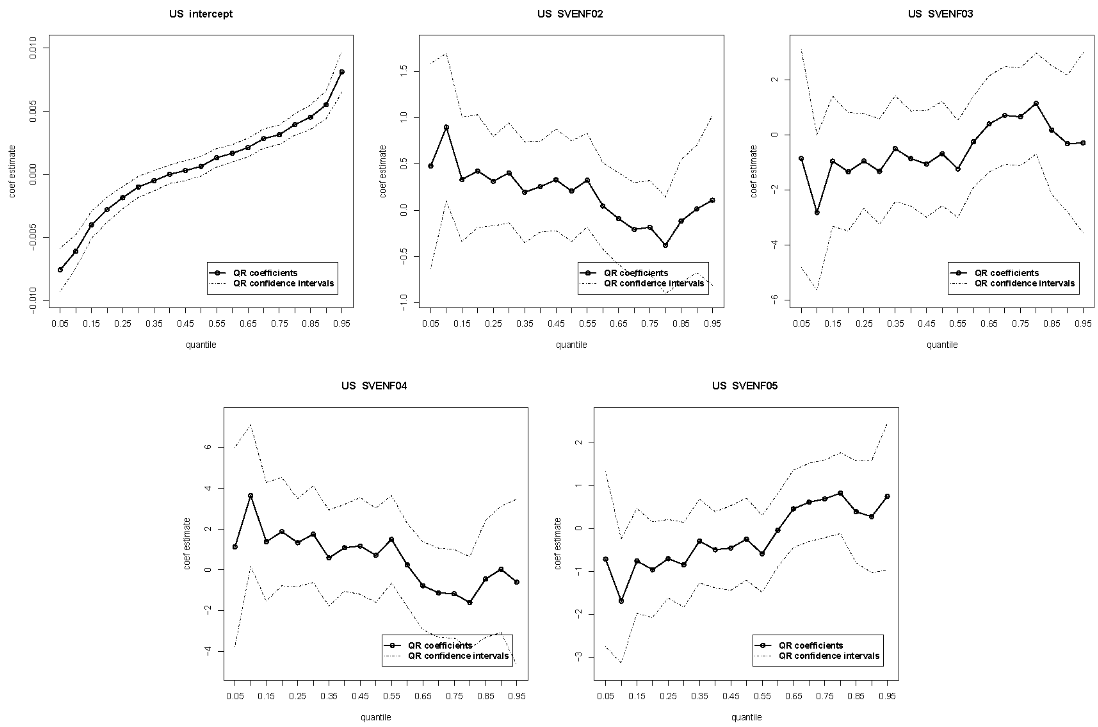
10-years maturity estimates

(c) TY



30-years maturity estimates

(d) US



Note: Parameter estimates with corresponding 95% confidence intervals are plotted in solid and dashed lines respectively.

3.6 Conclusion

In this chapter, we show that basic pricing equation (Cochrane, 2009) can be easily extended from the expected to the quantile utility framework. We transform quantile Euler equation of de Castro and Galvao (2018) into stochastic discount factor/pricing kernel representation and present link to the quantile factor asset pricing.

In the empirical application, we study the behavior of the quantiles of the US and German Treasury returns. Our analysis demonstrates a significant influence of the Realized Volatility and Realized Semivariance on the distribution of the returns. Specifically, Realized Volatility plays an important role in all but median quantiles of the return distribution, whereas Realized Semivariances also partly explain the median variation of two German Treasuries. In contrast, we fail to find such significant relationship for the forward rates.

From the estimation point of view, we adopt recently developed Smoothed Method of Moments estimator (de Castro et al., 2018) and traditional quantile regression (Koenker and Bassett Jr, 1978). In the single-factor quantile asset pricing model we illustrate proximity of both methods. This proximity is further applied in the multi-factor models where we rely solely on the quantile regression approach.

3.7 Appendix

Table 3.6: Descriptive statistic open-close returns - US & German Treasuries

	Mean	St.dev	Median	Min	Max	Skew	Kurt	JB	ARCH-LM	$\overline{\text{Volume}}$
TU	0.0000	0.0009	0.0000	-0.008	0.008	-0.024	7.98	9668.28	570.48	166697
FV	0.0001	0.0025	0.0001	-0.016	0.017	0.010	3.05	1415.91	396.94	492719
TY	0.0002	0.0038	0.0002	-0.020	0.033	0.101	3.10	1462.84	322.63	953719
US	0.0002	0.0065	0.0005	-0.029	0.047	0.077	2.12	685.47	301.90	289027
BZ	0.0000	0.0007	0.0000	-0.005	0.005	-0.142	6.68	5616.75	491.84	343470
BL	0.0001	0.0020	0.0001	-0.011	0.010	-0.060	3.20	1285.15	422.11	403728
BN	0.0001	0.0035	0.0002	-0.019	0.019	-0.122	2.41	737.60	284.82	704011
BX	0.0002	0.0076	0.0002	-0.040	0.035	-0.191	2.27	668.30	340.10	13195

Note: $\overline{\text{Volume}}$ denotes average daily volume of trades. Time span of US and German Treasuries are July 1,2003-November 30,2017 and October 1, 2005-November 30, 2017 respectively.

Table 3.7: Descriptive statistic selected forward rates

	Mean	St.dev	Median	Min	Max	Skew	Kurt	JB	ARCH-LM
SVENF02	0.0215	0.0134	0.0169	0.003	0.051	0.616	-0.89	350.56	3606.18
SVENF03	0.0270	0.0118	0.0239	0.006	0.051	0.172	-1.13	212.21	3601.48
SVENF04	0.0319	0.0111	0.0310	0.010	0.052	-0.134	-1.23	240.46	3598.10
SVENF05	0.0361	0.0110	0.0382	0.014	0.056	-0.227	-1.29	284.28	3597.56

Time span: July 1,2003-November 30,2017.

Table 3.8: Descriptive statistic Realized Measures

	Mean	St.dev	Median	Min	Max	Skew	Kurt	JB	ARCH-LM
					RV				
TU	0.0000	0.0000	0.0000	0.000	0.000	6.769	67.99	728653.24	1323.89
FV	0.0000	0.0000	0.0000	0.000	0.000	7.280	100.74	1571018.84	132.67
TY	0.0000	0.0000	0.0000	0.000	0.001	11.910	302.96	14003348.42	11.59
US	0.0000	0.0000	0.0000	0.000	0.002	16.876	569.64	49372998.31	3.07
BZ	0.0000	0.0000	0.0000	0.000	0.000	5.877	49.66	326734.21	813.17
BL	0.0000	0.0000	0.0000	0.000	0.000	4.205	28.13	108172.50	865.00
BN	0.0000	0.0000	0.0000	0.000	0.000	12.596	349.86	15436397.10	5.81
BX	0.0001	0.0001	0.0000	0.000	0.001	4.102	28.27	108732.91	622.70
					RVOL				
TU	0.0010	0.0005	0.0008	0.000	0.005	3.050	13.71	34167.50	1886.08
FV	0.0023	0.0011	0.0020	0.001	0.014	2.382	10.63	20561.85	1186.47
TY	0.0036	0.0016	0.0032	0.001	0.025	2.627	16.10	43487.57	974.88
US	0.0062	0.0023	0.0057	0.002	0.044	2.723	23.01	84808.44	779.43
BZ	0.0007	0.0004	0.0006	0.000	0.004	2.612	9.97	326734.21	1647.20
BL	0.0018	0.0008	0.0016	0.001	0.008	1.844	5.23	108172.50	1640.83
BN	0.0032	0.0012	0.0029	0.001	0.020	2.389	15.77	15436397.10	934.25
BX	0.0072	0.0029	0.0066	0.002	0.029	1.780	5.37	108732.91	1510.55
					RSV-P				
TU	0.0007	0.0004	0.0006	0.000	0.005	3.609	20.25	70107.23	1422.80
FV	0.0016	0.0008	0.0014	0.000	0.010	2.697	13.36	31474.21	960.15
TY	0.0025	0.0012	0.0022	0.001	0.016	2.514	12.05	25865.67	943.94
US	0.0044	0.0018	0.0040	0.002	0.023	2.211	10.13	18515.70	843.37
BZ	0.0005	0.0003	0.0004	0.000	0.004	3.011	14.37	30452.68	1265.93
BL	0.0013	0.0006	0.0011	0.000	0.006	2.100	7.22	8762.03	1291.95
BN	0.0022	0.0009	0.0020	0.001	0.015	2.465	15.52	33265.93	727.08
BX	0.0050	0.0021	0.0046	0.002	0.024	1.891	6.39	6928.07	1306.90
					RSV-N				
TU	0.0007	0.0004	0.0006	0.000	0.005	3.093	14.49	37642.92	1555.67
FV	0.0016	0.0008	0.0014	0.000	0.013	2.862	17.88	53435.57	723.78
TY	0.0025	0.0012	0.0022	0.001	0.024	3.595	37.94	226074.95	475.96
US	0.0043	0.0017	0.0040	0.002	0.039	3.758	49.38	378238.91	421.14
BZ	0.0005	0.0003	0.0004	0.000	0.003	2.637	10.27	16715.77	1464.02
BL	0.0013	0.0006	0.0011	0.000	0.006	1.960	6.01	6454.57	1385.40
BN	0.0022	0.0009	0.0020	0.001	0.014	2.370	14.04	27569.95	843.11
BX	0.0050	0.0020	0.0046	0.001	0.022	1.795	5.93	6028.92	1250.21

Note: Time span of US and German Treasuries Realized measures are July 1,2003-November 30,2017 and October 1, 2005-November 30, 2017 respectively.

