Dynamic multiple standardised quantiles for asset returns

Gelly Mitrodima[1] Jaideep Oberoi[2]

[1]Department of Statistics, London School of Economics
[2]School of Finance and Management, SOAS University of London

Outline

Outline for section 1

Overview

- We explore time variation in the shape of the conditional return distribution using a model of multiple quantiles under a semi-parametric framework
- We propose a joint model of scale (proxied by the interquartile range) and other quantiles standardised by the scale
- The model allows us to estimate the scale and shape of the distribution (both time-varying) in one step
- We find that once we effectively capture the dynamics of the scale, the time variation in the shape allows a simpler interpretation

Time-dependence beyond variance

- GARCH (1,1) with standard normal shocks (Bollerslev, 1986) captures a great deal of dependence in the conditional variance, but requires modification to better fit the tails of many return distributions (French et al., 1987; Bollerslev, 1987; Hansen and Lunde, 2005)
- Leverage effect (Schwert, 1989; Nelson, 1991; Glosten et al., 1993; Engle and Ng, 1993)
- Alternative distributions for the underlying shocks (Bollerslev, 1987; McNeil and Frey, 2000)
- Models with tails and long memory in volatility (Ding et al., 1993) or component models (Ding and Granger, 1996; Engle and Lee, 1999)
- Semiparametric GARCH (Engle and Gonzalez-Rivera, 1991; Kalli et al., 2014) and nonparametric stochastic volatility models (Jensen and Maheu, 2010; Delatola and Griffin, 2011)
- Continuous time stochastic volatility models for fat-tailed return distributions with leverage effects as well as jumps in returns and volatility in option pricing (Chernov et al., 2003; Bates, 2006)

Addressing challenges

Despite the above improvements, fitting the tails of the return distribution remains a difficult task.

- Gallant et al. (1991) show that the standardised error term has a time-varying conditional distribution
- As it is difficult to model the entire distribution, modelling efforts beyond the variance have focused on higher moments such as skewness and kurtosis *e.g.* Hansen (1994), Harvey and Siddique (1999), Backus et al. (1997), Christoffersen et al. (2013)
- Kim and White (2004) proposed the use of quantile based estimators (rather than estimators based on averages of past realised moments)
- White et al. (2010) proposed joint models of conditional quantiles to obtain robust estimates of conditional skewness and kurtosis [1]

[1] however, multiple quantiles cannot be satisfactory estimated using separate application of single quantile models at different probability levels since these estimates do not guarantee monotonicity of the quantile function

Addressing challenges: scale information separately obtained and then used to enhance quantile models

- Jeon and Taylor (2013) model value at risk (VaR) by incorporating option implied volatility in a CAViaR model
- Chen and Gerlach (2014) use intra-day data to capture volatility and tail risk for estimating expected shortfall (ES) in an auto-regressive expectile model
- Taylor (2008) estimates VaR and ES using expectiles

Our approach

- We use the information in the daily returns to model both the scale and the shape dynamics in a single estimation step
- We separately specify dynamics for a common time-varying scale and the quantiles standardised by this time-varying scale

Advantages

Our approach:

- naturally extends GARCH and stochastic volatility models and allows for a richer analysis of the conditional distribution via quantiles
- helps avoid quantiles crossing in the estimation procedure
- uses the scale as a common factor that builds dependence across quantiles in a parsimonious manner [2]
- In turn, this makes it easier to estimate the model with the quantiles in their standardised form
- achieves a robust scale and shape decomposition of the conditional return distribution
- allows simpler comparison to commonly used models for financial time series which directly model the conditional variance

[2] intuition: most of the variation in the shape of the conditional return distribution should be captured by a time-varying volatility process

