Lecture 3: Realized GARCH Models

Making Use of Realized Measures

Peter Reinhard Hansen

University of North Carolina

Big Data in Aarhus, September 2016

Outline

- GARCH in an hour or less.
- Realized GARCH Model
- Multivariate Realized GARCH

Heteroskedasticity

- You may know about "controlling" for Heteroskedasticity
- In regression model, $Y_t = X_t'\beta + \varepsilon_t$, the least squares estimator

$$\hat{\beta} - \beta = \left[\sum_{t=1}^{n} X_t X_t'\right]^{-1} \sum_{t=1}^{n} X_t \varepsilon_t,$$

- LLN

$$\frac{1}{n}\sum_{t=1}^{n} X_t X_t' \xrightarrow{p} \mathbb{E} X_t X_t',$$

- and CLT

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n} X_t \varepsilon_t \xrightarrow{d} N(0, [\mathbb{E}(X_t X_t' \varepsilon_t^2)]).$$

Heteroskedasticity

- By Slutsky

$$\sqrt{n}(\hat{\beta}-\beta)=\left[\frac{1}{n}\sum_{t=1}^{n}X_tX_t'\right]^{-1}\frac{1}{\sqrt{n}}\sum_{t=1}^{n}X_t\varepsilon_t \xrightarrow{d} N(0,\Sigma_\beta),$$

- with

$$\Sigma_\beta=\left[\mathbb{E}X_tX_t'\right]^{-1}\left[\mathbb{E}(X_tX_t'\varepsilon_t^2)\right]\left[\mathbb{E}X_tX_t'\right]^{-1}.$$

- Classical assumption (Homoskedasticity)

$$\mathbb{E}(X_tX_t'\varepsilon_t^2)=\mathbb{E}(X_tX_t')\mathbb{E}(\varepsilon_t^2),$$

- so that

$$\Sigma_\beta=\sigma_\varepsilon^2\left[\mathbb{E}X_tX_t'\right]^{-1}.$$

Heteroskedasticity

- Computing standard errors based on

$$\hat{\Sigma}_\beta = \left(\frac{1}{n}\sum X_t X_t'\right)^{-1} \frac{1}{n}\sum X_t X_t' \hat{\varepsilon}_t^2 \left(\frac{1}{n}\sum X_t X_t'\right)^{-1},$$

- Is accounting for the possibility that

$$\mathbb{E}(X_t X_t' \varepsilon_t^2) \neq \mathbb{E}(X_t X_t')\mathbb{E}(\varepsilon_t^2),$$

- or equivalently, that

$$\mathrm{var}(\varepsilon_t | X_t) = \mathbb{E}(\varepsilon_t^2 | X_t),$$

might depend on X_t.

Conditional Expectation

- A conditional mean
$$\mathbb{E}(Y|X),$$
implies
$$\mathbb{E}(Y|X) = g(X),$$
for some function g.
- So we could attempt to model the heteroskedasticy
$$\mathbb{E}(\varepsilon_t^2|X_t) = g(X_t),$$
with a suitable specification for g (instead of "accounting" for it).
- ARCH/GARCH models seek to model this variation.

Time Series

- Let $\{X_t\}$ be a time series, and let $\mathcal{F}_t$ be a filtration, to which X_t is adapted, i.e.$X_t \in \mathcal{F}_t$, X_t is $\mathcal{F}_t$-measurable.
- Time series models typically specifies distributional features of

$$X_t | \mathcal{F}_{t-1}.$$

- E.g. $AR(p)$

$$\mathbb{E}(X_t | \mathcal{F}_{t-1}) = \mu + \varphi_1 X_{t-1} + \cdots + \varphi_p X_{t-p}$$

- ARMA

$$\mathbb{E}(X_t | \mathcal{F}_{t-1}) = \mu + \varphi_1 X_{t-1} + \cdots + \varphi_p X_{t-p} - \theta_1 \varepsilon_{t-1} - \cdots - \theta_q \varepsilon_{t-q},$$

with $\mathcal{F}_t = \sigma(X_t, X_{t-1}, \ldots, \varepsilon_t, \varepsilon_{t-1}, \ldots)$.

GARCH Modeling

- Three components

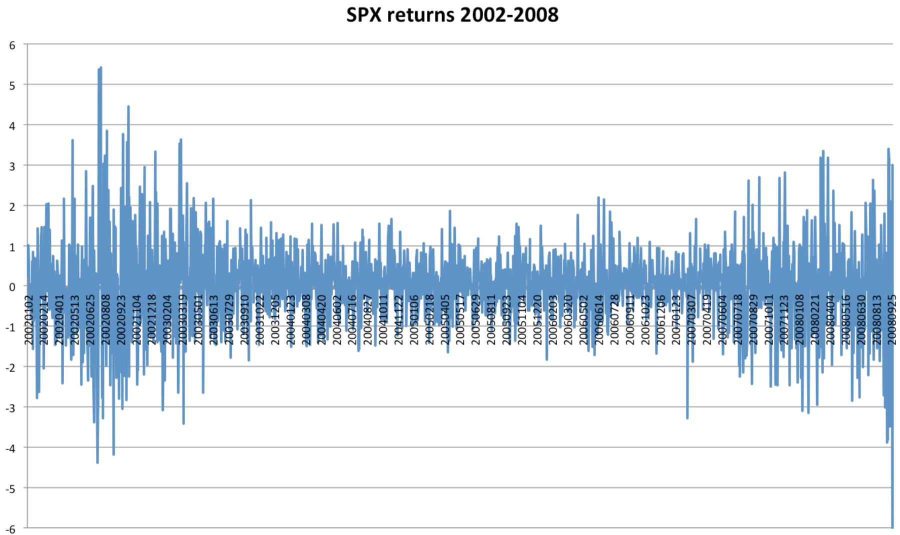
$$\begin{aligned} \mu_t &= \mathbb{E}(r_t|\mathcal{F}_{t-1}), \\ h_t &= \mathrm{var}(r_t|\mathcal{F}_{t-1}) \\ z_t = \frac{r_t - \mu_t}{\sqrt{h_t}} &\sim F, \end{aligned}$$

  for some distribution.

- Key is h_t. Modeled with a GARCH equation

$$h_t = g(\mathcal{F}_{t-1})$$

- Often $\mu_t = \mu$ (constant) and $z_t \sim iidN(0,1)$.


SPX returns 2002-2008
6
5
4
3
2
1
0
-1
-2
-3
-4
-5
-6
20020102
20020214
20020401
20020513
20020625
20020808
20020923
20021104
20021218
20030204
20030319
20030501
20030613
20030729
20030910
20031022
20031205
20040123
20040308
20040420
20040602
20040716
20040827
20041011
20041122
20050106
20050218
20050405
20050517
20050629
20050811
20050923
20051104
20051220
20060203
20060320
20060502
20060614
20060728
20060911
20061023
20061206
20070123
20070307
20070419
20070604
20070718
20070829
20071011
20071123
20080108
20080221
20080404
20080516
20080630
20080813
20080925

Engle's Autoregressive Conditional Heteroskedasticity

- If
$$\mathbb{E}(\varepsilon_t^2|r_{t-1}),$$
depends on r_{t-1}... let try to build a model.
- Conditional variance of ε_t:
$$h_t \equiv \mathbb{E}(\varepsilon_t^2|r_{t-1}),$$
- ARCH(1) Model
$$h_t = \omega + \alpha r_{t-1}^2.$$
- First applied to macroeconomic time series (UK inflation).