Conclusion

In the presented dissertation we show innovative approaches to risk modelling and empirical asset pricing. Specifically, we propose a method of obtaining more efficient estimates and forecasts of covariance matrices by generalizing already popular method (HAR); we demonstrate the possibility to identify common risk factors in the tails of panels of volatilities that drives the distribution of asset returns; and we formalize the basic quantile asset pricing equation and show its connection to factor asset pricing.

Chapter 1 introduces the Generalized Heterogeneous Autoregressive model intended for covariance matrix modelling and forecasting. We show that building a system of seemingly unrelated heterogenous autoregressions over Cholesky decomposed elements of realized covariance matrices delivers more accurate estimates and subsequent forecasts. Motivation to switch from Ordinary Least Squares to Generalized Least Squares estimation is the contemporaneous correlation in the residuals of the original HAR model. Using generalized least squares, we capture dependencies hidden in the residuals delivering more efficient estimates. In the empirical application, we study portfolios consisting of five, ten and fifteen stocks and we compare the performance of the GHAR against several benchmark models (HAR, Vector ARFIMA, Dynamic Conditional Correlation GARCH, RiskMetrics) during the period of the financial crisis. Results of our analysis suggest that GHAR provides more precise and more efficient covariance matrix forecasts and they translates to economic gains directly. Moreover, we study the economic benefit of estimating the realized covariance with more efficient multivariate realized kernel and sub-sampled realized covariance estimators.

Chapter 2 introduces the Panel Quantile Regression Model for Returns, an innovative approach of modelling commonalities in the quantiles of future returns using information from panels of realized measures and anticipated volatility. We build on the classical portfolio theory risk-return trade-off well documented in the literature. In the proposed approach, the panel of assets returns is modelled via ex-post and/or ex-ante volatility using panel quantile regression techniques. The penalized fixed effects estimator allows us to control the unobserved heterogeneity among financial assets and disentangle overall market risk into the systematic and idiosyncratic risks. In the empirical application, we study datasets containing the period of global financial crisis and compare the performance against several benchmark models (RiskMetrics, Univariate Quantile Regression Model for Returns, portfolio version of Univariate Quantile Regression Model for Returns). We document more accurate estimates of our model that also translates into better forecasting performance. Overall, Panel Quantile Regression Model for Returns is dynamically correctly specified, its forecasting performance is statistically better than those of benchmark models in the economically important quantiles and it provides us with direct economic gains according to Global Minimum Value-at-Risk Portfolio and the efficient frontiers of the Value-at-Risk Return trade-off evaluation criteria.

Chapter 3 formalizes the basic pricing equation in the quantile set-up and concentrates on the quantile pricing of bond futures contracts. We work with quantile preferences instead of expectations and study asset pricing of economic agents differing in their level of risk aversion. In particular, we extend the results of de Castro and Galvao (2018) who derive quantile Euler equation using properties of quantile preferences as defined in Manski (1988) and Rostek (2010). We show that the quantile Euler equation can be extended into a stochastic discount factor representation of the quantile asset pricing equation. Moreover, we present a link to the factor models. In the empirical application, we focus on quantile pricing of the two, five, ten and thirty years US and German government bond futures contracts from the Chicago Board of Trade and the EUREX exchanges using ex-post measures of asset uncertainty and US forward rates. Results of our analysis demonstrate a significant influence of the Realized Measures on the quantiles of the treasury returns. In particular, Realized Volatility plays an important role in all but the median quantiles of the return distribution, whereas Realized Semivariances also partly explain the median variation of two German Treasuries. In contrast, forward rates are of very limited use in quantile asset pricing.

References

Aielli, G. P. (2013, July). Dynamic conditional correlation: On properties and estimation. *Journal of Business & Economic Statistics 31*(3), 282–299.

Aït-Sahalia, Y., J. Fan, and D. Xiu (2010). High-frequency covariance estimates with noisy and asynchronous financial data. *Journal of the American Statistical Association 105*(492), 1504–1517.

Amendola, A. and G. Storti (2015). Model uncertainty and forecast combination in high-dimensional multivariate volatility prediction. *Journal of Forecasting 34*(2), 83–91.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and H. Ebens (2001). The distribution of realized stock return volatility. *Journal of Financial Economics 61*(1), 43–76.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2003). Modeling and forecasting realized volatility. *Econometrica 71*(2), 579–625.

Andersen, T. G., T. Bollerslev, and X. Huang (2011). A reduced form framework for modeling volatility of speculative prices based on realized variation measures. *Journal of Econometrics 160*(1), 176–189.

Ando, T. and J. Bai (2018). Quantile co-movement in financial markets: A panel quantile model with unobserved heterogeneity. *Journal of the American Statistical Association* (just-accepted), 1–35.

Artzner, P., F. Delbaen, J.-M. Eber, and D. Heath (1999). Coherent measures of risk. *Mathematical Finance 9*(3), 203–228.

Bache, S. H., C. Dahl, and J. Tang Kristensen (2008). Determinants of birthweight outcomes: Quantile regression based on panel data. *CREATES research paper 20*, 2008.

Bannouh, K., D. v. Dijk, and M. Martens (2009, September). Range-based covariance estimation using high-frequency data: The realized co-range. *Journal of Financial Econometrics 7*(4), 341–372.

Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard (2011). Multivariate realised kernels: consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading. *Journal of Econometrics 162*(2), 149–169.

Barndorff-Nielsen, O. E., S. Kinnebrock, and N. Shephard (2010). Measuring downside risk-realised semivariance. *Volatility and Time Series Econometrics: Essays in Honor of Robert Engle*.

Barndorff-Nielsen, O. E. and N. Shephard (2004a). Econometric analysis of realized covariation: High frequency based covariance, regression, and correlation in financial economics. *Econometrica 72*(3), 885–925.

Barndorff-Nielsen, O. E. and N. Shephard (2004b). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics 2*(1), 1–37.

Barndorff-Nielsen, O. E. and N. Shephard (2006). Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics 4*(1), 1–30.

Bassett, G. W., R. Koenker, and G. Kordas (2004). Pessimistic portfolio allocation and choquet expected utility. *Journal of Financial Econometrics 2*(4), 477–492.

Bauer, G. H. and K. Vorkink (2011). Forecasting multivariate realized stock market volatility. *Journal of Econometrics 160*(1), 93–101.

Baur, D. G., T. Dimpfl, and R. C. Jung (2012). Stock return autocorrelations revisited: A quantile regression approach. *Journal of Empirical Finance 19*(2), 254–265.

Bauwens, L., C. M. Hafner, and S. Laurent (2012). *Handbook of volatility models and their applications*, Volume 3. John Wiley & Sons.

Bauwens, L., S. Laurent, and J. V. Rombouts (2006). Multivariate GARCH models: a survey. *Journal of Applied Econometrics 21*(1), 79–109.

Berkowitz, J., P. Christoffersen, and D. Pelletier (2011). Evaluating value-at-risk models with desk-level data. *Management Science 57*(12), 2213–2227.

Billger, S. M. and C. Lamarche (2015). A panel data quantile regression analysis of the immigrant earnings distribution in the united kingdom and united states. *Empirical Economics 49*(2), 705–750.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics 31*(3), 307–327.

Bollerslev, T. (1990). Modelling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH model. *Review of Economics and Statistics 72*, 498–505.

Bollerslev, T., B. Hood, J. Huss, and L. H. Pedersen (2018). Risk everywhere: Modeling and managing volatility. *The Review of Financial Studies 31*(7), 2729–2773.

Bollerslev, T., S. Z. Li, and B. Zhao (2017). Good volatility, bad volatility, and the cross section of stock returns. *Journal of Financial and Quantitative Analysis*, 1–57.

Bonato, M. (2009). Estimating the degrees of freedom of the realized volatility wishart autoregressive model. *Available at SSRN 1357044*.

