Temporal Aggregation of GARCH Models: Conditional Kurtosis and Optimal Frequency

Thomas Breuer Martin Jandačka

Research Centre PPE, FH Vorarlberg, Dornbirn, Austria

SSEV
St. Gallen, 22 March, 2007

Outline

Motivation

Risk Management at Long Time Horizons

- Use high frequency data
  if time horizon > 1 / data frequency?
  If yes: need to aggregate
- Drost & Nijman:
  (Strong) GARCH does not aggregate to (strong) GARCH
  Weak GARCH does;
  but gives not enough information for risk management:
  only best linear predictor of mean and variance,
  no higher moments, no quantiles
- Thus: Aggregate strong GARCH - even if painful

Aggregation of strong GARCH

Assume ϵ is strong GARCH with density f.
Density of conditional m-period aggregated distribution $\epsilon_{(m)t} := \sum_{i=0}^{m-1} \epsilon_{t+i}$ is

$$f_{\epsilon_{(m)t}|I_{t-1}}(y) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} d\epsilon_t \dots d\epsilon_{t+m-2}$$

$$f_{\epsilon_{t+m-1}|I_{t+m-2}}\left(y - \sum_{i=1}^{m-1} \epsilon_{t+i-1}\right) \prod_{i=1}^{m-1} f_{\epsilon_{t+i-1}|I_{t+i-2}}(\epsilon_{t+i-1}).$$

Goal

- Calculate conditional vola and kurtosis of aggregate distribution $\epsilon_{(m)t}$
- Analyse long run behaviour of conditional vola and kurtosis
- If high frequency data is available for long run predictions: Determine optimal basic frequency to be aggregated.

Definition strong GARCH

Let $\{h_t, t \in \mathbb{Z}\}$ be defined as the stationary solution of

$$h_t = \psi + \alpha \epsilon_{t-1}^2 + \beta h_{t-1}.$$

A time series $\{\epsilon_t, t \in \mathbb{Z}\}$ is said to be generated by a *strong GARCH(1,1) process* if ψ, α, β can be chosen in such a way that

$$\xi_t \quad := \quad \epsilon_t / \sqrt{h_t} \sim D(0,1) \qquad i.i.d., \tag{1}$$

where $D(0,1)$ is some fixed distribution of errors with zero mean and unit variance.

Conditional variance of aggregated GARCH

ϵ_t a flow variable following a semi-strong GARCH process.
Then for aggregated $\epsilon_{(m)t} := \sum_{i=0}^{m-1} \epsilon_{t+i}$

$$Var(\epsilon_{(m)t}|I_{t-1}) = m\sigma^2 + \frac{1-(\alpha+\beta)^m}{1-(\alpha+\beta)}(h_t - \sigma^2),$$

where σ^2 is the unconditional variance of one period returns ϵ_t,
$\sigma^2 = \psi/(1-\alpha-\beta)$.

Conditional kurtosis of aggregated GARCH

Theorem

If innovations ξ_t are symmetric with finite fourth moment equal to κ, the conditional fourth moments of the aggregated $\epsilon_{(m)t}$ are

$$E(\epsilon_{(m)t}^4 | I_{t-1}) = B \sum_{i=0}^{m-1} A^i b,$$

where $B := (1, h_t, h_t^2, 0, 0)$, $b' = (0, 0, \kappa, 0, 6)$, *and*

$$A = \begin{pmatrix} 1 & \psi & \psi^2 & 0 & 0 \\ 0 & \alpha+\beta & 2\psi(\alpha+\beta) & 1 & \psi \\ 0 & 0 & \kappa\alpha^2 + 2\alpha\beta + \beta^2 & 0 & \kappa\alpha+\beta \\ 0 & 0 & 0 & 1 & \psi \\ 0 & 0 & 0 & 0 & \alpha+\beta \end{pmatrix}.$$

Long term behaviour of conditional kurtosis

Kurtosis of the conditional aggregated dist $\kappa_m := \frac{E(\epsilon^4_{(m)t}|I_{t-1})}{\left(E(\epsilon^2_{(m)t}|I_{t-1})\right)^2}$.

