Erasmus University Rotterdam

Erasmus School of Economics

Bachelor Thesis Econometrie en Operationele Research

Slope Homogeneity Testing and Forecasting Stock Returns using Quantile Regression

Name Student/Author	*Student Number*	*Supervisor*	*Second Assessor*
Dylan J. van Gemeren	477690	Y. Sun	A. M. Schnucker

Date Final Version: July 4, 2020

Abstract

It is often assumed for panel data that the slope coefficients are homogeneous across individuals. In this paper this assumption is tested for different quantiles by using Swamy-type tests for Quantile Regressions specifically. For this purpose, I use U.S. financial data from 56 firms during the period 1970-2011, including returns and several firm characteristics. The tests show empirical evidence to reject the homogeneity assumption for the tail quantiles. This indicates that it might be useful to pool data for estimating the central quantiles, while individually estimating the tail quantiles. I use these findings to construct efficient-flexible quantile forecasts for the period 2012-2019 to innovate on several Quantile Regression based point forecasting methods. However, using efficient quantile estimates or OLS actually performs better for forecasting. In addition I try to forecast the future quantile using both the quantile forecasts and the point forecasts. Nonetheless, it seems to be particularly hard to forecast tail quantiles in the future.

Contents

1 Introduction

Over the past few years the amount of available data on financial markets has grown rapidly. This has made it possible to research and model these markets more accurately than before. Among these models are panel data models that consider time series of subsequent observations for different individuals. Researchers that use these models often assume that all individuals are affected to the same extent by the variables of interest, while allowing for different intercepts, also called fixed effects. However, this so-called slope homogeneity assumption might be too strong in practice. Moreover, whether these slope coefficients are homogeneous might even depend on the given quantile of the dependent variable.

Although a lot of research has been conducted recently to test the homogeneity assumption for regular panel data models, little research has been conducted to test this assumption for more advanced panel data models. One such model is the Fixed-Effects Quantile Regression (FE-QR) model introduced by Koenker (2004). Similar to regular panel data models, the FE-QR model only allows for heterogeneity in the intercepts. In order to test the slope homogeneity assumption for the FE-QR model, Galvao et al. (2018) propose two Swamy-type tests. They also use these tests on U.S. financial data, containing the excess asset return and several firm characteristics, to find whether the slope coefficients of these firm characteristics are homogeneous. They find no empirical evidence to reject this assumption for the central quantiles, but in the upper and lower quantiles, they do find empirical evidence to reject the assumption of homogeneous slope coefficients. This indicates that especially during times of financial distress or financial booms, certain firm characteristics might have different effects on different firms' returns. This confirms the hypothesis that the assumption of homogeneous effects for financial data is probably too strong in general. Hence, allowing for more flexible quantile estimates per individual might be more appropriate in certain cases.

Besides helping to better understand slope homogeneity and slope heterogeneity in financial data, it might be interesting to see whether these results can be used to improve forecasting of the returns. Moreover, as Ma & Pohlman (2008) propose two return forecasting methods for quantile regressions specifically, based on different quantiles forecasts, it might be interesting to see whether I can improve on these methods and deliver more accurate results by making use of a efficiency-flexibility trade-off. Furthermore, it might be useful to forecast future quantiles of the firms' returns as a way for investors and financial institutions to possibly predict the business cycle.

All in all, the question central in this research is: "Can I detect slope heterogeneity of firm characteristics for the returns in tail quantiles, and to what extent can I use these findings to create reliable return point forecasts and future quantile forecasts?".

In this paper I use data similar to Galvao et al. (2018). The data consists of annual data of the returns and several firm characteristics from 56 U.S. firms for the period 1970-2019 and is obtained from Wharton Research Data Services (1993). This data will be split into an estimation sample and a forecast sample. I first use the FE-QR model to estimate the different firms' slope parameters individually for different quantiles and calculate the Swamy-type test statistics proposed by Galvao et al. (2018) to test the slope homogeneity assumption for this dataset for these respective quantiles. I find no empirical evidence for the central quantiles to reject the assumption of slope homogeneity, but in the tail quantiles on the other hand I do find empirical evidence to reject this assumption. This is in line with the findings of Galvao et al. (2018). These results are used to construct efficient-flexible quantile forecasts by making use of the efficiency-flexibility trade-off in panel data of homogeneous slope estimates versus heterogeneous slope estimates as described by Zhang et al. (2019), in an attempt to improve on two point forecasts methods, the Quantile Regression Alpha Distribution (QRAD) Location and Probability model, proposed by Ma & Pohlman (2008). These forecasts are then compared with simple OLS forecasts. However, these methods are not able to outperform the OLS model in terms of forecasting. Finally, I look at how well these forecasts are able to predict the future quantiles. I find that the OLS forecasts might be able to predict the future quantiles to some extent in the central qauntiles. However it seems particularly hard to forecast more extreme future quantiles.

The structure of the remainder of this paper is as follows. Section 2 gives a summary of the literature on Quantile Regression. Then, Section 3 elaborates on the methods used for implementing the models and assessing their performance and Section 4 discusses the data and shows some relevant characteristics of the data. The results are discussed in Section 5 and finally a discussion and summary of the research are give in Section 6 and 7.

2 Literature

2.1 Quantile Regression

Recently, due to the rapidly increasing amount of available data in finance, Quantile Regressions (QR) have become a popular topic for research. One of the first modern works on QR was by Koenker & Bassett Jr (1978), introducing the original QR model. QR can be seen as an extension of the linear regression model that allows to estimate the coefficients for a specific quantile. That is, instead of the constant coefficient estimates of linear regression models, coefficient estimates in the QR model depend on the quantile, meaning that these estimates are allowed to differ per quantile. Advantages of QR models over linear regressions models are that they allow for more flexibility by using the distribution of the dependent variable (Güloğlu et al. (2016)). Besides, according to Koenker & Hallock (2001) QR models might give a more complete picture of the dependency between several variables. These models are also more robust to outliers as the model differs per quantile (Mello & Perrelli (2003)). This is especially important in finance as these outliers occur relatively often as is argued by Van Dijk & Franses (2003).

2.2 Heterogeneity in Panel Data

Another interesting topic in finance recently, has been the usage of panel data. Panel data consists of a cross section of individual time series data, that is, sequences of successive observations over time for different firms, households or countries (Heij et al. (2004)). A common way of analyzing this type of data is by using panel models with fixed effects. These models allow for the intercepts, also called fixed effects, to differ per individual, whereas the effects of other variables are assumed to be homogeneous among individuals and the data will be pooled.

This assumption of homogeneous slope coefficients in panel data is commonly used in finance. However, this assumption might be somewhat strong in practice which might lead to certain problems. For example, Browning et al. (2007) conclude that there is more heterogeneity in the data than scientists often allow for, and that handling heterogeneity of the slope coefficients incorrectly might result in estimation errors. Moreover, Zhang et al. (2019) and Hsiao & Sun (2000) argue that there is a efficiency-flexibility trade-off when deciding between assuming slope homogeneity and allowing for heterogeneous effects, which affects the performance of the model. That is, homogeneous effects allow for pooling of the data and thus lead to more efficient estimates, whereas heterogeneous effects might better explain the true underlying process and are thus more flexible. Therefore, it is important to first test this assumption for the given dataset.

For this purpose, several tests have been developed over the past few years. For example, Phillips & Sul (2003) propose a Hausman (1978) type test, however this test was shown to be invalid in case of cross section independence by Pesaran & Yamagata (2008). Moreover, Pesaran & Yamagata (2008) develop a Swamy (1970)-type test as well as a standardized Swamy-type test for higher dimensions to test this assumption.

2.3 Quantile Regression Homogeneity Tests

However, relatively little research has yet been done to test slope homogeneity for QR models for panel data. Koenker (2004) introduces the FE-QR model in order to investigate the applications of QR for panel data specifically. However, similar to the panel models with fixed effects, the FE-QR model only assumes fixed effects to differ per individual. In the context of QR models, these fixed effects can be interpreted as location shifts for the conditional quantiles. The explanatory variables on the other hand, are again often assumed to have homogeneous effects.

Therefore, Galvao et al. (2018) propose two tests to test for homogeneity of slope coefficients across individuals for different quantiles in the FE-QR model. Similar to Pesaran & Yamagata (2008), they develop a Swamy-type test and standardized Swamy test and show that these tests possess the right properties to test this assumption.

Galvao et al. (2018) also apply their tests on U.S. financial data to test for slope homogeneity of several firm characteristics on excess return on assets. They find that for the central quantiles the homogeneity assumption can not be rejected, indicating that the effects of these firm characteristics are similar across the individual firms for these quantiles. For the higher and lower quantiles on the other hand, they find that these slope coefficients do significantly differ per firm, indicating that during times of financial distress or financial booms, firms' returns react differently depending on several firm characteristics. These results are in line with the earlier findings of for example Güloğlu et al. (2016). This boom and burst behavior of returns might be due to the underlying sentiment of the investors. Ni et al. (2015) found that the investor sentiment positively affects stock with relatively high returns heavily in the near future, whereas for stock with smaller returns they find a negative effect on the long run. Similarly, Giovannetti (2013) concludes that optimism corresponds with higher probabilities of high returns, and on the other hand pessimism with higher probabilities of lower returns.

2.4 Quantile Regression for Return Forecasting

Although return forecasting is often considered impossible, many models have been created in an attempt to do so, such as the Fama & French (1989) model. Welch & Goyal (2008) investigated the out-of-sample performance of some of these models, but concluded that most of them were not able to outperform the historical average forecast in terms of forecasting as they were often not stable enough. Nonetheless recent studies with improved forecasting strategies have claimed that it might actually be possible to outperform this benchmark (D. Rapach & Zhou (2013)). Among others, these strategies are based on models with economic restrictions (Ferreira & Santa-Clara (2011)) or combining forecasts (D. E. Rapach et al. (2010)), indicating that it might be possible to forecast future returns.

For that reason, it might be interesting to see whether it is also possible to forecast returns using QR. Ma & Pohlman (2008) propose two forecasting methods, QRAD Location and Probability respectively, that use the different forecasts per quantile to construct more robust point forecasts. Based on robust goodness-of-fit measures, they conclude that these methods outperform regular forecasting methods such as the mean and median forecasts under the right conditions, indicating that these methods might be useful for return forecasting.

