Estimation of conditional value at risk of returns during high volatility periods using cross-sectional quantile regression

Xenxo Vidal-Llana
juanjose.vidal@ub.edu

Montserrat Guillen
mguillen@ub.edu

Abstract

Evaluating Value at Risk for a firm's returns during turmoil periods is difficult because there is too much volatility in the market. We propose to estimate conditional Value at Risk and Expected Shortfall for a given firm return using quantile regression with cross-sectional information from other same-market firms. An illustration with data from the US market from 2000 to 2020 is included. Empirical results show that our model approach has some advantages over a CAViaR model, which estimates Value at Risk based on the analysis of the univariate time-series of the firm's returns without covariates. During crises like the Great Recession (2007-2011) and the Covid-19 pandemic (2020), Value at Risk at the 5% level are lower for our approach than for the CAViaR on average. However, the average Expected Shortfall in 2020 is higher for our approach than for the CAViaR. Implications of adopting quantile regression include an increase of the average reserves needed to cover extreme events due a lower average Expected Shortfall in comparison with the average obtained from a CAViaR approximation. However, during the Covid-19 pandemic, the average reserves would be lower with the quantile regression approximation. Identification of low risk firms and computer time reduction are also an advantage with the new method.

Keywords: value at risk, quantile regression, asset management

Classification codes: 210, 310

1 Introduction

When evaluating firm's stock returns tails are essential for investors and regulators to manage investment decisions and evaluate capital allocation, as they can provide information about the possibility of future losses of a market asset or a portfolio.

Firm returns' movements react differently during periods of high market volatility. A univariate time-series perspective analysis relies only on past information and does not include information about other firms in the same market. To evaluate the returns of a firm conditional on how the rest of the firms in the market behave, a cross sectional model is necessary. We propose using a cross-sectional quantile regression model to study tail returns for one firm using information from other firms in the same market, and we compare this approach with the CAViaR model, which analyses the time series of a single firm's returns.

Value at Risk (VaR) is an easy tool to summarise risks ([Bodnar et al., 1998]). Several studies have enhanced VaR estimation methods: freedom to choose a probability distribution ([Hull and White, 1998]); new improvements and approaches, like ARCH and GARCH models, that treat heteroskedasticity as a variable to be modelled ([Engle, 2001]); CoVaR, a measure for systemic risk for institutions under adverse situations ([Adrian and Brunnermeier, 2011]); and the CAViaR model, that estimates the tail with autoregressive processes ([Engle and Manganelli, 2004]).

In recent years, innovative ways of estimating VaR to improve risk strategies involving asset evaluation have been proposed by many authors (e.g: [Wang et al., 2018]; [Sahamkhadam et al., 2018]; [Lin et al., 2018]; [Kwon, 2019]; [Gribisch and Eckernkemper, 2019]; [Cai and Stander, 2020]; [Pei et al., 2021]; [Bodnar et al., 2021]).

Researchers have remarked the importance of studying high volatility periods, that can be devastating in terms of losses and lack of liquidity (e.g.: [Dias, 2016]; [Alexandridis and Hasan, 2020]; [Babalos et al., 2021]; [Belaid et al., 2021]). In spite of the difficulty to capture variability in high volatility periods, VaR estimation is a standard practice, i.e. Basel III.

2 Methodology

2.1 Quantile regression

Quantile regression (QR) aims to fit the quantile of the response variable given a set of covariates [Koenker and Bassett Jr, 1978, Koenker, 2017]. In our context, QR is a useful method to understand what influences the possibility of observing extreme returns (see also, [Uribe and Guillen, 2020] for other applications).

Let y_i be a random variable with probability distribution function F_i that depends on covariates $X_i' = \{X_{1i}, X_{2i}, \ldots, X_{ki}\}$ for $i = 1, \ldots, N$, N is the number of observations, so $0 \le F_i(y|X_i) \le 1$. We specify the α-*th conditional quantile* $(0 \le \alpha \le 1)$ as:

$$Q_{Y_i|X_i}(\alpha) = \beta_{(\alpha)0} + \beta_{(\alpha)1}X_{1i} + \beta_{(\alpha)2}X_{2i} + ... + \beta_{(\alpha)k}X_{ki} = X_i'\beta_{(\alpha)} \ , \tag{1}$$

with parameter estimates $\hat{\beta}_{(\alpha)} = \arg\min_{\beta} \mathrm{E}[\rho_\alpha(Y_i - X_i'\beta)]$ where $\rho_\alpha(u) = [\alpha - \mathbb{1}_{\{u<0\}}]u$, and $\mathbb{1}_{\{u<0\}}$ is the identity function, with value 1 when the subscript is true and 0 otherwise.