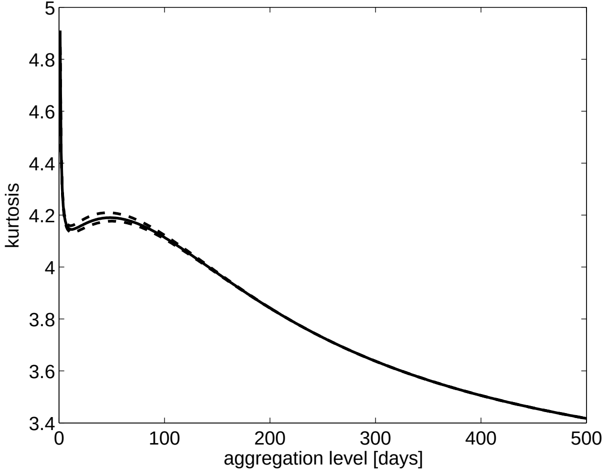
Corollary

$$\lim_{m\to\infty} \kappa_m = \begin{cases} 3 & \textit{if } \gamma < 1 \\ \infty & \textit{if } \gamma > 1 \\ 3+d & \textit{if } \gamma = 1 \end{cases},$$

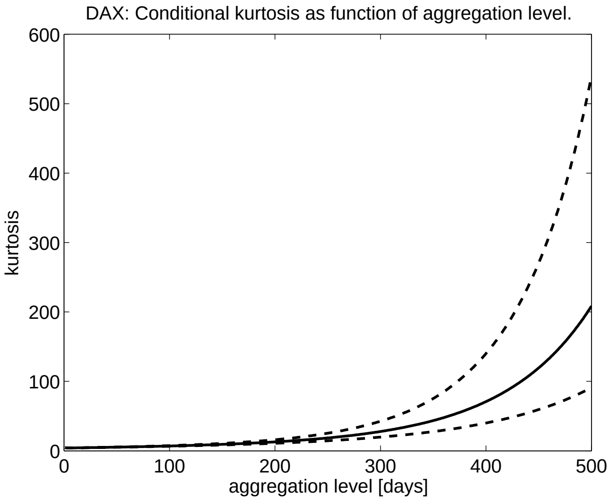
where $\gamma := \alpha^2\kappa + 2\alpha\beta + \beta^2$,
$d = \frac{\psi\kappa(\psi/\sigma^2+2(\alpha+\beta))}{2\sigma^2-3(\kappa\alpha+\beta)}$,
and κ *unconditional kurtosis of* ϵ_t.

Long run conditional kurtosis of DAX

GBP/EUR via DEM: Conditional kurtosis as function of aggregation level.



Long run conditional kurtosis of GBP/EUR



Optimal Frequency of GARCH Models: Two questions

1. At which frequency is the assumption of strong GARCH processes best justified?
2. When making forecasts over a long time horizon, should we use higher frequency data if available?

Estimate Optimal Frequency with Quasi Max Likelihood

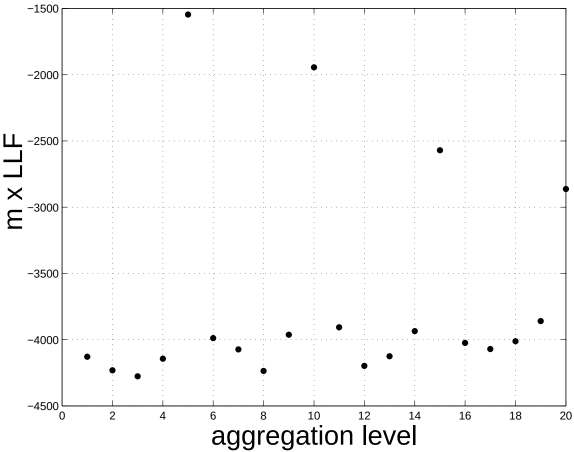
N observations of highest frequency
$n = \lfloor N/m \rfloor$ observations of m periods each.