Furthermore, Chen & Chen (2002) use QR to predict the Value-at-Risk (VaR) by using the different quantile forecasts. They conclude that these VaR calculations provides more accurate results than those with variance-covariance approaches. Additionally, Cech & Barunik (2017) and Baruník & Čech (2020) investigated the forecasting performance of QR models by means of VaR for more frequent data. They found that their respective forecasts were able to outperform traditional benchmarks such as the RiskMetrics model. Hence, QR might also be helpful for predicting future quantiles of the returns' distributions, which might be useful for investors.

All in all, this paper contributes to the current literature by investigating whether it is possible to forecast the returns with the QRAD methods by making use of the efficiency-flexibility trade-off. That is, the results of the homogeneity tests for different quantiles will be used to determine whether certain quantiles should be forecasted using the joint slope estimates or the individual slope estimates. In addition, this paper investigates the possibility of forecasting the future quantiles.

3 Methodology

3.1 The FE-QR Model

First of all, I start by showing the QR model needed to test the slope homogeneity assumption. Similar to Galvao et al. (2018), I use the FE-QR model proposed by Koenker (2004) to estimate the slope coefficients for different quantiles for all firms individually. That is

$$Q_{r_{i,t+1}}(\tau|x_{it}, \alpha_{i0}(\tau)) = \alpha_{i0}(\tau) + x_{it}^{'}\beta_{i0}(\tau) \tag{1}$$

where $i \in \{1, 2, ..., n\}$ denotes the firm, and $t \in \{1, 2, ...T-1\}$ denotes the respective time period. $Q_{r_{i,t+1}}(\tau|x_{it}, \alpha_{i0}(\tau))$ denotes quantile τ of the stock return of firm i at time t+1, given x_{it} and $\alpha_{i0}(\tau)$. Note the predictive character of this model due to the t+1 term for the returns. $\alpha_i(\tau)$ is the fixed effect of firm i in quantile τ, and $\beta_i(\tau)$ is the vector of slope coefficients for firm i in quantile τ. Note that this is different from the original FE-QR model, where it is assumed that the explanatory variables share the same slope coefficients across individuals. Lastly, the vector x_{it} is the vector containing the explanatory variables, and it is assumed that there is no relation between the explanatory variables and the fixed effects. Equivalently, the model can be written more concisely as:

$$Q_{r_{i,t+1}}(\tau|X_{it}) = X_{it}^{'}\theta_{i0}(\tau) \tag{2}$$

where $X_{it} = (1, x_{it}^{'})^{'}$ and $\theta_{i0}(\tau) = (\alpha_{i0}(\tau), \beta_{i0}^{'}(\tau))^{'}$. As discussed in Section 2.1, this model has the advantage over simple linear regression models that it gives a more complete picture of how the dependent variable interacts with the explanatory variables, as the coefficients are allowed to differ per quantile.

From here, the QR estimates $\hat{\theta}_i(\tau)$ can be calculated as in Koenker & Bassett Jr (1978)

$$\hat{\theta}_i(\tau) = argmin_{\theta_i(\tau) \in R^{k+1}} \frac{1}{T} \sum_{t=1}^{T} \rho_\tau(r_{it} - X_{it}^{'}\theta_i(\tau)) \tag{3}$$

where k is the number of explanatory variables, and $\rho_\tau(e) = e(\tau - I(e \leq 0))$, with I being the indicator function, the function assigning the weights to the errors. Note that these weight are asymmetrically assigned depending on the specific quantile. The estimated quantile slope coefficients can then be obtained from $\hat{\beta}_i(\tau) = \Xi\hat{\theta}_i(\tau)$, with $\Xi = [0_{kx1}|I_{kxk}]$. In general these estimates might be subject to the incidental parameter problem as described by Neyman & Scott (1948), meaning that these estimates might be inconsistent. Therefore, it is important to consider a larger time period so that these estimates become consistent and ensure asymptotic normality of the estimates.

I use the *qr_standard* function by Xu (2020) that makes use of the computationally more efficient interior point algorithm as described by Koenker & Ng (2005) to implement this model in MATLAB (1984) for $\tau \in \{0.05, 0.10...0.90, 0.95\}$.

3.2 Homogeneity Test

In addition, I use the FE-QR model to test the hypothesis of homogeneous slope coefficients across the firms for different quantiles τ. That is, I test the null hypothesis

$$H_0(\tau) : \beta_{i0}(\tau) = \beta_0(\tau), \ \forall i \tag{4}$$

against the alternative

$$H_a(\tau) : \beta_{i0} \neq \beta_{j0}, \ \exists \ i, j \tag{5}$$

for some fixed $\beta_0(\tau)$. As it is not possible to observe the actual $\beta_0(\tau)$, I have to estimate its value as well. For this purpose I use the efficient Minimum Distance (MD) estimator as proposed by Galvao & Wang (2015). This MD estimator is a weighted average of the respective slope estimates and can be calculated as

$$\hat{\beta}_{MD}(\tau) = (\sum_{i=1}^{n} \hat{V}_i^{-1}(\tau))^{-1} \sum_{i=1}^{n} \hat{V}_i^{-1}(\tau)\hat{\beta}_i(\tau) \tag{6}$$

where $\hat{V}_i(\tau) = \Xi\tilde{V}_i(\tau)\Xi'$, and $\tilde{V}_i(\tau)$ is an estimator of the covariance for the individual regression quantiles. Again, for the sake of consistent covariance estimates, it is important to consider larger time periods.

An appropriate estimator of the covariance for the individual regression quantiles according to Galvao et al. (2018) is

$$\tilde{V}_i(\tau) = \tilde{\Gamma}_i^{-1}\tilde{\Omega}_i\tilde{\Gamma}_i^{-1} \tag{7}$$

$$\tilde{\Gamma}_i(\tau) = \frac{1}{T}\sum_{t=1}^{T} K_{h_n}(\hat{e}_{it}(\tau))X_{it}X_{it}' \tag{8}$$

$$\tilde{\Omega}_i(\tau) = \frac{\tau(1-\tau)}{T}\sum_{t=1}^{T} X_{it}X_{it}' \tag{9}$$

where $\tilde{\Gamma}_i(\tau)$ is called Powell's kernel estimator, after the work of Powell (1991). It has to be noted that this estimator among others assumes that the data is independent across individuals as well as within individuals, which might be a somewhat strong assumption for real financial data. $\hat{e}_{it}(\tau)$ is the estimation error for quantile τ for firm i at time t. This error can be computed as $\hat{e}_{it}(\tau) = r_{i,t+1} - X_{it}'\hat{\theta}_i(\tau)$. Furthermore, I use the Gaussian kernel K_{h_n} with the Hall & Sheather (1988) bandwidth h_n

$$K_{h_n}(e) = \frac{1}{h_n}\frac{1}{\sqrt{2\pi}}exp(-0.5 * (\frac{e}{h_n})^2) \tag{10}$$

$$h_n = n^{-\frac{1}{3}} z_\alpha^{\frac{2}{3}} (\frac{1.5 * \phi^2(\Phi^{-1}(\tau))}{2(\Phi^{-1}(\tau))^2 + 1})^{\frac{1}{3}} \tag{11}$$

Here, $\phi(t)$ and $\Phi(t)$ denote the standard normal probability density function and standard normal cumulative density function respectively, and $\Phi(z_\alpha) = 1 - \alpha/2$ for test size α, as in Kato et al. (2012).

The significance of slope homogeneity is then assessed by the two tests proposed by Galvao et al. (2018). First a Swamy-type test

$$\hat{S}(\tau) = \sum_{i=1}^{n}(\hat{\beta}_i(\tau) - \hat{\beta}_{MD}(\tau))'(\frac{\hat{V}_i(\tau)}{T})^{-1}(\hat{\beta}_i(\tau) - \hat{\beta}_{MD}(\tau)) \tag{12}$$

and second a standardized Swamy test

$$\hat{\Delta}(\tau) = \sqrt{n}(\frac{\frac{1}{n}\hat{S}(\tau) - k}{\sqrt{2k}}) \tag{13}$$

Galvao et al. (2018) show that the Swamy-type test statistic is asymptotically chi-squared distributed as $\hat{S}(\tau) \overset{a}{\sim} \chi^2_{(n-1)*k}$. Recall that k is the number of explanatory variables. Additionally, they show that standardized Swamy test statistic can be approximated by a standard normal distribution, $\hat{\Delta}(\tau) \approx N(0,1)$. Note that the hypothesis of homogeneous slope coefficients gets rejected in case of $\hat{S}(\tau)$ and $\hat{\Delta}(\tau)$ being large.

3.3 Efficient Flexible Quantile Forecasting

After the evaluation of the homogeneity tests, I will forecast the returns of the QR model for all previously specified quantiles τ. To accommodate for the efficiency-flexibility trade-off as discussed by Zhang et al. (2019), I use the outcomes of the Swamy-type test to determine whether I should use the efficient or flexible parameter estimates for a certain quantile. That is, when the data appears to be homogeneous in a specific quantile I use the more efficient pooled estimates, whereas in case of heterogeneity I use the individual specific parameter estimates. Therefore, the efficient-flexible quantile estimates of the returns look as follows

$$\hat{r}_{i,t+1}(\tau) = \begin{cases} X'_{i,t}\hat{\theta}_i(\tau) & \text{if } p_{\hat{S}(\tau)} \leq \alpha \\ X'_{i,t}\tilde{\theta}_i(\tau) & \text{if } p_{\hat{S}(\tau)} > \alpha \end{cases} \tag{14}$$

where $\hat{\theta}_i(\tau)$ is the vector of individual coefficient estimates, $\hat{\theta}_i(\tau) = (\hat{\alpha}_i(\tau), \hat{\beta}_i(\tau))$ and $\tilde{\theta}_i(\tau_q)$ the vector of pooled coefficient estimates while allowing for different fixed effects per individual, $\tilde{\theta}_i(\tau) = (\tilde{\alpha}_i(\tau), \tilde{\beta}(\tau))$ for quantile τ_q. Furthermore, $p_{\hat{S}(\tau)}$ is the quantile's corresponding p-value of the Swamy test statistic and $\alpha = 0.05$ is the significance level.

3.4 Point Forecasting

Using these different quantile forecasts I then calculate different return point forecast. In this Section I discuss two Quantile Regression Alpha Distribution (QRAD) methods proposed by Ma & Pohlman (2008) to forecast the returns. Moreover, I innovate on these methods by using efficient-flexible quantile forecasts from Section 3.3 to hopefully improve results. Therefore, I first introduce the following notation: τ_q corresponding with the previously specified quantiles, with $q = \{1, 2, ..., Q\}$. Note that in this case Q=19, as I have split the quantiles in 0.05 intervals ranging form 0.05 to 0.95.