Generalized ARCH

- Bollerslev (1986). GARCH(1,1)

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

- GARCH(p,q)

$$h_t = \omega + \alpha_1 r_{t-1}^2 + \cdots + \alpha_q r_{t-q}^2 + \beta_1 h_{t-1} + \cdots + \beta_p h_{t-p}.$$

- How about estimation?

Time-Series MLE

- Likelihood for time series models.
- Key objective is the (log-) density for the observed data, i.e. $f_\theta(r_1, \ldots, r_T)$.
- Decompose into conditional densities:

$$f_\theta(r_1, \ldots, r_T) = \frac{f_\theta(r_1, \ldots, r_T)}{f_\theta(r_1, \ldots, r_{T-1})} \frac{f_\theta(r_1, \ldots, r_{T-1})}{f_\theta(r_1, \ldots, r_{T-2})} \cdots f_\theta(r_1),$$

- $$\frac{f_\theta(r_1, \ldots, r_t)}{f_\theta(r_1, \ldots, r_{t-1})} = f_\theta(r_t | r_1, \ldots, r_{t-1}) \quad = f_\theta(r_t | \mathcal{F}_{t-1}).$$

Towards Estimating GARCH

- Return equation

$$r_t = \mu + \sqrt{h_t} z_t,$$

where

$$z_t \sim \text{iid} N(0, 1).$$

- GARCH Equation

$$\begin{aligned} h_t &= \omega + \alpha r_{t-1}^2 + \beta h_{t-1} \qquad \text{or} \\ h_t &= \omega + \alpha (r_{t-1} - \mu)^2 + \beta h_{t-1} \end{aligned}$$

- Conditional distribution of returns given past, $r_t | \mathcal{F}_{t-1}$, is $N(\mu, h_t)$.

Maximum Likelihood Estimation of GARCH

- Conditional density is (recall $r_t = \mu + \sqrt{h_t} z_t$)

$$f(r_t|\mathcal{F}_{t-1}) = \frac{1}{\sqrt{2\pi h_t}} \exp\left\{-\frac{1}{2}\frac{(r_t-\mu)^2}{h_t}\right\} = \frac{1}{\sqrt{2\pi h_t}} \exp\left\{-\frac{1}{2}z_t^2\right\}.$$

- Likelihood function (density for all data)

$$L(\mu,\omega,\alpha,\beta) = \prod_{t=1}^{T} \frac{1}{\sqrt{2\pi h_t}} \exp\left\{-\frac{1}{2}\frac{(r_t-\mu)^2}{h_t}\right\}$$

- Where $h_t = \omega + \alpha r_{t-1}^2 + \beta h_{t-1}$.

Log-Likelihood for Simple GARCH

- So the log-likelihood

$$
\begin{aligned}
\log f_\theta(r_1,\ldots,r_T) &= \sum_{t=1}^{T} \log f_\theta(r_t|\mathcal{F}_{t-1}) \\
&\propto -\tfrac{1}{2}\sum_{t=1}^{T}(\log h_t + \frac{(r_t-\mu)^2}{h_t}).
\end{aligned}
$$

- where $h_t = h_t(\theta)$ depends on the unknown parameters $(\omega,\alpha,\beta)'$, through the GARCH equation

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

Starting Values

- GARCH Equation

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

- Need a starting value for h_t.
- In practice it is convenient to treat h_1 as an unknown parameter,
- i.e. maximize log-likelihood wrt.

$$\theta = (h_1, \mu, \omega, \alpha, \beta)'.$$

Likelihood Analysis of GARCH

- Same basic principle... but with new twists.
- For simplicity, take h_1 to be fixed and known and $\mu = 0$.
- Log-likelihood is

$$\ell(\theta) = -\tfrac{1}{2}\sum_{t=1}^{T}\left(\log h_t(\theta) + \frac{r_t^2}{h_t(\theta)}\right) \qquad \theta = (\omega, \alpha, \beta)'.$$

- Score

$$\frac{\partial \ell}{\partial \theta} = \sum_{t=1}^{T}\frac{\partial \ell_t}{\partial h_t}\frac{\partial h_t}{\partial \theta} = -\tfrac{1}{2}\sum_{t=1}^{T}\left(\frac{1}{h_t(\theta)} - \frac{r_t^2}{h_t^2(\theta)}\right)\frac{\partial h_t(\theta)}{\partial \theta}$$

Score of GARCH(1,1)

- $$\frac{\partial \ell}{\partial \theta} = -\tfrac{1}{2}\sum_{t=1}^{T}\left(\frac{h_t(\theta) - r_t^2}{h_t^2(\theta)}\right)\frac{\partial h_t(\theta)}{\partial \theta}$$

- Let's derive
$$\frac{\partial h_t(\theta)}{\partial \theta},$$
where $\theta = (\omega, \alpha, \beta)'$.

- What is
$$\frac{\partial h_t}{\partial \omega} = \frac{\partial(\omega + \alpha r_{t-1}^2 + \beta h_{t-1})}{\partial \omega} = ?$$

Score of GARCH(1,1)

- $$\frac{\partial h_t}{\partial \omega} = \frac{\partial(\omega + \alpha r_{t-1}^2 + \beta h_{t-1})}{\partial \omega} = 1 + \beta \frac{\partial h_{t-1}}{\partial \omega}.$$

- Similarly
$$\frac{\partial h_t}{\partial \alpha} = r_{t-1}^2 + \beta \frac{\partial h_{t-1}}{\partial \alpha}.$$

Score of GARCH(1,1)

- $$\frac{\partial h_t}{\partial \beta} = \frac{\partial(\omega + \alpha r_{t-1}^2 + \beta h_{t-1})}{\partial \beta} = h_{t-1} + \beta \frac{\partial h_{t-1}}{\partial \beta}.$$
- So that
$$\frac{\partial h_t}{\partial \theta} = \begin{pmatrix} 1 \\ r_{t-1}^2 \\ h_{t-1} \end{pmatrix} + \beta \frac{\partial h_{t-1}}{\partial \theta}.$$
- With
$$\frac{\partial h_1}{\partial \theta} = ??$$

Score of GARCH(1,1)

- Recursion. So

$$\frac{\partial h_t}{\partial \theta} = \sum_{j=1}^{t-1} \beta^{j-1} \begin{pmatrix} 1 \\ r_{t-j}^2 \\ h_{t-j} \end{pmatrix}.$$

- Scores are

$$\begin{pmatrix} s_{\omega t} \\ s_{\alpha t} \\ s_{\beta t} \end{pmatrix} = -\frac{1}{2}\left(\frac{h_t(\theta) - r_t^2}{h_t^2(\theta)}\right) \begin{pmatrix} \frac{1-\beta^{t-1}}{1-\beta} \\ \sum_{j=1}^{t-1} \beta^{j-1} r_{t-j}^2 \\ \sum_{j=1}^{t-1} \beta^{j-1} h_{t-j} \end{pmatrix}$$

Score MGD?

