Realising the future: forecasting volatility and dependence through the credit crunch with HEAVY models and HF data

Neil Shephard

Oxford-Man Institute
and
Department of Economics
University of Oxford

Joint with
Nathaniel Frank
Economics and OMI, Oxford
and
Kevin Sheppard
Economics and OMI, Oxford

April, 2008

Overview

- Computing high frequency volatility & dependence measures through the credit crunch
- HEAVY models bring together the intellectual lessons from GARCH with the modern literature on high frequency data
- Thesis: harnessing high frequency measures improves
  1. robustness to structural breaks
  2. speed of adjustment
- Methodological contributions
  1. Define and analyse HEAVY models
  2. Look at speeds of adjustment

Empirical challenges

HF challenges:

- Overwhelming data streams
- Data cleaning
- Measures need to be robust to market microstructure effects
- Irregularly spaced data
- Non-synchronous data

Generic challenges:

- Differential opening/close times
- Structural breaks for forecasting models

Univariate analysis: measurement

- Rvol: Realised volatility (with subsampling)
- Kvol: Realised kernel (tick by tick when possible), HAC type estimator, somewhat robust to noise
- Econometric literature on Rvol
  1. Andersen, Bollerslev, Diebold, Labys (2001, JASA)
  2. Barndorff-Nielsen and Shephard (2002, JRSSB)
- Econometric literature on Kvol
  1. Barndorff-Nielsen, Hansen, Lunde and Shephard (2008, Econometrica) [alternatives include preaveraging and multiscale methods]

Univariate realised quantities

Rvol=$\sqrt{RV}$ where

$$RV = \sum_j x_j^2, \quad x_j = X_{\tau_j} - X_{\tau_{j-d}},$$

typically computed skipping d ticks or using 1 minute returns. Can subsample.

KV (Realised kernel variance) takes $d = 1$ (all the data) and

$$KV = \sum_{h=-H}^{H} k\left(\tfrac{h}{H+1}\right)\gamma_h, \quad \gamma_h = \sum_{j=|h|+1}^{n} x_j x'_{j-|h|}, \tag{1}$$

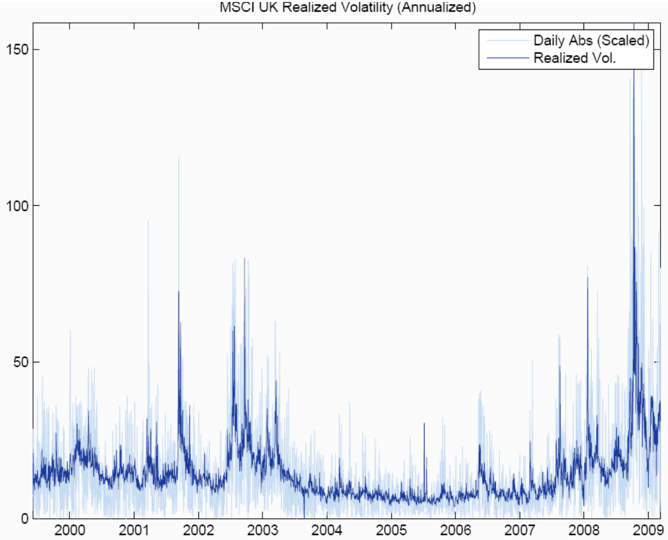
where $k(x)$ is the Parzen kernel function

$$k(x) = \begin{cases} 1 - 6x^2 + 6x^3 & 0 \leq x \leq 1/2 \\ 2(1-x)^3 & 1/2 \leq x \leq 1 \\ 0 & x > 1. \end{cases}$$

Report Kvol: $Kvol = \sqrt{KV}$. Uses refresh time or fixed interval (MSCI) like 1 minute.

Realised measure, RM: KV or RV.

Descriptive analysis



Here Bartlett kernel, 5 minute window, 30 second returns. Longer window for wide index, some slow components?