3.4.1 QRAD Location

The QRAD Location method assumes that, as returns are in general quite unpredictable, the future return will remain in the same quantile as the return of former period. That is

$$\hat{r}_{i,t+1,QRADL} = \begin{cases} \hat{r}_{i,t+1}(\tau_Q) & \text{if } r_{i,t} \geq \hat{r}_{i,t}(\tau_Q) \\ \hat{r}_{i,t+1}(\tau_{Q-1}) & \text{if } \hat{r}_{i,t}(\tau_Q) > r_{i,t} \geq \hat{r}_{i,t}(\tau_{Q-1}) \\ ... & \\ \hat{r}_{i,t+1}(0.5) & \text{otherwise} \\ ... & \\ \hat{r}_{i,t+1}(\tau_2) & \text{if } \hat{r}_{i,t}(\tau_1) < r_{i,t} \leq \hat{r}_{i,t}(\tau_2) \\ \hat{r}_{i,t+1}(\tau_1) & \text{if } r_{i,t} \leq \hat{r}_{i,t}(\tau_1) \end{cases} \tag{15}$$

Although this assumption might be somewhat strong, there is general evidence for stock performing better (worse) when is has performed well (bad) in the former period (Asness (1995)). Moreover, strategies based on this momentum have even proved to be profitable according to Jegadeesh & Titman (2011). Given that the probability of returns staying in the same quantile is relatively high, Ma & Pohlman (2008) show that these forecasts perform better than mean and median forecasts. They argue that there exists a trade-off between the number of different quantiles taken into consideration, the forecasting accuracy and the probability of the returns staying in the same quantile.

3.4.2 QRAD Probability

The second approach, the QRAD Probability method, on the other hand does not share the disadvantage of the QRAD Location method of relying on the assumption of returns remaining in the same quantile. Instead, this method assigns probabilities of being in a certain quantile and uses these probabilities to compute a point forecast. As the quantiles should correspond with the distribution of the returns, the most straightforward way of assigning probabilities to the quantiles is to assign the probability consistent with the underlying distribution. This means that I assign probabilities $p = [0.05, ..., 0.05, 0.10, 0.05, ..., 0.05]$ to the corresponding quantile, with only the median getting assigned a probability of 0.10. From here I calculate the expected excess return as

$$\hat{r}_{i,t+1,QRADP} = \sum_{q=1}^{Q} p_q \hat{r}_{i,t+1}(\tau_q) \tag{16}$$

Again, it can be shown that under certain conditions, these forecasts outperform the traditional mean and median forecasts. In addition, Ma & Pohlman (2008) show that the two approaches asymptotically yield the same results. However, even assigning probabilities consistent with the underlying distribution to predict the future returns might be a somewhat strong assumption, as the probability of being in a certain quantile might well be time-variant. That is, during times of financial distress (booms) the probability of being in the lower (higher) quantiles is probably underestimated (overestimated) by this method.

3.4.3 Goodness of Fit

I then compare the relative forecasting performance of these methods with each other and a simple OLS model. For this purpose I consider two forecasting accuracy measures that are commonly used, namely the Mean Squared Prediction Error, $MSPE = \frac{1}{(M-T)*n} \sum_{t=T}^{M-1} \sum_{i=1}^{n} (\hat{r}_{i,t+1} - r_{i,t+1})^2$, and the more robust Mean Absolute Prediction Error, $MAPE = \frac{1}{(M-T)*n} \sum_{t=T}^{M-1} \sum_{i=1}^{n} |\hat{r}_{i,t+1} - r_{i,t+1}|$. I also give the differences in performance significance by means of the Diebold-Mariano test (Diebold & Mariano (2002)):

$$DM = \frac{\bar{d}}{\sqrt{Var(\hat{d}_{it})/(M-T)*n}} \tag{17}$$

Here d_{it} is the difference between the squared prediction errors for firm i at time t, $d_{it} = e^2_{i,t,k} - e^2_{i,t,l}$, where $e_{i,t,k} = \hat{r}_{i,t+1,k} - r_{i,t+1}$, and $k, l \in \{QRADL, QRADP, OLS\}$, and $\bar{d}$ is the mean of the d_{it}. It holds asymptotically that $DM \overset{a}{\sim} N(0,1)$.

3.5 Forecasting Future Quantiles

Finally, it might be useful for investors to be able to predict the business cycle as well. As quantile correspond with certain parts of the distribution, these quantiles might also be indicators of the business cycle. Therefore, it might be interesting to see to what extent I am able to forecast the future quantiles. For this purpose I use the same point forecasts as described in Section 5.2, implying that $k \in \{QRADL, QRADP, OLS\}$, and the quantile forecasts from Section 3.3 to forecast the future quantile as follows.

$$\hat{\tau}_{i,t+1,k} = \begin{cases} \tau_Q & \text{if } \hat{r}_{i,t+1,k} \geq \hat{r}_{i,t+1}(\tau_Q) \\ \tau_{Q-1} & \text{if } \hat{r}_{i,t+1}(\tau_Q) > \hat{r}_{i,t+1,k} \geq \hat{r}_{i,t+1}(\tau_{Q-1}) \\ ... & \\ 0.5 & \text{otherwise} \\ ... & \\ \tau_2 & \text{if } \hat{r}_{i,t+1}(\tau_1) < \hat{r}_{i,t+1,k} \leq \hat{r}_{i,t+1}(\tau_2) \\ \tau_1 & \text{if } \hat{r}_{i,t+1,k} \leq \tilde{r}_{i,t+1}(\tau_1) \end{cases} \tag{18}$$

That is, if the forecasted return falls in a certain quantile, that particular quantile will be the future quantile forecast $\hat{\tau}_{i,t+1,k}$. Therefore, this way of quantile forecasting can be seen as an additional way of evaluating the particular point forecasts.

Furthermore, to assess whether these quantile estimates do correctly cover the true distribution of the future returns, I will briefly test for correct unconditional coverage as well. I will do so only for the lower quantiles, as these are often considered most important by investors. For this reason, I apply the Likelihood Ratio (LR) test for unconditional coverage from Shephard & Sheppard (2010). That is, I test the null hypothesis $H_0 : p = \tau_q$ against the alternative $H_a : p = \pi$, with $\pi \neq \tau_q$. The LR test statistic can be calculated as

$$LR = -2ln(\frac{(1-\tau_q)^{T0}\tau_q^{T1}}{(1-\hat{\pi})^{T0}\hat{\pi}^{T1}}) \tag{19}$$

Here $\hat{\pi} = \frac{T1}{T0+T1}$, with $T1 = \sum_{t=T+1}^{M}\sum_{i=1}^{n} I(r_{t+1} < \tilde{r}_{i,t+1}(\tau_q))$, and $T0 = (M-T)*n - T1$. It holds asymptotically that $LR \overset{a}{\sim} \chi_1^2$.

4 Data

The data is obtained from Wharton Research Data Services (1993) from the CRSP/Compustat Merged dataset, and contains yearly data of several firm characteristics and the returns from the period 1970-2019, thus including 50 years of data. This dataset is split in an estimation sample (1970-2011) and forecast sample (2012-2019). The variables included in this research are the following: Monthly Returns (r_m), Total Assets (AT), Long-Term Debt ($DLTT$), Short-Term Debt (DLC), Annual Close Price ($PRCC$), Common Shares Outstanding ($CSHO$), Income

Before Extraordinary Items (IB), Interest Expenses ($XINT$), Income Taxes (TXT), Preferred Stock Liquidating Value ($PSTKL$), Depreciation and Amortization (DP) and lastly Property, Plant & Equipment ($PPENT$). Additionally, I include annualized 1-month U.S. bond rates as the Risk-Free Rate (RF) and the Monthly Market Return (MR_m) on the S&P 500 composite index from WRDS. Lastly, I include the Consumer Price Index (CPI) from U.S. Bureau of Labor Statistics (1884). From here I can then calculate the firm characteristics similar to Galvao et al. (2018):

-Annual excess return on assets: $r = \Pi_{m=January}^{December}(1 + r_m) - (1 + RF)$,

-Market debt ratio: $MDR = \frac{DLTT+DLC}{DLTT+DLC+PRCC*CSHO}$,

-Profitability: $EBIT = \frac{IB+XINT+TXT}{AT}$,

-Market-to-book ratio: $MBR = \frac{DLTT+DLC+PSTKL+PRCC*CSHO}{AT}$,

-Depreciation and amortization as a proportion of total assets: $DA = \frac{DP}{AT}$,

-Logarithmic asset size (measured in 1983 dollar): $LNA = ln(\frac{AT*1{,}000{,}000}{PCI(1983)})$,

-Fixed assets as a proportion of total assets: $PPE = \frac{PPENT}{AT}$,

-Market excess return: $MKR = \Pi_{m=January}^{December}(1 + MR_m) - 1 - RF$

As is usual with financial data, observations corresponding with regulated utilities, financial firms and nonprofit firms are not considered in this analysis. This means that observation corresponding with firms with SIC-codes 4900-4999, 6000-6999 and 9000+ are removed from this dataset. Furthermore, firms with too high growth rates are not considered in this analysis. That is, if the value of Property, Plant & Equipment, the Sales or the Book Value of Assets has grown by over 100 percent in a certain year, the corresponding firm will be omitted from the sample. In similar fashion, those firms with missing and/or negative Property, Plant & Equipment and Sales values are omitted, as well as those firms with Total Asset Book Values smaller then $10,000,000 as measured in 1983 dollars.

In this paper, I consider three different model specifications with different sets of explanatory variables similar to Galvao et al. (2018). First, the three variable specification, including the variables $x_{it}^3 = \{MBR, LNA, MKR\}$ corresponding with the pricing factors in the Fama & French (1992) model. Second, the six variable specification, including all firm characteristics as explanatory variables, $x_{it}^6 = \{EBIT, MBR, DA, LNA, PPE, MDR\}$. Lastly, the seven variable specification including all variables, $x_{it}^7 = \{MDR, EBIT, MBR, DA, LNA, PPE, MKR\}$.

4.1 Data Characteristics

The final sample consists of 56 firms ($n = 56$) with 49 years of data ($T + M = 49$) due to the predictive nature of the model, resulting in 2744 total observations. Summary statistics of the data are shown in Table 1. Note that the average yearly excess return is about 10,7% indicating that the firms included in the sample seem to perform relatively well.