Motivation: Scale and shape decomposition

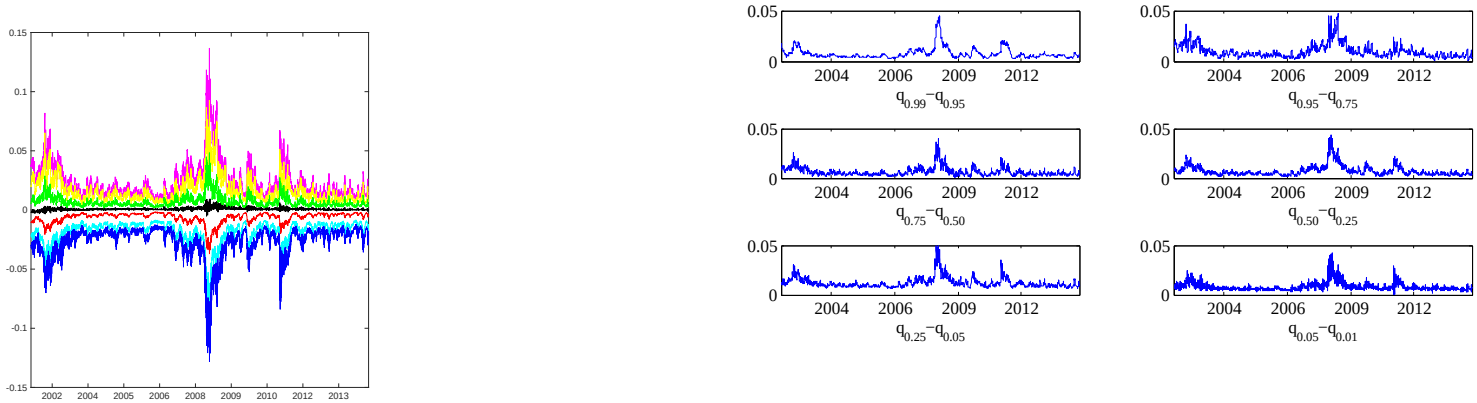

Figure: SPX from January 01, 2001 to December 31, 2014. (a) Separately estimated conditional quantiles (b) Differences between adjacent conditional quantiles

Contributions

- We present some novel dynamic multiple standardised quantile (DMSQ) models. Under a semiparametric framework estimation and prediction is fast and simple to implement
- We provide the asymptotic theory for a general class of semiparametric DMSQ models. [3]
- Drawing on Fissler and Ziegel (2016), our third contribution is to address the problem of elicitability directly. This introduces new directions in multiple quantile modeling and prediction
- Our forth and final contribution is an application of our new models and estimation methods on four equity indices over a long period ranging from January 2001 to November 2024. We compare these new models with existing methods from the literature

[3] This theory is an extension of results for VaR presented in Weiss (1991) and Engle and Manganelli (2004), and draws on identification results in Fissler and Ziegel (2016) and results for M-estimators in Newey and McFadden (1994)

Outline for section 2

DMSQ models

The model can be expressed in terms of the conditional quantiles of the return distribution at probability levels $\theta_1, \theta_2, \ldots, \theta_K$ using the transformations $q_{\theta_1,t} = q_{\theta_2,t} + s_t$ and $q_{\theta_k,t} = s_t q^{(z)}_{\theta_k,t}$ for $k = 2, \ldots, K$ leading to the following model for $q_{\theta_1,t}, q_{\theta_2,t}, \ldots, q_{\theta_K,t}$

$$s_t = u^{(s)} + \sum_{i=1}^{M} \beta_i^{(s)} s_{t-i} + \sum_{j=1}^{L} \gamma_j^{(s)} \, l_j \left(y_1, \ldots, y_{t-1}\right) \tag{1}$$

$$q_{\theta_1,t} = q_{\theta_2,t} + s_t \tag{2}$$

$$q_{\theta_k,t} = s_t \left(u^{(z)}_{\theta_k} + \sum_{i=1}^{M} \beta^{(z)}_{\theta_k,i} \frac{q_{\theta_k,t-i}}{s_{t-i}} + \sum_{j=1}^{L} \gamma^{(z)}_{\theta_k,j} \, l_j \left(\frac{y_1}{s_1}, \ldots, \frac{y_{t-1}}{s_{t-1}} \right) \right) \tag{3}$$

where $k = 2, \ldots, K, M$ and L are the orders of the model, $u^{(s)}$, $\beta_i^{(s)}$, and $\gamma_i^{(s)}$ are scalars, $u^{(z)}_{\theta_k}$ and $\gamma^{(z)}_{\theta_k,j}$ are $(K-1)$-dimensional vector, $\beta^{(z)}_{\theta_k,i}$ is a $((K-1) \times (K-1))$-dimensional matrix with the non-diagonal equal to zero and $l_j \left(y_1, \ldots, y_{t-1}\right)$ represents a function of the history of the return series

DMSQ-SAV

Choosing $M = 1$, $L = 1$, $\theta_1 = 75\%$, $\theta_2 = 25\%$ and $l_1(y_1, \ldots, y_{t-1}) = |y_{t-1}|$ leads to a generalisation of the symmetric absolute value (SAV) CAViaR model Engle and Manganelli (2004), and has the form

$$s_t = u^{(s)} + \beta^{(s)} s_{t-1} + \gamma^{(s)} |y_{t-1}| \tag{4}$$

$$q_{\theta_1,t} = q_{\theta_2,t} + s_t \tag{5}$$

$$q_{\theta_k,t} = s_t \left(u^{(z)}_{\theta_k} + \beta^{(z)}_{\theta_k} \frac{q_{\theta_k,t-1}}{s_{t-1}} + \gamma^{(z)}_{\theta_k} \frac{|y_{t-1}|}{s_{t-1}} \right), \quad k = 3, \ldots, K \tag{6}$$

where we drop 1 from $\beta^{(z)}_{\theta_k,1}$ and $\gamma^{(z)}_{\theta_k,1}$ for ease of reading

