

Der Open-Access-Publikationsserver der ZBW – Leibniz-Informationzentrum Wirtschaft  
*The Open Access Publication Server of the ZBW – Leibniz Information Centre for Economics*

Baur, Dirk G.; Dimpfl, Thomas; Jung, Robert C.

Working Paper

## Stock return autocorrelations revisited: A quantile regression approach

University of Tübingen working papers in economics and finance, No. 24

**Provided in cooperation with:**

Eberhard Karls Universität Tübingen

Suggested citation: Baur, Dirk G.; Dimpfl, Thomas; Jung, Robert C. (2012) : Stock return autocorrelations revisited: A quantile regression approach, University of Tübingen working papers in economics and finance, No. 24, urn:nbn:de:bsz:21-opus-59769 , <http://hdl.handle.net/10419/54730>

**Nutzungsbedingungen:**

Die ZBW räumt Ihnen als Nutzerin/Nutzer das unentgeltliche, räumlich unbeschränkte und zeitlich auf die Dauer des Schutzrechts beschränkte einfache Recht ein, das ausgewählte Werk im Rahmen der unter

→ <http://www.econstor.eu/dspace/Nutzungsbedingungen> nachzulesenden vollständigen Nutzungsbedingungen zu vervielfältigen, mit denen die Nutzerin/der Nutzer sich durch die erste Nutzung einverstanden erklärt.

**Terms of use:**

*The ZBW grants you, the user, the non-exclusive right to use the selected work free of charge, territorially unrestricted and within the time limit of the term of the property rights according to the terms specified at*

→ <http://www.econstor.eu/dspace/Nutzungsbedingungen>  
*By the first use of the selected work the user agrees and declares to comply with these terms of use.*

University of Tübingen  
Working Papers in  
Economics and Finance

No. 24

Stock Return Autocorrelations Revisited:  
A Quantile Regression Approach

by

Dirk G. Baur, Thomas Dimpfl & Robert C. Jung

Faculty of Economics and Social Sciences  
[www.wiwi.uni-tuebingen.de](http://www.wiwi.uni-tuebingen.de)



# Stock Return Autocorrelations Revisited: A Quantile Regression Approach

Dirk G. Baur

University of Technology Sydney, Australia

Thomas Dimpfl\*

Universität Tübingen, Germany

Robert C. Jung

Universität Erfurt, Germany

October 29, 2011

## Abstract

The aim of this study is to provide a comprehensive description of the dependence pattern of stock returns by studying a range of quantiles of the conditional return distribution using quantile autoregression. This enables us in particular to study the behavior of extreme quantiles associated with large positive and negative returns in contrast to the central quantile which is closely related to the conditional mean in the least-squares regression framework. Our empirical results are based on 30 years of daily, weekly and monthly returns of the stocks comprised in the Dow Jones Stoxx 600 index. We find that lower quantiles exhibit positive dependence on past returns while upper quantiles are marked by negative dependence. This pattern holds when accounting for stock specific characteristics such as market capitalization, industry, or exposure to market risk.

*JEL classification:* C22, G14

*Keywords:* stock return distribution; quantile autoregression; overreaction and underreaction

---

\*Corresponding author: Wirtschafts- und Sozialwissenschaftliche Fakultät, Mohlstraße 36, 72074 Tübingen, Germany; Tel. +49 70 71 29 76 417; Mail: thomas.dimpfl@uni-tuebingen.de

# 1 Introduction

A widely studied time series property of stock returns is their autocorrelation which is a particularly important feature as it implies, if present, predictability of stock returns. In empirical studies a wide range of results is reported including: positive (and possibly significant) autocorrelation, no autocorrelation, as well as negative (and possibly significant) autocorrelation. Campbell et al. (1997), for example, report significant positive autocorrelation for daily, weekly and monthly stock index returns calculated from the CRSP dataset (Center for Research in Security Prices, comprising all NYSE, AMEX and NASDAQ common stocks) for a sample period spanning from 1962 to 1994. They find daily autocorrelations to be slightly higher (around 0.2 for a value-weighted index and around 0.3 for an equally-weighted index) than weekly and monthly return autocorrelations. Lo and MacKinlay (1990) also find mostly positive autocorrelation in daily stock returns which they attribute to non-synchronous trading or non-trading. Campbell et al. (1993) investigate the relationship between trading volume and serial correlation in daily returns of individual stocks and indices. When controlling for volume they find stock returns to be negatively autocorrelated. Lewellen (2002) studies momentum and autocorrelation of stock returns based on the CRSP database for a sample period from 1941 to 1999 on a monthly basis and finds that stock returns are generally negatively autocorrelated, albeit the correlation is not statistically significant in most cases.

A common feature of all these studies is that they confine the analysis to the conditional mean of the return distribution to examine the influence of a lagged return on the current return. In contrast to this, we intend to provide a comprehensive description of the full return distribution conditioned on a lagged return. For this purpose, we use the quantile regression framework as introduced by Koenker and Bassett (1978) and examine the influence of a lagged return on all quantiles of the current return. Quantile regression has been successfully applied in various fields of empirical finance. An early contribution is Engle and Manganelli (2004) who develop a quantile regression based method for Value-at-Risk estimation. Feng et al. (2008) and Ma and Pohlman (2008) introduce quantile momentum measures to construct momentum portfolios in asset management. Chuang et al. (2009) examine the dynamic relationship of the quantiles of stock market index

returns and trading volume and find causal effects of volume to be heterogeneous across quantiles.

This paper proposes a novel approach to investigate the predictability of the various parts of the conditional return distribution in a linear, autoregressive framework. In particular, we expect the autoregression to be different over the various quantiles. Our expectation is based upon various findings in the theoretical as well as the empirical finance literature. For example, Campbell et al. (2008) have provided empirical evidence that cross-correlations between stock return indices vary systematically across quantiles. Veronesi (1999) develops an intertemporal, rational expectations equilibrium model of asset pricing. He finds that stock markets overreact to bad news in good times and underreact to good news in bad times. This state dependence could be captured by the different quantiles of the conditional return distribution: a good (bad) state can be associated with upper (lower) quantiles. This suggests that the impact of lagged returns is different across quantiles. The behavioral finance literature has developed models which explain the empirical phenomena of “momentum” and “reversal” which are attributed to under- and overreaction of stock prices to good or bad news (see, for example, Barberis et al., 1998; Lewellen, 2002). As the consequences of under- and overreaction for the stock price are different, this may lead to differing dependence patterns across the return distribution.

Our empirical findings are based on a sample of daily, weekly and monthly returns of 600 major European stocks comprised in the Dow Jones Stoxx 600 index from 1979 to 2009 and support the conjectures outlined above. We find a distinctive, s-shaped dynamic pattern across quantiles of the conditional stock return distribution. In particular, we find lower quantiles to exhibit positive dependence with past returns while upper quantiles are marked by negative dependence. Typically, we find no or only very weak dependence for central quantiles. This pattern is robust with respect to different data frequencies (daily, weekly, monthly) and is found to be the same for individual stocks as well as the Dow Jones Stoxx 600 index and portfolios sorted according to size, market risk and industry, respectively. We also investigate the importance of the sign and the size of a previous period’s return on the current return distribution. When conditioning on extreme past returns we find that positive dependence is more pronounced in lower quantiles while it disappears in upper quantiles. With respect to the sign of past returns we find that

positive previous day's returns lead to strong positive correlation with today's positive returns and marked negative correlation with today's negative returns. The opposite pattern is visible for past negative returns.

The remainder of the paper is structured as follows. Section two provides a brief introduction to the quantile regression method and describes the data used for the analysis. Section three presents the estimation results and discusses the findings while Section four summarizes the findings.

## 2 Models and Data

### 2.1 The Quantile Autoregressive Model

Ordinary least squares (OLS) regression based on the linear model  $y = \beta'x + \varepsilon$  focuses on the estimation of the conditional mean of the dependent variable  $y$  given the explanatory variable(s)  $x$ . In contrast, quantile regression as introduced by Koenker and Bassett (1978) provides a technique to estimate conditional quantiles of  $y$  given one or more explanatory variables. Inter alia, it allows for a more precise description of the tails of the distribution of  $y$ . Moreover, it is robust to heteroskedasticity, skewness and leptokurtosis which are common features of financial data.