In order to evaluate the performance of a QR model, a scoring function is defined. It aims to measure the discrepancy between a predicted and an observed value. Let $\tilde{y}_i(\alpha) := X_i'\hat{\beta}_{(\alpha)}$ be the fitted quantile α for observation i, and y_i the observed value for observation i, we define the score following [Koenker and Hallock, 2001] as:

$$Q_\alpha^0 = \frac{1}{N}\sum_{i=1}^{N}\left[\alpha - \mathbb{1}_{\{y_i \le \tilde{y_i}(\alpha)\}}\right](y_i - \tilde{y_i}(\alpha)) \ . \tag{2}$$

Note that Q_α^0 is a weighted average of the absolute distance between the observed value and the fitted quantile. The lower the value of Q_α^0, the better the approximation.

The Expected Shortfall for a level α is defined as:

$$ES_\alpha(Y_i|X_i) = E[Y_i|Y_i \le VaR_\alpha(Y_i|X_i)] \ . \tag{3}$$

The Expected Shortfall (ES), also known as Tail Conditional Expectation (TCE), Conditional Tail Expectation (CTE) or Tail Value at Risk (TVaR), is a risk measure that approximate the expected loss conditioned to surpassing VaR.

2.2 CAViaR model

We will use the CAViaR model to analyse a univariate time series of the returns of one firm, where y_t denotes the return in time t. This model was proposed by [Engle and Manganelli, 2004] and its general model specification is the following for a given α level:

$$f_t(\beta) = \beta_0 + \sum_{i=1}^{q} \beta_i f_{t-i}(\beta) + \sum_{j=q+1}^{r} \beta_j l(y_{t-j}) \ , \tag{4}$$

where y_{t-j} is the observed return in time $t-j$. $f_t(\beta)$ is an abbreviation for $f_t(y_{t-1}, \beta_\alpha)$, which is the quantile α at t of the distribution of returns, that depends on the observed variables from previous periods. $\beta_\alpha = (\beta_0, \ldots, \beta_r)$ is the vector of parameters to be estimated. [Engle and Manganelli, 2004] also incorporate a lag function $l(\cdot)$, in order to link observed values to the information set. The CAViaR model that we use in this study is called Indirect GARCH(1,1), which defines the quantile as:

$$f_t(\beta) = (\beta_1 + \beta_2 f_{t-1}^2(y_{t-2}, \beta_\alpha) + \beta_3 y_{t-1}^2)^{1/2} \ . \tag{5}$$

2.3 Cross-Sectional Quantile Regression versus CAViaR for conditional VaR estimation of returns

In Cross-Sectional Quantile Regression (CSQR) we fix a point in time t, and we adjust a QR model for the returns using firms' characteristics. This allows us to characterize the state of the market.

Once the QR is estimated, we can compare the results of a CSQR model and the corresponding CAViaR model. Note that the two approaches are essentially different. CSQR uses all firms observed at a given point in time t and assumes that quantiles depend on firm's characteristics. CAViaR assumes that quantiles of a firm's returns depend on past returns. Based on the CSQR and the CAViaR approaches, we obtain two different estimates of VaR for each firm at every point in time t.

There are some advantages of using CSQR: 1) The computational requirements to calculate a quantile regression model at t is lower than to adjust a CAViaR model (Equation (5)) for each firm. 2) The CSQR model uses covariates, which allows us to include exogenous characteristics and to predict return's quantiles for external firms that were not initially in our dataset. 3) Finally, time series models need a minimum observational period, while our approach does not need previous observations.