DMSQ-AS

We choose $l_1(y_1, \ldots, y_{t-1}) = y_{t-1}^+$ and $l_1(y_1, \ldots, y_{t-1}) = y_{t-1}^-$ where the notation employed is: $y^+ = \max(y, 0)$ and $y^- = -\min(y, 0)$. This leads to a generalisation of the asymmetric slope (AS) CAViaR model Engle and Manganelli (2004), which has the form

$$s_t = u^{(s)} + \beta^{(s)} s_{t-1} + \gamma^{(s)} y_{t-1}^+ + \delta^{(s)} y_{t-1}^- \tag{7}$$

$$q_{\theta_1,t} = q_{\theta_2,t} + s_t \tag{8}$$

$$q_{\theta_k,t} = s_t \left(u_{\theta_k}^{(z)} + \beta_{\theta_k}^{(z)} \frac{q_{\theta_k,t-1}}{s_{t-1}} + \gamma_{\theta_k}^{(z)} \frac{y_{t-1}^+}{s_{t-1}} + \delta_{\theta_k}^{(z)} \frac{y_{t-1}^-}{s_{t-1}} \right), \quad k = 2, \ldots, K \tag{9}$$

Elicitability

A risk measure (statistical functional) is elicitable if there exists a loss function such that the risk measure is the solution to minimising the expected loss *e.g.* VaR is elicitable using the tick loss function.
Thus, we need to show that the new DMSQ models have such loss function.
We show that $T = (q_{\theta_1}, \cdots, q_{\theta_6}, s)$ is elicitable and identifiable. To do so, we follow similar steps to Fissler and Ziegel (2021), who show that $T' = (q_{\theta_1}, q_{\theta_2}, \text{RVaR}_{\theta_1,\theta_2})$ ($\text{RVaR}_{\theta_1,\theta_2}$ is the range VaR) is elicitable under weak regularity conditions.

Consistent scoring rule for DMSQ

We show that the following class of loss functions (indexed by the functions G_r and H) is consistent for $Q = [q_{\theta_1}, \cdots, q_{\theta_6}]$, and s. That is, minimising the expected loss using any of these loss functions returns the true Q and s.

$$L_0(y, Q, s; \theta, G_r, G) = \sum_{r=1}^{(K+1)-1} (\mathbf{1}_{\{y \leq q_r\}} - \theta_r) G_r(q_r) - \mathbf{1}_{\{y \leq q_r\}} G_r(y) \quad (10)$$

$$+ H'(s)\left(s + L_{\theta_2}(q_2, y) - L_{\theta_1}(q_1, y)\right) \quad (11)$$

$$- H(s) + a(y) \quad (12)$$

where the tick loss function $L_{\theta_r} = (\mathbf{1}_{\{y \leq q_r\}} - \theta_r)(q_r - y)$

Normalised distance loss function

We choose $G_r(q_r) = q_r$ and $H(s) = |s - c|y||$, where $c = 1.349\sqrt{(\pi/2)}$. Thus, the normalised distance loss function, L_{ND}, is a consistent loss function for (Q, s)

$$L_{ND}(y, Q, s; \theta) = \sum_{r=1}^{6} (\mathbf{1}_{\{y \leq q_r\}} - \theta_r)(q_r - y) \tag{13}$$

$$+ \frac{s - c|y|}{|s - c|y||}(s + L_{\theta_2}(q_2, y) - L_{\theta_1}(q_1, y)) \tag{14}$$

$$- |s - c|y|| \tag{15}$$

We choose to work with the absolute distances of the scale s from the interquantile of the (relevant) standard normal distribution.

Proposition

Under the assumption that the conditional quantiles of interest can be negative and positive while the IQR is strictly positive, the loss differences generated by the proposed loss function are homogeneous of degree zero iff $G_r(q_r) = q_r$ and $H(s) = \left|s - 1.349\sqrt{(\pi/2)}|y|\right|$.

Outline for section 3

Estimation of DMSQ models

The parameters of the DMSQ models are estimated by minimising the normalised distance loss function, L_{ND}. Let a sample of observations $y_1, \ldots, y_n$ generated by the model

$$y_t = Q_t(\beta^0)s_t(\beta^0), t = 1, \ldots, n \tag{16}$$

The vector of unknown parameters to be estimated is $\beta^0 \in B \subset \mathbb{R}^p$. Let

$$L_{ND,n}(\beta) = n^{-1} \sum_{t=1}^{n} L_{ND}(y_t, Q_t(\beta), s_t(\beta); \theta)$$

then the unknown parameters are estimated by minimising

$$\hat{\beta}_n = \arg\min_{\beta \in B} L_{ND,n}(\beta) \tag{17}$$

at fixed probability levels $\theta = [75\%, 25\%, 1\%, 5\%, 95\%, 99\%]$.

Consistency

Theorem 1

$\hat{\beta}_n \xrightarrow{p} \beta^{\mathbf{0}}$ *as* $n \to \infty$

For the financial time series considered here, various uniform laws of large numbers can be applied to verify the assumptions. We also consider assumptions that are standard in time series parameter inference.