For the purpose of this study we fit three different quantile regression models. The basic model is simply a first-order quantile autoregressive model (QAR(1), see Koenker and Xiao, 2006). The second and the third models are extensions thereof designed to capture possible asymmetric influences of lagged returns in the spirit of threshold GARCH models (see, for example, Glosten et al., 1993).

More precisely, we fit the following QAR(1) model:

$$Q(\tau|\mathcal{F}_{t-1}) = \alpha_i(\tau) + \beta_i(\tau)r_{t-1,i} \quad (1)$$

where  $r_{t,i}$  is the return of stock  $i$  on day  $t$ ,  $Q(\tau|\mathcal{F}_{t-1})$  denotes the  $\tau$ -th conditional quantile of stock  $i$ 's return and  $\mathcal{F}_{t-1}$  is the lagged past return of the respective stock. The

coefficients  $\beta_i(\tau)$  may be interpreted as quantile specific autoregressive parameters which are the focus of this study.

We extend the basic model specification in Equation (1) to specifically account for the size of the lagged return. The model is specified as

$$Q(\tau|\mathcal{F}_{t-1}) = \alpha_i(\tau) + \beta_i(\tau)r_{t-1,i} + \gamma_i(\tau)r_{t-1,i}I(|r_{t-1,i}| > r^q) \quad (2)$$

where the indicator variable  $I(|r_{t-1,i}| > r^q)$  is equal to one if the return of stock  $i$  in period  $t-1$  exceeds a certain threshold  $r^q$  and zero otherwise.  $r^q$  is chosen to be the 95% quantile of the unconditional distribution of absolute returns of the respective stock. We will refer to this specification as 'Size Model'. Its purpose is to assess the influence of both extreme positive as well as extreme negative previous day's returns (measured by  $\gamma_i(\tau)$ ).

Another extension of the basic model seeks to capture the role of the sign of the previous period's return (henceforth 'Sign Model'). It reads as follows:

$$Q(\tau|\mathcal{F}_{t-1}) = \alpha_i(\tau) + \beta_i(\tau)r_{t-1,i} + \delta_i(\tau)r_{t-1,i}I(r_{t-1,i} < 0). \quad (3)$$

The indicator variable  $I(r_{t-1,i} < 0)$  is equal to one if the previous day's return is negative and zero otherwise.

It is straightforward to estimate all three models using the standard optimization routine of Koenker and Bassett (1978) as implemented, for example, in R. Asymptotic standard errors are estimated using bootstrap methods. As we use time series data, we perform a block bootstrap with fixed block length of 25 observations and 600 replications (see, for example, Chernick, 2008, ch. 5).

We seek to contrast the autoregressive parameter estimates of the quantile autoregression models with those implied by a standard conditional mean approach. We therefore also estimate linear regression models corresponding to Equations (1), (2) and (3) using OLS.

## 2.2 Data - Summary and Descriptive Statistics

Our data cover closing prices of 600 European stocks which are used as of March 2009 to construct the Dow Jones Stoxx 600 index, as well as the index itself. The individual

companies which make up the index are based in Austria, Belgium, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Luxembourg, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland and the United Kingdom. The sample period spans roughly 32 years from January 1, 1987, to March 18, 2009. The maximum number of observations is  $T = 5794$  for daily data,  $T = 1158$  for weekly data and  $T = 265$  for monthly data. The sample size for individual stocks may, however, be lower due to the fact that they were only listed after January 1987.

Returns used in this study are calculated as logarithmic stock price differences or logarithmic stock index differences. Daily returns are calculated as close to close returns. Weekly returns are measured from Wednesday close to Wednesday close the following week. If on Wednesday the market was closed for trading, we use the last available closing value to construct the weekly return. Monthly returns are calculated as the return between the last closing prices in two consecutive months. Key statistics of the return data are summarized in Table 1. The average return across all stocks considered is close to zero for daily, weekly as well as monthly data. Extreme negative returns are found to be larger in absolute value compared to positive ones. This may not be surprising, however, due to the fact that our sampling period covers various crisis periods, in particular the global financial crisis which started in 2007.

<Insert Table 1 about here>

## 3 Empirical Results

### 3.1 Results for the Basic QAR(1) Model

The first model studied is the basic QAR(1) model given in Equation (1) for daily, weekly and monthly data. The parameter estimates  $\hat{\beta}_i(\tau)$  for all stocks  $i$  using daily data are summarized by means of box-plots for selected quantiles in the top left panel of Figure 1. For lower quantiles, we find in general positive autoregressive parameter estimates whereas for the upper quantiles, parameter estimates tend to be negative. Coefficient estimates for central quantiles as well as the OLS estimate are close to zero.

<Insert Figure 1 about here>



We find that in general lower and upper quantiles are significantly different from zero while central quantiles tend not to be. More precisely, autoregressive coefficient estimates for very extreme (1 and 99%) quantiles are less likely to be statistically different from zero than the remaining quantiles (2, 5, 10, and 90, 95, 98%).

We also perform various  $t$ -tests for each stock to test whether the estimated parameters in selected quantiles  $\tau$  are different from those in the median and from those in the corresponding  $(1 - \tau)$ -quantile. It turns out that due to the high variance in the 1 percent quantile, the autoregressive parameter estimate in the 1% quantile is only significantly different from the one in the 99% quantile in 20 out of 600 cases or different from the median estimate in 9 cases. The variation in the 99% quantile is only slightly lower such that the estimates are different from the median estimate in 12% of all cases. The 2 and 98 percent quantiles are statistically different for 24% of all stocks and different from the median in 21% or 15% of all cases, respectively. The 5 and 95 percent quantiles are statistically different for 34% of all stocks and different from the median in 21% or 25% of all cases, respectively. The 10 and 90 percent quantiles are statistically different for 31% of all stocks and different from the median in 5% or 38% of all cases, respectively.

As is also evident from Figure 1, the average estimated autoregressive parameter is decreasing from lower to upper quantiles. A possible explanation for this pattern is that the reaction of today's return on its most recent history changes with respect to the state of the individual stock, i.e. which side of the return distribution is affected. In order to illustrate the different impact of lagged returns on the current return, consider Figure 2 which shows the average regression lines of all individual daily stock returns for three different  $\tau$  values ( $\tau = 0.02$ ,  $\tau = 0.50$ , and  $\tau = 0.98$ ) together with the OLS regression line.

<Insert Figure 2 about here>

The Figure shows that the regression line for  $\tau = 0.50$  is virtually flat and almost identical with the OLS regression line. This implies that on average yesterday's return has no impact on today's return irrespective of its sign and size. However, as we move away from the 50% quantile towards estimates in the tails of the return distribution, the impact of lagged returns changes markedly. Consider the on average upward sloping regression line which corresponds to the positive autoregressive parameter for  $\tau = 0.02$ . Low quantiles of

today's return distribution contain negative returns which we associate with a bad state. If the stock is in a bad state today the dependence of today's return on yesterday's return is positive. Consider the following two situations: if yesterday's return was negative, today's negative return is even lower due to the positive autoregressive coefficient. Correspondingly, previous day's positive returns alleviate today's bad state to some extent. In contrast, the on average downward sloping regression line corresponds to the negative autoregressive parameter for  $\tau = 0.98$ . Upper quantiles of today's return distribution which are marked by positive returns are associated with a good state of the stock. So there is some reversal effect if the previous day's return was negative while a positive previous day's return dampens an extreme positive return today.

For weekly and monthly data the dependence patterns across quantiles are similar to the one found in daily data as is evident from the graphs in the top right and bottom left panels of Figure 1. Significance of the autoregressive parameters of the different quantiles is in the same range for weekly data as found for daily data while for monthly data significance is in general weaker.