Bonato, M., M. Caporin, and A. Ranaldo (2013). Risk spillovers in international equity portfolios. *Journal of Empirical Finance 24*, 121–137.

Boudt, K., J. Cornelissen, and C. Croux (2012, November). Jump robust daily covariance estimation by disentangling variance and correlation components. *Computational Statistics & Data Analysis 56*(11), 2993–3005.

Canay, I. A. (2011). A simple approach to quantile regression for panel data. *The Econometrics Journal 14*(3), 368–386.

Caporin, M. and M. McAleer (2013). Ten things you should know about the dynamic conditional correlation representation. *Econometrics 1*(1), 115–126.

Cappiello, L., B. Gérard, A. Kadareja, and S. Manganelli (2014). Measuring comovements by regression quantiles. *Journal of Financial Econometrics 12*(4), 645–678.

Čech, F. and J. Baruník (2017). On the modelling and forecasting of multivariate realized volatility: generalized heterogeneous autoregressive (ghar) model. *Journal of Forecasting 36*(2), 181–206.

Chambers, C. P. (2007). Ordinal aggregation and quantiles. *Journal of Economic Theory 137*(1), 416–431.

Chiriac, R. and V. Voev (2011). Modelling and forecasting multivariate realized volatility. *Journal of Applied Econometrics 26*(6), 922–947.

Clements, M. P., A. B. Galvão, and J. H. Kim (2008). Quantile forecasts of daily exchange rate returns from forecasts of realized volatility. *Journal of Empirical Finance 15*(4), 729–750.

Cochrane, J. H. (2009). *Asset Pricing:(Revised Edition)*. Princeton university press.

Cochrane, J. H. and M. Piazzesi (2005). Bond risk premia. *American Economic Review 95*(1), 138–160.

Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics 7*(2), 174–196.

Corsi, F., S. Mittnik, C. Pigorsch, and U. Pigorsch (2008). The volatility of realized volatility. *Econometric Reviews 27*(1-3), 46–78.

Covas, F. B., B. Rump, and E. Zakrajšek (2014). Stress-testing us bank holding companies: A dynamic panel quantile regression approach. *International Journal of Forecasting 30*(3), 691–713.

Dahl, C. M., D. Le Maire, and J. R. Munch (2013). Wage dispersion and decentralization of wage bargaining. *Journal of Labor Economics 31*(3), 501–533.

Damette, O. and P. Delacote (2012). On the economic factors of deforestation: What can we learn from quantile analysis? *Economic Modelling 29*(6), 2427–2434.

Daníelsson, J., B. N. Jorgensen, G. Samorodnitsky, M. Sarma, and C. G. de Vries (2013). Fat tails, var and subadditivity. *Journal of Econometrics 172*(2), 283–291.

de Castro, L., A. F. Galvao, D. M. Kaplan, and X. Liu (2018). Smoothed GMM for quantile models. Working paper, available at `https://faculty.missouri.edu/~kaplandm`.

de Castro, L. I. and A. F. Galvao (2018). Dynamic quantile models of rational behavior. SSRN Scholarly Paper ID 2932382.

Diebold, F. X. and R. S. Mariano (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics 13*, 253–263.

Dufrenot, G., V. Mignon, and C. Tsangarides (2010). The trade-growth nexus in the developing countries: A quantile regression approach. *Review of World Economics 146*(4), 731–761.

Elst, V. H. and D. Veredas (2015, January). Smoothing it out: Empirical and simulation results for disentangled realized covariances. SSRN Scholarly Paper ID 2490335, Social Science Research Network, Rochester, NY.

Engle, R. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics 20*(3), 339–350.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, 987–1007.

Engle, R. F. and K. F. Kroner (1995). Multivariate simultaneous generalized ARCH. *Econometric Theory 11*(01), 122–150.

Engle, R. F. and S. Manganelli (2004). Caviar: Conditional autoregressive value at risk by regression quantiles. *Journal of Business & Economic Statistics 22*(4), 367–381.

Engle, R. F. and K. Sheppard (2001). Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH. Technical report, National Bureau of Economic Research, NBER Working Paper No. 8554.

Epps, T. W. (1979). Comovements in stock prices in the very short run. *Journal of the American Statistical Association 74*(366), 291–296.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance 25*(2), 383–417.

Fama, E. F. and R. R. Bliss (1987). The information in long-maturity forward rates. *American Economic Review 77*(4), 680–692.

Fama, E. F. and K. R. French (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics 33*(1), 3–56.

Feng, G., S. Giglio, and D. Xiu (2019). Taming the factor zoo: A test of new factors. *Fama-Miller Working Paper*, 17–04.

Foster-McGregor, N., A. Isaksson, and F. Kaulich (2014). Importing, exporting and performance in sub-saharan african manufacturing firms. *Review of World Economics 150*(2), 309–336.

French, K. R., G. W. Schwert, and R. F. Stambaugh (1987). Expected stock returns and volatility. *Journal of Financial Economics 19*(1), 3–29.

Galvao, A. and K. Kato (2015). Smoothed quantile regression for panel data. *Available at SSRN 2606188*.

Galvao, A. F. (2011). Quantile regression for dynamic panel data with fixed effects. *Journal of Econometrics 164*(1), 142–157.

Galvao, A. F., T. Juhl, G. Montes-Rojas, and J. Olmo (2017). Testing slope homogeneity in quantile regression panel data with an application to the cross-section of stock returns. *Journal of Financial Econometrics 16*(2), 211–243.

Galvao, A. F. and G. Montes-Rojas (2015). On bootstrap inference for quantile regression panel data: A monte carlo study. *Econometrics 3*(3), 654–666.

Galvao, A. F. and G. V. Montes-Rojas (2010). Penalized quantile regression for dynamic panel data. *Journal of Statistical Planning and Inference 140*(11), 3476–3497.

Galvao, A. F. and L. Wang (2015). Efficient minimum distance estimator for quantile regression fixed effects panel data. *Journal of Multivariate Analysis 133*, 1–26.

Giacomini, R. and I. Komunjer (2005). Evaluation and combination of conditional quantile forecasts. *Journal of Business & Economic Statistics 23*(4), 416–431.

Giovannetti, B. C. (2013). Asset pricing under quantile utility maximization. *Review of Financial Economics 22*(4), 169–179.

Gonzalo, J., J. J. Dolado Lobregad, L. Chen, et al. (2017). Quantile factor models. Technical report, Working paper - Universidad Carlos III de Madrid. Departamento de Economía.

Gouriéroux, C., J. Jasiak, and R. Sufana (2009). The wishart autoregressive process of multivariate stochastic volatility. *Journal of Econometrics 150*(2), 167–181.

Graham, B. S., J. Hahn, A. Poirier, and J. L. Powell (2015, March). Quantile regression with panel data. Working Paper 21034, National Bureau of Economic Research.

Gürkaynak, R. S., B. Sack, and J. H. Wright (2007). The us treasury yield curve: 1961 to the present. *Journal of Monetary Economics 54*(8), 2291–2304.

Hansen, P. R., A. Lunde, and J. M. Nason (2011). The model confidence set. *Econometrica 79*(2), 453–497.

Harding, M. and C. Lamarche (2009). A quantile regression approach for estimating panel data models using instrumental variables. *Economics Letters 104*(3), 133–135.

Harding, M. and C. Lamarche (2014). Estimating and testing a quantile regression model with interactive effects. *Journal of Econometrics 178*(1), 101–113.

Harvey, C. R., Y. Liu, and H. Zhu (2016). ... and the cross-section of expected returns. *Review of Financial Studies 29*(1), 5–68.

Harvey, D., S. Leybourne, and P. Newbold (1997). Testing the equality of prediction mean squared errors. *International Journal of forecasting 13*(2), 281–291.

Hautsch, N., L. M. Kyj, and R. C. A. Oomen (2012, June). A blocking and regularization approach to high-dimensional realized covariance estimation. *Journal of Applied Econometrics 27*(4), 625–645.

Hayashi, T. and N. Yoshida (2005). On covariance estimation of non-synchronously observed diffusion processes. *Bernoulli 11*(2), 359–379.

Jin, X. and J. M. Maheu (2013). Modeling realized covariances and returns. *Journal of Financial Econometrics 11*(2), 335–369.

Jorion, P. (2007). *Value at risk: the new benchmark for controlling market risk 3rd ed.* McGraw Hill Professional.

Kato, K., A. F. Galvao, and G. V. Montes-Rojas (2012). Asymptotics for panel quantile regression models with individual effects. *Journal of Econometrics 170*(1), 76–91.

Kempf, A. and C. Memmel (2006). Estimating the global minimum variance portfolio. *Schmalenbach Business Review 58*(4).

Klomp, J. and J. De Haan (2012). Banking risk and regulation: Does one size fit all? *Journal of Banking & Finance 36*(12), 3197–3212.