3 Data and characteristics in the Cross-Sectional Quantile Regression model

Our information database has 204 characteristics for 26,298 different firms between 1990 and 2020 in the US market. Data have a monthly frequency. We combined firm's data with their returns, obtained via CRSP. Our baseline specification of CSQR uses the following seven variables: size of the firm (MC), book-to-market ratio (BM), operating profitability (OP), growth rate of investment (INV), 12 months momentum (MOM), liquidity of the firm (LIQ) and market beta ($beta$). Firm's size (MC) has been constructed using CRSP information like in [Uribe Gil et al., 2021], the other factors have been retreived from [Chen and Zimmermann, 2020] dataset. The chosen variables correspond to the standard magnitudes used to price average returns using cross-sectional factor firm characteristics. MC, BM, OP and INV are suggested in [Fama and French, 2020], MOM, LIQ and $beta$ have been added following the discussions in [Campbell, 2017] and [Malkiel, 2019]. Because of the CAViaR requirements, we had to restrict our results to firms that had information covering the whole observational period (438 firms), but our proposed model uses and is able to evaluate all firms (26,298 firms).

4 Cross-Sectional Quantile Regression model for calculating VaR: a comparison with CAViaR

In general, high volatility periods may have a big impact on predictions and turn historical analyses into a mess, as extreme observations incorporate an extra variability to our estimates. In this section we discuss the differences between the CSQR model and CAViaR for VaR estimation.

Our dataset includes two crisis periods, the Great Recession (2007-2011) and the Covid-19 pandemic (2020). Our results are presented only from January 2000 to December 2020, as we used the initial 10 years to capture the autoregressive momentum needed for estimating CAViaR models. Here we analyse the 0.05 level, as we want to study losers in terms of returns. The same methodology can be implemented for any quantile level. Code and results are available from the authors.

Figure 1 shows the average predictions for the 438 firms each time for both CSQR model (purple)

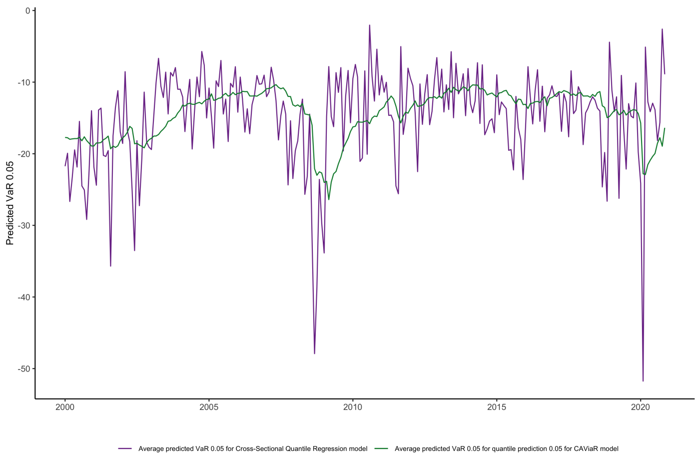

Figure 1: Average predicted Value at Risk for quantile 0.05 for Cross-Sectional Quantile Regression model (purple) and CAViaR model (green), from 2000 to 2020, for 438 firms.

and CAViaR model (green). The CSQR model creates a more volatile series of predictions due to the usage of other returns and covariates, which creates a richer perspective when calculating VaR. We note how the CAViaR presents a delay on the fitted VaR in comparison with the CSQR model as, for example for The Great Recession, the decrease on the predicted VaR is delayed between four and seven months later. We also remark that the CAViaR usually produces less extreme VaR estimates than the CSQR model, as for high volatility periods the CAViaR model creates higher VaR for the 0.05 level than for the CSQR model, meaning that with the CSQR we detect higher amount of risk, therefore an increase of the reserves needed for turmoil periods.

Figure 2 shows the observed returns for an individual firm (firm number 24010) in grey points. The lines show the predictions for the 0.05 quantile using CSQR (purple) and CAViaR (green). We consider an exception a return lower than a 0.05 fitted quantile. This should occur only 5% of the times. Exceptions with the same color as the corresponding model refer to returns that are lower than the fitted VaR for that model but not for the other model, namely, purple crosses mean returns that we would consider exceptions under the CSQR model but not under the CAViaR model. Green crosses mean that under the CAViaR model we would consider those returns as exceptions but not under the CSQR model. Red crosses mean that those returns are considered exceptions for both models. We see how during the Great Recession (2007-2011) and during the Covid-19 pandemic (2020) the volatility of the returns increases significantly. If we draw our attention to the quantile predictions during the Great Recession, we see that the returns for this period are larger than the fitted quantiles for