Asymptotic normality

Theorem 2

$$\sqrt{n}\boldsymbol{A}_n^{-1/2}\boldsymbol{D}_n(\hat{\beta}_n - \beta^{\mathbf{0}}) \xrightarrow{d} N(0, I) \text{ as } n \to \infty$$

where

$$\boldsymbol{D}_n = \mathbb{E}\left[f_t(Q_t(\beta^0)|\mathcal{F}_{t-1})\nabla' Q_t(\beta^0)\nabla Q_t(\beta^0) + \frac{s_t(\beta^0) - c|y_t|}{\left|s_t(\beta^0) - c|y_t|\right|}\nabla' s_t(\beta^0)\nabla s_t(\beta^0)\right]$$

$$\boldsymbol{A}_n = \mathbb{E}[h_t(\beta^0)h_t(\beta^0)']$$

$$h_t(\beta) = \frac{\partial L_{ND}(y_t, Q_t(\beta), s_t(\beta); \theta)}{\partial \beta}$$

$$= \nabla' Q_t(\beta^0)\sum_{r=1}^{6}\left(\mathbf{1}_{\{y_t \le q_{r,t}\}} - \theta_r\right) + \sum_{r=1}^{6}\left(\mathbf{1}_{\{y_t \le q_{r,t}\}} - \theta_r\right)(q_{r,t}(\beta) - y_t)$$

$$+ \frac{s_t(\beta^0) - c|y_t|}{\left|s_t(\beta^0) - c|y_t|\right|}(\nabla' s_t(\beta^0) + \nabla' L_{\theta_2} - \nabla' L_{\theta_1})$$

$$+ \frac{s_t(\beta^0) - c|y_t|}{\left|s_t(\beta^0) - c|y_t|\right|}\nabla' s_t(\beta^0)$$

Asymptotic normality cont.

$$\nabla' L_{\theta_2} = \nabla' q_{2,t}(\beta^0)(\mathbf{1}_{\{y_t \leq q_{2,t}\}} - \theta_2) + (\mathbf{1}_{\{y_t \leq q_{2,t}\}} - \theta_2)(q_{2,t}(\beta) - y_t)$$
$$\nabla' L_{\theta_1} = \nabla' q_{1,t}(\beta^0)(\mathbf{1}_{\{y_t \leq q_{1,t}\}} - \theta_1) + (\mathbf{1}_{\{y_t \leq q_{1,t}\}} - \theta_1)(q_{1,t}(\beta) - y_t)$$

Asymptotic covariance matrix

Theorem 3

$\hat{\boldsymbol{A}}_n - \boldsymbol{A}_n \xrightarrow{p} \mathbf{0}$ *and* $\hat{\boldsymbol{D}}_n - \boldsymbol{D}_n \xrightarrow{p} \mathbf{0}$ *where*

$$\hat{\boldsymbol{D}}_n = n^{-1} \sum_{t=1}^{n} \Big[\frac{1}{2c_n} \mathbf{1}\Big(|y_t - Q_t(\hat{\beta}_n)| < c_n \Big) \nabla' Q_t(\hat{\beta}_n) \nabla Q_t(\hat{\beta}_n) + \frac{s_t(\hat{\beta}_n) - c|y_t|}{\Big| s_t(\hat{\beta}_n) - c|y_t| \Big|} \nabla' s_t(\hat{\beta}_n) \nabla s_t(\hat{\beta}_n) \Big]$$

$$\hat{\boldsymbol{A}}_n = n^{-1} \sum_{t=1}^{n} h_t(\hat{\beta}_n) h_t(\hat{\beta}_n)'$$

Algorithm

- The proposed loss function can be highly multi-modal for multi-quantile time series models. The estimation issue of quantiles (single and multiple) is well known in the literature see *e.g.* Patton et al. (2019); Engle and Manganelli (2004); White et al. (2012)
- We apply a combination of estimation strategies
  - to obtain "warm" starting values by considering the properties of the conditional quantiles (stationarity and quantile ordering) (Mitrodima and Oberoi, 2023)
  - to allow for some smoothing (Patton et al., 2019)

This procedure leads to improved estimates and is particularly well suited to our model.

DMSQ-SAV: "warm" starting values

- Small set of possible values (4 sets for 2 values) for $\beta^{(z)}_{\theta_k,i}$ and $\beta^{(s)}_i$ between 0.5 and 0.9
- Associated with each set of βs possible values for $\gamma^{(z)}_{\theta_k,i}$ are $\{-0.2, -0.1, -0.02, -0.01, 0\}$ if $\theta_k < 0.5$ or $\{0, 0.01, 0.02, 0.1, 0.2\}$ if $\theta_k > 0.5$
- $\gamma^{(s)}_i$ and $\gamma^{(z)}_{0.5,i}$, both positive and negative values from the above sets are chosen
- $u^{(z)}_{\theta_k}$ we assume the unconditional mean of $q_{\theta_k,t}/s_t$ is equal to the corresponding quantile of a standard normal distribution (although, other distribution could be used)
- The initial value for $u^{(z)}_{\theta_k}$ is $u^{(z)}_{\theta_k} = (1 - \beta^{(z)}_{\theta_k,1})\overline{q_{\theta_k}} - \gamma^{(z)}_{\theta_k,1} E\left[\frac{|y_t|}{s_t}\right]$, where $\overline{q_{\theta_k}}$ is the quantile of standard normal distribution at probability level θ_k
- We also set $E\left[\frac{|y_t|}{s_t}\right] = \frac{1}{1.347}\sqrt{\frac{2}{\pi}}$. A similar strategy is applied to the IQR, whereby $u^{(z)}_t$ is chosen so that the unconditional mean of s_t is equal to the corresponding sample IQR