- $$s_t(\theta) = \begin{pmatrix} s_{\omega t} \\ s_{\alpha t} \\ s_{\beta t} \end{pmatrix} = -\frac{1}{2}\left(\frac{h_t(\theta) - r_t^2}{h_t^2(\theta)}\right)\frac{\partial h_t}{\partial \theta}.$$

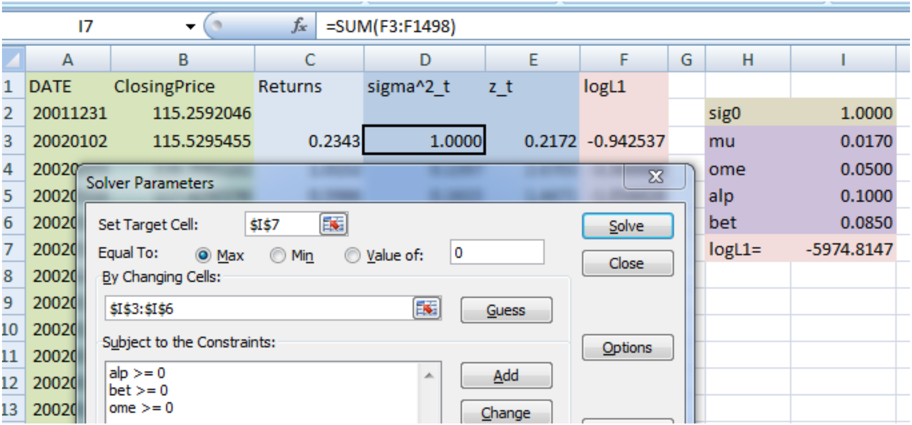
- $$\mathbb{E}[s_t|\mathcal{F}_{t-1}] = -\frac{1}{2}\mathbb{E}\left[\left(\frac{h_t(\theta) - r_t^2}{h_t^2(\theta)}\right)\frac{\partial h_t}{\partial \theta}\right] =?$$

MLE of GARCH(1,1) with Excel

- Maximize
$$-\frac{1}{2}\sum_{t=1}^{T}\left[\log(2\pi)+\log h_t+\frac{(r_t-\mu)^2}{h_t}\right],$$
with respect to θ, subject to
- GARCH equation
$$h_t=\omega+\alpha r_{t-1}^2+\beta h_{t-1}.$$
- Excel Solver can do it

F3 =-1/2*(LN(2*PI())+LN(D3)+E3^2)

	A	B	C	D	E	F	G	H	I
1	DATE	ClosingPrice	Returns	sigma^2_t	z_t	logL1			
2	20011231	115.2592046						sig0	1.0000
3	20020102	115.5295455	0.2343	1.0000	0.2172	-0.942537		mu	0.0170
4	20020103	116.7083182	1.0152	0.1397	2.6703	-3.500066		ome	0.0500
5	20020104	117.4090196	0.5986	0.1615	1.4471	-1.054424		alp	0.1000
6	20020107	116.7663866	-0.5488	0.0975	-1.8118	-1.396535		bet	0.0850
7	20020108	116.1734286	-0.5091	0.0903	-1.7507	-1.249252		logL1=	-5974.8147
8	20020109	115.766295	-0.3511	0.0854	-1.2599	-0.482167			
9	20020110	115.8233766	0.0493	0.0708	0.1213	0.397622			
10	20020111	114.9055372	-0.7956	0.0561	-3.4302	-5.362098			
11	20020114	114.2135714	-0.6040	0.1208	-1.7868	-1.458521			


I7
=SUM(F3:F1498)
DATE
ClosingPrice
Returns
sigma^2_t
z_t
logL1
20011231
115.2592046
20020102
115.5295455
0.2343
1.0000
0.2172
-0.942537
sig0 1.0000
mu 0.0170
ome 0.0500
alp 0.1000
bet 0.0850
logL1= -5974.8147
Solver Parameters
Set Target Cell: $I$7
Equal To: Max Min Value of: 0
By Changing Cells:
$I$3:$I$6
Subject to the Constraints:
alp >= 0
bet >= 0
ome >= 0
Solve
Close
Guess
Options
Add
Change

JOURNAL OF APPLIED ECONOMETRICS
J. Appl. Econ. **20**: 873–889 (2005)
Published online 30 March 2005 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/jae.800

A FORECAST COMPARISON OF VOLATILITY MODELS: DOES ANYTHING BEAT A GARCH(1,1)?

PETER R. HANSEN[a]* AND ASGER LUNDE[b]

[a] *Department of Economics, Brown University, Providence, USA*
[b] *Department of Information Science, Aarhus School of Business, Denmark*

SUMMARY

We compare 330 ARCH-type models in terms of their ability to describe the conditional variance. The models are compared out-of-sample using DM–$ exchange rate data and IBM return data, where the latter is based on a new data set of realized variance. We find no evidence that a GARCH(1,1) is outperformed by more sophisticated models in our analysis of exchange rates, whereas the GARCH(1,1) is clearly inferior to models that can accommodate a leverage effect in our analysis of IBM returns. The models are compared with the test for superior predictive ability (SPA) and the reality check for data snooping (RC). Our empirical results show that the RC lacks power to an extent that makes it unable to distinguish 'good' and 'bad' models in our analysis.

GARCH ZOO I

ARCH:	$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i u_{t-i}^2$
GARCH:	$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i u_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2$
IGARCH	$\sigma_t^2 = \omega + u_{t-1}^2 + \sum\limits_{i=2}^{p} \alpha_i (u_{t-i}^2 - u_{t-1}^2) + \sum\limits_{j=1}^{q} \beta_j (\sigma_{t-j}^2 - u_{t-1}^2)$
Taylor/Schwert:	$\sigma_t = \omega + \sum_{i=1}^{p} \alpha_i \lvert u_{t-i} \rvert + \sum_{j=1}^{q} \beta_j \sigma_{t-j}$
A-GARCH:	$\sigma_t^2 = \omega + \sum_{i=1}^{p} \left[\alpha_i u_{t-i}^2 + \gamma_i u_{t-i}\right] + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2$
NA-GARCH:	$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \left(u_{t-i} + \gamma_i \sigma_{t-i}\right)^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2$
V-GARCH:	$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \left(e_{t-i} + \gamma_i\right)^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2$
Thr.-GARCH:	$\sigma_t = \omega + \sum\limits_{i=1}^{p} \alpha_i \left[(1-\gamma_i) u_{t-i}^{+} - (1+\gamma_i) u_{t-i}^{-}\right] + \sum\limits_{j=1}^{q} \beta_j \sigma_{t-j}$
GJR model:	$\sigma_t^2 = \omega + \sum_{i=1}^{p_1} \left[\alpha_i + \gamma_i I_{\{u_{t-i}^2 > 0\}}\right] u_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2$
NGARCH:	$\sigma_t^\delta = \omega + \sum_{i=1}^{p} \alpha_i \lvert u_{t-i} \rvert^\delta + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^\delta$
log-GARCH:	$\log(\sigma_t) = \omega + \sum_{i=1}^{p} \alpha_i \lvert e_{t-i} \rvert + \sum_{j=1}^{q} \beta_j \log(\sigma_{t-j})$

GARCH ZOO II

EGARCH: $\log(\sigma_t^2) = \omega + \sum_{i=1}^{p} \left[\alpha_i e_{t-i} + \gamma_i \left(|e_{t-i}| - E|e_{t-i}|\right)\right] + \sum_{j=1}^{q} \beta_j \log(\sigma_{t-j}^2),$