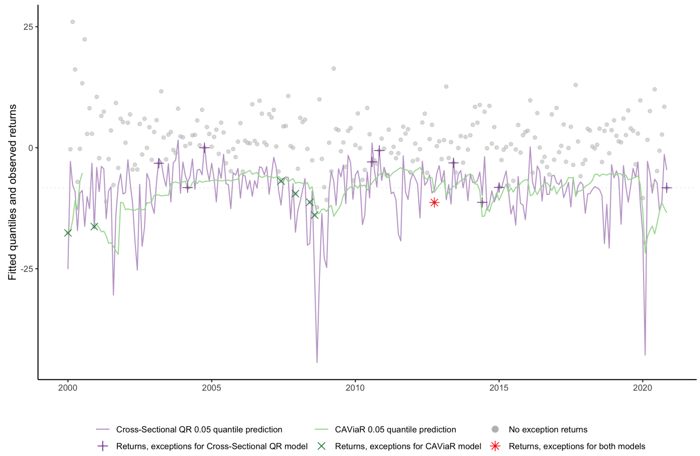

Figure 2: Comparison between the evolution of returns and predictions for quantile 0.05 for both Cross-Sectional Quantile Regression model (purple) and CAViaR model (green) for firm 24010 from 2000 to 2020, monthly data.

the CSQR model. Contrarily, the CAViaR model identifies four points as exceptions, marked in the graph as green crosses. These differences are an example of a localized potentially biased estimation of extreme returns during high volatility periods.

The example presented in Figure 2 shows how our method is robust against turmoil periods. The same happens for almost all our sample firms. More examples of comparisons for other firms between both models can be seen in Figure 6 in the Appendix.

4.1 Scoring the models and distribution of exceptions

We evaluate quantile estimates. In Table 1 we present the score in Equation (2) for all months and for the two crisis periods for the CSQR model and the CAViaR model, for 438 firms. In parenthesis, we present the total number of months when a model (the model in the corresponding row) has had a lower score than the other model.

In Table 1, the scoring gives the CSQR model a better approximation to the 0.05 quantile during high volatility periods because the CSQR model has lower scores 12,110 times while CAViaR had lower scores only 11,431 times. In general terms this would be unnoticed, as overall the CAViaR models score better 64,279 times compared to only 52,919 times for the CSQR model. We argue that the CSQR outperforms CAViaR in crises due to the usage of market data, thus it seems that the CSQR model is able to adapt to a situation of high volatility periods easier than the CAViaR.

$Q^0_{0.05}$	All (2000-2020)	Great Recession (2007-2011)	Covid-19 pandemic (2020)
Cross-sectional QR	0.97 (52,919)	1.02 (12,110)	1.13 (3,155)
CAViaR	0.88 (64,279)	1.05 (11,431)	1.31 (2,134)

Table 1: Scoring values at the 0.05 quantile for each model and number of months that the model had a lower (better) score than the other for 438 firms, by periods, .

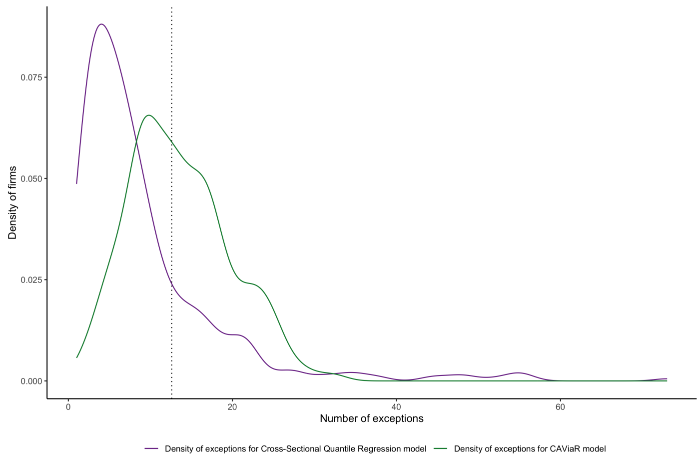

Figure 3: Kernel plot of the number of exceptions by firm for Cross-Sectional Quantile Regression model (purple) and CAViaR models (green) from 2000 to 2020 for quantile 0.05 (438 firms). The dotted line marks value 12.6 (expected number of exceptions, 5% of 21 years, 12 months each).