Univariate predictive models

Traditional GARCH models

$$\mathrm{Var}\left(r_t|\mathcal{F}_{t-1}^{LF}\right) = \sigma_t^2 = c + \alpha_G r_{t-1}^2 + \beta_G \sigma_{t-1}^2,$$

where $\mathcal{F}_{t-1}^{LF}$ is low frequency past data.
Focus on: "HEAVY models" (High frEquency bAsed VolatilitY models)

$$\begin{aligned}\mathrm{Var}\left(r_t|\mathcal{F}_{t-1}^{HF}\right) &= h_t = d + \alpha RM_{t-1} + \beta h_{t-1}, \quad \alpha \geq 0, \quad \beta \in [0,1)\,, \\ \mathrm{E}\left(RM_t|\mathcal{F}_{t-1}^{HF}\right) &= \mu_t = f + \alpha_R RM_{t-1} + \beta_R \mu_{t-1}, \quad \alpha_R + \beta_R \in [0,1)\,.\end{aligned}$$

where $\mathcal{F}_{t-1}^{HF}$ is high frequency data. Could include r_{t-1}^2 but typically tests out. Useful to focus directly on the above model.

- β momentum parameter, $\alpha_R + \beta_R$ persistence

Name these equations: HEAVY-r and HEAVY-RM.
Heavy models allow multistep ahead volatility forecasts — our focus.

MSCI Canada

of days: 2003, from 13 Feb 2001 to 27th Feb 2009

GARCH-X

$$h_t = \omega_X + \alpha_X RM_{t-1} + \gamma_X r_{t-1}^2 + \beta_X h_{t-1}$$

ω_X	α_X	γ_X	β_X	LL
2.934 (.002)	.444 (.000)	.002 (.838)	.778 (.000)	-10948.9

HEAVY-RM Tracking

$$\mu_t = \overline{RM}\,(1 - \alpha_R - \beta_R) + \alpha_R RM_{t-1} + \beta_R \mu_{t-1}$$

α_R	β_R	$\alpha_R + \beta_R$
.340 (.005)	.649 (.000)	.989

HEAVY-r

$$h_t = d + \alpha RM_{t-1} + \beta h_{t-1}$$

d	α	β	LL
2.949 (.001)	.448 (.000)	.778 (.000)	-10948.9

GARCH

$$h_t = c + \alpha_G r_{t-1}^2 + \beta_G h_{t-1}$$

c	α_G	β_G	$\alpha_G + \beta_G$	LL
1.264 (.004)	.074 (.000)	.914 (.000)	.988	-11028.2

- 6 out of 33 fitted models had a significant γ_J coefficients, all but one minor coefficients. 24 were exactly zero.
- a lot of the data we report starts in 1996.

Tracking or not

Enforce stationarity, use a moment estimator of the long run mean.

HEAVY-RM

$$\mu_t = f + \alpha_R RM_{t-1} + \beta_R \mu_{t-1}$$

f	α_R	β_R	$\alpha_R + \beta_R$
.722 (.005)	.342 (.000)	.649 (.000)	.991

HEAVY-RM Tracking

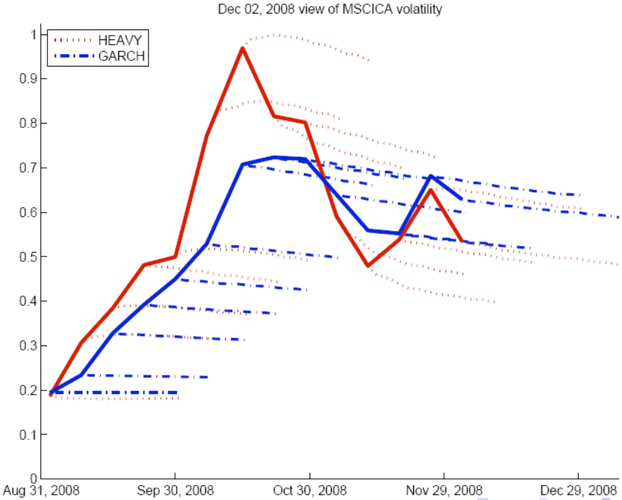
$$\mu_t = \overline{RM}\,(1 - \alpha_R - \beta_R) + \alpha_R RM_{t-1} + \beta_R \mu$$

α_R	β_R	$\alpha_R + \beta_R$
.340 (.005)	.649 (.000)	.989

Does not make a lot of difference to the dynamics in the cross-section.
Slightly less persistence.
Helpful in terms of long-run forecasting.
Experience: if one mean tracks the RM then the HEAVY-r is roughly variance tracked. Can be exactly enforced easily of course.