AP-GARCH: $\sigma^\delta = \omega + \sum_{i=1}^{p} \alpha_i \left[|u_{t-i}| - \gamma_i u_{t-i}\right]^\delta + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^\delta$

H-GARCH: $\sigma_t^\delta = \omega + \sum_{i=1}^{p} \alpha_i \delta \sigma_{t-i}^\delta \left[|e_t - \kappa| - \tau\left(e_t - \kappa\right)\right]^\nu + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^\delta$

GQ-ARCH: $\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i u_{t-i} + \sum_{i=1}^{p} \alpha_{ii} u_{t-i}^2 + \sum_{i<j}^{p} \alpha_{ij} u_{t-i} u_{t-j} + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2$

Aug -GARCH:

$$\sigma_t^2 = \begin{cases} |\delta\phi_t - \delta + 1|^{1/\delta} & \text{if } \delta \neq 0 \\ \exp\left(\phi_t - 1\right) & \text{if } \delta = 0 \end{cases}$$

$$\begin{aligned} \phi_t = {} & \omega + \sum_{i=1}^{p} \left[\alpha_{1i} \left|u_{t-i} - \kappa\right|^\nu + \alpha_{2i} \max(0, \kappa - u_{t-i})^\nu\right] \phi_{t-j} \\ & + \sum_{i=1}^{p} \left[\alpha_{3i} f\left(\left|u_{t-i} - \kappa\right|, \nu\right) + \alpha_{4i} \frac{\max(0, \kappa - u_{t-i})^\nu - 1}{\nu}\right] \phi_{t-j} \\ & + \sum_{j=1}^{q} \beta_j \phi_{t-j}^2 \end{aligned}$$

Modern Volatility Models use Realized Measures

- Conventional GARCH models take information from returns to adjust the value of the conditional variance σ_t^2.
- High-Frequency data. Today we have high-frequency data (thousands of observation per day per stock)
  - HF data yields far more accurate measures of volatility than, say, squared return.

GARCH Is Updating Slowly: A Thought Experiment

- Suppose a sudden change in the true latent volatility

$$\sqrt{\mathrm{var}(r_t|\mathcal{F}_{t-1})} = \begin{cases} 20\% & t \leq T \\ 40\% & t > T. \end{cases}$$

- GARCH(1,1)

$$\begin{aligned} h_t &= \omega + \beta h_{t-1} + \alpha r_{t-1}^2 \\ &= \omega + \beta(\omega + \beta h_{t-2} + \alpha r_{t-2}^2) + \alpha r_{t-1}^2 \\ &\vdots \\ &= \tfrac{1-\beta^t}{1-\beta}\omega + \beta^t h_0 + \alpha \sum_{j=0}^{t-1} \beta^j r_{t-j-1}^2. \end{aligned}$$

GARCH Is Updating Slowly: A Thought Experiment

- Typical GARCH estimates $\alpha + \beta = 1$ (e.g. $\alpha = 0.05$ and $\beta = 0.95$) and $\omega = 0$. Implies

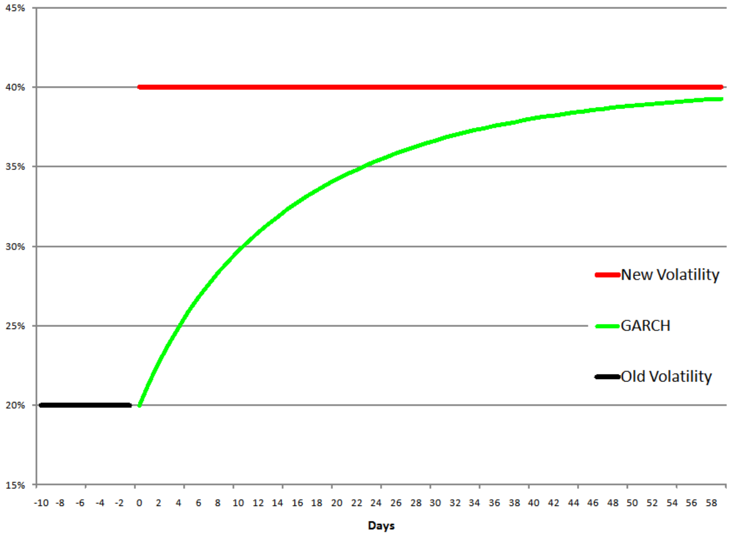
$$h_t = \alpha \sum_{j=0}^{\infty} \beta^j r_{t-j-1}^2.$$

- So in our thought experiment

$$\mathbb{E}(h_{T+h}) = \alpha \sum_{j=0}^{\infty} \beta^j \mathbb{E}(r_{T+h-j}^2) = \alpha \sum_{j=0}^{h-1} \beta^j (0.4)^2 + \alpha \sum_{j=h}^{\infty} \beta^j (0.2)^2.$$

- How long does it take for h_t to reach the true variance?

GARCH is Slow



GARCH Model with Realized Measure

- $$h_t = \omega + \alpha r_{t-1}^2 + \beta h_{t-1} + \gamma x_{t-1},$$

- where x_t is a realized measure of volatility.
  - Realized Variance
  - Realized Kernel
  - MC estimator
  - .... etc.

Realized GARCH: A Complete Model of Returns and Realized Measures of Volatility*

Peter Reinhard Hansen† Zhuo (Albert) Huang Howard Howan Shek

March 11, 2010

Abstract

GARCH models have been successful in modeling financial returns. Still, much is to be gained by incorporating a realized measure of volatility in these models. In this paper we introduce a new framework for the joint modeling of returns and realized measures of volatility. The Realized GARCH framework nests most GARCH models as special cases and is, in many ways, a natural extension of standard GARCH models. We pay special attention to linear and log-linear Realized GARCH specifications. This class of models has several attractive features. It retains the simplicity and tractability of the classical GARCH framework; it implies an ARMA structure for the conditional variance and realized measures of volatility; and models in this class are parsimonious and simple to estimate. A key feature of the Realized GARCH framework is a measurement equation that relates the observed realized measure to latent volatility. This equation facilitates a simple modeling of the dependence between returns and future volatility that is commonly referred to as the leverage effect. An empirical application with DJIA stocks and an exchange traded index fund shows that a simple Realized GARCH structure leads to substantial improvements in