Estimation approach

- The procedure for finding initial values exploits the use of standardised quantiles [4]
- Following Engle and Manganelli (2004), we initialise the quantile and interquartile range series at their empirical counterparts (using the first 300 observations)
- For smoothing the indicator function we replace this by a Logistic function, since the smoothing function of the Logistic function converges to the indicator function as the smoothing parameter $\nu \to \infty$. Here we choose $\nu = 5$ and $\nu = 20$ (Patton et al., 2019)
- The smoothed objective function is differentiable and so we first use gradient-based numerical search algorithms (*e.g.* fminunc or fmincon) for the set of initial parameter values for $\nu = 5$ and then for $\nu = 20$
- Separate optimisers of the loss function are then run for the obtained initial parameter values ("fminsearch"), and the largest value parameter is reported as the estimate

[4] Our estimation approach is similar in spirit to the variance targeting strategy explained in Christoffersen (2003), but it is only used to initialise the optimisation, not

Outline for section 4

Data and models

- We use data provided by Datastream
- We analyse the S&P500 index (SPX), along with Russell 2000 (RUT), Russell 1000 (RUI), and NASDAQ 100 (NDX)
- The complete data set consists of daily log returns from January $01, 2001$ to November $20, 2024$
- For comparison, we estimate a GARCH(1,1)[5] and the autoregressive conditional density model (ARCD) by Hansen (1994)[6] as well as our proposed scale-shape models DMSQ-SAV and DMSQ-AS for the $1\%, 5\%, 25\%, 50\%, 75\%, 95\%$ and 99% probability levels.

[5] Monte Carlo simulation: standardise residuals by GARCH(1,1) and draw from them randomly and use them as returns in a new GARCGH(1,1) of length 1000.

[6] Hansen (1994) proposed an explicit model of time-varying higher moments, where he considers a parsimonious density function which generalises the standard normal to generate heavy tails, skewness, or bi-modality