Table 1: Summary Statistics 1970-2019 Data

	r	MDR	EBIT	MBR	DA	LNA	PPE	MKR
Mean	0.107	0.231	0.125	1.506	0.046	18.154	0.365	0.036
Median	0.093	0.181	0.120	1.137	0.043	18.051	0.332	0.065
Maximum	4.286	0.966	0.492	9.955	0.168	21.277	0.924	0.296
Minimum	-0.888	0.000	-0.449	0.040	0.000	16.122	0.035	-0.400
Std. Dev.	0.335	0.183	0.074	1.110	0.021	1.040	0.188	0.167
Jarque-Bera	40859	879	935	13543	1783	123	170	211
Observations	2744	2744	2744	2744	2744	2744	2744	2744

Moreover, Figure 1 shows the distribution of the annual returns for the sample. It is worth noting that there are quite a few large outliers in the right tail of the distribution, with the maximum annual return being 4.286. The distribution of the excess returns is also very non-normally distributed as is indicated by the Jarque-Bera test statistic for normality in Table 1 as well. Furthermore, Figure 1 shows again the overall good performance of these stocks with a large majority of the returns being positive.

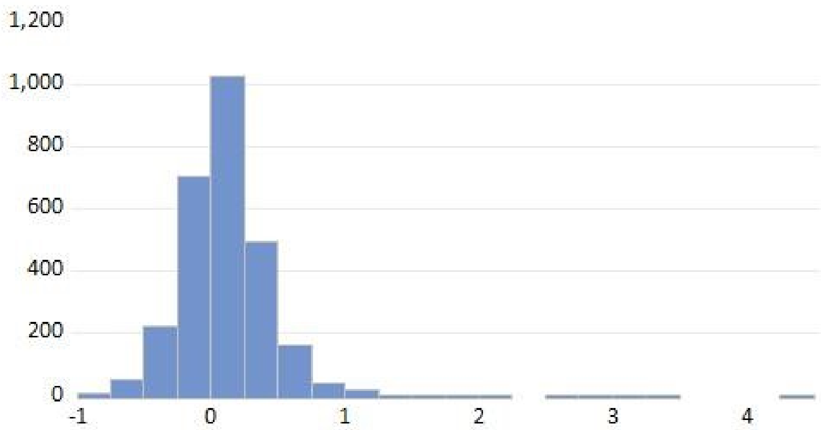

Figure 1: Histogram of the excess returns for the period 1970-2019

5 Results

5.1 Homogeneity Testing

In this Section I assess the results of the homogeneity tests for the three different model specifications of the FE-QR model from Section 4. Recall that I use the data from the period 1970-2011 for this purpose, meaning that T=41. Hence these quantile estimates might be subject to the incidental parameter problem. Therefore these results should be looked at with some caution. For the interested reader, the joint quantile estimates are shown and briefly discussed in the Appendix.

5.1.1 Three Variable Specification

I start the analysis by evaluating the slope homogeneity test results from the three-variable specification model, including the market-to-book ratio (MBR), the logarithmic asset size (LNA) and the excess market return (MKR). The p-values of the joint slope homogeneity tests are shown in Table 2 for different quantiles. Both the joint Swamy and joint standardized Swamy test find no empirical evidence to reject the hypothesis of slope homogeneity in the central quantiles at a 5% significance level. That is, between the 0.25 to 0.75 quantile of the returns' distributions, the effects of these three variables on the excess returns do not significantly differ per individual. Therefore, it is probably more efficient to pool the data for estimating these quantiles.

Table 2: p-values for Swamy ($\hat{S}$) and standardized Swamy ($\hat{\Delta}$) test for slope homogeneity of the three variable FE-QR model

	Joint		**MBR**		**LNA**		**MKR**	
τ	$\hat{S}$	$\hat{\Delta}$	$\hat{S}$	$\hat{\Delta}$	$\hat{S}$	$\hat{\Delta}$	$\hat{S}$	$\hat{\Delta}$
0.05	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.10	0.000	0.000	0.000	0.000	0.122	0.142	0.104	0.118
0.15	0.000	0.000	0.000	0.000	0.121	0.140	0.821	0.844
0.20	0.015	0.017	0.002	0.001	0.460	0.523	0.893	0.901
0.25	0.662	0.728	0.058	0.060	0.617	0.672	0.993	0.987
0.30	0.999	0.998	0.375	0.435	0.994	0.988	0.999	0.997
0.35	1.000	1.000	0.757	0.792	0.999	0.997	1.000	0.999
0.40	1.000	1.000	0.860	0.874	0.999	0.996	1.000	1.000
0.45	1.000	1.000	0.742	0.779	0.996	0.991	1.000	1.000
0.50	1.000	1.000	0.634	0.687	0.995	0.989	1.000	1.000
0.55	1.000	1.000	0.732	0.771	0.980	0.972	1.000	1.000
0.60	1.000	0.999	0.713	0.755	0.949	0.946	1.000	0.998
0.65	0.997	0.996	0.306	0.361	0.970	0.964	0.996	0.991
0.70	0.767	0.816	0.032	0.029	0.900	0.906	0.918	0.920
0.75	0.176	0.227	0.003	0.001	0.615	0.670	0.867	0.880
0.80	0.000	0.000	0.000	0.000	0.069	0.074	0.097	0.110
0.85	0.000	0.000	0.000	0.000	0.002	0.001	0.000	0.000
0.90	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.95	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

For the more extreme quantiles on the other hand, I do find significant evidence to reject the hypothesis of equal slope coefficients. That is, the effects of the two firm characteristics and the market return seem to differ significantly across firms when the returns are relatively low or high, often corresponding with times of financial bursts and booms. These results are similar to those of Galvao et al. (2018). As the three variables correspond with the factors from Fama & French (1992), this indicates that their model does not correctly capture the dynamics for the excess return in the more extreme quantiles. This result also suggests that it might be better not to pool the data during times of financial bursts and booms, and allow for individual specific slope coefficients, as pooling the data leads to false estimates for at least some firms.

Finally, I look at the marginal homogeneity test results for the different explanatory variables individually, to gain a further insights into the heterogeneous effects of certain variables. In general the patterns of the marginal tests are very similar to those of the joint test, with the rejection of slope homogeneity in the extremer quantiles. However, there are some minor differences for the different variables. For example, for the logarithmic asset size and the excess market return I only find empirical evidence to reject the hypothesis of marginal homogeneous effects for the 0.05 quantile in the left tail and 0.85 quantile and higher. For the market-to-book ratio on the other hand, the marginal effects already seem to be heterogeneous from the 0.20 quantile and lower, and the 0.70 quantile and higher. Hence, this variable seems to exhibit more heterogeneity than the other two variables. Therefore, from an efficiency point of view, it might be better to use the pooled estimates of the other two variables, and the individual estimates of the market-to-book ratio in those quantiles where only the market-to-book ratio marginal test rejects the slope homogeneity assumption.

5.1.2 Six and Seven Variable Specification

In this Section I discuss both the six and seven variable specification of the FE-QR model. Recall that the six variable specification considers the variables market-debt ratio (MDR), operating profitability ($EBIT$), market-to-book ratio (MBR), depreciation and amortization ratio to total assets (DA), the logarithmic asset size (LNA) and the proportion of fixed assets (PPE), and the seven variable specification includes the same variables as well as the excess market return (MKR). p-values for both the joint slope homogeneity tests and the individual tests for the seven variable specification are shown in Table 3. The results for the six variable specification are shown in the Appendix due to the large similarities between the two. Nonetheless, in this Section I will also discuss some differences between the two specifications.

In general, the results of these two specifications are quite similar to those of the three variable specification. That is, again I find no empirical evidence to reject the joint hypothesis of slope homogeneity in the central quantiles, whereas in the upper and lower tails of the firms' excess returns distributions I do find significant evidence to reject this hypothesis at a 5% significance level. This shows again that effects of the different firm characteristics on the returns differ significantly per individual especially during times of financial booms and financial distress. However, these specifications already show signs of slope heterogeneity at the 0.25 quantile and 0.70 quantile respectively, indicating that adding variables to the model might negatively impact the joint homogeneity assumption. Again, it is probably more efficient to use pooled data for estimating the central quantiles and use the more flexible individual estimates for the tail quantiles.

As for the marginal slope homogeneity tests for the seven variable specification, I find that similar to the three variable specification the market excess return shows the least heterogeneous marginal effects. Quite remarkably however, compared to the three variable specification the marginal effects of the market-to-book ratio are not significantly different per individual for a wider range of the returns' distribution, whereas the opposite holds for the logarithmic asset size. A possible explanation is that some part of the heterogeneous (homogeneous) effects of the other variables was captured by by these variables in the three variable specification model.

Finally, by comparing the six and seven variable specification directly, I find that the inclu-

Table 3: p-values for Swamy ($\hat{S}$) and standardized Swamy ($\hat{\Delta}$) test for slope homogeneity of the seven variable FE-QR specification

	Joint		**MDR**		**EBIT**		**MBR**		**DA**		**LNA**		**PPE**		**MKR**	
τ	$\hat{S}$	$\hat{\Delta}$	$\hat{S}$	$\hat{\Delta}$	$\hat{S}$	$\hat{\Delta}$	$\hat{S}$	$\hat{\Delta}$	$\hat{S}$	$\hat{\Delta}$	$\hat{S}$	$\hat{\Delta}$	$\hat{S}$	$\hat{\Delta}$	$\hat{S}$	$\hat{\Delta}$
0.05	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.10	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.030	0.028
0.15	0.000	0.000	0.000	0.000	0.000	0.000	0.021	0.018	0.000	0.000	0.000	0.000	0.003	0.001	0.274	0.324
0.20	0.000	0.000	0.035	0.034	0.000	0.000	0.112	0.129	0.002	0.001	0.000	0.000	0.011	0.008	0.666	0.715
0.25	0.000	0.000	0.287	0.339	0.029	0.026	0.218	0.259	0.112	0.129	0.008	0.005	0.169	0.199	0.895	0.902
0.30	0.530	0.636	0.291	0.343	0.277	0.327	0.188	0.222	0.269	0.319	0.219	0.260	0.367	0.427	0.965	0.959
0.35	0.984	0.989	0.333	0.390	0.468	0.531	0.304	0.358	0.656	0.707	0.432	0.495	0.673	0.721	0.999	0.995
0.40	1.000	1.000	0.407	0.469	0.722	0.763	0.560	0.620	0.912	0.915	0.742	0.780	0.861	0.875	0.999	0.996
0.45	1.000	1.000	0.523	0.585	0.601	0.657	0.489	0.551	0.920	0.922	0.647	0.699	0.863	0.877	1.000	0.997
0.50	1.000	1.000	0.626	0.680	0.683	0.729	0.553	0.613	0.973	0.966	0.708	0.751	0.961	0.956	0.999	0.997
0.55	1.000	1.000	0.330	0.387	0.598	0.654	0.378	0.438	0.954	0.949	0.445	0.508	0.919	0.921	1.000	0.998
0.60	0.968	0.979	0.414	0.477	0.605	0.661	0.504	0.566	0.963	0.957	0.309	0.364	0.472	0.535	0.998	0.994
0.65	0.345	0.451	0.147	0.173	0.257	0.304	0.325	0.381	0.834	0.854	0.152	0.179	0.265	0.314	0.983	0.976
0.70	0.001	0.001	0.025	0.021	0.115	0.132	0.088	0.098	0.671	0.719	0.074	0.080	0.129	0.150	0.842	0.861
0.75	0.000	0.000	0.001	0.000	0.022	0.018	0.022	0.018	0.407	0.469	0.025	0.022	0.000	0.000	0.476	0.539
0.80	0.000	0.000	0.000	0.000	0.051	0.052	0.004	0.002	0.009	0.006	0.002	0.001	0.000	0.000	0.071	0.076
0.85	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.90	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.95	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

sion of the market return as a variable in the model seems to have a relatively large impact on the Swamy test statistics. Although the joint hypothesis is rejected for exactly the same quantiles in both specifications, in general the respective p-values in the six variable specification are lower. This makes sense as I already found that the market return shows the least heterogeneous marginal effects for the different quantiles, hence not including this variable might take away some of the homogeneity in the model. All in all, the results of both the six and seven variable specification are again in line with the findings of Galvao et al. (2018).