Koenker, R. (2004). Quantile regression for longitudinal data. *Journal of Multivariate Analysis 91*(1), 74–89.

Koenker, R. and G. Bassett Jr (1978). Regression quantiles. *Econometrica 46*(1), 33–50.

Kyj, L. M., B. Ostdiek, and K. Ensor (2010). Covariance estimation in dynamic portfolio optimization: a realized single factor model. In *American Finance Association 2010 Atlanta Meetings Paper*.

Lamarche, C. (2008). Private school vouchers and student achievement: A fixed effects quantile regression evaluation. *Labour Economics 15*(4), 575–590.

Lamarche, C. (2010). Robust penalized quantile regression estimation for panel data. *Journal of Econometrics 157*(2), 396–408.

Lamarche, C. (2011). Measuring the incentives to learn in colombia using new quantile regression approaches. *Journal of Development Economics 96*(2), 278–288.

Lee, J.-S., G.-L. Huang, C.-T. Kuo, and L.-C. Lee (2012). The momentum effect on chinese real estate stocks: Evidence from firm performance levels. *Economic Modelling 29*(6), 2392–2406.

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics*, 13–37.

Longerstaey, J. and M. Spencer (1996). RiskMetricsTM - technical document. *Morgan Guaranty Trust Company of New York: New York*.

Manski, C. F. (1988). Ordinal utility models of decision making under uncertainty. *Theory and Decision 25*(1), 79–104.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance 7*(1), 77–91.

Matano, A. and P. Naticchioni (2011). Wage distribution and the spatial sorting of workers. *Journal of Economic Geography*, lbr013.

Merton, R. C. (1973). An intertemporal capital asset pricing model. *Econometrica 41*(5), 867–887.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica 59*(2), 347–370.

Patton, A., T. Bollerslev, and R. Quaedvlieg (2017). Realized semicovariances: Looking for signs of direction inside the covariance matrix. *Economic Research Initiatives at Duke (ERID) Working Paper*.

Patton, A. J. and K. Sheppard (2015). Good volatility, bad volatility: Signed jumps and the persistence of volatility. *Review of Economics and Statistics 97*(3), 683–697.

Powell, D. and J. Wagner (2014). The exporter productivity premium along the productivity distribution: evidence from quantile regression with nonadditive firm fixed effects. *Review of World Economics 150*(4), 763–785.

Rabemananjara, R. and J.-M. Zakoian (1993). Threshold arch models and asymmetries in volatility. *Journal of Applied Econometrics 8*(1), 31–49.

Ross, S. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory 13*(3), 341–360.

Rostek, M. (2010). Quantile maximization in decision theory. *Review of Economic Studies 77*(1), 339–371.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance 19*(3), 425–442.

Soroka, S. N. (2006). Good news and bad news: Asymmetric responses to economic information. *The journal of Politics 68*(2), 372–385.

Toomet, O. (2011). Learn english, not the local language! ethnic russians in the baltic states. *American Economic Review 101*(3), 526–531.

Voev, V. and A. Lunde (2007). Integrated covariance estimation using high-frequency data in the presence of noise. *Journal of Financial Econometrics 5*(1), 68–104.

White, H., T.-H. Kim, and S. Manganelli (2015). Var for var: Measuring tail dependence using multivariate regression quantiles. *Journal of Econometrics 187*(1), 169–188.

You, W.-H., H.-M. Zhu, K. Yu, and C. Peng (2015). Democracy, financial openness, and global carbon dioxide emissions: Heterogeneity across existing emission levels. *World Development 66*, 189–207.

Zellner, A. (1962). An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of the American Statistical Association 57*(298), 348–368.

Zhang, L. (2011, January). Estimating covariation: Epps effect, microstructure noise. *Journal of Econometrics 160*(1), 33–47.

Zhang, L., P. A. Mykland, and Y. Ait-Sahalia (2005). A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association 100*(472), 1394–1411.

Zhang, Y.-J., H.-R. Peng, Z. Liu, and W. Tan (2015). Direct energy rebound effect for road passenger transport in china: A dynamic panel quantile regression approach. *Energy Policy 87*, 303–313.

Žikeš, F. and J. Baruník (2016). Semi-parametric conditional quantile models for financial returns and realized volatility. *Journal of Financial Econometrics 14*(1), 185–226.

Response to referees reports

I am grateful to all opponents for the discussion, valuable suggestions and all points raised. The opponents' reports helped to improve dissertation significantly. The response to individual comments is reported below.

prof. Fredj Jawadi

Q: *The introduction to the thesis needs to be improved to clarify the motivation behind the study. Why should we focus on these topics related to Risk Modelling and Empirical Asset Pricing?*
A: Thank you for this general comment. In the final version of dissertation, we have improved the introduction by more detailed description of our motivation and main findings of each paper.

Q: *The literature review in all three studies is rather short and does not make the contribution of these new studies to the related literature clear enough.*
A: We have tried to balanced the length of the literature review sections according to the comments/suggestions we have received from the referees in the peer-review process.

Q: *In Chapter 1, given that the author uses HFD, why was it not possible to augment the GHAR with jump proxies and continuous volatility?*
A: GHAR is a general way of covariance matrix modelling and can accommodate any proxy of the covariance matrix. In our work, we consider in literature already well established methods such as Realized Covariance (Barndorff-Nielsen and Shephard, 2004a), Sub-Sampled Realized Covariance (Zhang et al., 2005) and Multivariate Realized Kernels Covariance estimator (Barndorff-Nielsen et al., 2011). Certainly, covariance matrix estimates obtained from Realized Co-Range (Bannouh et al., 2009), Two Scale Realized Covariance (Zhang, 2011), jumps robust covariance estimator (Boudt et al., 2012), block-wise Multivariate Realized Kernels (Hautsch et al., 2012) and many other can be subsequently modeled by the GHAR, however, it is not the purpose of the paper to do the horse-race between various measures of realized covariation.

Q: *In Chapter 1, is it possible to generalize the approach even for non-synchronous data?*

A: Yes - if we use estimator of the realized covariation that can handle non-synchronous data, we can model this estimates by GHAR.

Q: *Why does the volatility proxy or estimate differ in Chapter 1 and Chapter 2?*
A: The first chapter focuses on the multivariate volatility modelling and adopt various estimates of the realized covariation, while the second chapter adopts univariate volatility measures. Since the development of new measures of realized covariation is challenging, not all univariate volatility measures have multivariate counterparts, e.g. Realized Semivariances were introduced in Barndorff-Nielsen et al. (2010) while Realized Semicovariances (Patton et al., 2017) are still subject to research.

Q: *In Chapter 2, I would prefer the Harvey et al. (1997) test to the Diebold and Mariano (1995) test to evaluate the forecasting performance of the estimated model.*
A: According to Harvey et al. (1997) modified version of Diebold–Mariano statistic is of following form:

$$S_1^* = \left[\frac{n + 1 - 2h + n^{-1}h(h-1)}{n}\right]^{1/2} S_t,$$

where S_t is original Diebold–Mariano statistic, h is the forecasting horizon and n is the length of the evaluated sample. In our application we consider one–step–ahead forecast, i.e. $h = 1$, and out–off-sample length is 1613 days, i.e. $n = 1613$. Hence modified test statistic is $S_1^* = 0.9997 * S_t$, which does not change results of our analysis.

Q: *In Chapter 2, which test was applied to use fixed effect assumption?*
A: Thank you for this excellent question. Currently, we are not aware of any reliable test for fixed effects that would be directly applicable in the panel quantile regression set-up. Panel quantile regression techniques are relatively new in the literature and are still subject to research.

Q: *In Chapter 3, I would like to see the results of the normality test, the ARCH test in Table 3.7, etc.*
A: We have added results of the tests to the tables.

Q: *Page 14. What do you mean by Economic Criteria?*
A: We evaluate covariance forecasts in terms of economic gains. In contrast to statistical evaluation where numbers of loss functions can be used to study unbiasedness of the forecasts, in the economic evaluation we directly compare profitability of the studied models. In our work, we consider Markowitz-like optimization where we study risk–return trade-off by plotting efficient frontiers; and Global Minimum Variance Portfolio, where we set assets weights in a way that total risk of the portfolio is minimized.

Q: *The thesis includes a short general introduction and while it provides a concise introduction to the three essays, it does not contain a conclusion. The conclusion always helps to give an overview of the different essays, recalls the assumption of the results and provides some criticisms, limitations and suggests potential future extensions.*

A: We have added short conclusion summarizing our main findings to the text.

Minor Comments:

Q: *Page 104, the reference of De Castro and Galvao appears incomplete. Needs checking.*
A: Corrected.

Q: *Page 6, page 40. Replace "paper" by "chapter". The same remark applies for all chapters.*
A: Corrected.