As can be seen from Figure 1 (top right), for weekly data autoregressive coefficient estimates in the lower quantiles are similar to the ones obtained for daily data. In the upper quantiles, however, the average autoregressive coefficients are clearly lower than those based on daily data. Moreover, the dispersion across stocks of the estimated autoregressive parameters is considerably larger than that for daily data, in particular in the very extreme quantiles. With respect to monthly data, the autoregressive estimates are even more pronounced (larger coefficients in lower quantiles and even smaller coefficients in upper quantiles). The variation of cross-sectional (stock) estimates increases further with coefficients exceeding 0.6 in the lower tail and  $-0.3$  in the upper tail. This may be attributed to the fact that with lower sampling frequencies less and less observations are available for the quantile regression of the individual time series. Estimation of the extreme quantiles is particularly affected by this fact.

The results described so far indicate a pronounced s-shaped pattern of the autoregressive coefficient estimates which is stable across return frequencies. For daily as well as weekly or monthly data we find varying autoregressive parameters across different quantiles of the conditional return distribution: lower quantiles exhibit on average positive dependence on past returns while upper quantiles are marked by negative dependence. Our results for

the upper quantiles are in line with the findings of Lewellen (2002) who finds negative autocorrelation for monthly returns. He attributes the negative autocorrelation to an overreaction of investors to macroeconomic news in a good state. In this sense, our results would suggest an underreaction to macroeconomic news if the stocks are in a bad state. This interpretation is in line with the theoretical results of Veronesi (1999).

In contrast to the quantile regression method, OLS estimation of a linear autoregressive model corresponding to Equation (1) delivers a single autoregressive parameter only which is of course constant across all quantiles. For daily data, the OLS estimate of the autoregressive parameter is found to be slightly positive on average (see Figure 1, top left panel), albeit usually not (individually) statistically significant. For weekly data, we find that the average OLS autoregressive parameter has a negative sign (see Figure 1, top right panel), but is still not significant for most stocks. For monthly data, we find autoregressive parameter estimates to be on average positive (see Figure 1, bottom left panel) and significantly different from zero in more than 50% of all stocks considered.

A sensible question to ask when considering the pattern of the parameter estimates for the individual stocks is whether idiosyncratic factors determine its shape. To answer this question we consider a broad portfolio such that individual factors are diversified away. We therefore estimate the QAR(1) model for the Dow Jones Stoxx 600 index which is a value-weighted portfolio of 600 stocks. Our quantile regression results based on daily data are displayed in the top left panel of Figure 3. We find a pattern which is very similar to the (average) estimates of the individual stocks presented in Figure 1.

In Figure 3 the dots represent estimates of the slope coefficients  $\beta_i(\tau)$  in the QAR(1) model for  $\tau$  ranging from 0.01 to 0.99. The shaded area characterizes 95% confidence bands for the estimates for all quantiles. Clearly, the lower quantiles are marked by positive (and significant) dependence while the upper tail exhibits significantly negative dependence on past returns. For the central quantiles estimated autoregressive coefficients are small and in general not significant. OLS estimation on the other hand would suggest that there is no autoregression in the returns of the Dow Jones Stoxx 600 index for daily data.

The correspondence of the dynamic patterns found for individual stocks and for the index suggests that idiosyncratic factors seem not to be the driving force behind these results. This finding is also supported by additional quantile regression results based on a QAR(1)

model in Equation (1) augmented by the market return, represented by the return of the Dow Jones Stoxx 600 index.<sup>1</sup>

<Insert Figure 3 about here>

The top right and bottom left panels of Figure 3 present the autoregressive coefficient estimates for the index returns based on weekly and monthly data, respectively. The graphs depict a pattern which is very similar to the one found for the average individual stock return estimates.

### 3.2 Robustness of the dependence pattern

So far we have analyzed the autoregressive structure without accounting for the characteristics of individual stocks. In order to differentiate the results obtained, we proceed in two different ways. First, we construct five or ten equally-weighted portfolios of firms sorted according to size, market beta and industry, respectively, and estimate the QAR(1) model for the various portfolios. Second, we sort the stock specific quantile regression results  $\beta_i(\tau)$  according to size, market beta and industry, respectively, in five or ten groups. The first strategy is intended to provide information about the average development of a group of stocks with similar characteristics and allows for comparison with the results obtained for the Dow Jones Stoxx 600 index as well as with the results found by other studies which generally focus on portfolios. With the second strategy we seek to reveal the variation of the dynamic pattern within the different groups. These results will be represented in boxplots which allows for easy comparison with the unsorted estimation results as represented in Figure 1. The set of results presented below is restricted to daily data due to the fact that for weekly and monthly data qualitatively similar results are obtained.

We construct the five equally-weighted portfolios according to market capitalization (MCAP) distinguishing companies with an MCAP less than 1,000 million Euro (57 companies), between 1,000 and 2,000 million Euro (182 companies), 2,000 and 10,000 million Euro (256 companies), 10,000 and 50,000 million Euro (90 companies) and companies with an MCAP

---

<sup>1</sup>Detailed results are available from the authors upon request.

exceeding 50,000 million Euro (15 companies). Figure 4 summarizes the estimation results obtained from the QAR(1) model for the five portfolios.

It clearly emerges from all panels of Figure 4 that the autoregressive pattern found above holds also true for the sorted MCAP portfolios. We find only a slight decrease of the autoregressive estimates in the lower quantiles when moving from low MCAP to large MCAP portfolios. The autoregressive estimates decrease roughly by one half for quantiles lower than 5%. In contrast upper quantiles are less sensitive to MCAP.

Sorting the individual autoregressive coefficient estimates of the QAR(1) model according to MCAP into five groups of 120 stocks each leads to the boxplots in Figure 5. It is evident that overall a similar pattern as for the unsorted data is (as depicted in the top left panel of Figure 1) is present. We therefore conclude that the identified dynamic pattern is not related to the market capitalization of the individual stocks. Moreover, as the stocks contained in these five groups are homogeneous with respect to their MCAP, we can seek to interpret the width of the boxes in the various panels of Figure 5. Moving from the top left panel down to the bottom left panel it seems that autoregressive estimates for companies with high MCAP are more homogeneous, i.e. the variation of the estimates is lower than for companies with low MCAP.

<Insert Figure 4 about here>

<Insert Figure 5 about here>

A second way to sort the coefficient estimates is by industry. We use SIC Industry Sector Codes on the highest level to group our findings. Figure 6 presents five selected industry groups: Oil & Gas, Industrials, Consumer Services, Financials, and Technology. The results for the remaining industries (Basic Materials, Consumer Goods, Health Care, Financials, and Technology) are qualitatively similar and therefore not reported here. Again, the dependence pattern seems to be unaffected by affiliation of the individual stock to a certain industry.

<Insert Figure 6 about here>

A third factor considered here is the market risk of a stock as captured by its beta-coefficient. We estimate the individual stock's exposure to the market in a basic market

model framework. We then sort the individual stocks into five equal groups with different beta. Average beta for these groups ranges from 0.42 to 1.25, i.e. the first group exhibits only weak exposure to market risk while the fifth group strongly reacts to the market. Again, the pattern of decreasing autoregressive parameter estimates across quantiles is not altered. However, it seems that the dispersion within each of these groups is somewhat more pronounced for lower betas.

<Insert Figure 7 about here>

Finally, we investigate the role of fat tailed distributions or GARCH-type effects (or their combination) on our quantile autoregression model<sup>2</sup>. Campbell et al. (2008) analyze the impact of fat tails on cross-country stock market return correlations and find that there is less support for an increase in conditional correlations under alternative distributional assumptions accounting for fat tails like a bivariate Student- $t$  distribution. We therefore conducted a Monte Carlo simulation to gain insight into the effects of these time series properties on quantile autoregression. For this purpose, we generated data from an AR(1)-GARCH(1,1) model using different parametrizations and either normal or Student- $t$  distributed innovations. Quantile regression estimation of these data revealed no deviations from the simulated autoregressive parameter across the entire range of quantiles.<sup>3</sup> We therefore conclude that neither fat tails nor conditional heteroskedasticity per se generate significant differences across quantiles and are thus not causal to the asymmetric pattern we document.

### 3.3 Results for the Extended QAR(1) Models

In this section we consider the extensions of the QAR(1) model as proposed in Equations (2) and (3) analyzing the influence of the size and the sign, respectively, of a lagged return on the conditional return distribution.