5.2 Point Forecasting

In this Section I discuss the performance of the different forecasting models QRADL, QRADP and OLS for the seven variable specification, where the first two methods are based on the efficient-flexible estimates by using results of the Swamy test statistics as discussed in Section 3.3. The forecasting performances of these methods for the three and six variable specification are not discussed in this Section as these results are quite similar, but the results are shown in Tables 14-21 in the Appendix. Recall that the forecast period includes the forecasts for all 56 firms for the period 2012-2019.

I start the analysis by comparing some measures of forecasting accuracy, namely the MSPE and MAPE, which are shown in Table 4. Note that the simple OLS model outperforms the QRADL and QRADP method proposed by Ma & Pohlman (2008), based on both lower values for the MSPE and MAPE. According to the Diebold-Mariano test, all differences are significant. This indicates that these methods actually have relatively little forecasting power compared to OLS, and hence might not be very useful for the purpose of return forecasting with firm characteristics. In addition, the QRADP model also performs significantly better than the QRADL model based on the same measures, even though the QRADL model has a lower bias. All in all, both the assumption of returns staying in the same quantile and the assumption of all quantiles being assigned equal probabilities are probably too strong.

One possible explanation for the the better forecasting performance of the OLS model is that during the forecasting period returns may have been relatively stable, as the period 2012-2019 did not involve any major crisis (Conerly (2019)). Moreover, the returns have been relatively high, as the negative biases of the different models indicate. Thus, assigning equal probabilities to all different quantiles as in the QRADP model might have been an incorrect assumption for this particular period, thereby negatively affecting the point forecasts. Besides, the OLS model uses more efficient estimates which might improve the performance of its forecasts as well.

To see where the strength of these methods lies, their relative in-sample performances is examined as well. That is, I am interested whether the in-sample predictions of the QRADL and QRADP method for the period 1970-2011 are better than those of OLS. Therefore, I calculate and compare the MSE and MAE, similar to the forecasting measures. The results are shown in Table 5, and again all differences are significant. I find that for the purpose of in-sample predictions the QRADP method actually seems to perform better than OLS with both a lower MSE and MAE. Hence, assigning probabilities to different quantiles in order to calculate a weighted average seems to give more accurate in-sample predictions than using OLS.

Table 4: Forecasting Performance Measures 1970-2011

Model	Bias	MSPE	MAPE
OLS	-0.044	0.067	0.195
QRADL	-0.083	0.216	0.313
QRADP	-0.093	0.123	0.257

Table 5: In Sample Performance Measures 2012-2019

Model	Bias	MSE	MAE
OLS	0.000	0.118	0.234
QRADL	-0.001	0.174	0.273
QRADP	-0.005	0.091	0.213

As the efficiency of the OLS model was one possible explanation for the better forecasting performance of OLS compared to QRADP and QRADL, I decide to calculate the QRADP forecasts with efficient and flexible quantile forecasts separately to gain further insights in the effects of using efficient or flexible quantile forecasts. Recall that the former QRADP forecasts were constructed by using the efficient or flexible quantile forecasts for a specific quantile depending on the results of the Swamy tests (see Section 3.3). This means that for the separate efficient and flexible QRADP forecasts, I either use the efficient quantile forecasts or the flexible quantile forecasts to calculate the QRADP forecasts. The results are shown in Tables 6 and 7.

Note that in this case, the QRADP method with the efficient estimates is better in terms of forecasting based on both the MSPE and MAPE, whereas the QRADP method with flexible estimates has a better in-sample performance. Moreover, the efficient QRADP model performs much more similar to the plain OLS model in terms of both forecasting and in-sample performance, although the Diebold-Mariano tests still confirms that there is a significant difference in their respective performances. On the other hand, the QRADP model with flexible estimates actually seems to significantly outperform all other models in terms of in-sample performance. Hence, allowing for flexibility of estimating different slope parameters improves the in-sample performance, but for the purpose of point forecasting it is probably better to use more efficient models, such as the QRADP model with efficient estimates or the OLS model.

Table 6: Forecasting Performance Measures 1970-2011

Model	Bias	MSPE	MAPE
QRADP	-0.093	0.123	0.257
QRADP (efficient)	-0.070	0.075	0.205
QRADP (flexible)	-0.103	0.205	0.326

Table 7: In Sample Performance Measures 2012-2019

Model	Bias	MSE	MAE
QRADP	-0.005	0.091	0.213
QRADP (efficient)	-0.010	0.114	0.235
QRADP (flexible)	-0.002	0.085	0.205

5.3 Forecasting Future Quantiles

In this Section, I will use the different point forecasts from Section 5.2 and analyze their respective abilities to forecast the future quantiles. Therefore, this Section can be seen as an additional way of analyzing these models' respective forecasting performances. Note that I distinguish between forecasted future quantiles, the future quantile that I predict the future return to be in next period, and quantile forecasts, the predicted future returns for a specific quantile as explained in Section 3.3. The predicted future quantiles for the different point forecasting methods and different quantile forecasts are shown in Figures 4-6 in the Appendix.

Figure 2 displays the number of observations of the true return that fall in a particular quantile. I will refer to these quantiles as the 'realized quantiles'. These quantiles are based on the efficient-flexible, efficient and flexible quantile forecasts respectively. That is, the true values of the returns are compared to the corresponding quantile forecasts to decide in which quantile the future return is according to those particular quantile forecasts. Figure 2 shows that the efficient quantile estimates predict the future quantiles the most consistent, as the realized quantiles are relatively evenly distributed. For the (efficient)-flexible estimates on the other hand, many of the returns fall in the highest or lowest quantile respectively. This presumably has to do with the monotonicity problem found in both these quantile forecasts. That is, due to their flexibility it is possible that these methods actually forecast lower returns in for example the 0.10 quantile than in the 0.05 quantile, which goes against the definition of these quantiles. Moreover, the LR-test rejects correct unconditional coverage of the 0.05 quantile for both the (efficient-)flexible quantile forecasts, whereas correct unconditional coverage is not rejected for the efficient quantile forecasts. This shows that the efficient quantile forecasts are probably the best to consider for the purpose of forecasting future quantiles. Furthermore, for the efficient quantile forecasts I observe less true returns forecasted in the lower quantiles in contrast to the other two forecasting methods, which is more in line with the fact that the period 2012-2019 was relatively stable and did not include any major crisis, as discussed in the previous Section. Hence, I conclude that the efficient quantile forecasts are probably the most appropriate to consider for the purpose of quantile forecasting, hence I will mainly focus on these quantile forecasts in the remainder of this analysis

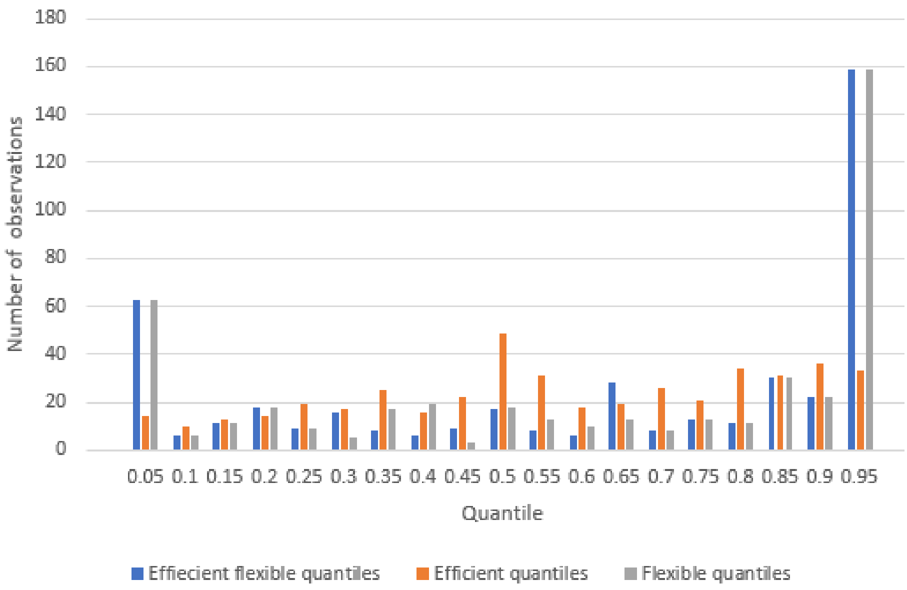

Figure 2: Realized return per quantile for the period 2012-2019 for different quantile forecasting methods

Now I turn to the evaluation of the different future quantile forecasts. The differences between the forecasted OLS, QRADL, and QRADP future quantiles and the 'realized quantiles' based on the efficient quantile forecasts are shown in Figure 3. From this Figure it can easily be seen that OLS forecasts the future quantile correctly more often than the other two methods. Besides, its maximum and minimum errors are smaller than for the other two methods as well, indicating that this method gives more reliable future quantile forecasts. In addition, based on this Figure, it seems again as if the QRADP method outperforms the QRADL method for the same reasons as why the OLS model outperforms the other two methods. This was to be expected as Section 5.2 already pointed out that the forecasting performance of OLS was better than the forecasting performance QRADP and QRADL respectively.