Q: *Page 88, check the text for typos.*
A: Corrected.

Q: *Page 93. You need to adjust the size of the graphs.*
A: Corrected.

prof. Evžen Kočenda

Q: *In order to obtain parameter estimates, optimization problem (2.2) is solved. Part of the optimization problem is the penalty term lambda that influences precision of the estimates alpha and beta. It is found that the choice of the lambda does not affect precision of beta estimates. The wording suggests that the choice of the lambda affects precision of alpha estimates. Would the precision of the alpha estimates change if the lambda was not set arbitrarily (as is the case in the paper) but based on a theoretical approach of the lambda selection? From the formal point, the sentence structure could be polished at places, and some typos are remaining in the text.*
A: We have rewritten whole section 2.3 and added paragraphs about the theoretical selection of the penalty parameter λ.

Q: *Some references are incomplete.*
A: Corrected.

Q: *Abbreviations of the bonds of different length are provided in Table 1. These abbreviations are again used throughout the text and in Tables 2 and 3. Abbreviations in tables should be changed to year-length labels of bonds because when a reader gets to Table 2 and Table 3, abbreviations are already forgotten.*
A: Year-length information have been added to all tables and figures.

Q: *The author notes that the signs of the beta "coefficients are consistent with classical risk-return trade-off (e.g., Value-at-Risk)", meaning that there are "negative coefficients for quantiles below the median and positive coefficients for quantiles above the median." Empirical results then show a stronger influence in upper quantiles than in lower quantiles for the US bonds, but reverse result is shown for the German bonds. Theoretically, coefficients at median should be zero. Is this true? It might be (based on the Figures 1 and 2), but a more direct numerical assessment would be welcome given the asymmetry in empirical results between upper and lower quantiles. The asymmetries might also be more elaborated on. Results of the analysis where the Realized Semivariances are used bring even more asymmetries. One sort is due to the use of the Realized Semivariance instead of the Realized Variance. Another sort (dominance of results at different quantiles for different bonds) cannot be explained with the different variance/semivariance measure only. It would be nice to see some, at least tentative, economic reasons for these asymmetries. Are they due to the differences in trading rules, issuing structure, liquidity, quantity of bonds issued, etc.?*
A: Thank you for this discussion. The median coefficients should be zero (or statistically not different from zero) according to Efficient Market Hypothesis (Fama, 1970) which states that returns should be unpredictable since prices follow random-walk. The asymmetries in the tail behavior of the conditional return distribution when Realized Volatility serve as risk factor are partly due to differences in the estimation procedures. When we look at the Method of Moments estimates, there is higher absolute influence in the upper quantiles in all the maturities of both US and German Treasuries. In contrast, the quantile regression estimates indicate higher absolute influence in the lower quantiles

in all the maturities of both US and German Treasuries except 2-years US bond. In case of multi-factor analysis with Realized Semivariances being risk factors, the asymmetries can be partly explained by the differences in Realize Measures and different risk assessment by the Realized Semivariance. We agree with prof. Kočenda that institutional set-up (trading rules, issuing structure, etc.) will be important source of asymmetry and we will concentrate on it more in our future research.

To improve the understanding of our results we have added following text to dissertation: *"Asymmetric influence documented in this multi-factor specification has many potential sources. Besides difference of the Realized Volatility and Realized Semivariance, we attribute it also to the institutional differences of the US and EU markets, e.g. different trading hours (US market 23 hours a day vs 14 hours EU market), investors' perception of US Treasury futures being well established investment instruments (history of 30-years US Tresury futures date back to October 1982 while 30-year German Treasury futures were introduced in September 2005), etc.. Last but not least, liquidity connected to different trading hours also plays a role."*

Q: *The author states that the theory suggests that lower frequency returns exhibit certain degree of predictability while this is not valid for higher frequency returns. It would be useful to cite the source(s) to better back the claim.*
A: We have added references to the text.

Q: *Finally, careful editing would benefit the final version of the essay as it contains some errors in sentences structure and typos.*
A: Corrected.

Dr. Michael Ellington

Comments on the Frontmatter of the Thesis

Q: *The abstract requires some rewording in order to enhance its readability. My recommended changes can be found on the paper copy you receive at the pre-defence.*
A: Corrected.

Q: *Following the Contents, please insert a List of Tables and List of Figures followed by the page number each are located.*
A: Corrected.

Q: *Introduction, page 2, paragraph 2, sentence 5. Please reword, ending the sentence with in further work does not read well.*
A: Corrected.

Q: *The author might also consider finishing the thesis with a "Conclusion" that briefly summarises each paper followed by two or three paragraphs discussing directions for future research. In my opinion, having the frontmatter and an introduction with no concluding section is not logical.*
A: We have added short conclusion summarizing our main findings to the text.

Comments for: On the Modelling and Forecasting Multivariate Realized Volatility: Generalized Heterogeneous Autoregressive (GHAR) Model.

Q: *Section 1.5, page 16 concisely explains how the global minimum variance portfolio is obtained as cumulative measures. However, there lacks detail on how the annualised versions are obtained. Please clarify, and state the sample used as it is unclear if they are for the whole out of sample period, or not.*
A: Corrected. Explanation of procedure added in text: "For annualized GMVP calculation we use annualized realized covariance of the whole out–off–sample period calculated as $RCOV_{annualized} = \frac{\sum_{i=1}^{T} RCOV_i}{\frac{T}{250}}$."

Q: *How are the assets chosen for the 5 and 10 asset portfolios? Given the 15 stocks, there are 10 combinations of 5 asset portfolios and 45 combinations of 10 asset portfolios. I realise that 45 combinations of the 10 asset portfolios is inefficient, but examining the forecasting performance of all possible combinations of 5 asset portfolios might be worth considering and reporting within the appendix.*
A: We have chosen assets included in portfolios according to market capitalization. Details of procedure are following:

- we dived assets into three groups according to their market capitalization, e.g. first group contains assets with the highest market capitalization
- the portfolio consisting of five stocks contains assets from the first group
- the portfolio consisting of ten stocks contains assets from the first and second group

- the portfolio consisting of fifteen stocks contains all assets in our dataset.

Since there is $\binom{15}{5} = \frac{15!}{(15-5)!5!} = 3003$ and $\binom{15}{10} = \frac{15!}{(15-10)!10!} = 3003$ possible asset combinations to create five and ten assets portfolios respectively we stick to simple "expanding" portfolio as described before.

Q: *With regards the model confidence set, is $\alpha = 0.95$ or $\alpha = 0.05$? This needs to be clarified. I would also like to see a justification for using this to promote the benefits of the GHAR model. In Table 1.3, the HAR model delivers a lower RMSE across MRK, RCOV, and Sub-Sampled RCOV. This is even more apparent for the 10 and 15 asset portfolios in Table 1.6 and 1.7; where in the latter, the GHAR model only enters the 95% model confidence set for RCOV 5min. To me this suggests that the HAR model delivers a better forecast than all models, and therefore this is the model that should be used; at least for your data and sample. What happens if you examine the 99% model confidence set, does the GHAR model remain in the set of the best forecasting models for 1 step ahead forecasts?*
A: We set *alpha* to 0.05 thus in notation of Hansen et al. (2011) we have 95% MCS, i.e. $\widehat{\mathcal{M}}^*_{1-\alpha} = \widehat{\mathcal{M}}^*_{95\%}$. However, to make our results directly comparable with Chiriac and Voev (2011) we stick to their notation and report our results as 100*α MCS, i.e. 5% MCS. Since the Model Confidence Set is analogous to a confidence interval for a parameter estimates, all models included in the 95% MSC will be included also in the 99% MCS - Table 7, p. 490 in Hansen et al. (2011) shows that all models in 75% MSC are also in 90% MCS.

Q: *When looking at forecasting a longer time horizon, all models are in the model confidence set for a portfolio of 5 stocks. To me, this test only allows you to drop certain forecasting models, whilst being unable to distinguish which model should be used. I realise that the differences in RMSEs are negligible, but with vast sums of wealth at stake, practitioners would like to minimise losses.*
A: Practitioners are mainly involved in applications that invole inverse of the covariance matrix forecasts. Since the RMSE is a statistical loss function that measures deviation of the forecast (not inverse of forecast) from the proxy, it does not tells us how the inverse of covariance matrix will look like and whether it exists. On the other hand, in the economic evaluation we work with inverse of covariance matrices and we get direct comparison of economic benefits. Disadvantage of statistical evaluation in high-dimensional problems is summarized in Bauwens et al. (2012) - "Virtually, all evaluation of covariance forecasts in high dimensional problems utilize "economic" loss functions, as opposed to statistical loss functions. The obvious rational for this choice is that the goal of high dimensional forecasts is to provide improvements over simple estimators such as the rolling window covariance estimator. A less obvious reason is that unbiased forecasts are often undesirable since other considerations, namely that the forecast is well conditioned and invertible, are more important. Moreover, unbiasedness of the covariance does not translate into unbiasedness of the inverse".