The first extension as given in Equation (2) explicitly considers the impact of the size of a previous period's return. Figure 8 presents the estimation results of the 'Size Model' based on daily data. The estimation results for the autoregressive coefficients  $\beta_i(\tau)$  are presented

---

<sup>2</sup>We thank an anonymous referee for this suggestion.

<sup>3</sup>The simulation study is performed in R and can be obtained from the authors upon request.

in the upper panel of Figure 8. In comparison to the top left panel of Figure 1 we find the average autoregressive coefficient to be about twice as high as in the basic QAR(1) model in the lower quantiles. In the upper quantiles average autoregressive estimates are close to zero. Consequently, their s-shaped pattern disappears. The estimation results for the  $\gamma_i(\tau)$  coefficients are presented in the lower panel of Figure 8. The additional impact of large (positive or negative) returns on the conditional return distribution is found to be inverted u-shaped. When considering the combined effect of lagged returns, i.e. the sum of the estimates for  $\beta_i(\tau)$  and  $\gamma_i(\tau)$  for each quantile, again the s-shaped pattern as documented for the basic QAR(1) model emerges. We therefore conclude that this particular shape is driven mainly by extreme, lagged returns. Results for weekly and monthly data are qualitatively similar and therefore not reported, but can be obtained from the authors upon request.

<Insert Figure 8 about here>

The 'Sign Model' as given in Equation (3) considers the possibility that lagged positive and negative returns exhibit a different impact on the conditional return distribution. Estimation results based on daily data are presented in Figure 9 where the top panel contains  $\beta_i(\tau)$  and the bottom panel  $\delta_i(\tau)$  estimates.

As Figure 9 illustrates, the inclusion of negative returns as a second explanatory variable completely alters the pattern of autoregressive estimates. Instead of an s-shaped pattern we find a pronounced inverted s-shaped pattern for the  $\beta_i(\tau)$  estimates. Considering only negative returns, however, retains the previously documented s-shaped pattern. The results imply that investors react differently to positive and negative lagged returns and that the reaction also depends on the state of the return, i.e. the quantile. For example, positive lagged returns and positive current states (upper quantiles), and negative lagged returns and negative current states (lower quantiles) display positive dependence while positive lagged returns and negative states, and negative lagged returns and positive states display negative dependence on past returns. In other words, when the lagged return and the current state are aligned (both positive or negative), returns are positively correlated across time while the opposite is true if the lagged returns and current states are not aligned. The former is consistent with the notion of investor underreaction at  $t - 1$  while the latter is consistent with investor overreaction at  $t - 1$ . Again, results for weekly and

monthly data are qualitatively similar and therefore not reported, but can be obtained from the authors upon request.

<Insert Figure 9 about here>

### 3.4 Specification Issues

The estimation results presented above are based on a sample period of more than 30 years. To test for the stability of the regression parameters over time we split the sample in various ways. The results obtained hold qualitatively for different sampling periods. For example, for daily data restricted to the time span January 2000 to March 2009, we obtain the following results for the basic QAR(1) model. The size of the estimated  $\beta_i(\tau)$  parameters are almost unchanged in the lower quantiles of the conditional return distribution while in the upper quantiles the parameter estimates are slightly more pronounced. The OLS estimation results on the other hand now shows on average a negative autoregressive coefficient which is in contrast to the previous findings. This would suggest a fundamental change of the dynamic structure. Employing quantile regression, however, reveals that the relationship did not change systematically as the pattern of the estimated parameters remains stable. It is the extremes in the distribution which are affected by changing the time span while the median seems to be unswayed.

## 4 Summary

We employ quantile regression analysis to investigate the predictability of different parts of the distribution of stock returns. We find that the autoregressive parameters in a first order quantile autoregressive model in general follow a decreasing pattern over the quantiles of the conditional return distribution: negative returns (i.e. lower quantiles) are generally marked by positive dependence while positive returns (i.e. upper quantiles) exhibit in general negative dependence on past returns. We find such a pattern to hold when accounting for market capitalization, industry classification and exposure to market risk. A more detailed analysis accounting for the marginal effect of negative returns in the previous period shows that negative returns exhibit a stronger influence across the



whole distribution than positive returns. In addition, large negative returns are found to be the driving force behind the s-shaped pattern of autoregressive parameter estimates. We hope that the results in this paper will revive the investigation of the predictability of stock returns.

## **Acknowledgments**

We thank Emiliós Galariotis, Joachim Grammig, an anonymous associate editor and an anonymous referee for their comments which helped to significantly improve the paper. Financial support from the Fritz Thyssen Stiftung and the German Research Foundation (DFG) is gratefully acknowledged.

## References

- Barberis, N., Shleifer, A., Vishny, R., 1998. A model of investor sentiment. *Journal of Financial Economics* 49 (3), 307–343.
- Campbell, J. Y., Grossman, S. J., Wang, J., 1993. Trading volume and serial correlation in stock returnstrading volume and serial correlation in stock returns. *The Quarterly Journal of Economics* 108 (4), 905–939.
- Campbell, J. Y., Lo, W., MacKinley, A. C., 1997. *The Econometrics of Financial Markets*. Princeton University Press, Princeton, New Jersey.
- Campbell, R. A., Forbes, C. S., Koedijk, K. G., Kofman, P., 2008. Increasing correlations or just fat tails? *Journal of Empirical Finance* 15 (2), 287–309.
- Chernick, M. R., 2008. *Bootstrap Methods: a guide for practitioners and researchers*, 2nd Edition. John Wiley & Sons, Inc., Hoboken, New Jersey.
- Chuang, C.-C., Kuan, C.-M., Lin, H.-Y., 2009. Causality in quantiles and dynamic stock return-volume relations. *Journal of Banking & Finance* 33 (7), 1351–1360.
- Engle, R. F., Manganelli, S., 2004. Caviar: conditional autoregressive value at risk by regression quantiles. *Journal of Business & Economic Statistics* 22 (4), 367–381.
- Feng, Y., Chen, R., Basset, G. W., 2008. Quantile momentum. *Statistics and its interface* 1, 243–254.
- Glosten, L. R., Jagannathan, R., Runkle, D. E., 1993. On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance* 48 (5), 1779–1801.
- Koenker, R., Bassett, Jr., G., 1978. Regression quantiles. *Econometrica* 46 (1), 33–50.
- Koenker, R., Xiao, Z., 2006. Quantile autoregression. *Journal of the American Statistical Association* 101 (475), 980–990.
- Lewellen, J., 2002. Momentum and autocorrelation in stock returns. *The Review of Financial Studies* 15 (2), 533–564.