We draw now our attention to the exceptions. In Figure 3 we observe the kernel of the number of exceptions for all firms that can be compared in the CAViaR and the CSQR models (438 firms). For CAViaR models, we expect to have a very high density overlapping the dotted line, because this model adjusts quantiles in a time-series perspective, leaving approximately a $\alpha\%$ of observed values as exceptions. This corresponds to a benchmark equal to 12.6 in our case (21 years $\times$ 12 months each $\times$ 0.05 quantile). We note how, for the CSQR model, the density is more spread and has less exceptions than the benchmark value. By locating firms in Figure 3, we can identify firms that have more exceptions than expected, and consider them under-performing. The CSQR model has no assumptions on the number of exceptions, which makes CSQR ideal for identifying well-performing firms in terms of returns.

We assume, for both tested models, that exceptions do not depend on the moment in time.

Thus, a non-equal distribution of exceptions over time indicates a bad performance of the model. We performed a Kolmogorov-Smirnoff test for the duration between exceptions for both CSQR and CAViaR models, where the null hypotesis tests equal distribution of durations with an exponential duration with parameter equal to the average of durations, and the alternative hypotesis assumes a different distribution. Results indicate that for the great majority of firms the null hypotesis was not rejected. For the few firms that the null hypotesis was rejected, the CAViaR model represents a greater number of cases than the CSQR model. In order to find stronger evidence of these conclusions, an expansion of the dataset, time-series wise, is recommended.

4.2 Comparison of Value at Risk and Expected Shortfall estimation

In Figure 4 we present the average VaR for all 438 firms splitted by periods calculated using CSQR and CAViaR models.

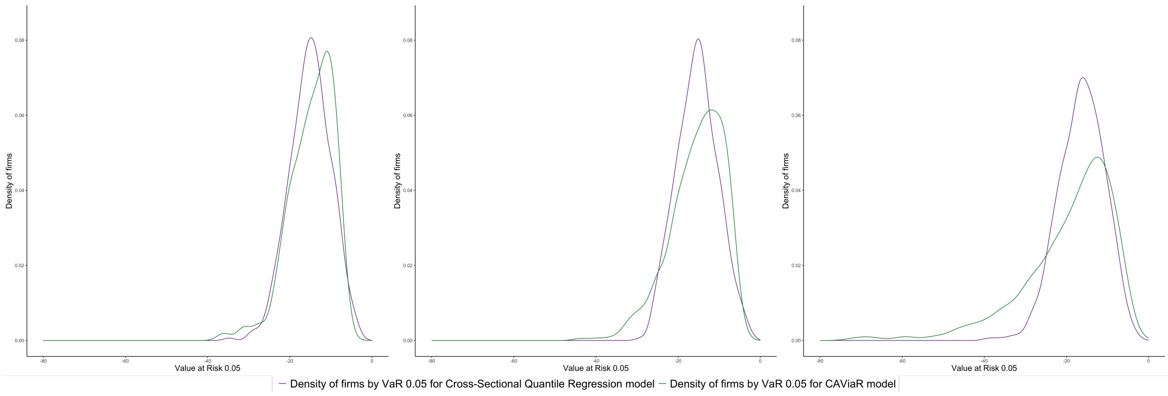

Figure 4: Kernel of estimated Value at Risk using Cross-Sectional Quantile Regression model (purple) and CAViaR model (green) for quantile 0.05, for the whole period (2000 to 2020, left), the Great Recession (2007 to 2012, middle) and Covid-19 pandemic (2020, right), for 438 firms.

We note how the VaR for quantile level 0.05 calculated with the CSQR model is, on average, lower than the VaR calculated using CAViaR model for all periods (left subfigure). For high volatility periods (center and right subfigures), the CAViaR model seems to have a group of firms with a fitted VaR lower that the VaR fitted with the CSQR. So the density of CAViaR adjusted VaRs is above the density of the CSQR adjusted VaRs on the left of the subfigure.

We calculate the ES in order to quantify the average loss beyond quantiles and we also compate the results under the two approaches. The empirical ES is obtained as the average of returns of exceptions, i.e. returns below the predicted VaR at level 0.05. ES indicates the magnitude of loss that is needed to cover average losses beyond the VaR. For quantile level 0.05 and for each firm, we have estimated the ES using the CSQR model and the CAViaR model.