Conditional density: unstandardised quantile estimates for SPX

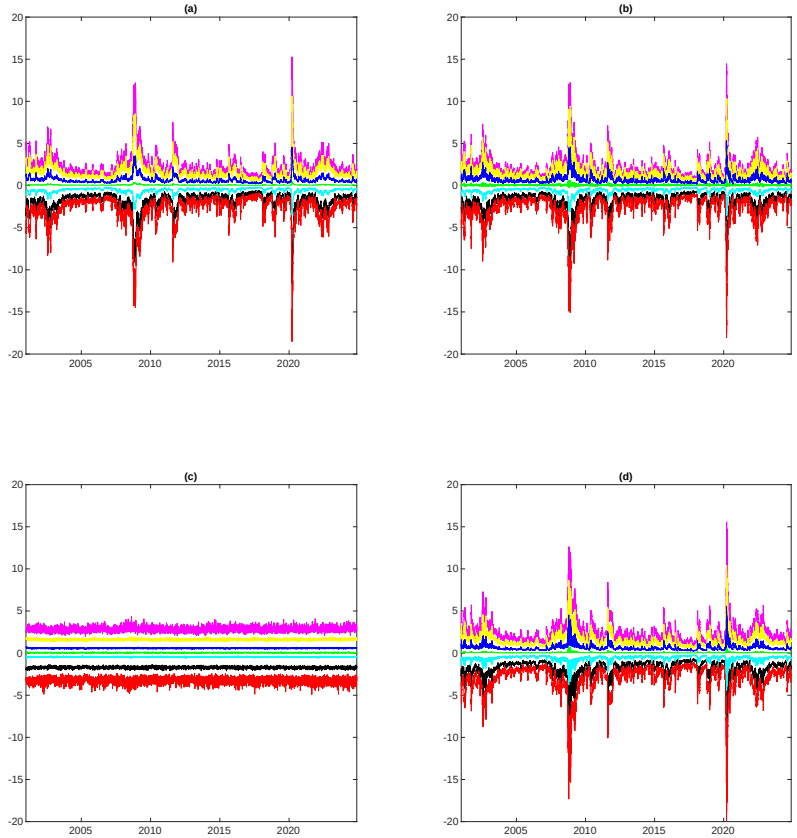

Figure: (a) DMSQ-SAV, (b) DMSQ-AS, (c)GARCH, (d)ACRD

Conditional density: standardised quantile estimates for SPX

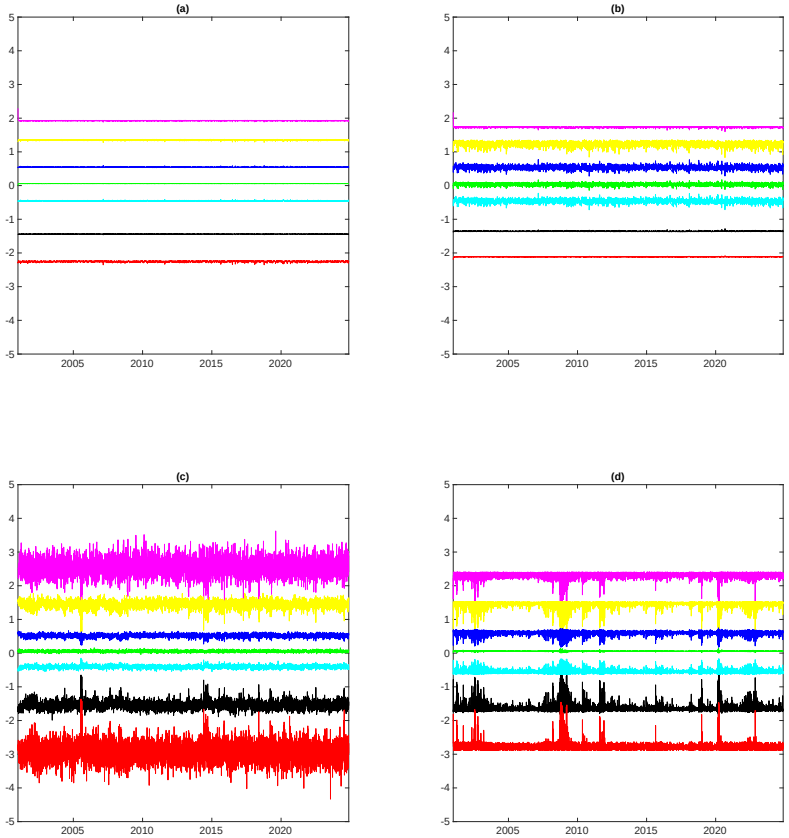

Figure: (a) DMSQ-SAV, (b) DMSQ-AS, (c)GARCH, (d)ACRD

Robust estimators of higher moments

The estimates of the conditional skewness and kurtosis can be calculated from the estimated conditional quantiles using the following formulae suggested by White et al. (2010) and Kim and White (2004), who follow the statistics literature on robust estimators of higher moments. The daily conditional skewness is estimated as

$$\frac{q_{0.75,t} + q_{0.25,t} - 2 \times q_{0.50,t}}{q_{0.75,t} - q_{0.25,t}}$$