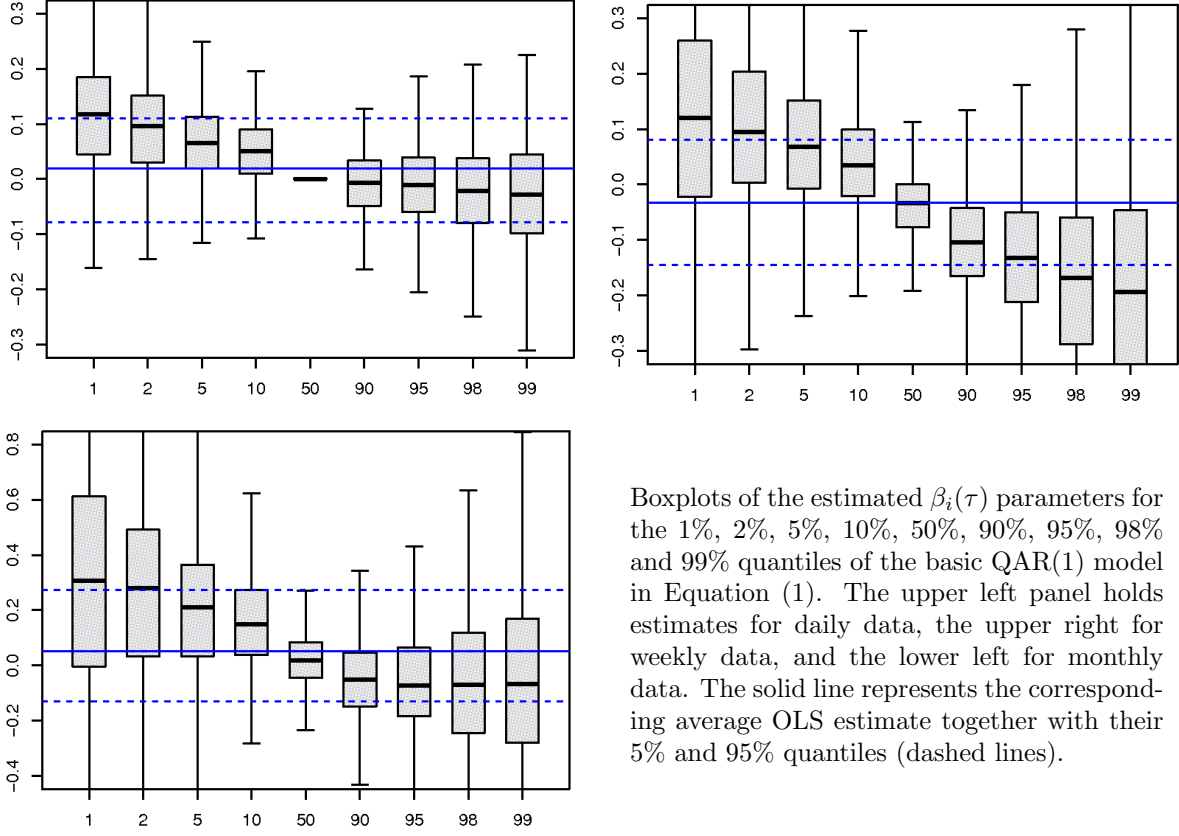
- Lo, A. W., MacKinlay, A. C., 1990. An econometric analysis of nonsynchronous trading. *Journal of Econometrics* 45 (1-2), 181–211.
- Ma, L., Pohlman, L., 2008. Return forecasts and optimal portfolio construction: A quantile regression approach. *European Journal of Finance* 14 (5), 409–425.
- Veronesi, P., 1999. Stock market overreactions to bad news in good times: a rational expectations equilibrium model. *The Review of Financial Studies* 12 (5), 975–1007.

Table 1: Descriptive Statistics

Stock Returns	Daily	Weekly	Monthly
Max Number of Observations	5794	1157	265
Median Number of Observations	4744	949	217
Min Number of Observations	148	29	6
Average Return	0.0001	0.0004	0.0015
Average Return Std. Dev.	0.0228	0.0509	0.1010
Minimum	-1.7635	-1.7635	-1.6449
Avg. 5% Quantile	-0.0335	-0.0774	-0.1609
Avg. Median	0.0000	0.0001	0.0064
Avg. 95% Quantile	0.0341	0.0752	0.1449
Maximum	1.3683	1.3683	1.2656
Index Returns			
Number of observations	5794	1157	265
Average Index Return	0.0001	0.0006	0.0026
Index Return Std. Dev.	0.0112	0.0247	0.0498
Minimum	-0.0888	-0.2425	-0.3040
Maximum	0.0941	0.1243	0.1353

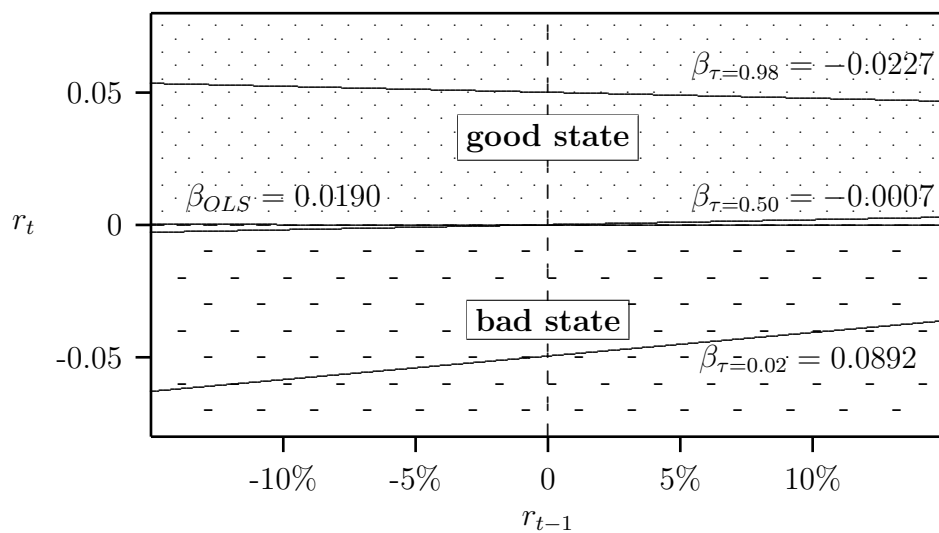
The table provides summary descriptive statistics for the individual stocks (upper panel) and the Dow Jones STOXX 600 index (lower panel) for the sample period 1987 to 2009.

Figure 1: Basic QAR(1) Estimation Results



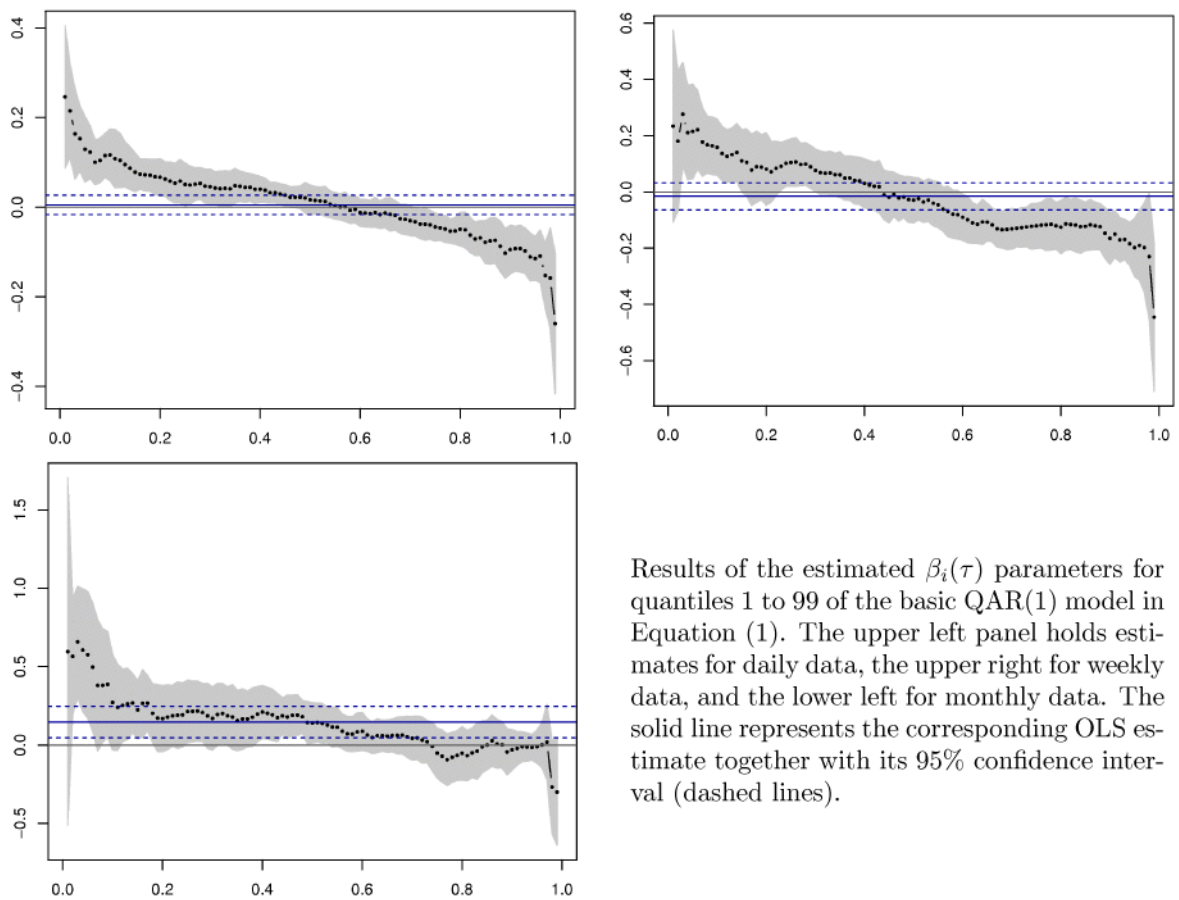
Boxplots of the estimated  $\beta_i(\tau)$  parameters for the 1%, 2%, 5%, 10%, 50%, 90%, 95%, 98% and 99% quantiles of the basic QAR(1) model in Equation (1). The upper left panel holds estimates for daily data, the upper right for weekly data, and the lower left for monthly data. The solid line represents the corresponding average OLS estimate together with their 5% and 95% quantiles (dashed lines).