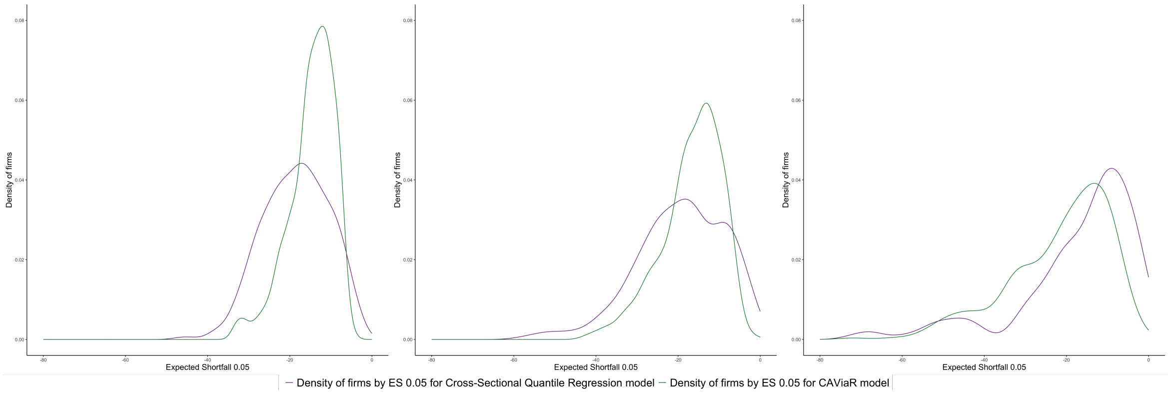

Figure 5: Kernel of estimated empirical ES using Cross-Sectional Quantile Regression model (purple) and CAViaR model (green) for quantile 0.05, for the whole period (2000 to 2020, left), the Great Recession (2007 to 2012, middle) and Covid-19 pandemic (2020, right), for 438 firms.

In Figure 5 we observe the difference between empirical ES for the CSQR model and the CAViaR model, separated by periods. For the whole period (left subfigure), in general, average ES for the CAViaR model are greater than average ES from the CSQR model, meaning that on average the requirements of capital should be increased when using the CSQR model instead of a time-series perspective as in the CAViaR approach. We note that, for the high volatility periods (center and right subfigures), the CSQR model predicts average ES levels similar to those of the CAViaR model. In the Covid-19 period (right subfigure), capital requirements would be lower with the CSQR model in comparison with the CAViaR approach, because the density of average estimated ES has shifted to the right.

The CAViaR model overestimates ES during high volatility periods in comparison with the CSQR as seen in Figure 1, and does not capture the effect of those market situations, while the CSQR model is able to adapt to those situations, as seen in Figure 5. Choosing one model or the other is not equivalent for investment decisions.

5 Conclusions

We proposed a Cross-Sectional Quantile Regression model using seven firm characteristics in order to evaluate the return's tail behaviour. For each firm, we focus on the predicted VaR using CSQR and compare it with the VaR estimated using a CAViaR model.

We included two high volatility periods, the Great Recession (2007-2011) and the Covid-19 outbreak (2020). The CSQR model shows lower estimated quantile 0.05 over high volatility periods compared to a CAViaR approach. Scoring results show a better performance for the CSQR model for quantile 0.05 during both periods of turmoil and a lower number of exceptions.

With CSQR, the VaR and empirical ES are, on average, lower than the calculated using CAViaR models. For high volatility periods, the ES using CSQR increases, surpassing the ES calculated with the CAViaR for the Covid-19 pandemic. This translates, from a risk management point of view, in an increase of reserves if CSQR models are used instead of CAViaR models, due to a lower predicted VaR for quantile 0.05. But, from an investor perspective, it is easier to identify firms that have a larger losers tail with the CSQR model than with the CAViaR model during high volatility periods because CSQR produces estimates that are more spread.

Using a CSQR model can be an enhancement for evaluating company returns, and can provide an improvement in calculations of reserves during turmoil periods and allows quantile approximation for out-of-sample firms.

References

[Adrian and Brunnermeier, 2011] Adrian, T. and Brunnermeier, M. K. (2011). CoVaR. Technical report, National Bureau of Economic Research. https://www.nber.org/papers/w17454.

[Alexandridis and Hasan, 2020] Alexandridis, A. K. and Hasan, M. S. (2020). Global financial crisis and multiscale systematic risk: Evidence from selected European stock markets. *International Journal of Finance & Economics*, 25(4):518–546.