This Figure also shows that these models are able to forecast the future quantile at least to some extent, as for example OLS forecasts over 100 observations (approximately) right, which is about 1/4 of the total observations. Besides, only a few of the OLS future quantile forecasts show absolute errors larger than 0.4, indicating that these errors stay relatively small. However, on average OLS estimates the quantile about 0.05 too low, which probably originates from the problem that the OLS model forecasts too low returns in general, as discussed in Section 5.2. In addition, the mean absolute error is about 0.22, which seems quite large for good quantile forecasts as the maximum possible error is 0.9. Furthermore, the distribution of the efficient OLS quantile forecasts indicates that OLS is unable to forecast extreme quantiles (see the Appendix for details). As the quantiles corresponding with down-side risk are often considered the most interesting by for example investors, this indicates that forecasting future quantiles in this manner may not be very useful in practice.

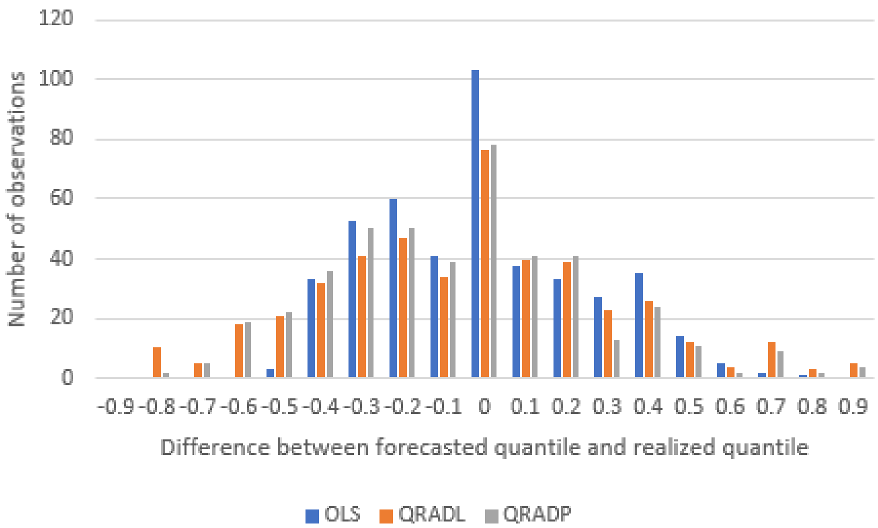

Figure 3: Difference between the forecasted quantile and the realized quantile for efficient quantile forecasts.
(Note that only every 0.10 quantile difference is considered, meaning that the 0.05 quantiles are rounded up)

6 Discussion

In this research I first tested for slope homogeneity of several firm characteristics in different quantiles of the firms' returns using the FE-QR model and Swamy-type tests as proposed by Galvao et al. (2018). I found empirical evidence for the existence of slope heterogeneity for both the higher and lower quantiles. For the central quantiles on the other hand, I found no such evidence. This result is in line with the findings of Galvao et al. (2018) and indicates that different firms' returns are impacted differently by their firm characteristics especially during times of financial booms and bursts, which might be due to underlying investor sentiments. This means that for the central quantiles the pooling of data is probably more appropriate as this results in more efficient estimates, whereas for the tail quantiles it might be better to use the individual slope estimates to have more flexibility in the model and to accommodate for slope heterogeneity across firms, as a way to make use of the efficiency-flexibility trade-off described by Zhang et al. (2019).

I then used these results to construct efficient-flexible quantile forecasts. With these quantile forecasts I calculated point forecasts with the QRADL and QRADP methods based on the work of Ma & Pohlman (2008), and compared these forecasts with the OLS forecasts. The OLS model outperforms both the QRADL and QRADP model in terms of forecasting for this dataset. By further exploring the implications of using efficient and flexible quantile forecasts and constructing the QRADP model based on these respective quantile forecasts separately, I found that the QRADP model with efficient estimates actually performed better than the models with more flexible estimates, although still worse than OLS. This might have been due to the problem of the (efficient-)flexible quantile forecasts lack of monotonicity for some of the quantile forecasts. Another possibility is that the efficient forecasts are better in forecasting relatively stable periods, such as the 2012-2019 period. Lastly, I found that all these models forecast the future returns too low, which is presumably due to the returns during this period being relatively high as the economy was doing good and did not include any major crisis.

After the point forecasting, I tried to forecast the future quantiles. Again I noted that using the (efficient-)flexible quantile forecasts lead to a problem, as due to the lack of monotonicity relatively many observations fell in the lowest and highest quantile. Thus, using the efficient quantile estimates for this purpose seems to be more reasonable, as the 'realized returns' are more smoothly distributed. I found that OLS predicts about 25% of the time the right quantile in this case, and the maximum error stays relatively low. However, on average it underpredicts the true quantile, again probably because of the healthy state of the economy, and makes an average absolute error of 0.22, which seems relatively large. Moreover, the OLS model seems to be unable to predict the extreme quantiles, which presumably are the most interesting quantiles for investors, indicating that this method might not be very useful in practice.

However, the data considered in this research contained some possible errors that might have affected the results in this paper. First of all, the excess returns used in this paper are relatively large, averaging about 10% compared to for example an average excess return 0.3% of Galvao et al. (2018). This might be partially due to some kind of survivorship bias, as the sample of 50 years might favor larger firms with higher returns. Nonetheless, even with a survivorship bias, returns should probably not be this high as Galvao et al. (2018) used a sample including 42 years of data as well. Furthermore, in case of survivorship bias, Barber & Lyon (1997) showed that this should not significantly affect the estimated effect of financial variables on returns. In addition, as discussed by Galvao et al. (2018), the FE-QR model is subject to the incidental parameter problem, meaning that these estimates might be slightly biased, which could have affected all the results throughout the paper. Therefore, some caution should be taken when using these results.

7 Conclusion

All in all, in this research I tried to answer the question "Can I detect slope heterogeneity of firms characteristics for the returns in tail quantiles, and to what extent can I use these findings to create reliable return point forecasts and future quantile forecasts?". I have shown that there indeed seems to be slope heterogeneity in tail quantiles for the effects of several firm characteristics on the returns. However, innovating on the return point forecast methods proposed by Ma & Pohlman (2008) by using efficient-flexible quantile forecasts based on the Swamy test results did not improve forecasting performance. Besides, I found that the simple OLS model actually performs better in terms of forecasting than the quantile based forecasting methods. Finally, for future quantile forecasting I found that it might be hard to forecast the extreme future quantiles.

For future research it might be interesting to see whether using efficient-flexible quantile forecasts can be useful for forecasting in other applications, or in cases where the monotonicity problem does not occur. Additionally, it might be useful to consider some other forecasting models than QRADL and QRADP that make better use of the available information about the future quantiles. Furthermore, it might be interesting to see whether using the marginal Swamy test results of the individual firm characteristics, instead of the joint Swamy test results, is able to improve this way of constructing efficient-flexible quantile forecasts in an attempt to make even better use of the efficiency-flexibility trade-off.

References

Asness, C. S. (1995). The power of past stock returns to explain future stock returns. *Available at SSRN 2865769*.

Barber, B. M., & Lyon, J. D. (1997). Firm size, book-to-market ratio, and security returns: A holdout sample of financial firms. *The Journal of Finance*, *52*(2), 875–883.

Baruník, J., & Čech, F. (2020). Measurement of common risks in tails: A panel quantile regression model for financial returns. *Journal of Financial Markets*, 100562.

Browning, M., Carro, J., et al. (2007). Heterogeneity and microeconometrics modeling. *Econometric Society Monographs*, *43*, 47.

Cech, F., & Barunik, J. (2017). Measurement of common risk factors: A panel quantile regression model for returns. *Available at SSRN 3028488*.

Chen, M.-Y., & Chen, J.-E. (2002). Application of quantile regression to estimation of value at risk. *Review of Financial Risk Management*, *1*(2), 15.

Conerly, B. (2019). *U.s. economy is stable: The great moderation returns.* Retrieved from `https://www.forbes.com/sites/billconerly/2019/07/19/us-economy-is-stable-the-great-moderation-returns/#64fac9884675` (accessed: 27.06.2020)

Diebold, F. X., & Mariano, R. S. (2002). Comparing predictive accuracy. *Journal of Business & economic statistics*, *20*(1), 134–144.

Fama, E. F., & French, K. R. (1989). Business conditions and expected returns on stocks and bonds. *Journal of financial economics*, *25*(1), 23–49.

Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. *the Journal of Finance*, *47*(2), 427–465.

Ferreira, M. A., & Santa-Clara, P. (2011). Forecasting stock market returns: The sum of the parts is more than the whole. *Journal of Financial Economics*, *100*(3), 514–537.

Galvao, A. F., Juhl, T., Montes-Rojas, G., & Olmo, J. (2018). 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., & Wang, L. (2015). Efficient minimum distance estimator for quantile regression fixed effects panel data. *Journal of Multivariate Analysis*, *133*, 1–26.

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

Güloğlu, B., Uyar, S. G. K., & Uyar, U. (2016). Dynamic quantile panel data analysis of stock returns predictability. *International Journal of Economics and Finance*, *8*(2).

Hall, P., & Sheather, S. J. (1988). On the distribution of a studentized quantile. *Journal of the Royal Statistical Society: Series B (Methodological)*, *50*(3), 381–391.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica: Journal of the econometric society*, 1251–1271.

Heij, C., Heij, C., de Boer, P., Franses, P. H., Kloek, T., van Dijk, H. K., et al. (2004). *Econometric methods with applications in business and economics.* Oxford University Press.

Hsiao, C., & Sun, B.-H. (2000). To pool or not to pool panel data. In *Panel data econometrics: Future directions.*

Jegadeesh, N., & Titman, S. (2011). Momentum. *Annu. Rev. Financ. Econ.*, *3*(1), 493–509.

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

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

Koenker, R., & Bassett Jr, G. (1978). Regression quantiles. *Econometrica: journal of the Econometric Society*, 33–50.

Koenker, R., & Hallock, K. F. (2001). Quantile regression. *Journal of economic perspectives*, *15*(4), 143–156.

Koenker, R., & Ng, P. (2005). A frisch-newton algorithm for sparse quantile regression. *Acta Mathematicae Applicatae Sinica*, *21*(2), 225–236.

Ma, L., & Pohlman, L. (2008). Return forecasts and optimal portfolio construction: a quantile regression approach. *The European Journal of Finance*, *14*(5), 409–425.

MATLAB. (1984). The mathworks, r2018a, software available at `https://nl.mathworks.com/products/matlab.html`.