Figure 2: Basic QAR(1) Model Regression Quantiles (Daily Data)



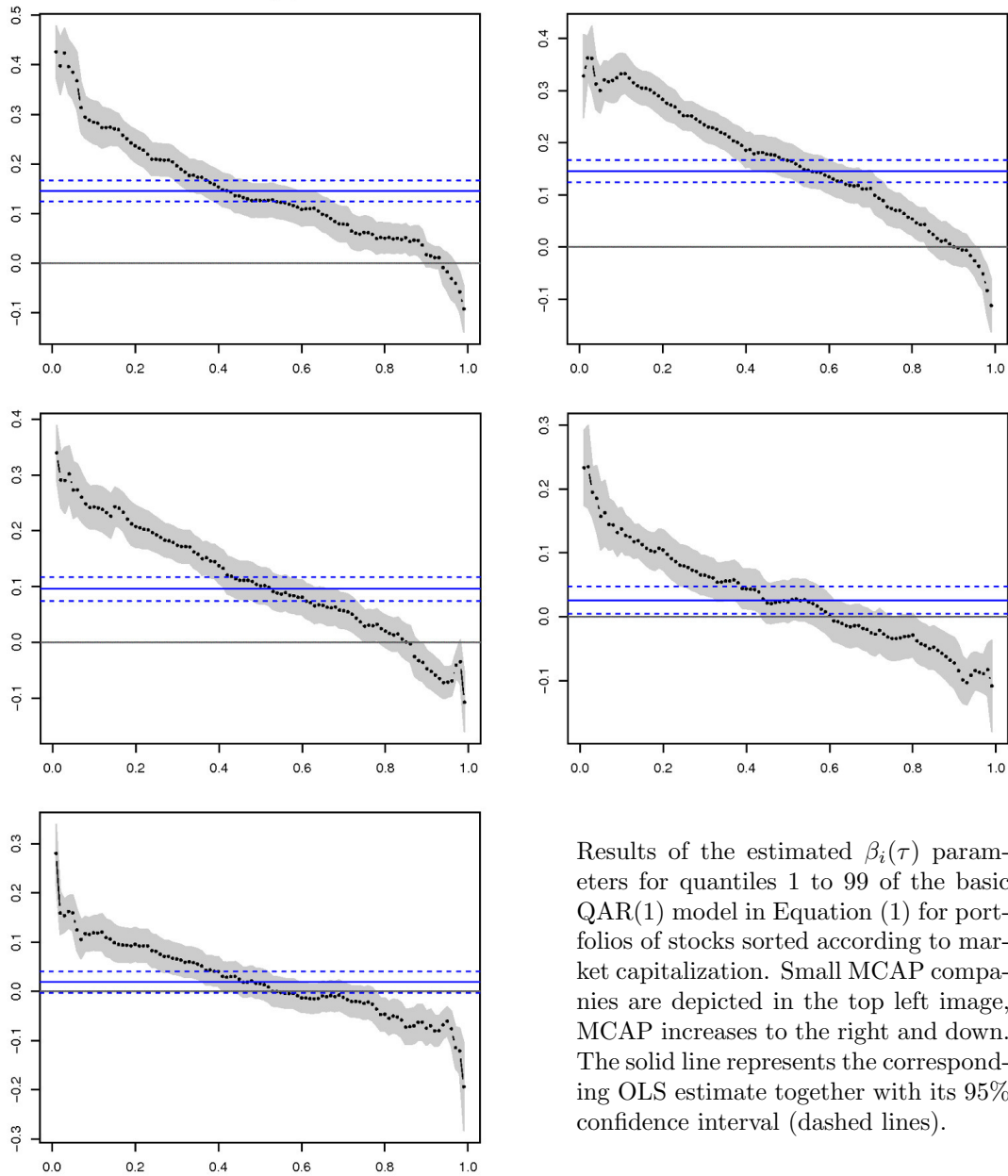
Average regression lines for three different quantiles ( $\tau = 0.02$ ,  $\tau = 0.50$  and  $\tau = 0.98$ ) of the basic QAR(1) model in Equation (1) along with the corresponding average OLS regression line.

Figure 3: Stoxx 600 Index Estimation Results



Results of the estimated  $\beta_i(\tau)$  parameters for quantiles 1 to 99 of the basic QAR(1) model in Equation (1). The upper left panel holds estimates for daily data, the upper right for weekly data, and the lower left for monthly data. The solid line represents the corresponding OLS estimate together with its 95% confidence interval (dashed lines).

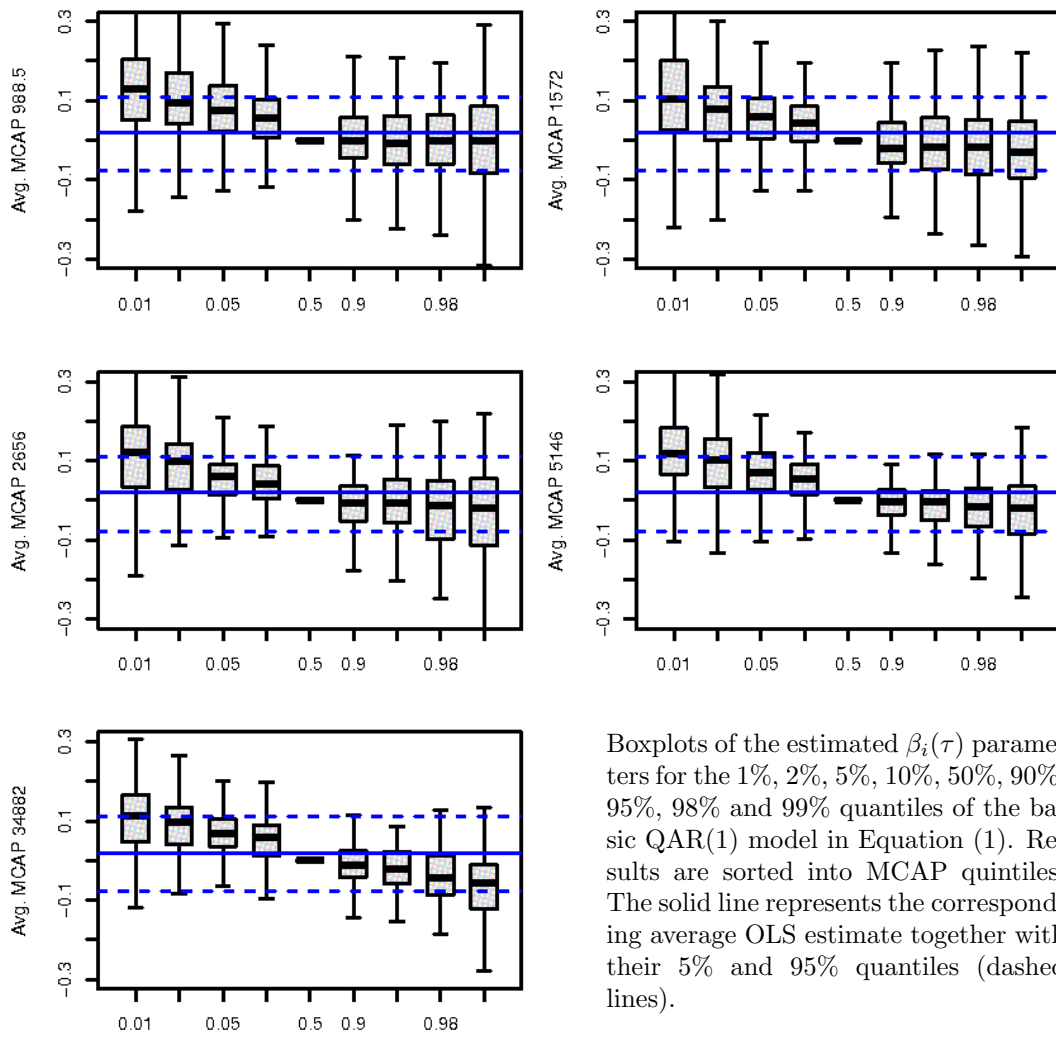
Figure 4: MCAP Portfolio Estimation Results (Daily Data)



Results of the estimated  $\beta_i(\tau)$  parameters for quantiles 1 to 99 of the basic QAR(1) model in Equation (1) for portfolios of stocks sorted according to market capitalization. Small MCAP companies are depicted in the top left image, MCAP increases to the right and down. The solid line represents the corresponding OLS estimate together with its 95% confidence interval (dashed lines).