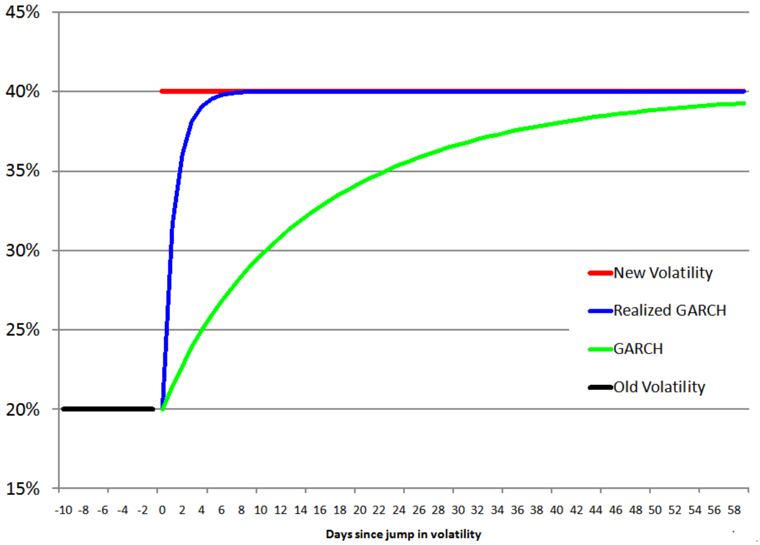
GARCH-X Model

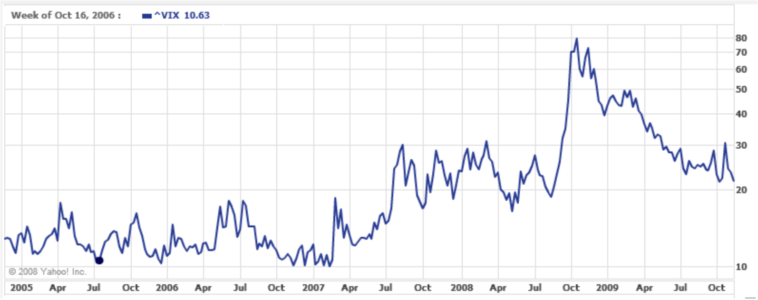
- Engle (2002), and many others

$$h_t = \omega + \alpha r_{t-1}^2 + \beta h_{t-1} + \gamma x_{t-1}.$$

  x_t is a realized measure of volatility (e.g. RV)
- Huge improvement in the empirical fit.
- Typically
  - $\hat{\gamma} \simeq 0.5.$
  - $\hat{\alpha} \simeq 0.$ (ARCH parameter becomes insignificant)

GARCH with Realized Measure is Fast





- Volatility prediction implied by options prices.
- Lots of **time variation**, sometimes **rapid changes**.
- Figure displays the implied volatility-prediction for the next month.
- We should expect even greater variation in daily volatility.

Completing the GARCH-X

GARCH-X is Incomplete

- Data (r_t, x_t), but model only specifies $r_t | r_{t-1,} x_{t-1}, \ldots$
- Simple case

$$\begin{aligned} r_t &= \sqrt{h_t} z_t. \\ h_t &= \omega + \alpha r_{t-1}^2 + \beta h_{t-1} + \gamma x_{t-1}. \end{aligned}$$

- Need a Model for x_t.

Realized GARCH

Realized GARCH: Simple Case

- GARCH-X structure

$$r_t = \sqrt{h_t} z_t$$
$$h_t = \omega + \alpha r_{t-1}^2 + \beta h_{t-1} + \gamma x_{t-1}$$

- **Measurement Equation** completes the model

$$x_t = \xi + \varphi h_t + \mathrm{Error}_t.$$

  - x_t is noisy measurement of QV_t
  - QV_t is h_t + volatility shock.

Measurement Equation

- Measurement Equation

$$x_t = \xi + \varphi h_t + \tau(z_t) + u_t,$$

where the leverage function is

$$\tau(z) = \tau_1 z + \tau_2(z^2 - 1)$$

- Captures the joint dependence between
  - return shocks, z_t
  - volatility shocks, $\tau(z_t) + u_t$.
- $\tau(z)$ is called a **Leverage Function**

Logarithmic Specification

- Logarithmic specification is preferred

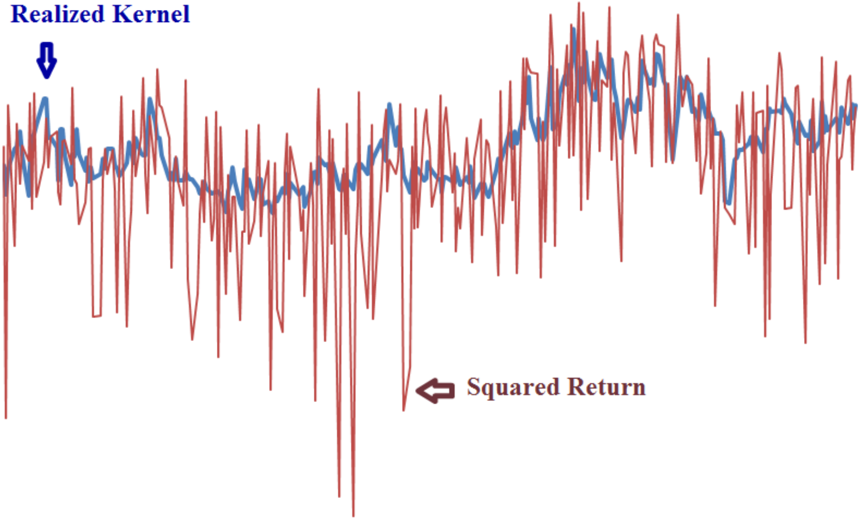
$$\begin{aligned}\log h_t &= \omega + \beta \log h_{t-1} + \gamma \log x_{t-1}.\\ \log x_t &= \xi + \varphi \log h_t + \tau(z_t) + u_t.\end{aligned}$$

- Drawback of conventional LogGARCH (zero returns) is not relevant here, because x_t is positive.
- Can be rewritten in a form that is similar to EGARCH

Key Features of Realized GARCH

- Empirical Features
  - **Easy** to estimate.
  - Captures return-volatility dependence (**leverage effect**).
  - Properties of multiperiod returns (**skewness and kurtosis**)
  - **Outperforms** conventional GARCH
- Theoretical Features (elegant mathematical structure)
  - **Parsimonious**
  - **Tractable** analysis (quasi maximum likelihood).
  - Induced **simple ARMA structure** for both x and h
- Natural extension of conventional GARCH

Squared Return is a Noisy Signal of Volatility



Estimation and Inference

QMLE

- Gaussian specification adopted:

$$z_t \sim \mathrm{iid} N(0,1) \qquad u_t \sim \mathrm{iid} N(0,\sigma_u^2).$$

- The quasi-log likelihood function is

$$\ell(r,x;\theta) = -\frac{1}{2}\sum_{t=1}^{n}[\log(h_t) + r_t^2/h_t + \log(\sigma_u^2) + u_t^2/\sigma_u^2].$$

- Partial log-likelihood

$$\ell(r;\theta) = -\frac{1}{2}\sum_{t=1}^{n}[\log(h_t) + r_t^2/h_t].$$

- Can be compared to Conventional GARCH.

Empirical Analysis

Data

- Dow Jones Industrial Average stocks and SPY (ETF).
  - 2002-01-01 to 2007-12-31 as in-sample data and
  - 2008-01-01 to 2008-08-31 as out-of-sample.
- For x, we use the realized kernel (RK) by BHLS (2008)
  - $x_t \approx h_t$ with open-to-close returns
  - $x_t < h_t$ (on average) with close-to-close returns

Linear Model (SPY Open-to-Close)

- GARCH Equation

$$h_t = \underset{(0.05)}{0.09} + \underset{(0.16)}{0.29}\, h_{t-1} + \underset{(0.18)}{0.63}\, x_{t-1}$$

- Measurement Equation

$$x_t = -\underset{(0.09)}{0.05} + \underset{(0.19)}{1.01}\, h_t + \underbrace{-\underset{(0.02)}{0.02}\, z_t + \underset{(0.01)}{0.06}\,(z_t^2 - 1)}_{\tau(z)} + u_t$$

- Standard deviation of u_t: $\hat{\sigma}_u = \underset{(0.05)}{0.51}$.