Mello, M., & Perrelli, R. (2003). Growth equations: a quantile regression exploration. *The Quarterly Review of Economics and Finance*, *43*(4), 643–667.

Neyman, J., & Scott, E. L. (1948). Consistent estimates based on partially consistent observations. *Econometrica: Journal of the Econometric Society*, 1–32.

Ni, Z.-X., Wang, D.-Z., & Xue, W.-J. (2015). Investor sentiment and its nonlinear effect on stock returns—new evidence from the chinese stock market based on panel quantile regression model. *Economic Modelling*, *50*, 266–274.

Pesaran, M. H., & Yamagata, T. (2008). Testing slope homogeneity in large panels. *Journal of econometrics*, *142*(1), 50–93.

Phillips, P. C., & Sul, D. (2003). Dynamic panel estimation and homogeneity testing under cross section dependence. *The Econometrics Journal*, *6*(1), 217–259.

Powell, J. L. (1991). Estimation of monotonic regression models under quantile restrictions. *Nonparametric and semiparametric methods in Econometrics*, 357–384.

Rapach, D., & Zhou, G. (2013). Forecasting stock returns. In *Handbook of economic forecasting* (Vol. 2, pp. 328–383). Elsevier.

Rapach, D. E., Strauss, J. K., & Zhou, G. (2010). Out-of-sample equity premium prediction: Combination forecasts and links to the real economy. *The Review of Financial Studies*, *23*(2), 821–862.

Shephard, N., & Sheppard, K. (2010). Realising the future: forecasting with high-frequency-based volatility (heavy) models. *Journal of Applied Econometrics*, *25*(2), 197–231.

Swamy, P. A. (1970). Efficient inference in a random coefficient regression model. *Econometrica: Journal of the Econometric Society*, 311–323.

U.S. Bureau of Labor Statistics. (1884). *Cpi for all urban consumers (cpi-u).* Retrieved from `https://data.bls.gov/pdq/SurveyOutputServlet` (accessed: 12.05.2020)

Van Dijk, D., & Franses, P. H. (2003). Selecting a nonlinear time series model using weighted tests of equal forecast accuracy. *Oxford Bulletin of Economics and Statistics*, *65*, 727–744.

Welch, I., & Goyal, A. (2008). A comprehensive look at the empirical performance of equity premium prediction. *The Review of Financial Studies*, *21*(4), 1455–1508.

Wharton Research Data Services. (1993). *Crsp/compustat merged fundamentals annual.* Retrieved from `https://wrds-www.wharton.upenn.edu`

Xu, P. (2020). *Quantile regression.* Retrieved from `https://www.github.com/zjph602xtc/Quantile_reg` (accessed: 24.05.2020)

Zhang, Y., Wang, H. J., & Zhu, Z. (2019). Quantile-regression-based clustering for panel data. *Journal of Econometrics*, *213*(1), 54–67.

Appendix

Six Variable Specification Homogeneity Tests Results

Table 8: p-values for Swamy ($\hat{S}$) and standardized Swamy ($\hat{\Delta}$) test for slope homogeneity of the six variable FE-QR specification

	Joint		**MDR**		**EBIT**		**MBR**		**DA**		**LNA**		**PPE**	
τ	$\hat{S}$	$\hat{\Delta}$	$\hat{S}$	$\hat{\Delta}$	$\hat{S}$	$\hat{\Delta}$	$\hat{S}$	$\hat{\Delta}$	$\hat{S}$	$\hat{\Delta}$	$\hat{S}$	$\hat{\Delta}$	$\hat{S}$	$\hat{\Delta}$
0.05	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.10	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.15	0.000	0.000	0.000	0.000	0.000	0.000	0.006	0.004	0.002	0.001	0.000	0.000	0.005	0.002
0.20	0.000	0.000	0.014	0.010	0.000	0.000	0.098	0.111	0.008	0.005	0.000	0.000	0.002	0.000
0.25	0.000	0.000	0.250	0.296	0.018	0.014	0.054	0.056	0.224	0.266	0.001	0.000	0.055	0.057
0.30	0.163	0.229	0.051	0.053	0.448	0.511	0.107	0.122	0.583	0.641	0.049	0.050	0.370	0.430
0.35	0.817	0.871	0.233	0.277	0.845	0.862	0.315	0.370	0.948	0.945	0.145	0.170	0.821	0.843
0.40	0.996	0.996	0.318	0.374	0.938	0.936	0.503	0.566	0.962	0.957	0.404	0.466	0.849	0.866
0.45	1.000	1.000	0.595	0.652	0.944	0.941	0.690	0.736	0.995	0.989	0.585	0.643	0.969	0.963
0.50	1.000	1.000	0.627	0.681	0.764	0.798	0.845	0.863	0.988	0.981	0.540	0.601	0.954	0.949
0.55	0.997	0.997	0.326	0.382	0.752	0.788	0.741	0.779	0.977	0.970	0.382	0.443	0.849	0.866
0.60	0.912	0.939	0.213	0.252	0.618	0.673	0.872	0.884	0.992	0.986	0.201	0.239	0.689	0.735
0.65	0.325	0.422	0.030	0.027	0.741	0.778	0.467	0.530	0.957	0.953	0.076	0.083	0.464	0.527
0.70	0.000	0.000	0.000	0.000	0.373	0.433	0.115	0.132	0.805	0.831	0.003	0.001	0.064	0.068
0.75	0.000	0.000	0.000	0.000	0.543	0.603	0.012	0.009	0.000	0.000	0.000	0.000	0.000	0.000
0.80	0.000	0.000	0.000	0.000	0.371	0.431	0.006	0.003	0.000	0.000	0.000	0.000	0.000	0.000
0.85	0.000	0.000	0.000	0.000	0.008	0.005	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.90	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.95	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Joint Quantile Estimates

In this Section, I briefly look at the joint slope estimates for some of the different quantiles to see how these estimates differ per quantiles. I start by evaluating the three variable specification. These joint quantile slope estimates are shown in Table 9. Note that these estimates might not be appropriate for all quantiles, as especially the tail qauntiles exhibit heterogeneity (see Section 5.1). I find that the market-to-book ratio in general has a negative effect on the returns which remains fairly constant over the quantiles. This indicates that firms with a relatively high market value are more likely to perform worse next period and vice versa. This might be due to investors knowing that the firm is actually overpriced on the market which will make the price drop. Furthermore, the market returns have no significant effect at 5% significance level for the 0.25 and 0.50 quantiles. For the higher quantiles, I see that in general it has a negative impact on the firms' returns. This is somewhat counter intuitive, as one might expect that high market returns lead to more confidence among investors, hence to higher returns for individual firms as well. However, as these effects differ significantly among firms for the higher quantiles according to the Swamy tests, this does not mean that this effect is negative for all firms

Table 9: Quantile estimates for the three variable specification for the period 1970-2011

τ	0.10	0.25	0.50	0.75	0.90
MBR	-0.070	-0.048	-0.052	-0.039	-0.044
	(0.000)	(0.000)	(0.000)	(0.000)	(0.010)
LNA	-0.024	0.002	-0.001	-0.030	-0.026
	(0.036)	(0.838)	(0.952)	(0.011)	(0.259)
MKR	0.172	-0.014	-0.072	-0.173	-0.310
	(0.000)	(0.733)	(0.065)	(0.000)	(0.000)

Note: p-values are between brackets and based on the Wald test

Now I briefly discuss the joint slope estimates for the different quantiles as shown in Table 11 and 10 for both the six and seven variable specification respectively. I find that the market returns again negatively affect the returns significantly for the higher quantiles, whereas they do not affect the returns significantly for the lower central quantiles. Furthermore, by comparing the six variable specification estimates and the seven variable specification estimates it can be seen that when omitting the market return as a variable, the other estimates change quite a bit. This indicates that some of its effects are then contained in the other variables, which might add to the heterogeneity, which was found more in the six variable specification then in the six variable specification.

Table 10: Quantile estimates for the six variable specification for the period 1970-2011

tau	0.10	0.25	0.50	0.75	0.90
MDR	0.287	0.386	0.312	0.427	0.771
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
EBIT	-0.109	0.118	-0.207	-0.124	0.195
	(0.492)	(0.499)	(0.118)	(0.411)	(0.316)
MBR	-0.068	-0.041	-0.037	-0.022	-0.035
	(0.000)	(0.000)	(0.000)	(0.029)	(0.008)
DA	2.301	2.298	1.155	1.850	2.173
	(0.003)	(0.007)	(0.076)	(0.013)	(0.023)
LNA	-0.004	-0.016	-0.037	-0.056	-0.088
	(0.779)	(0.319)	(0.003)	(0.000)	(0.000)
PPE	0.119	-0.264	-0.347	-0.371	-0.495
	(0.351)	(0.059)	(0.001)	(0.002)	(0.002)

Note: p-values are between brackets and
based on the Wald test

Table 11: Quantile estimates for the seven variable specification for the period 1970-2011

tau	0.10	0.25	0.50	0.75	0.90
MDR	0.315	0.386	0.285	0.390	0.686
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
EBIT	-0.063	0.123	-0.265	-0.276	-0.087
	(0.658)	(0.493)	(0.048)	(0.071)	(0.680)
MBR	-0.060	-0.041	-0.037	-0.008	-0.007
	(0.000)	(0.001)	(0.000)	(0.412)	(0.600)
DA	2.462	2.405	1.290	1.650	2.525
	(0.000)	(0.006)	(0.048)	(0.026)	(0.014)
LNA	-0.014	-0.019	-0.034	-0.052	-0.074
	(0.298)	(0.260)	(0.006)	(0.000)	(0.000)
PPE	0.144	-0.292	-0.355	-0.308	-0.600
	(0.200)	(0.042)	(0.001)	(0.011)	(0.000)
MKR	0.158	0.024	-0.049	-0.166	-0.276
	(0.000)	(0.644)	(0.210)	(0.000)	(0.000)

Note: p-values are between brackets and
based on the Wald test

Point Forecasting

Diebold-Mariano Test Statistics seven variable specification

Table 12: Diebold-Mariano test statistic and corresponding p-values for the forecasting period 2012-2019 for the seven variable specification

Model	DM-statistic	p-value
OLS-QRADL	-4.623	0.000
OLS-QRADP	-5.728	0.000
QRADP-QRADL	-3.746	0.000
OLS-QRADP (eff)	-4.279	0.000
OLS-QRADP (flex)	-6.427	0.000
QRADL-QRADP (eff)	4.409	1.000
QRADL-QRADP (flex)	0.633	0.737
QRADP-QRADP (eff)	5.169	1.000
QRADP-QRADP (flex)	-6.419	0.000
QRADP (eff)-QRADP (flex)	-6.124	0.000