and the kurtosis as

$$\frac{q_{0.99,t} - q_{0.01,t}}{q_{0.75,t} - q_{0.25,t}} - 3.45$$

Conditional skewness for SPX

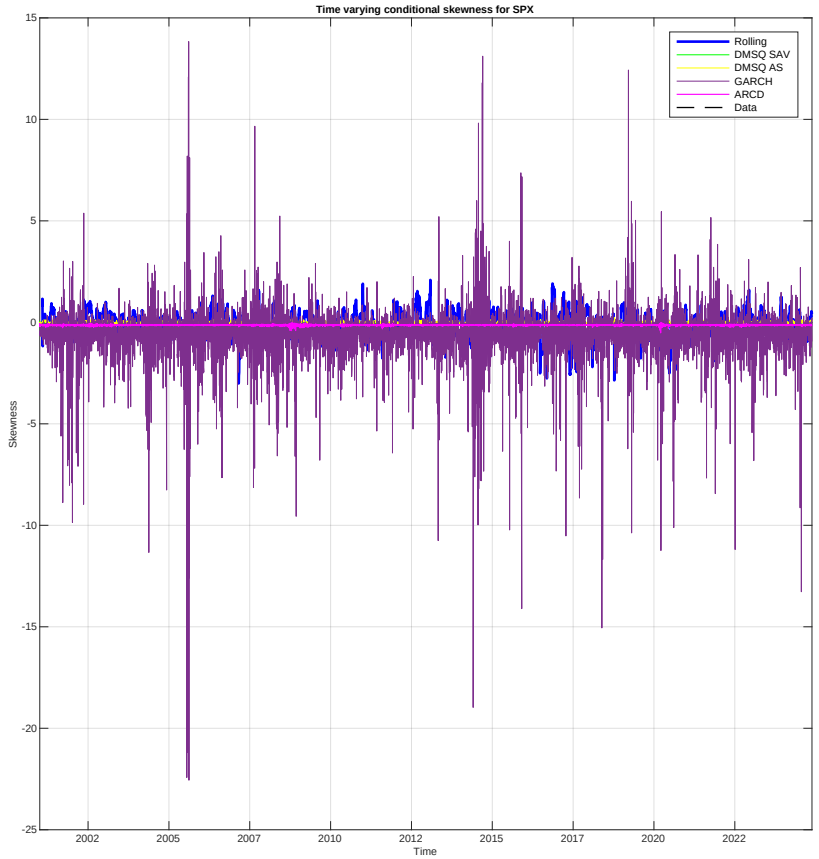

Figure: Skewness

Conditional kurtosis for SPX

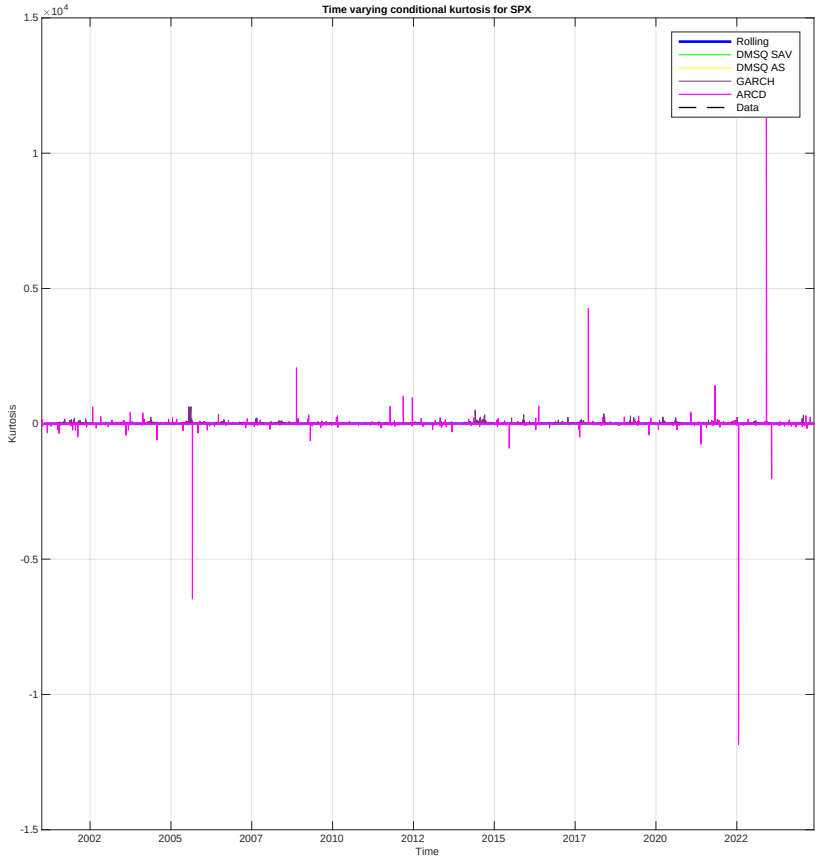

Figure: Kurtosis

Hit ratios

One way to evaluate the fit or the performance of quantile models is to compare their in-sample hit ratios (coverage). We report in-sample hit ratios ($\frac{1}{T}\sum_{t=1}^{T} I(y_t < q_{\theta,t})$) for each of the joint quantile models and their benchmarks for SPX.

Table: In-sample hit ratios ($\frac{1}{T}\sum_{t=1}^{T} I(y_t < q_{\theta,t})$) for SPX

θ	**0.99**	**0.95**	**0.50**	**0.05**	**0.01**	**0.25**	**0.75**
DMSQ SAV	**0.9903**	**0.9499**	0.4963	**0.0511**	**0.0117**	0.2422	**0.7505**
DMSQ AS	0.9888	**0.9501**	0.4846	0.0519	0.0120	0.2397	0.7508
ARCD	0.9752	0.9176	0.4962	0.0661	0.0195	0.2287	0.7485
GARCH	0.9842	0.9441	**0.4993**	0.0588	0.0145	**0.2543**	0.7457

Note: The sample ranges from January 01, 2001 to November 20, 2024. With bold we indicate the best performing model.

Outline for section 5

Conclusion

- A joint multi-quantile time series model is proposed for the conditional distribution of asset returns, which separately specifies
  - A process for the scale, proxied by the interquartile range
  - A process for the standardised conditional quantiles, which captures the distributional shape
- This structure mirrors standard financial time series models (*e.g.* GARCH) where returns = scale$\times$standardised shock
- Key benefits:
  - A robust estimate of time-varying scale
  - Time-varying shape estimate, offering better insight into return distribution evolution
  - The decomposition helps address quantile crossing, and the algorithm improves estimation

Conclusion

- Results show
  - Scale captures most dynamics, leaving the shape relatively stable
  - When scale is modelled properly, simple dynamics suffice to represent the standardised tails
- Implication: Complex volatility models are useful, but complicated distributional assumptions for shocks may be unnecessary

References I

Backus, D., Foresi, S., Li, K., and Wu, L. (1997). Accounting for Biases in Black-Scholes. Technical Report 30, CRIF Working Paper series.

Bates, D. S. (2006). Maximum likelihood estimation of latent affine processes. *Review of Financial Studies*, 19(3):909–965.

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

Bollerslev, T. (1987). A conditionally heteroskedastic time series model for speculative prices and rates of return. *The Review of Economics and Statistics*, 69(3):542–47.

Chen, C. and Gerlach, R. (2014). Bayesian expected shortfall forecasting incorporating the intraday range. *Journal of Financial Econometrics*, 14:128–158.

References II

Chernov, M., Gallant, R. A., Ghysels, E., and Tauchen, G. (2003). Alternative models of stock price dynamics. *Journal of Econometrics*, 116(1-2):225–257.

Christoffersen, P. (2003). *Elements of Financial Risk Management*. Academic Press, London.