Figure 5: Basic QAR(1) Estimation Results (Daily Data) sorted by MCAP



Boxplots of the estimated  $\beta_i(\tau)$  parameters for the 1%, 2%, 5%, 10%, 50%, 90%, 95%, 98% and 99% quantiles of the basic QAR(1) model in Equation (1). Results are sorted into MCAP quintiles. The solid line represents the corresponding average OLS estimate together with their 5% and 95% quantiles (dashed lines).

Figure 6: Basic QAR(1) Estimation Results (Daily Data) sorted by industries

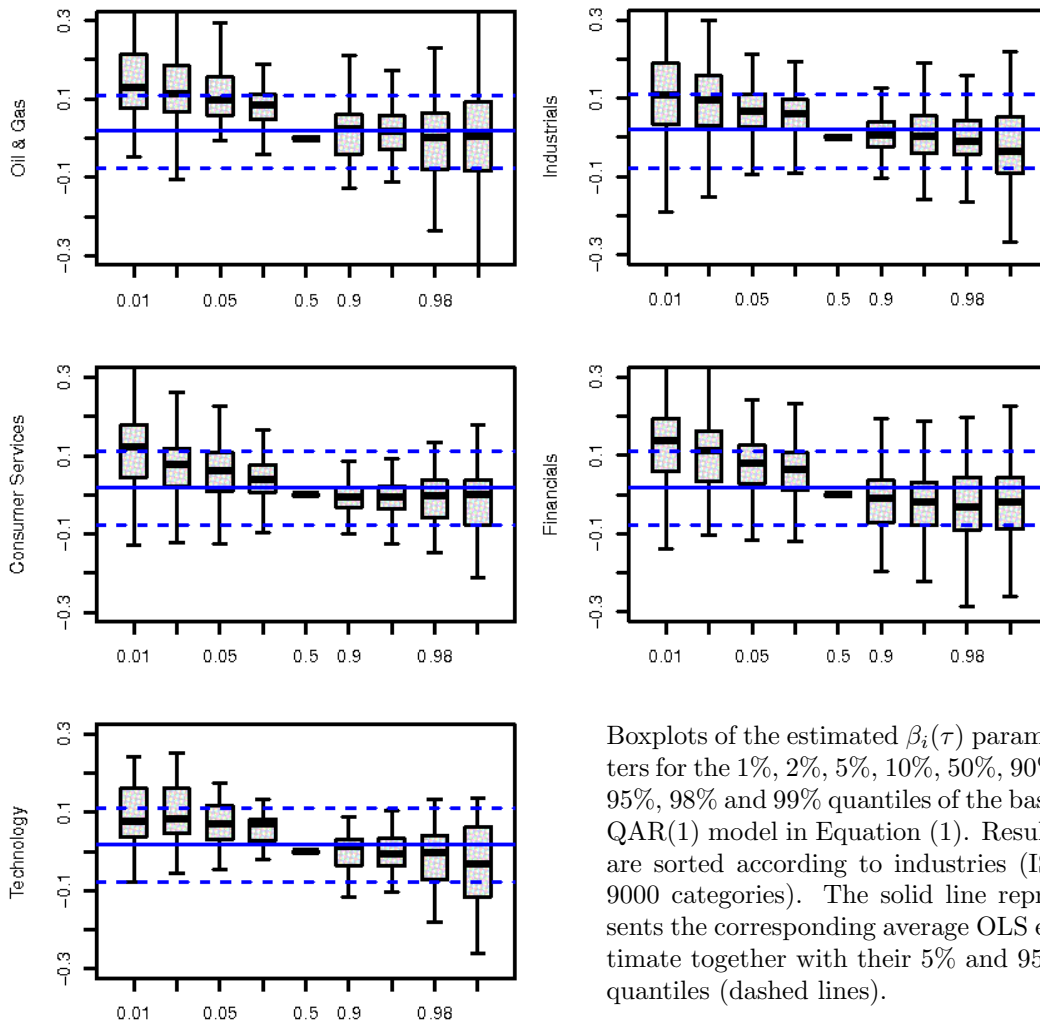
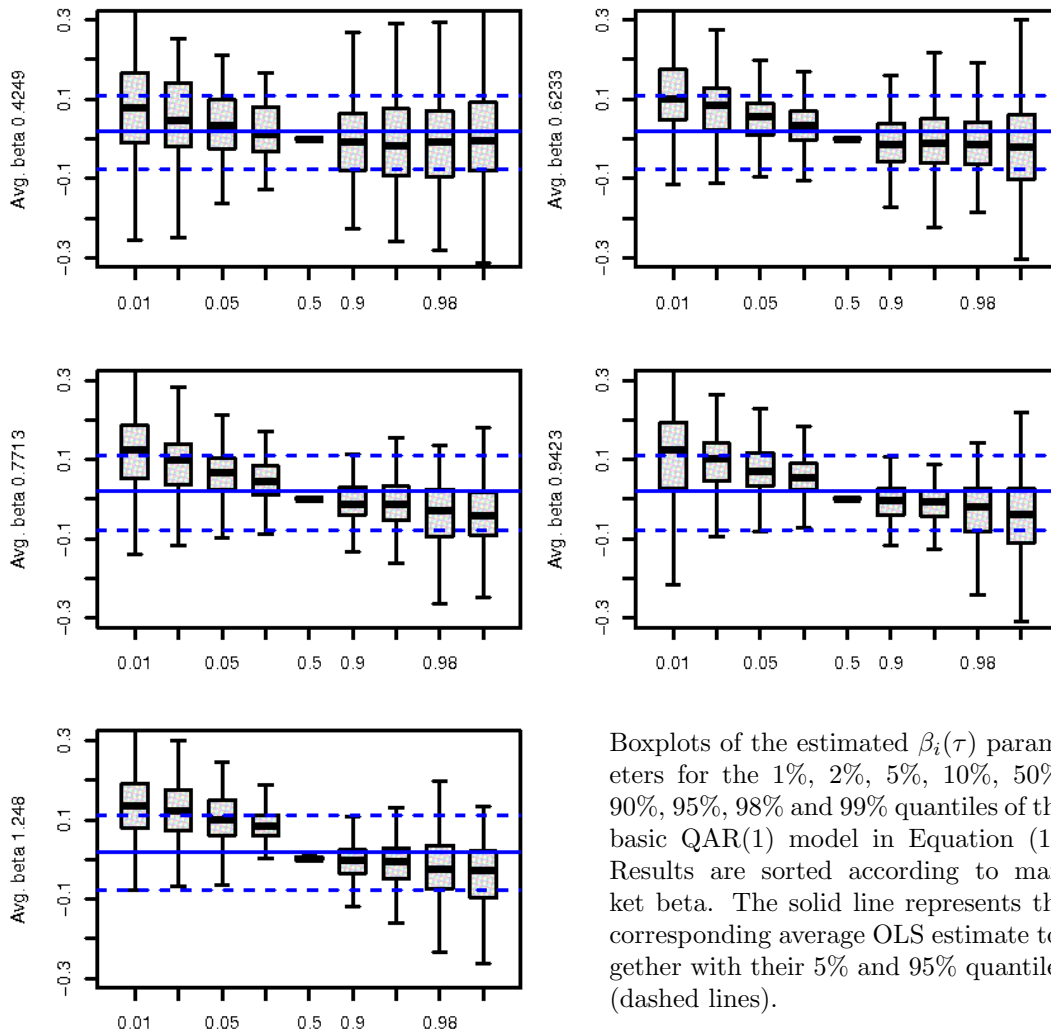
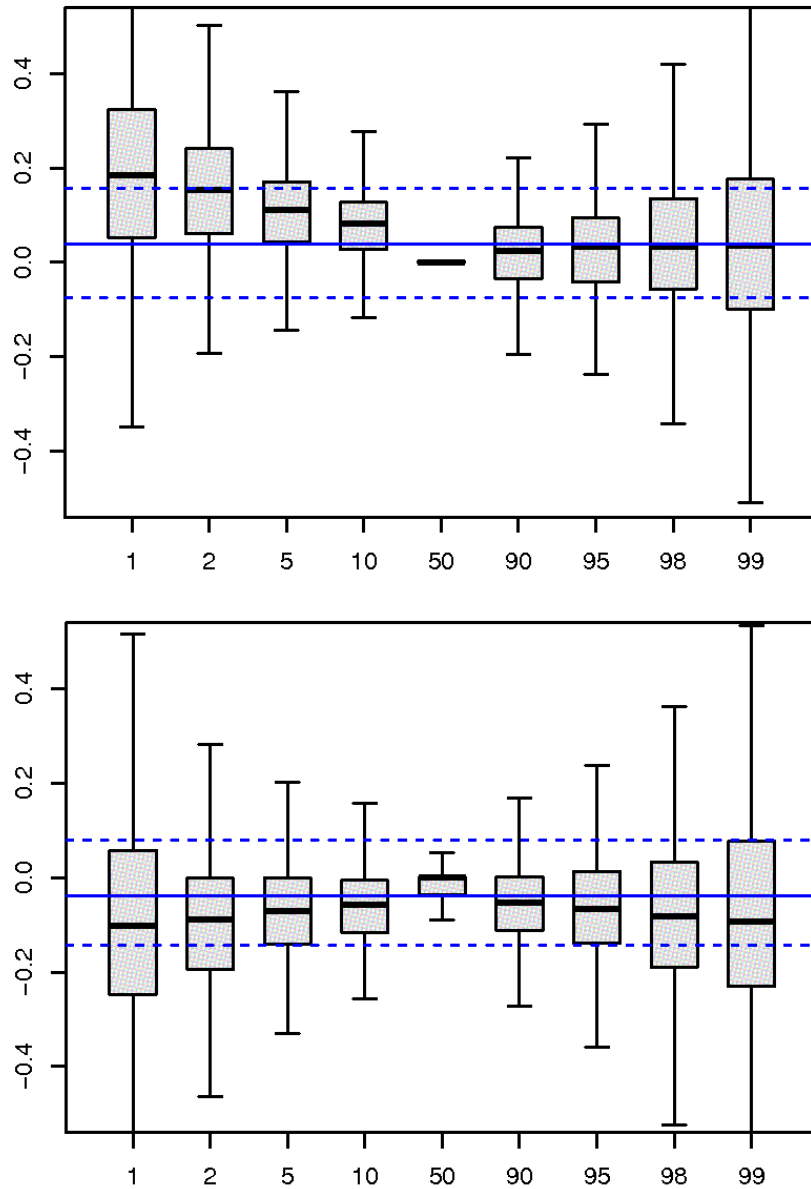