Linear Model (SPY Close-to-Close)

- GARCH Equation

$$h_t = \underset{(0.04)}{0.07} + \underset{(0.15)}{0.29}\, h_{t-1} + \underset{(0.25)}{0.87}\, x_{t-1}$$

- Measurement Equation

$$x_t = +\underset{(0.08)}{0.00} + \underset{(0.14)}{0.74}\, h_t + \underbrace{-\underset{(0.02)}{0.07}\, z_t + \underset{(0.01)}{0.03}\, (z_t^2 - 1)}_{\tau(z)} + u_t$$

- Standard deviation of u_t: $\hat{\sigma}_u = \underset{(0.06)}{0.51}$.

Log-Linear Model (SPY Open-to-Close)

- GARCH Equation

$$\log h_t = \underset{(0.02)}{0.04} + \underset{(0.05)}{0.70} \log h_{t-1} + \underset{(0.04)}{0.45} \log x_{t-1} - \underset{(0.06)}{0.18} \log x_{t-2}$$

- Measurement Equation

$$\log x_t = \underset{(0.05)}{-0.18} + \underset{(0.07)}{1.04} \log h_t + \underbrace{\underset{(0.01)}{-0.07} z_t + \underset{(0.01)}{0.07} (z_t^2 - 1)}_{\tau(z)} + u_t$$

- Standard deviation of u_t: $\hat{\sigma}_u = \underset{(0.08)}{0.38}$.
- Persistence Parameter

$$\hat{\pi} = \hat{\beta} + (\hat{\gamma}_1 + \hat{\gamma}_2)\hat{\varphi} = 0.986$$

Log-Linear Model (SPY Close-to-Close)

- GARCH Equation

$$\log h_t = \underset{(0.02)}{0.11} + \underset{(0.05)}{0.72} \log h_{t-1} + \underset{(0.06)}{0.48} \log x_{t-1} - \underset{(0.07)}{0.21} \log x_{t-2}$$

- Measurement Equation

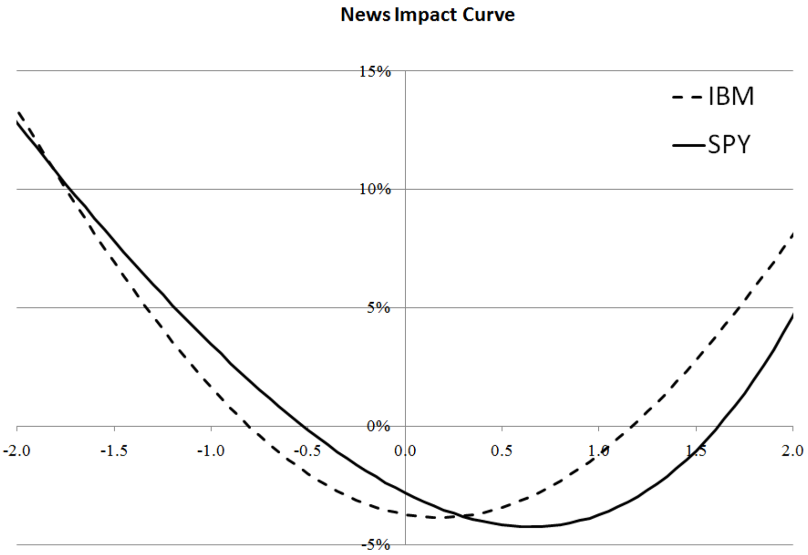
$$\log x_t = -\underset{(0.06)}{0.42} + \underset{(0.10)}{1.00} \log h_t + \underbrace{-\underset{(0.01)}{0.11}\, z_t + \underset{(0.01)}{0.04}(z_t^2 - 1)}_{\tau(z)} + u_t$$

- Standard deviation of u_t: $\hat{\sigma}_u = \underset{(0.08)}{0.38}$.
- Persistence Parameter

$$\hat{\pi} = \hat{\beta} + (\hat{\gamma}_1 + \hat{\gamma}_2)\hat{\varphi} = 0.987$$

Return and Volatility Shocks
(Leverage Effect)

Estimated News Impact Curve



Log Linear Model (SPY Open-to-Close)

Panel B: Auxiliary Statistics

Model	G(1,1)	RG(1,1)	RG(1,2)	RG(2,1)	RG(2,2)
π	.99	.98	.99	.98	1
ρ		-0.18	-0.18	-0.16	-0.19
ρ^-		-0.33	-0.32	-0.32	-0.35
ρ^+		0.12	0.12	0.13	0.13
$\ell(r)$	-1752.7	-1712.0	-1710.3	-1711.4	-1712.3

- Significant "Leverage Effect"

$$\begin{aligned}\rho^- &= \mathrm{corr}\{\tau(z_t) + u_t, z_t | z_t < 0\},\\ \rho^+ &= \mathrm{corr}\{\tau(z_t) + u_t, z_t | z_t > 0\}.\end{aligned}$$

Log Linear Model (SPY Open-to-Close)

Panel B: Auxiliary Statistics

Model	G(1,1)	RG(1,1)	RG(1,2)	RG(2,1)	RG(2,2)
π	.99	.98	.99	.98	1
ρ		-0.18	-0.18	-0.16	-0.19
ρ^-		-0.33	-0.32	-0.32	-0.35
ρ^+		0.12	0.12	0.13	0.13
$\ell(r)$	-1752.7	-1712.0	-1710.3	-1711.4	-1712.3

- Realized GARCH dominates GARCH in terms of partial log-likelihood.

Log Linear Realized GARCH(1,2) for DJIA Stocks

	ω	β	γ_1	γ_2	ξ	φ	σ_u	τ_1	τ_2	π	ρ	ρ^-	ρ^+
AA	0.03	0.77	0.33	-0.14	-0.07	1.15	0.40	-0.04	0.09	0.98	-0.08	-0.32	0.24
AIG	0.02	0.74	0.45	-0.21	-0.06	1.02	0.45	-0.02	0.04	0.98	-0.06	-0.17	0.08
AXP	0.05	0.70	0.38	-0.12	-0.16	1.08	0.43	-0.02	0.10	0.99	-0.05	-0.30	0.25
BA	0.02	0.82	0.31	-0.17	-0.13	1.22	0.39	-0.03	0.09	0.99	-0.09	-0.36	0.26
BAC	0.00	0.78	0.51	-0.29	0.00	0.99	0.42	-0.04	0.08	0.99	-0.09	-0.31	0.21
C	-0.02	0.74	0.45	-0.19	0.09	0.99	0.39	-0.03	0.09	0.99	-0.07	-0.31	0.24
CAT	0.03	0.82	0.37	-0.22	-0.14	1.07	0.38	-0.03	0.09	0.99	-0.08	-0.32	0.27
CVX	0.03	0.71	0.33	-0.14	-0.09	1.32	0.39	-0.08	0.08	0.97	-0.19	-0.35	0.14
DD	-0.01	0.77	0.37	-0.17	0.08	1.08	0.40	-0.05	0.08	0.98	-0.13	-0.35	0.20
GE	0.00	0.81	0.38	-0.19	0.01	0.98	0.41	-0.01	0.08	0.99	-0.02	-0.26	0.25
GM	0.06	0.84	0.39	-0.24	-0.32	1.02	0.47	-0.01	0.12	0.99	-0.01	-0.33	0.31
HD	0.01	0.79	0.39	-0.20	0.00	1.01	0.41	-0.05	0.09	0.99	-0.13	-0.37	0.20
IBM	0.00	0.74	0.41	-0.15	0.01	0.94	0.39	-0.04	0.08	0.98	-0.09	-0.32	0.24
INTC	0.02	0.87	0.46	-0.33	-0.11	1.03	0.36	-0.02	0.07	1.00	-0.05	-0.24	0.22
MSFT	-0.01	0.79	0.44	-0.22	0.08	0.92	0.38	-0.03	0.08	0.99	-0.08	-0.31	0.24
WMT	-0.02	0.80	0.37	-0.19	0.12	1.04	0.39	-0.01	0.09	0.99	-0.02	-0.29	0.30
XOM	0.03	0.71	0.34	-0.12	-0.10	1.26	0.38	-0.08	0.08	0.98	-0.20	-0.37	0.15
SPY	0.04	0.70	0.45	-0.18	-0.18	1.04	0.38	-0.07	0.07	0.99	-0.17	-0.32	0.13
Average	0.01	0.79	0.41	-0.21	-0.02	1.04	0.41	-0.03	0.09	0.99	-0.08	-0.30	0.22