[Babalos et al., 2021] Babalos, V., Caporale, G. M., and Spagnolo, N. (2021). Equity fund flows and stock market returns in the USA before and after the global financial crisis: a VAR-GARCH-in-mean analysis. *Empirical Economics*, 60(2):539–555.

[Belaid et al., 2021] Belaid, F., Ben Amar, A., Goutte, S., and Guesmi, K. (2021). Emerging and advanced economies markets behaviour during the COVID-19 crisis era. *International Journal of Finance & Economics*. https://onlinelibrary.wiley.com/doi/full/10.1002/ijfe.2494.

[Bodnar et al., 1998] Bodnar, G. M., Hayt, G. S., and Marston, R. C. (1998). 1998 Wharton survey of financial risk management by US non-financial firms. *Financial Management*, 27(4):70–91.

[Bodnar et al., 2021] Bodnar, T., Lindholm, M., Thorsén, E., and Tyrcha, J. (2021). Quantile-based optimal portfolio selection. *Computational Management Science*, 18:299–324.

[Cai and Stander, 2020] Cai, Y. and Stander, J. (2020). The threshold GARCH model: Estimation and density forecasting for financial returns. *Journal of Financial Econometrics*, 18(2):395–424.

[Campbell, 2017] Campbell, J. Y. (2017). *Financial decisions and markets: a course in asset pricing.* Princeton University Press. Princeton, USA.

[Chen and Zimmermann, 2020] Chen, A. Y. and Zimmermann, T. (2020). Open source cross-sectional asset pricing. *Available at SSRN 3604626.*

[Dias, 2016] Dias, A. (2016). The economic value of controlling for large losses in portfolio selection. *Journal of Banking & Finance*, 72:S81–S91.

[Engle, 2001] Engle, R. (2001). GARCH 101: The use of ARCH/GARCH models in applied econometrics. *Journal of Economic Perspectives*, 15(4):157–168.

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

[Fama and French, 2020] Fama, E. F. and French, K. R. (2020). Comparing cross-section and time-series factor models. *The Review of Financial Studies*, 33(5):1891–1926.

[Gribisch and Eckernkemper, 2019] Gribisch, B. and Eckernkemper, T. (2019). Intraday Conditional Value at Risk: A Periodic Mixed-Frequency GAS Approach. *Available at SSRN 3713896.*

[Hull and White, 1998] Hull, J. and White, A. (1998). Value at risk when daily changes in market variables are not normally distributed. *Journal of Derivatives*, 5:9–19.

[Koenker, 2017] Koenker, R. (2017). Quantile regression: 40 years on. *Annual Review of Economics*, 9:155–176.

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

[Koenker and Hallock, 2001] Koenker, R. and Hallock, K. F. (2001). Quantile regression. *Journal of Economic Perspectives*, 15(4):143–156.

[Kwon, 2019] Kwon, J. H. (2019). Tail risk and the consumption CAPM. *Finance Research Letters*, 30:69–75.

[Lin et al., 2018] Lin, E. M., Sun, E. W., and Yu, M.-T. (2018). Systemic risk, financial markets, and performance of financial institutions. *Annals of Operations Research*, 262(2):579–603.

[Malkiel, 2019] Malkiel, B. G. (2019). *A random walk down Wall Street: the time-tested strategy for successful investing.* WW Norton & Company. New York, USA. London, UK.

[Pei et al., 2021] Pei, Z.-T., Wang, X.-S., Xu, Y.-H., and Yue, X.-Y. (2021). A worst-case risk measure by G-VaR. *Acta Mathematicae Applicatae Sinica, English Series*, 37(2):421–440.

[Sahamkhadam et al., 2018] Sahamkhadam, M., Stephan, A., and Östermark, R. (2018). Portfolio optimization based on GARCH-EVT-Copula forecasting models. *International Journal of Forecasting*, 34(3):497–506.

[Uribe and Guillen, 2020] Uribe, J. M. and Guillen, M. (2020). *Quantile Regression for Cross-Sectional and Time Series Data: Applications in Energy Markets Using R.* Springer.

[Uribe Gil et al., 2021] Uribe Gil, J. M., Guillén, M., and Vidal-Llana, X. (2021). Rethinking Asset Pricing with Quantile Factor Models. *IREA–Working Papers, 2021, IR21/04.*

[Wang et al., 2018] Wang, J.-N., Du, J., and Hsu, Y.-T. (2018). Measuring long-term tail risk: Evaluating the performance of the square-root-of-time rule. *Journal of Empirical Finance*, 47:120–138.