Q: *All Figures and Tables should be self contained. It is difficult to follow what the Figures and Tables are showing without detailed notes to accompany them.*
A: Corrected.

Comments for: Measurement of Common Risk Factors: A Panel Quantile Regression Model for Returns.
Q: *On page 44, there is a discussion of the penalty term λ within equation 2.2 reported on page 43. As stated, you arbitrarily pick this term and have conducted some robustness analysis regarding this. You also mention that there is a theoretical approach to selecting the value of this parameter, please clarify this method. Furthermore have you also tried this approach of selecting λ outlined in Galvao and Montes-Rojas (2010)?*
A: We have rewritten whole section 2.3 and added paragraphs about the theoretical selection of the penalty parameter λ.

Q: *There are typos in equations 2.3, 2.4, and 2.5. All the realised volatility measures are missing the i subscripts.*
A: Corrected.

Q: *Page 45, you refer to equation 3.28, yet there is no equation 3.28 in this paper. Please refer to correct equation.*
A: Corrected.

Q: *On page 46, you state you decided to use a Value at Risk (VaR) framework because forecasts we obtain from the Panel Quantile Regression Model for Returns are by definition semi-parametric VaRs. To the unfamiliar reader it is unclear that this is the case. Therefore you should explicitly state why they are semi-parametric VaRs in a footnote, or within the main text.*
A: We have added footnote wit following text *"According to Jorion (2007) p.17 "Value–at–Risk describes the quantile of the projected distribution of gains an losses over the targeted horizon." Since the VaR is a quantile of returns, and we model quantiles of returns directly by panel quantile regression, we therefore obtain semi-parametric VaR estimates."*

Q: *In Section 2.5 you conduct a simulation study to examine the in and out of sample fit of the model under a variety of different error distributions. Could you please clarify the number of stocks you include within the simulation study. Presumably it is 29 to maintain consistency with your empirical application.*
A: Corrected - we have added to main text that we are simulating 29 time series.

Q: *Page 53, paragraph 1, final sentence you state that multivariate random numbers are more homogeneous than the univariate case. I realise that you impose a covariance structure which induces dependence, but does this really translate into an increase in the degree of homogeneity? This seems to conflict the final paragraph on page 53, where you state that PQR outperform UQR in more heterogeneous data created by univariate error*

distributions. Currently these paragrahs are slightly confusing to the reader. Please clarify what you mean, and reword accordingly.
A: In our approach, we apply panel quantile regression to capture unobserved heterogeneity among the data. If there is no/limited heterogeniety in the data present, benefits of using PQR will be limited and the univariate quantile regression should provide better fit. When we simulate 29 time series using univariate error distribution, for all the series errors will be independent. In contrast, when we simulate data from the multivariate error distribution with the given covariance structure we impose certain degree of dependence in the data. Therefore, the data simulated from univariate error distribution might show higher degree of unobserved heterogeneity as we document in the Monte-Carlo experiment.

Q: *Why do you use only 29 stocks based on market capitalisation and liquidity? Can you please clarify the criteria adopted that filters out stocks based on market capitalisation and liquidity? I realise that you will need sufficient liquidity in order to adequately fit the model, but does this not mean that the proposed model will not work well for stocks with lower liquidity levels?*
A: We have chosen 29 stocks with the highest market capitalization at the NYSE in our original application that have full day of continuous trading without periods of trade interruption. As a robustness check we have added new section that study almost all constituents of the S&P 500 index, i.e. we have analyzed 496 assets.

Q: *Page 58, paragraph 1. This does not read well. Please reword.*
A: Corrected.

Q: *On page 59, you relate the findings of your in-sample analysis to new asymmetry. I would like to see a discussion of what the implications of this news asymmetry is for investors. I would also like to see your results linked to those within the literature. Obviously this phenomenon is not new so it should be trivial to find papers that relate well with your findings.*
A: Thank you for the comment. Actually, we did not find new asymmetry, we just document the validity of the so called "bad news effect". As stated by Soroka (2006) "There is a growing body of work suggesting that responses to positive and negative information are asymmetric *that negative information has a much greater impact on individuals' attitudes than does positive information ...*". In our work, positive semivariance is formed by the positive news, i.e. positive returns, and the negative semivariance is formed by the negative news, i.e. negative returns. Since our results show higher influence of the negative semivariance on the quantiles of future returns we document positive/negative news asymmetry. The same logic was previously applied in building asymmetric GARCH models, e.g. TGARCH(Rabemananjara and Zakoian, 1993), EGARCH (Nelson, 1991) etc..

Q: *In Figures 2.1, 2.2, and 2.3 there are* ○ *on each of the plots that are not defined. What are these? they should be noted in the legend and the notes.*

A: Circles, $\circ$, represents outliers in the boxplots.

Q: *Page 60, paragraph 2, sentence 2: you describe unconditional coverage of 89.9% might be better described as perfect fit for PQR-RSV. No model is perfect, please remove this. I understand what you are saying, but following sentence stressing that you cannot reject the null hypothesis of correct unconditional coverage is sufficient.*
A: Corrected.

Q: *When you forecast using the PQR model, are you assuming that the stock will remain in is current quantile? presumably this is case since when fitting the model* $\alpha(\tau)$ *and* $\beta(\tau)$ *are used to generate the forecast.*
A: We do not explicitly assume that the stock will remain in a certain quantile, we forecast conditonal return distribution and we condition these forecasts on the past volatility.

Q: *Section 2.6.3, paragraph 3, sentence 2, reword to Except for the median, the PQR-RV model performs best in all quantiles.*
A: Corrected.

Q: *For the Value at Risk application. As stated in the text the %VaR are the forecasts of the asset returns in a given quantile. Is the portfolio VaR for a given quantile therefore made up of assets that belong only to this quantile at the given period or all 29 assets? If so, is the covariance matrix therefore just the realised covariances of these assets on a given day?*
A: In this application, portfolio is always composed of the 29 assets, and for each day we use Relaized Covariance estimates.

Q: *As stated in section 2.6.3 the Mean-Variance analysis is carried out using annualised returns and variance. It is not clear what sample is being used to span Figure 2.4. Please clarify exactly how this optimisation is computed. For instance, is the model estimated using the full sample (i.e. July 1, 2005 to December 31, 2015? which means the efficient frontier is calculated at December 31 2015 and then annualised?).*
A: We have added details of annualised returns and VaR calculation to the text. The sample used in the comparison is whole out-of-sample period.

Q: *It might be worth reporting, within the appendix, the efficient frontiers calculated at the following dates:*

1. *December 31, 2006*
2. *December 31, 2008*
3. *December 31, 2010*

These periods correspond to boom, bust and recovery periods. In each case my recommendation would be to estimate the model using all data available (i.e. July 1, 2005- end date).

A: Unfortunately, the suggested estimation procedure will not work in our set-up as we are using rolling window procedure and always make one-step-ahead forecasts, these forecasts are further annualized and used as input for efficient frontier computation.

Q: *Should the global minimum VaR presented in Table 2.5 correspond to the global minimum VaR figures presented in Figure 2.4? If not, why not?*
A: In our set-up, no. Since we do not allow short-selling in the Markowitz like optimization, the results of GMVaRP and efficient frontiers are not going to be exactly the same.

Q: *Within section 2.6.3, I would like to see a discussion of the implications these results have for investors from an economic perspective. It seems to me that the major benefits are realised within this section.*
A: The direct implication of our results is the better Value-at-Risk - return trade-off, hence, the economic agent facing restrictions/limits on the riskiness of her position might be better off using our approach. The other implications (e.g. building a trading strategy) depend on the context of application of our general approach and are beyond the scope of the presented paper. For example, the trading strategy will require incorporation of the transaction costs, limits on the risky position, short-selling restriction and many other consideration not connected to our original intention of modelling tail risk in panels of returns via volatility.

Q: *Table 2.6, page 66. Please provide notes for the table of descriptive statistics. What sample is used to compute these statistics?*
A: Corrected.

Comments for: Dynamic Quantile Model for Bond Pricing.
Q: *In equations 3.14-3.23, Ω_t is not defined. While I realise this is the information set, it should be defined prior to paragraph 3 sentence 1 on page 80 where you discuss conditioning only on returns. I would also consider using F_t to denote the information set so as not to confuse readers with $\overline{\Omega}$.*
A: We decided to stick to Ω_t notation for information set because we use $\mathcal{F}$ to denote general random variable. We have defined Ω_t right after equation 3.15.