Log Linear Realized GARCH(1,2) for DJIA Stocks

	ω	β	γ_1	γ_2	ξ	φ	σ_u	τ_1	τ_2	π	ρ	ρ^-	ρ^+
AA	0.03	0.77	0.33	-0.14	-0.07	1.15	0.40	-0.04	0.09	0.98	-0.08	-0.32	0.24
AIG	0.02	0.74	0.45	-0.21	-0.06	1.02	0.45	-0.02	0.04	0.98	-0.06	-0.17	0.08
AXP	0.05	0.70	0.38	-0.12	-0.16	1.08	0.43	-0.02	0.10	0.99	-0.05	-0.30	0.25
BA	0.02	0.82	0.31	-0.17	-0.13	1.22	0.39	-0.03	0.09	0.99	-0.09	-0.36	0.26
BAC	0.00	0.78	0.51	-0.29	0.00	0.99	0.42	-0.04	0.08	0.99	-0.09	-0.31	0.21
C	-0.02	0.74	0.45	-0.19	0.09	0.99	0.39	-0.03	0.09	0.99	-0.07	-0.31	0.24
CAT	0.03	0.82	0.37	-0.22	-0.14	1.07	0.38	-0.03	0.09	0.99	-0.08	-0.32	0.27
CVX	0.03	0.71	0.33	-0.14	-0.09	1.32	0.39	-0.08	0.08	0.97	-0.19	-0.35	0.14
DD	-0.01	0.77	0.37	-0.17	0.08	1.08	0.40	-0.05	0.08	0.98	-0.13	-0.35	0.20
GE	0.00	0.81	0.38	-0.19	0.01	0.98	0.41	-0.01	0.08	0.99	-0.02	-0.26	0.25
GM	0.06	0.84	0.39	-0.24	-0.32	1.02	0.47	-0.01	0.12	0.99	-0.01	-0.33	0.31
HD	0.01	0.79	0.39	-0.20	0.00	1.01	0.41	-0.05	0.09	0.99	-0.13	-0.37	0.20
IBM	0.00	0.74	0.41	-0.15	0.01	0.94	0.39	-0.04	0.08	0.98	-0.09	-0.32	0.24
INTC	0.02	0.87	0.46	-0.33	-0.11	1.03	0.36	-0.02	0.07	1.00	-0.05	-0.24	0.22
MSFT	-0.01	0.79	0.44	-0.22	0.08	0.92	0.38	-0.03	0.08	0.99	-0.08	-0.31	0.24
WMT	-0.02	0.80	0.37	-0.19	0.12	1.04	0.39	-0.01	0.09	0.99	-0.02	-0.29	0.30
XOM	0.03	0.71	0.34	-0.12	-0.10	1.26	0.38	-0.08	0.08	0.98	-0.20	-0.37	0.15
SPY	0.04	0.70	0.45	-0.18	-0.18	1.04	0.38	-0.07	0.07	0.99	-0.17	-0.32	0.13
Average	0.01	0.79	0.41	-0.21	-0.02	1.04	0.41	-0.03	0.09	0.99	-0.08	-0.30	0.22

Log Linear Realized GARCH(1,2) for DJIA Stocks

	ω	β	γ_1	γ_2	ξ	φ	σ_u	τ_1	τ_2	π	ρ	ρ^-	ρ^+
AA	0.03	0.77	0.33	-0.14	-0.07	1.15	0.40	-0.04	0.09	0.98	-0.08	-0.32	0.24
AIG	0.02	0.74	0.45	-0.21	-0.06	1.02	0.45	-0.02	0.04	0.98	-0.06	-0.17	0.08
AXP	0.05	0.70	0.38	-0.12	-0.16	1.08	0.43	-0.02	0.10	0.99	-0.05	-0.30	0.25
BA	0.02	0.82	0.31	-0.17	-0.13	1.22	0.39	-0.03	0.09	0.99	-0.09	-0.36	0.26
BAC	0.00	0.78	0.51	-0.29	0.00	0.99	0.42	-0.04	0.08	0.99	-0.09	-0.31	0.21
C	-0.02	0.74	0.45	-0.19	0.09	0.99	0.39	-0.03	0.09	0.99	-0.07	-0.31	0.24
CAT	0.03	0.82	0.37	-0.22	-0.14	1.07	0.38	-0.03	0.09	0.99	-0.08	-0.32	0.27
CVX	0.03	0.71	0.33	-0.14	-0.09	1.32	0.39	-0.08	0.08	0.97	-0.19	**-0.35**	**0.14**
DD	-0.01	0.77	0.37	-0.17	0.08	1.08	0.40	-0.05	0.08	0.98	-0.13	-0.35	0.20
GE	0.00	0.81	0.38	-0.19	0.01	0.98	0.41	-0.01	0.08	0.99	-0.02	-0.26	0.25
GM	0.06	0.84	0.39	-0.24	-0.32	1.02	0.47	-0.01	0.12	0.99	-0.01	-0.33	0.31
HD	0.01	0.79	0.39	-0.20	0.00	1.01	0.41	-0.05	0.09	0.99	-0.13	-0.37	0.20
IBM	0.00	0.74	0.41	-0.15	0.01	0.94	0.39	-0.04	0.08	0.98	-0.09	-0.32	0.24
INTC	0.02	0.87	0.46	-0.33	-0.11	1.03	0.36	-0.02	0.07	1.00	-0.05	-0.24	0.22
MSFT	-0.01	0.79	0.44	-0.22	0.08	0.92	0.38	-0.03	0.08	0.99	-0.08	-0.31	0.24
WMT	-0.02	0.80	0.37	-0.19	0.12	1.04	0.39	-0.01	0.09	0.99	-0.02	-0.29	0.30
XOM	0.03	0.71	0.34	-0.12	-0.10	1.26	0.38	-0.08	0.08	0.98	-0.20	**-0.37**	**0.15**
SPY	0.04	0.70	0.45	-0.18	-0.18	1.04	0.38	-0.07	0.07	0.99	-0.17	**-0.32**	**0.13**
Average	0.01	0.79	0.41	-0.21	-0.02	1.04	0.41	-0.03	0.09	0.99	-0.08	-0.30	0.22

Out-of-Sample Analysis

- $n \simeq 1,500$ and $m \simeq 187$.
- Out-of-sample likelihood ratio statistic is asymptotically distributed as

$$\sqrt{\frac{n}{m}}\{\ell_i(r,x)-\ell_j(r,x)\} \stackrel{d}{\to} Z_1'Z_2, \quad \text{as} \quad m,n\to\infty \quad \text{with} \quad \frac{m}{n}\to 0,$$

  where Z_1 and Z_2 are independent $Z_i \sim N_k(0,I)$.
- Two-sided critical values can be inferred from the distribution of $|Z_1'Z_2|$.
  $k=1:$ 2.25 and 3.67, are the 5% and 1% critical values.
  $k=2$: 3.05 and 4.83.