A flavour of things to come

- Sequential out of sample 66 day volatility forecasts.
- Parameter estimation window of 3 years, Sept 1 2005 - Sept 1 2008.
- Go through the Lehmann's bankrupcy

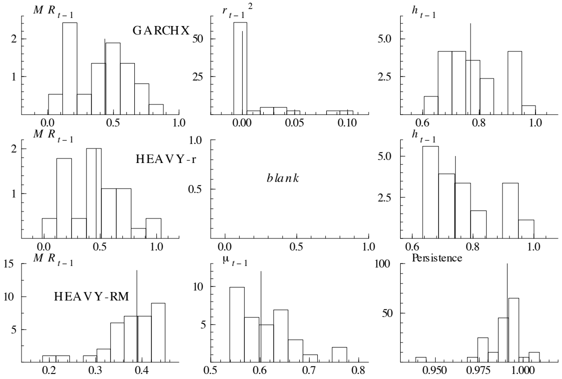


Taking these models to the data to 27th Feb 2009

	α_X	γ_X	β_X	α	β	α_R	β_R
DJI	0.479	0.000	0.698	0.479	0.698	0.412	0.566
CAC 40	0.480	0.000	0.691	0.480	0.691	0.372	0.617
FTSE 100	0.563	0.000	0.665	0.563	0.665	0.422	0.573
DAX	0.439	0.000	0.675	0.439	0.675	0.448	0.544
Nasdaq 100	0.498	0.025	0.725	0.607	0.696	0.432	0.558
MSCI Mexico	0.574	0.024	0.743	0.653	0.738	0.335	0.652
MSCI World	0.275	0.000	0.822	0.275	0.822	0.310	0.679
Nikkei 250	0.452	0.006	0.777	0.481	0.769	0.341	0.645
Russell 2000	0.206	0.104	0.820	0.925	0.671	0.381	0.628
S&P 500	0.362	0.000	0.773	0.362	0.773	0.417	0.566
$/E	*0.050*	*0.000*	*0.949*	*0.050*	*0.949*	*0.239*	*0.752*
$/£	*0.166*	*0.000*	*0.820*	*0.166*	*0.820*	*0.280*	*0.698*
$/Y	*0.155*	*0.000*	*0.806*	*0.155*	*0.806*	*0.383*	*0.561*

- Typical likelihood improvement of going from traditional GARCH to HEAVY-r is around 200. Currencies about 30, less good models.
- Russell 2000, small caps. By far the worst case in our cross-section.

Cross-section of 33 series, data up to 27th Feb 2009



Black line is the medium, so is the Engle pooled cross-sectional estimator (MacGyver). Studied in univariate context by Cavit Pavel (2009). In the γ_J case it is exactly zero.

Default, non-stationary parameters

A Riskmetrics type default would be: $d = f = 0$, $\alpha_R = 0.4$, $\beta_R = 0.6$ and $\beta = 0.7$.
We would take

$$\alpha = (1 - \beta) \frac{\mathrm{E}(r_t^2)}{\mathrm{E}(RM)}.$$

replacing ratios by empirical versions. Use a different default for currencies.

Multistep ahead forecasts

Write $v_{t+s|t-1} = \left(\mathrm{Var}\left(r_{t+s}|\mathcal{F}^{HF}_{t-1}\right), \mathrm{E}\left(RM_{t+s}|\mathcal{F}^{HF}_{t-1}\right)\right)'$ for $s \geq 0$, then

$$
\begin{aligned}
v_t &= a + \begin{pmatrix} \alpha & 0 \\ 0 & \alpha_R \end{pmatrix} RM_{t-1} + \begin{pmatrix} \beta & 0 \\ 0 & \beta_R \end{pmatrix} v_{t-1}, \quad a = \begin{pmatrix} d \\ f \end{pmatrix}, \\
&= \begin{pmatrix} d \\ f \end{pmatrix} + \begin{pmatrix} \alpha & 0 \\ 0 & \alpha_R \end{pmatrix} (RM_{t-1} - \mu_{t-1}) + B v_{t-1}.
\end{aligned}
$$

implying

$$
v_{t+s|t-1} = (I + B + \ldots + B^s) a + B^{s+1} v_{t-1}, \quad B = \begin{pmatrix} \beta & \alpha \\ 0 & \alpha_R + \beta_R \end{pmatrix}.
$$

Write $\vartheta = (\alpha_R + \beta_R)$. It has two roots β and $\alpha_R + \beta_R$ (e.g. Golub and van Loan (1989, p. 333)). Further

$$
B^J = \begin{pmatrix} \beta^J & \alpha\left(\vartheta^{J-1} + \vartheta^{J-2}\beta + \ldots + \beta^{J-1}\right) \\ 0 & \vartheta^J \end{pmatrix}.
$$