Figure 7: Basic QAR(1) Estimation Results (Daily Data) sorted by beta



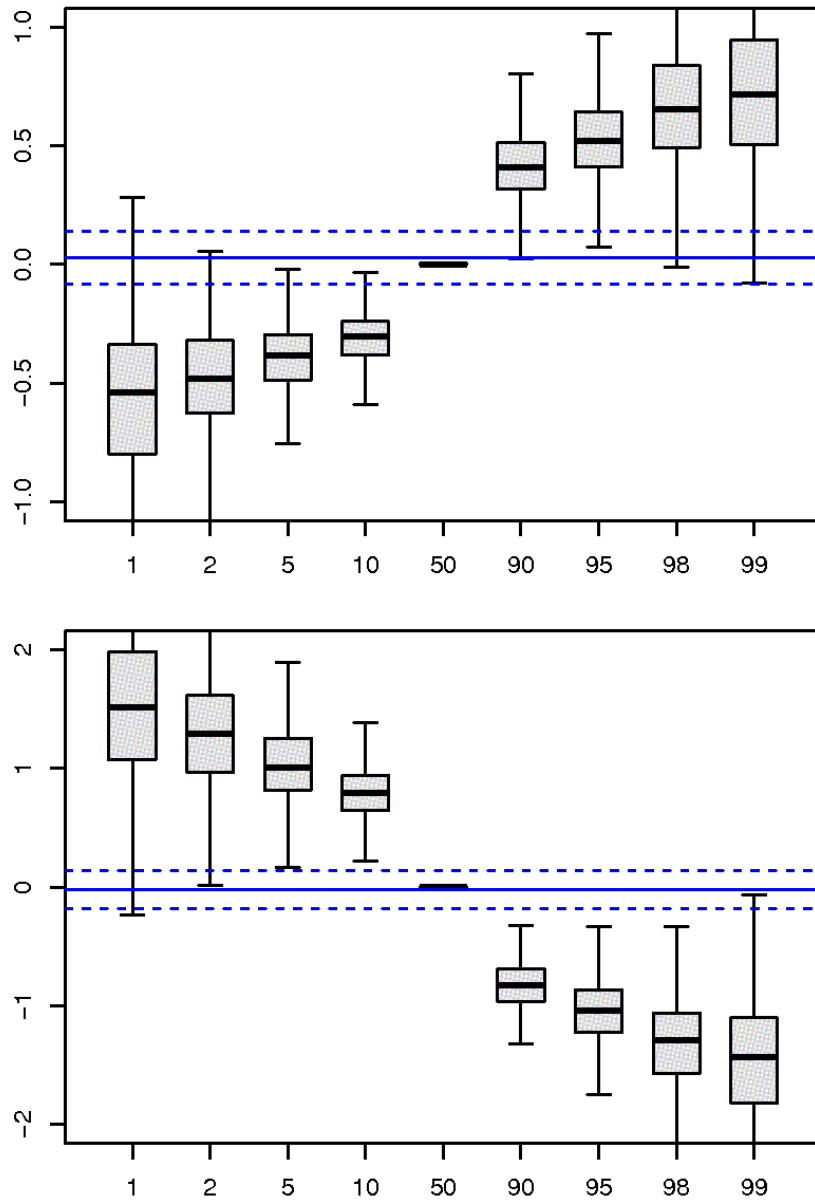
Boxplots of the estimated  $\beta_i(\tau)$  parameters for the 1%, 2%, 5%, 10%, 50%, 90%, 95%, 98% and 99% quantiles of the basic QAR(1) model in Equation (1). Results are sorted according to market beta. The solid line represents the corresponding average OLS estimate together with their 5% and 95% quantiles (dashed lines).

Figure 8: 'Size Model' Estimation Results (Daily Data)



Boxplots of the estimated  $\beta_i(\tau)$  (upper panel) and  $\gamma_i(\tau)$  (lower panel) parameters for the 1%, 2%, 5%, 10%, 50%, 90%, 95%, 98% and 99% quantiles of the extend QAR(1) model in Equation (2) measuring the impact of extreme past returns. The solid line represents the corresponding average OLS estimate together with their 5% and 95% quantiles (dashed lines).

Figure 9: 'Sign Model' Estimation Results (Daily Data)



Boxplots of the estimated  $\beta_i(\tau)$  (upper panel) and  $\delta_i(\tau)$  (lower panel) parameters for the 1%, 2%, 5%, 10%, 50%, 90%, 95%, 98% and 99% quantiles of the extend QAR(1) model in Equation (3) measuring the impact of negative past returns. The solid line represents the corresponding average OLS estimate together with their 5% and 95% quantiles (dashed lines).