Appendix

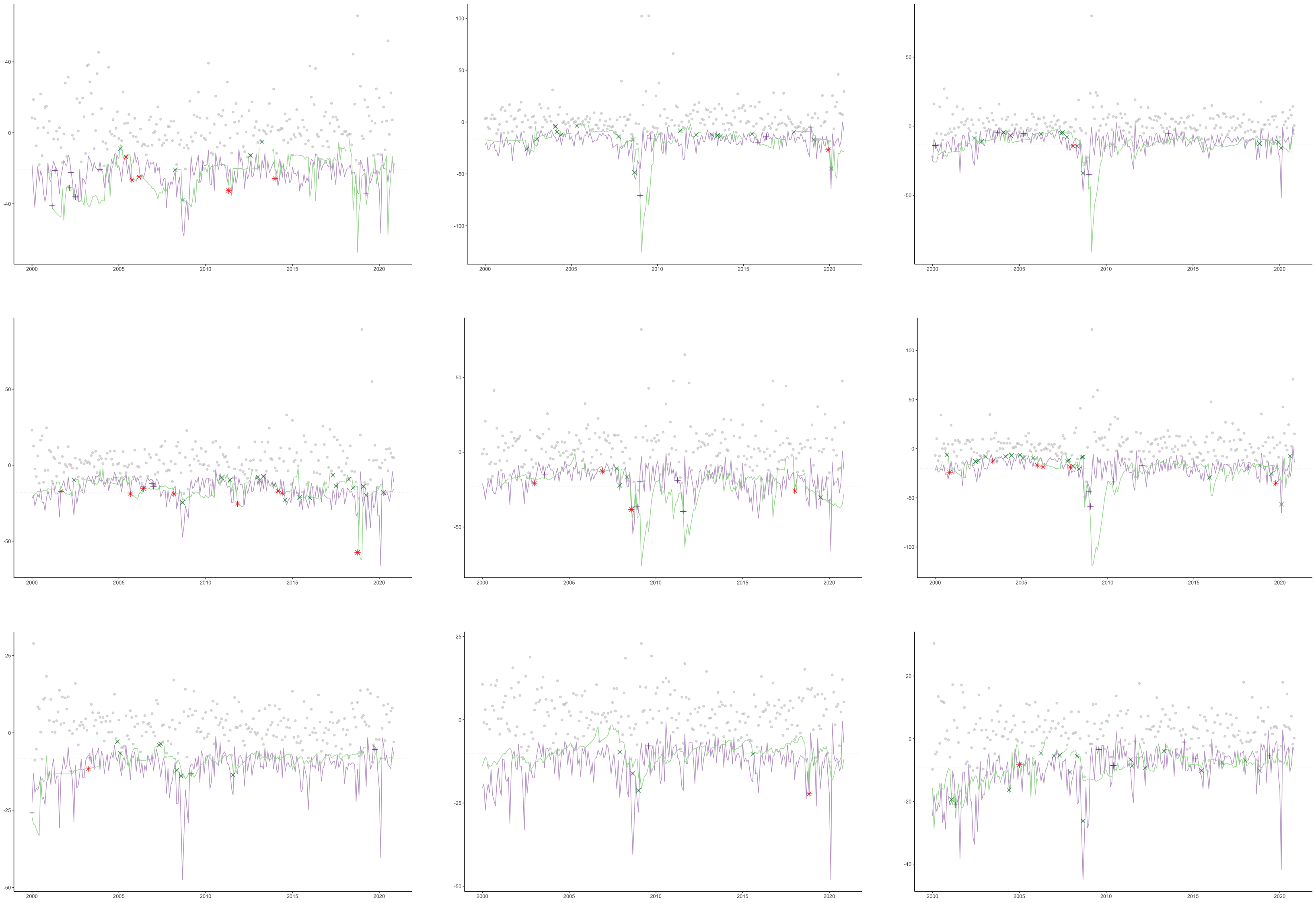

Figure 6: Comparison between the evolution of returns and predictions for quantile 0.05 for both Cross-Sectional Quantile Regression model (purple) and CAViaR model (green) for firms 10550, 17137, 21573, 45728, 51263, 54704, 57568, 61313 and 62092 from 2000 to 2020, monthly

.