Q: *Have you considered models that are over-identified? I realise that you state this is computationally demanding, but in my opinion this is not an excuse and I imagine a referee for a journal with impact factor would ask for this.*
A: We have considered also over-identified models with varying number of lagged Realized Volatility serving as instruments and we obtained qualitatively and also almost quantitatively identical results, therefore we are not presenting it explicitly in the text.

Q: *The results of your single factor models using the Method of Moments Estimator (MM) and standard Quantile Regression (QR) across both bond markets are qualitatively similar. Given that you then move on to multi-factor models estimated only using conventional QR methods, I think the MM estimator of de Castro et al. (2018) can relegated to the*

appendix, as well as the discussion of the GMM estimator in detail. To me, it is important to be consistent with the estimation method of your single and multi factor models. The MM estimator of your single factor models, as well as an investigation of over-identified models can be provided in the appendix as a robustness exercise. To me, the contribution of your paper is not the estimation method, but the modification of the theoretical results of de Castro and Galvao (2018) to an asset pricing framework.
A: Thank you for this comment. While we agree that our contribution is the modification of the theoretical results of de Castro and Galvao (2018) it is also the empirical application using proper estimation method. As it is common in the Classical Asset Pricing to work within GMM framework because it offers you great flexibility, we want to show that currently also quantile version of the estimator is available, although it has some limits. Therefore, we show that using "proper" (GMM) method we obtain results that are very similar to less complicated standard quantile regression framework.

Q: *Page 88, you state that the median anomalies require future research since they are in violation of the Efficient Market Hypothesis. Can you please clarify how and why this is the case? What are the implications for this violation?*
A: Efficient Market Hypothesis states that returns should be unpredictable since prices follow random-walk. In our case, however, we document that for certain assets median coefficient estimates are statistically different from zero which might translates to theoretical predictability of median returns.

Q: *Following up from the above comment, It is not satisfactory to state that this requires further research without specifying exactly what could/should be done. I think a useful extension to this project would be a to devise a trading strategy that takes advantage of this violation. If this is indeed correct, a footnote stating how further research could be conducted should be included.*
A: While we agree that natural extension of this project might be a trading strategy utilizing advantages of quantile asset pricing, it is well beyond the scope of this paper to formulate it.

Q: *In Section 3.5.2 you present results from multi-factor models using the QR estimation method only. If you do not relegate the MM (and recommended GMM) results of single factor models to the Appendix, please justify why you are using only QR methods here.*
A: In the multi-factor model we rely on the quantile regression only since the implementation of multiple moment conditions in quantiles is still the open question in the literature.

Q: *Page 92, you state that comparing the total dispersion of RS+, RS- that US Treasuries seem to be more heterogeneous relative to German Treasuries. There is no economic reasoning behind why this might be the case. There should be a few sentences spent convincing the reader of possible justification for this. Does the literature find results supporting your findings?*
A: We have tried to identify possible source the heterogeneity and we have added explanation to the text. We will elaborate on this issue in our future research.

Q: *When you estimate models in the spirit of Cochrane and Piazzesi (2005), your results directly contradict their findings. You address this by stating it is the structure of the dataset. can you confirm this by converting your results to a monthly frequency and re-estimating the models to make sure?*
A: Our dataset converted to monthly frequency results in 174 months. We have re-estimated our multifactor model and find out that three-year and five-year forward rates might by useful in modeling condition bond returns distributions. Our results thus partly match findings of Cochrane and Piazzesi (2005).

Table 3.9: US treasuries - QR forward rates - monthly frequency

	0.05	0.1	0.25	0.5	0.75	0.9	0.95	0.05	0.1	0.25	0.5	0.75	0.9	0.95
				TU							FV			
const	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)
SVENF02	-0.45	-0.18	-0.44	-0.3	-0.06	-0.05	-0.06	-0.86	-1.15	-0.73	-0.7	-0.27	-1.02	-1.85
	(-0.83)	(-0.55)	(-0.67)	(-0.53)	(-0.45)	(-0.77)	(-1.03)	(-2.36)	(-1.85)	(-1.37)	(-1.28)	(-1.76)	(-3.24)	(-3.06)
SVENF03	1.57	0.53	**1.45**	**1.03**	0.29	0.13	0.35	**2.79**	3.48	**2.44**	**2.46**	1.17	**3.77**	**7.15**
	(-1.39)	(-1.41)	(0.11)	(0.24)	(-0.41)	(-2.63)	(-3.91)	(-0.03)	(-5.35)	(1.16)	(-0.48)	(-1.94)	(-1.83)	(-3.47)
SVENF04	-1.86	-0.57	-1.73	-1.28	-0.46	-0.17	-0.61	-2.77	-3.64	-2.88	-3.02	-1.75	-5.01	-9.76
	(-3.66)	(-2.44)	(-2.64)	(-2.28)	(-2.03)	(-4.16)	(-5.75)	(-5.54)	(-8.06)	(-6.06)	(-4.79)	(-7.43)	(-15.74)	(-17.83)
SVENF05	0.69	0.19	**0.71**	**0.56**	0.25	0.11	0.36	0.71	1.25	**1.18**	**1.29**	0.9	**2.3**	**4.48**
	(-0.66)	(-0.28)	(0.09)	(0.17)	(-0.21)	(-0.75)	(-1.74)	(-0.59)	(-2.57)	(0.28)	(0.01)	(-0.78)	(0.31)	(-1.52)
				TY							US			
const	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(-0.01)	(-0.01)	(0)	(0)	(0)	(0)
SVENF02	-0.64	0.03	-0.83	-1.16	0.16	-1.01	-2.73	**3.59**	-0.02	-0.28	-2.91	-1.58	-2.36	-2.9
	(-1.76)	(-1.67)	(-2.45)	(-2.21)	(-2.87)	(-4.46)	(-5.62)	(-1.56)	(-2.22)	(-2.31)	(-4.08)	(-4.79)	(-6.6)	(-6.23)
SVENF03	1.74	-0.63	**2.74**	**3.98**	-0.3	**3.86**	**10.21**	-13.93	-1.83	0.56	**10.1**	**6.8**	**9.53**	**11.76**
	(-5.43)	(-9.4)	(-0.65)	(-0.63)	(-2.28)	(-1.12)	(-3.23)	(-18.13)	(-19.95)	(-6.31)	(-2.17)	(-0.6)	(1.45)	(2.2)
SVENF04	-0.87	1.56	-3.06	-4.81	0.01	-5.62	-13.69	**20.34**	4.85	-0.05	-11.94	-9.68	-13.93	-16.91
	(-5.27)	(-6.91)	(-10.93)	(-10.63)	(-12.49)	(-20.88)	(-26.62)	(-2.63)	(-8.99)	(-8.97)	(-16.97)	(-18.91)	(-31.11)	(-32.29)
SVENF05	-0.4	-1.01	1.15	**2.02**	0.18	**2.85**	**6.23**	-10.07	-2.94	-0.23	**4.81**	**4.52**	**6.85**	**8.08**
	(-3.75)	(-3.76)	(-1.24)	(0.35)	(-0.85)	(1.36)	(-0.26)	(-12.13)	(-10.97)	(-3.62)	(-1.18)	(0.74)	(2.6)	(2.96)

Q: *I would like to see a 5 and a 6-factor model combining realised volatility with the four forward rates, and realised semi-variances with the four forward rates. Provided the coefficients for volatility are consistent with your single factor models, you would be able to state the following: At higher frequencies volatility drives bond prices after controlling for forward rates. Therefore, the implication for investors possibly re-balancing on a daily basis and optimising quantiles of their utility function is that volatility matters.*
A: Below we provide results of the 5 and 6 factor models. Overall we can see that results of our previous analysis where we estimate quantile regressions using realized measures only do not change much. There are some minor differences, but they do not change our previous conclusions.