The optimal parameters $\theta = (m, \psi, \alpha, \beta)$ are those which maximise

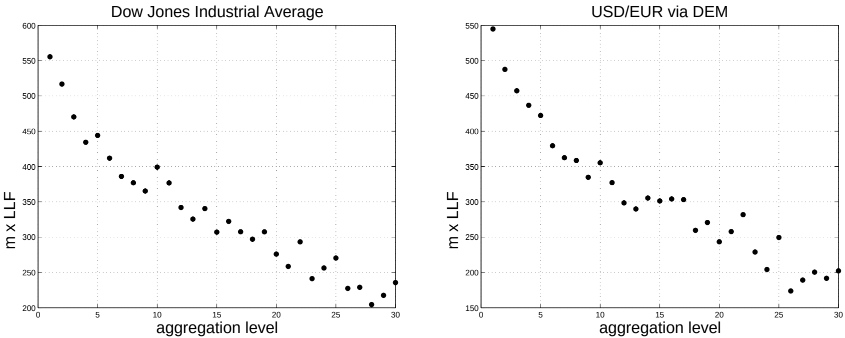
$$\begin{aligned} \theta^* &= \mathrm{argmax}_\theta \; m \cdot LLF(\theta) \\ &= \operatorname*{arg\,max}_{m,\psi,\alpha,\beta>0} m \sum_{j=1}^{n} \ln f_\theta \left(\epsilon_{(m)j} | I_{(m)(j-1)} \right) . \end{aligned}$$

Test of procedure

On data simulated to be GARCH at aggregation $m = 5$.



Optimal frequency of DAX and EUR/USD



Strong GARCH has higher likelihood at high frequencies than at low frequencies.

Evaluate long term density forecast capacity of models

Models:

- Basic frequencies 1d, 2d, 5d, 10d, 20d, 30d, 60d aggregated to 60d
- Residual distribution normal, Student t, or EVT (80%body from historical simulation, tails fitted to generalised Pareto)

Which model produces best 60 day density forecasts from daily data?

Test procedure 1

Goal: Evaluate out of sample density forecast of various models

- 60 day returns r_t from some true conditional densities $f_t(.)$.
- Model produces 60 day conditional density forecasts $p_t(.)$
- $z_t := \int_{-\infty}^{r_t} p_t(u)du$ uniformly i.i.d. if true density f_t agrees with predicted density p_t.
- $n_t := \Phi^{-1}(z_t)$ are N(0,1) i.i.d. if true density f_t agrees with predicted density p_t.
- Perform various tests of normality and iid on n_t.

Test procedure 2

- Kolmogorov-Smirnov on n_t.
- Joint Wald test for $\beta_0 = 0, \beta_1 = 0, \gamma_0 = 1$ and $\gamma_1 = 0$ in

$$\begin{aligned} n_t &= \beta_0 + \beta_1 n_{t-1} + u_t \\ n_t^2 &= \gamma_0 + \gamma_1 n_{t-1}^2 + v_t \end{aligned}$$

- Jarque-Bera for skewness and kurtosis of n_t
- χ^2 test for unifomity of z_t in five subintervals of [0,1].
- χ^2 test in four quadrants of 2-dim plot of adjacent obervations n_t.

Test results on 19 market time series

Model	Test 1 (KS) # accept.	Test 2 (JB) # accept.	Test 3 (W) # accept.	Test 4 (χ^2-1d) # accept.	Test 5 (χ^2-2d) # accept.
1d_G_N	19	0	0	0	0
1d_G_t	0	0	0	0	0
1d_G_EVT	19	0	0	0	0
2d_G_N	19	0	0	0	0
2d_G_t	0	0	0	0	0
2d_G_EVT	19	0	0	0	0
5d_G_N	19	0	0	0	0
5d_G_t	0	0	0	0	0
5d_G_EVT	19	0	0	0	0
10d_G_N	19	0	0	0	0
10d_G_t	0	0	0	0	0
10d_G_EVT	19	0	0	0	0
20d_G_N	19	0	0	0	17
20d_G_t	19	0	1	0	17
20d_G_EVT	19	0	0	0	16
30d_G_N	19	19	19	19	19
30d_G_t	19	0	0	19	0
30d_G_EVT	19	19	19	19	19
60d_G_N	19	19	19	19	19
60d_G_t	19	0	0	7	0
60d_G_EVT	19	19	8	16	19

Interpretation of Tests

- Student distributed residual fail Tests 2, 3, and 5 at all frequencies.
- Aggregating models for high frequency data is worse than discarding the high frequency data.

And here are the main points again ...

1. We know how to calculate conditional kurtosis and variance of temporal aggregations of strong GARCH processes.
2. Conditional kurtosis of highly aggregated models goes to infinity - in some domain of GARCH parameters.
3. Without aggregation, high frequency models have higher likelihood than low frequency models.
4. For long term density forecasts do not aggregate high frequency models - rather discard h.f. data.