Impulse responses

$$v_{t+s|t-1} = (I + B + ... + B^s)a + B^{s+1}v_{t-1}, \quad B = \begin{pmatrix} \beta & \alpha \\ 0 & \alpha_R + \beta_R \end{pmatrix}.$$

Of course the impulse response of a shock to RM_{t-1} is linear

$$\frac{\partial \mathrm{Var}\left(r_{t+s}|\mathcal{F}_{t-1}^{HF}\right)}{\partial RM_{t-1}} = \alpha\left(\vartheta^{s-1} + \vartheta^{s-2}\beta + ... + \beta^{s-1}\right), \quad \vartheta = \alpha_R + \beta_R.$$

Notice there are 3 dynamic parameters, not 2 as with GARCH models. Corresponding result for GARCH is

$$\frac{\partial \mathrm{Var}\left(r_{t+s}|\mathcal{F}_{t-1}^{LF}\right)}{\partial y_{t-1}^2} = \alpha_G \vartheta_G^{s-1}, \quad \vartheta_G = \alpha_G + \beta_G,$$

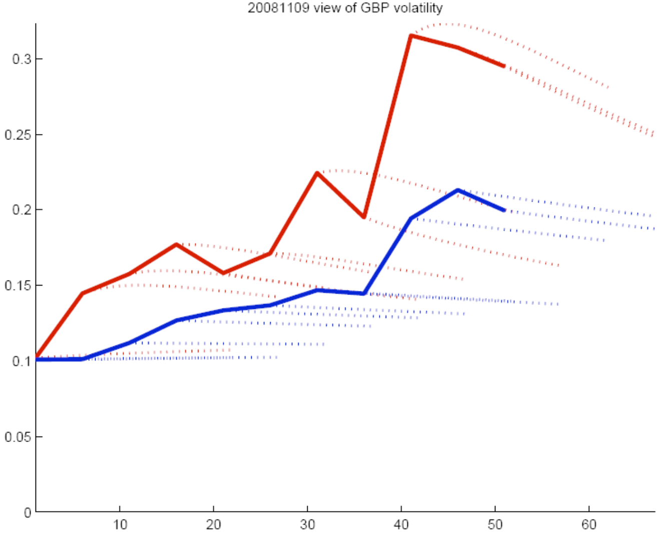
hence is an exponential damp down.

Out of sample multistep forecasts

- Sequential out of sample 66 day volatility forecasts.
- Parameter estimation window of 3 years, Sept 1 2005 - Sept 1 2008.
- Go through the Lehmann's bankrupcy

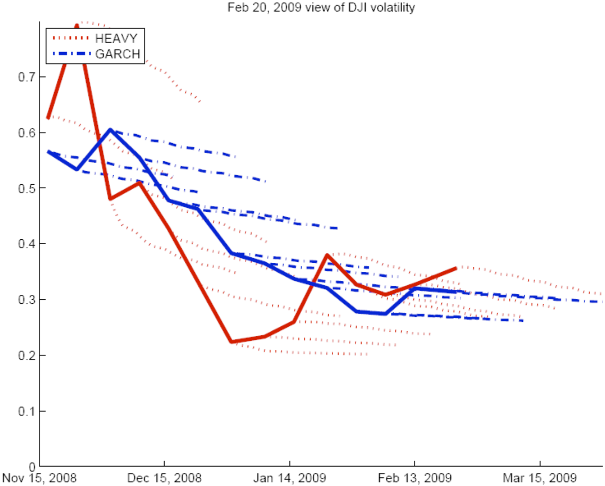
Sterling-dollar

Red: HEAVY model, blue: GARCH



Notice the HEAVY model predicts momentum and then mean reversion.

DJI's strongly falling vol



Shows the momentum effect of the HEAVY model.

Sequential 20 day cor forecasts through September 2008

scalar BEKK, scalar HEAVY, side by side

Conclusion

- HEAVY models bring together the intellectual lessons from GARCH with the modern literature on high frequency data
- Rely on a library of realised measures, once computed (non-trivial!) HEAVY models are easy to work with
- We have compared HEAVY models to GARCH models
- Econometrically HEAVY models outperform in sample across all assets we have looked at so long as the realised measure is not daft
- They give better out of sample performance
- The difference is very large during the Lehmann's month
- HEAVY model display momentum and mean reversion in theory and practice.
- Realised kernels gave slightly more reliable analysis than subsampled RV (HEAVY-r's likelihood is typically similar, a handful had 30 point improvement).