Table 13: Diebold-Mariano test statistic and corresponding p-values for the in-sample period 1970-2011 for the seven variable specification

Model	DM-statistic	p-value
OLS-QRADL	-6.054	0.000
OLS-QRADP	7.899	1.000
QRADP-QRADL	-8.568	0.000
OLS-QRADP (eff)	3.138	0.999
OLS-QRADP (flex)	6.687	1.000
QRADL-QRADP (eff)	6.670	1.000
QRADL-QRADP (flex)	8.539	1.000
QRADP-QRADP (eff)	-6.576	0.000
QRADP-QRADP (flex)	3.539	1.000
QRADP (eff)-QRADP (flex)	5.738	1.000

Point forecasting for Three and Six Variable Specification

Table 14: Forecasting performance measures three variable specification

Model	Bias	MSPE	MAPE
OLS	-0.051	0.067	0.196
QRADL	-0.079	0.110	0.251
QRADP	-0.086	0.075	0.207
QRADP (efficient)	-0.060	0.069	0.199
QRADP (flexible)	-0.126	0.104	0.243

Table 15: In sample performance measures three variable specification

Model	Bias	MSE	MAE
OLS	0.000	0.118	0.240
QRADL	-0.010	0.197	0.300
QRADP	-0.004	0.109	0.232
QRADP (efficient)	-0.010	0.117	0.236
QRADP (flexible)	-0.003	0.105	0.227

Table 16: Diebold-Mariano test statistic and corresponding p-values for the forecasting period 2012-2019 for the three variable specification

Model	DM-statistic	p-value
OLS-QRADL	-6.227	0.000
OLS-QRADP	-3.360	0.000
QRADP-QRADL	-5.610	0.000
OLS-QRADP (eff)	-1.589	0.056
OLS-QRADP (flex)	-6.254	0.000
QRADL-QRADP (eff)	6.122	1.000
QRADL-QRADP (flex)	0.917	0.820
QRADP-QRADP (eff)	3.005	0.999
QRADP-QRADP (flex)	-6.954	0.000
QRADP (eff)-QRADP (flex)	-6.135	0.000

Table 17: Diebold-Mariano test statistic and corresponding p-values for the in-sample period 1970-2011 for the three variable specification

Model	DM-statistic	p-value
OLS-QRADL	-9.398	0.000
OLS-QRADP	4.257	1.000
QRADP-QRADL	-9.812	0.000
OLS-QRADP (eff)	1.963	0.975
OLS-QRADP (flex)	4.364	1.000
QRADL-QRADP (eff)	9.935	1.000
QRADL-QRADP (flex)	9.540	1.000
QRADP-QRADP (eff)	-3.411	0.000
QRADP-QRADP (flex)	3.033	0.999
QRADP (eff)-QRADP (flex)	3.674	1.000

Table 18: Forecasting performance measures six variable specification

Model	Bias	MSPE	MAPE
OLS	-0.036	0.068	0.197
QRADL	-0.095	0.213	0.319
QRADP	-0.095	0.124	0.257
QRADP (efficient)	-0.071	0.075	0.206
QRADP (flexible)	-0.098	0.202	0.323

Table 19: In sample performance measures six variable specification

Model	Bias	MSE	MAE
OLS	0.000	0.119	0.240
QRADL	0.004	0.183	0.282
QRADP	-0.005	0.093	0.215
QRADP (efficient)	-0.010	0.115	0.235
QRADP (flexible)	-0.002	0.088	0.208

Table 20: Diebold-Mariano test statistic and corresponding p-values for the forecasting period 2012-2019 for the six variable specification

Model	DM-statistic	p-value
OLS-QRADL	-5.296	0.000
OLS-QRADP	-5.587	0.000
QRADP-QRADL	-4.141	0.000
OLS-QRADP (eff)	-3.869	0.000
OLS-QRADP (flex)	-6.608	0.000
QRADL-QRADP (eff)	5.046	1.000
QRADL-QRADP (flex)	0.632	0.736
QRADP-QRADP (eff)	5.140	1.000
QRADP-QRADP (flex)	-6.922	0.000
QRADP (eff)-QRADP (flex)	-6.350	0.000

Table 21: Diebold-Mariano test statistic and corresponding p-values for the in-sample period 1970-2011 for the six variable specification

Model	DM-statistic	p-value
OLS-QRADL	-6.821	0.000
OLS-QRADP	7.482	1.000
QRADP-QRADL	-8.855	0.000
OLS-QRADP (eff)	3.527	1.000
OLS-QRADP (flex)	6.449	1.000
QRADL-QRADP (eff)	7.406	1.000
QRADL-QRADP (flex)	8.740	1.000
QRADP-QRADP (eff)	-6.208	0.000
QRADP-QRADP (flex)	3.327	1.000
QRADP (eff)-QRADP (flex)	5.496	1.000

Value-at-Risk unconditional coverage

Table 22: Unconditional coverage Likelihood Ratio test for Value-at-Risk efficient-flexible quantile estimates

quantile	$\hat{\pi}$	LR statistic	p-value
0.05	0.031	3.096	0.000
0.10	0.154	12.684	0.000
0.15	0.179	2.732	0.098

Table 23: Unconditional coverage Likelihood Ratio test for Value-at-Risk efficient quantile estimates

quantile	$\hat{\pi}$	LR statistic	p-value
0.05	0.031	3.805	0.051
0.10	0.054	12.696	0.000
0.15	0.0826	18.574	0.000

Table 24: Unconditional coverage Likelihood Ratio test for Value-at-Risk flexible quantile estimates

quantile	$\hat{\pi}$	LR statistic	p-value
0.05	0.141	53.096	0.000
0.10	0.154	12.684	0.000
0.15	0.179	2.732	0.098

Distributions of forecasted quantiles

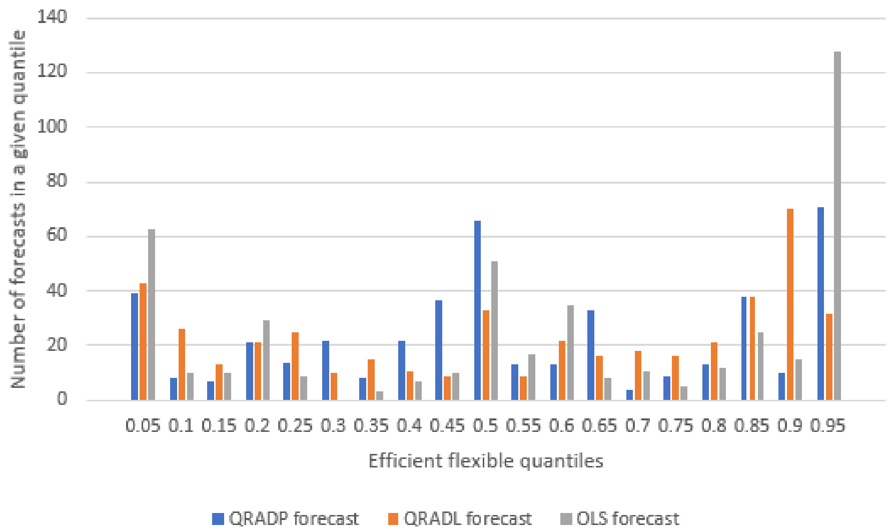

Figure 4: Forecasted efficient-flexible quantiles

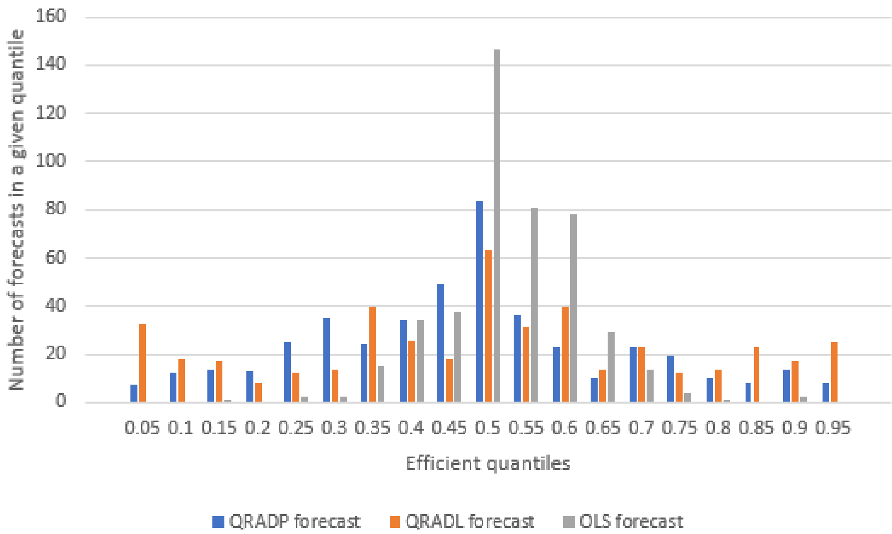

Figure 5: Forecasted efficient quantiles

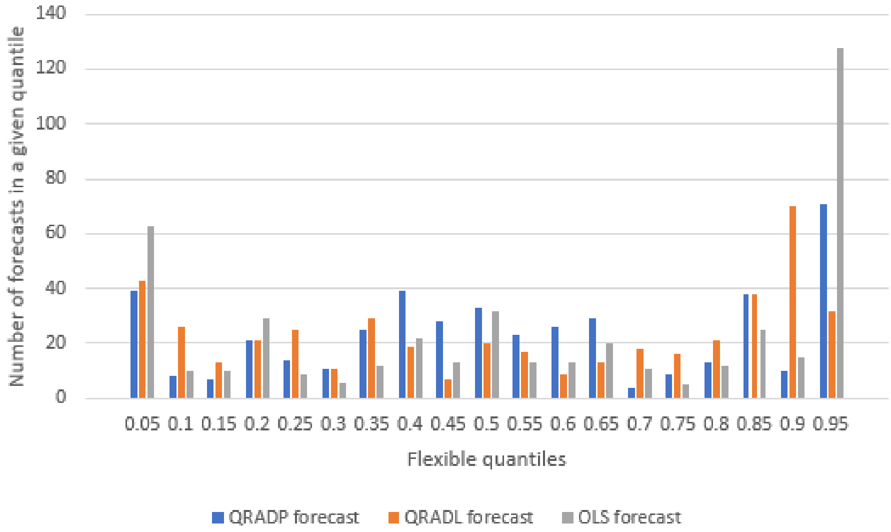

Figure 6: Forecasted flexible quantiles