Table 3.10: US treasuries - QR forward rates + Realized Volatility - 5 factor model

	0.05	0.1	0.25	0.5	0.75	0.9	0.95	0.05	0.1	0.25	0.5	0.75	0.9	0.95
	TU							FV						
const	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)
SVENF02	0.12	0.07	0.04	0.01	0	0.01	-0.04	0.41	0.26	0.03	0.15	0.16	0.2	-0.04
	(0.08)	(0.05)	(0.04)	(0.04)	(0.04)	(0.06)	(0.07)	(0.25)	(0.16)	(0.12)	(0.11)	(0.12)	(0.15)	(0.2)
SVENF03	-0.4	-0.22	-0.14	-0.07	-0.01	-0.06	0.11	-1.3	-0.74	-0.02	-0.56	-0.62	-0.75	0.02
	(0.26)	(0.16)	(0.13)	(0.12)	(0.13)	(0.19)	(0.22)	(0.85)	(0.54)	(0.41)	(0.38)	(0.42)	(0.5)	(0.69)
SVENF04	0.4	0.21	0.15	0.08	0.02	0.11	-0.1	1.44	0.74	-0.07	0.62	0.7	0.86	-0.02
	(0.31)	(0.19)	(0.15)	(0.15)	(0.15)	(0.23)	(0.27)	(1.02)	(0.65)	(0.49)	(0.46)	(0.5)	(0.6)	(0.83)
SVENF05	-0.14	-0.07	-0.06	-0.03	0	-0.04	0.05	-0.56	-0.27	0.04	-0.22	-0.23	-0.29	0.06
	(0.12)	(0.08)	(0.06)	(0.06)	(0.06)	(0.09)	(0.11)	(0.42)	(0.27)	(0.2)	(0.19)	(0.2)	(0.25)	(0.34)
RV	-1.47	-1.05	-0.44	0.11	0.69	1.2	1.32	-1.51	-1.08	-0.47	-0.03	0.46	0.93	1.17
	(0.13)	(0.11)	(0.06)	(0.08)	(0.08)	(0.1)	(0.11)	(0.16)	(0.1)	(0.07)	(0.08)	(0.08)	(0.09)	(0.08)

	0.05	0.1	0.25	0.5	0.75	0.9	0.95	0.05	0.1	0.25	0.5	0.75	0.9	0.95
	TY							US						
const	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)
SVENF02	0.63	0.23	-0.04	0.2	0.12	0.01	0.06	0.67	0.33	0.1	0.26	-0.25	-0.28	0.33
	(0.33)	(0.23)	(0.2)	(0.18)	(0.18)	(0.23)	(0.31)	(0.56)	(0.42)	(0.28)	(0.33)	(0.29)	(0.42)	(0.48)
SVENF03	-2.03	-0.83	0.2	-0.71	-0.27	0.07	-0.23	-2.47	-1.01	-0.33	-0.94	0.9	1.15	-1.01
	(1.14)	(0.79)	(0.69)	(0.65)	(0.63)	(0.81)	(1.07)	(1.96)	(1.45)	(0.98)	(1.15)	(1.01)	(1.51)	(1.69)
SVENF04	2.33	1.01	-0.29	0.75	0.04	-0.3	0.19	3.24	1.05	0.39	1.01	-1.28	-1.72	0.79
	(1.36)	(0.97)	(0.83)	(0.78)	(0.75)	(0.98)	(1.3)	(2.39)	(1.79)	(1.2)	(1.41)	(1.24)	(1.87)	(2.06)
SVENF05	-0.95	-0.42	0.12	-0.25	0.12	0.24	0.02	-1.42	-0.36	-0.19	-0.35	0.63	0.85	-0.11
	(0.55)	(0.4)	(0.34)	(0.32)	(0.31)	(0.41)	(0.54)	(0.99)	(0.75)	(0.5)	(0.59)	(0.52)	(0.78)	(0.85)
RV	-1.38	-1.19	-0.59	-0.07	0.49	0.98	1.24	-1.52	-1.24	-0.86	-0.16	0.48	1.06	1.2
	(0.15)	(0.1)	(0.1)	(0.09)	(0.08)	(0.11)	(0.12)	(0.17)	(0.11)	(0.09)	(0.09)	(0.09)	(0.12)	(0.15)

Table 3.11: US treasuries - QR forward rates + Realized Semivariances - 6 factor model

	0.05	0.1	0.25	0.5	0.75	0.9	0.95	0.05	0.1	0.25	0.5	0.75	0.9	0.95
	TU							FV						
const	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)
SVENF02	0.13	0.06	0.03	0.02	0.01	-0.02	-0.06	0.4	0.3	0.06	0.16	0.17	0.15	-0.01
	(0.07)	(0.05)	(0.04)	(0.04)	(0.04)	(0.05)	(0.06)	(0.21)	(0.15)	(0.12)	(0.11)	(0.11)	(0.17)	(0.2)
SVENF03	-0.42	-0.19	-0.12	-0.08	-0.04	0.04	0.2	-1.26	-0.91	-0.11	-0.59	-0.67	-0.57	-0.14
	(0.24)	(0.16)	(0.13)	(0.12)	(0.13)	(0.18)	(0.21)	(0.74)	(0.51)	(0.41)	(0.39)	(0.39)	(0.6)	(0.68)
SVENF04	0.42	0.18	0.13	0.1	0.06	-0.02	-0.21	1.4	0.96	0.04	0.65	0.77	0.64	0.22
	(0.29)	(0.19)	(0.15)	(0.14)	(0.15)	(0.21)	(0.25)	(0.89)	(0.62)	(0.49)	(0.46)	(0.47)	(0.71)	(0.82)
SVENF05	-0.15	-0.06	-0.05	-0.04	-0.02	0.01	0.1	-0.55	-0.36	0.01	-0.23	-0.27	-0.21	-0.05
	(0.12)	(0.08)	(0.06)	(0.06)	(0.06)	(0.09)	(0.1)	(0.36)	(0.25)	(0.2)	(0.19)	(0.19)	(0.29)	(0.34)
RSV P	-0.71	-0.76	-0.4	0.13	0.82	1.32	1.47	-1.01	-0.67	-0.38	-0.08	0.63	1.04	1.27
	(0.26)	(0.2)	(0.13)	(0.17)	(0.14)	(0.19)	(0.19)	(0.27)	(0.16)	(0.13)	(0.14)	(0.13)	(0.19)	(0.19)
RSV N	-1.37	-0.8	-0.26	0.05	0.18	0.44	0.5	-1.16	-0.89	-0.31	0.05	0.1	0.44	0.3
	(0.14)	(0.21)	(0.11)	(0.15)	(0.14)	(0.11)	(0.12)	(0.23)	(0.17)	(0.12)	(0.14)	(0.1)	(0.18)	(0.09)
	TY							US						
const	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)
SVENF02	0.55	0.21	-0.09	0.21	0.07	0.03	0.05	0.87	0.36	0.09	0.24	-0.28	-0.25	0.16
	(0.32)	(0.22)	(0.19)	(0.18)	(0.18)	(0.24)	(0.32)	(0.54)	(0.4)	(0.27)	(0.33)	(0.29)	(0.39)	(0.47)
SVENF03	-1.8	-0.76	0.35	-0.74	-0.12	0.03	-0.2	-3.37	-1.16	-0.3	-0.83	0.95	1.04	-0.38
	(1.08)	(0.76)	(0.66)	(0.64)	(0.62)	(0.84)	(1.11)	(1.88)	(1.41)	(0.96)	(1.15)	(1.01)	(1.39)	(1.65)
SVENF04	2.06	0.93	-0.45	0.79	-0.11	-0.26	0.15	4.49	1.27	0.36	0.85	-1.29	-1.53	-0.02
	(1.28)	(0.93)	(0.8)	(0.78)	(0.74)	(1)	(1.35)	(2.3)	(1.74)	(1.18)	(1.4)	(1.23)	(1.73)	(2.02)
SVENF05	-0.84	-0.39	0.19	-0.27	0.17	0.23	0.04	-1.98	-0.46	-0.17	-0.27	0.61	0.74	0.24
	(0.52)	(0.38)	(0.33)	(0.32)	(0.3)	(0.41)	(0.56)	(0.96)	(0.74)	(0.49)	(0.58)	(0.52)	(0.73)	(0.84)
RSV P	-0.98	-0.78	-0.61	-0.03	0.52	1.05	1.14	-0.64	-0.77	-0.63	-0.11	0.39	1.19	1.53
	(0.25)	(0.16)	(0.16)	(0.16)	(0.13)	(0.18)	(0.2)	(0.27)	(0.16)	(0.13)	(0.15)	(0.15)	(0.19)	(0.26)
RSV N	-1.04	-0.95	-0.26	-0.06	0.26	0.44	0.57	-1.54	-1.05	-0.63	-0.14	0.34	0.39	0.27
	(0.23)	(0.19)	(0.17)	(0.16)	(0.12)	(0.15)	(0.13)	(0.23)	(0.22)	(0.16)	(0.16)	(0.15)	(0.12)	(0.26)

Q: *All Figures and Tables in paper 3 should be self-contained. Therefore, I recommend providing a more detailed set of notes associated to each one to enhance the readability of this paper. Specifically for all figures in this paper, please state whether the % confidence intervals.*
A: Corrected.

Q: *Tables presenting the descriptive statistics should state the sample used.*
A: Corrected.

Comments for Bibliography **Q:** *There are some inconsistencies within your bibliography. For example, on page 102 Journal of Financial Econometrics is cited as Journal of Financial Econometrics and Journal of financial econometrics. In other cases, papers published in Econometrica sometimes read Econometrica: Journal of the Econometric Society or Econometrica. There are other typos and inconsistencies outlined in the paper copy that will be provided to you at the pre-defence.*
A: Corrected.