Christoffersen, P., Heston, S., and Jacobs, K. (2013). Capturing option anomalies with a variance-dependent pricing kernel. *Review of Financial Studies*, 26(8):1963–2006.

Delatola, E.-I. and Griffin, J. E. (2011). Bayesian nonparametric modelling of the return distribution with stochastic volatility. *Bayesian Analysis*, 6(4):901–926.

Ding, Z. and Granger, C. W. (1996). Modeling volatility persistence of speculative returns: A new approach. *Journal of Econometrics*, 73(1):185 – 215.

References III

Ding, Z., Granger, C. W., and Engle, R. F. (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, 1(1):83 – 106.

Engle, R. F. and Gonzalez-Rivera, G. (1991). Semiparametric ARCH models. *Journal of Business & Economic Statistics*, 9(4):345–359.

Engle, R. F. and Lee, G. (1999). A permanent and transitory component model of stock return volatility. In *Cointegration, Causality and Forecasting: A festschrift in honor of Clive W.J. Granger*, pages 475–497. Oxford University Press.

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

Engle, R. F. and Ng, V. K. (1993). Measuring and testing the impact of news on volatility. *The Journal of Finance*, 48(5):1749–1778.

References IV

Fissler, T. and Ziegel, J. F. (2016). Higher order elicitability and OsbandâĂŹs principle. *The Annals of Statistics*, 44(4):1680–1707.

Fissler, T. and Ziegel, J. F. (2021). On the elicitability of range value at risk. *Statistics & Risk Modeling*, 38(1-2):25–46.

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

Gallant, R. A., Hsieh, D. A., and Tauchen, G. E. (1991). On fitting a recalcitrant series: The pound/dollar exchange rate, 1974-1983. In Barnett, W. A., Powell, J., and Tauchen, G. E., editors, *Nonparametric and Semiparametric Methods in Economics and Statistics*, pages 199–2401. Cambridge University Press, Cambridge.

Glosten, L. R., Jagannathan, R., and 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.

References V

Hansen, B. E. (1994). Autoregressive conditional density estimation. *International Economic Review*, 35:705–730.

Hansen, P. R. and Lunde, A. (2005). A forecast comparison of volatility models: does anything beat a garch(1,1)? *Journal of Applied Econometrics*, 20(7):873–889.

Harvey, C. R. and Siddique, A. (1999). Autoregressive conditional skewness. *Journal of Financial and Quantitative Analysis*, 34(4):465–487.

Jensen, M. J. and Maheu, J. M. (2010). Bayesian semiparametric stochastic volatility modeling. *Journal of Econometrics*, 157:306–316.

Jeon, J. and Taylor, J. W. (2013). Using CAViaR models with implied volatility for value-at-risk estimation. *Journal of Forecasting*, 32(1):62–74.

References VI

Kalli, M., Walker, S. G., and Damien, P. (2014). Modelling the conditional distribution of daily stock index returns: an alternative bayesian semi-parametric model. *Journal of Business and Economic Statistics*, 31:371–383.

Kim, T.-H. and White, H. (2004). On more robust estimation of skewness and kurtosis. *Finance Research Letters*, 1(1):56–73.

McNeil, A. J. and Frey, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. *Journal of Empirical Finance*, 7(3ÃćâĆňâĂIJ4):271–300. Special issue on Risk Management.

Mitrodima, G. and Oberoi, J. (2023). CAViaR models for Value-at-Risk and Expected Shortfall with long range dependency features. *Journal of the Royal Statistical Society Series C: Applied Statistics*.

Nelson, D. B. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica*, 59(2):347–70.

References VII

Newey, W. K. and McFadden, D. (1994). Large sample estimation and hypothesis testing. In Engle, R. F. and McFadden, D. L., editors, *Handbook of Econometrics*, volume 4, pages 2111–2245. Elsevier.

Patton, A. J., Ziegel, J. F., and Chen, R. (2019). Dynamic semiparametric models for expected shortfall (and value-at-risk). *Journal of Econometrics*, 211(2):388–413.

Schwert, G. W. (1989). Why does stock market volatility change over time? *The Journal of Finance*, 44:1115–1153.

Taylor, J. (2008). Estimating value at risk and expected shortfall using expectiles. *Journal of Financial Econometrics*, 6:231–252.

Weiss, A. A. (1991). Asymmetric garch models and financial time series. In Fomby, T. B. and Seo, T. K., editors, *Nonlinear Models*, volume 11 of *Advances in Econometrics*, pages 95–109. JAI Press.

References VIII

White, H., Kim, T.-H., and Manganelli, S. (2010). Modeling autoregressive conditional skewness and kurtosis with multi-quantile CAViaR. In *Volatility and Time Series Econometrics: Essays in Honour of Robert Engle, ed. by M. Watson, T. Bollerslev and J. Russell*. Oxford University Press.

White, H., Kim, T.-H., and Manganelli, S. (2012). Var for var: Measuring tail dependence using multivariate regression quantiles. Working papers 2012rwp-45, Yonsei University, Yonsei Economics Research Institute.