Out-of-Sample Analysis

	RG(1,1)	RG(1,2)	RG(2,1)	RG(2,2)	RG(2,2)†	RG(2,2)*
AA	6.9	4.5	6.4	0	21.9	0.1
AIG	15.7	-0.2	6.0	0	25.5	12.5
AXP	1.2	2.7	1.5	0	13.3	0.2
BA	-1.4	0.4	-2.6	0	24.5	0.0
BAC	8.0	-0.7	-0.3	0	56.5	-2.8
C	1.3	-2.7	-2.5	0	-0.1	-7.2
CAT	1.7	-1.0	-4.2	0	9.0	2.2
CVX	5.3	0.0	2.2	0	20.8	-0.1
DD	4.6	4.1	3.0	0	-7.1	6.9
DIS	4.8	4.8	3.7	0	24.3	0.0
GE	14.1	0.6	5.9	0	9.3	0.6
GM	17.0	-0.2	6.8	0	7.2	14.6
HD	2.2	-1.0	-1.0	0	28.9	0.5
IBM	3.3	-0.1	1.0	0	25.0	-0.2
INTC	0.9	0.1	-3.4	0	36.8	19.1
MSFT	8.3	4.2	7.1	0	17.0	4.1
WMT	-7.2	-3.8	-7.9	0	28.7	9.4
XOM	5.7	-0.1	1.0	0	27.9	0.6
SPY	6.3	-1.2	1.9	0	24.4	1.4
Average	4.4	0.1	1.1	0	23.0	2.1

Out-of-Sample: Partial Likelihood

- Out-of-sample **partial** likelihood ratio

$$2\{\max_i \ell_i(r|x) - \ell_j(r|x)\}.$$

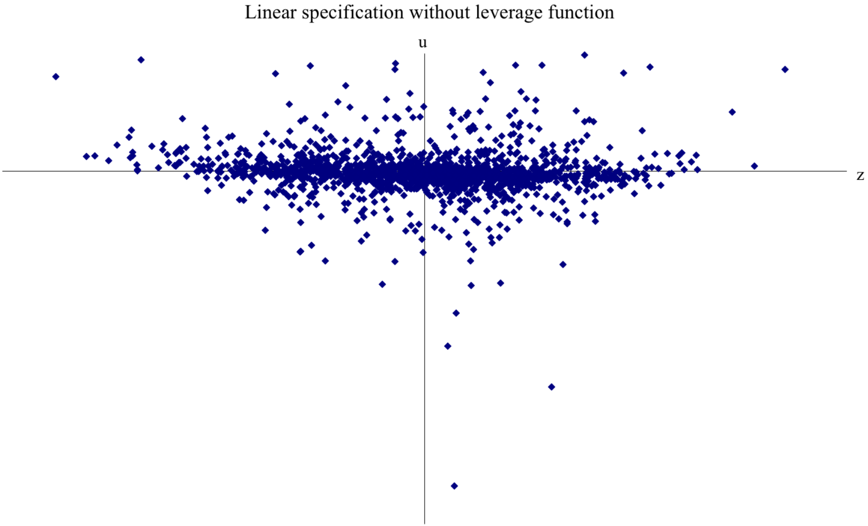
- "Best" will have a zero.
- Enables comparison with conventional GARCH.
- GARCH maximizes the in-sample partial log-likelihood.

Out-of-Sample Analysis (Partial)

	G11	RG(1,1)	RG(1,2)	RG(2,1)	RG(2,2)	RG(2,2)†	RG(2,2)*
AA	4.6	3.3	1.4	2.6	0.0	0.6	0.7
AIG	56.1	9.3	5.9	7.4	6.0	0.0	7.2
AXP	24.0	0.0	0.3	0.1	1.3	1.7	1.4
BA	1.1	0.6	1.8	1.7	0.0	0.3	0.0
BAC	147.9	4.1	0.6	0.0	1.5	19.7	0.9
C	26.9	0.3	0.9	0.5	1.1	0.0	0.4
CAT	47.3	0.1	0.8	0.0	1.0	1.6	1.5
CVX	30.1	0.6	0.3	0.2	0.3	0.0	0.3
DD	19.2	1.2	1.5	1.5	1.2	1.4	0.0
DIS	35.4	0.0	2.0	1.4	0.8	1.1	0.8
GE	41.6	0.9	0.0	0.6	0.0	0.3	0.7
GM	57.5	1.4	0.0	0.2	0.0	0.3	0.9
HD	45.5	0.9	0.0	0.1	0.2	2.1	0.4
IBM	12.0	0.4	0.5	0.7	0.1	0.3	0.0
INTC	133.1	1.7	0.5	2.0	0.0	0.1	1.8
MSFT	37.1	0.1	0.5	0.4	0.0	0.2	0.1
WMT	30.5	0.0	0.6	0.0	1.3	1.4	2.6
XOM	21.6	0.0	0.5	0.3	0.5	0.5	0.5
SPY	40.8	0.8	0.6	0.7	0.0	2.5	1.3
Average	33.5	1.2	1.0	1.0	0.8	1.6	1.0

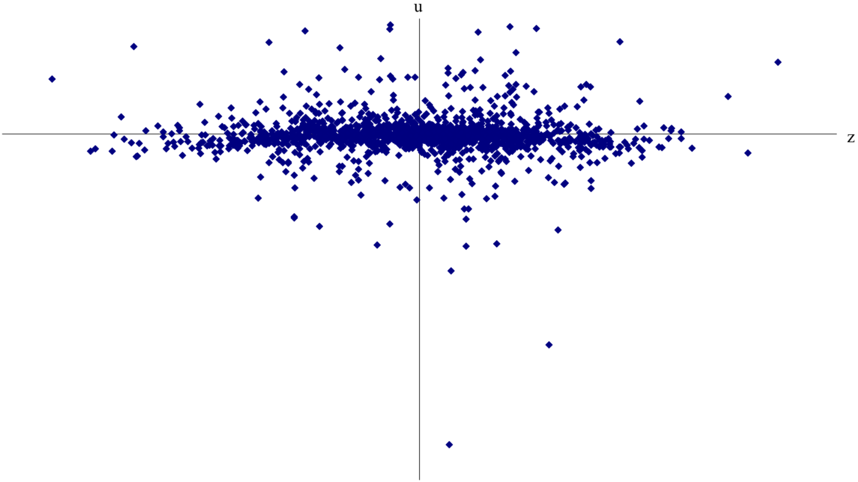
Misspecification Analysis

Lin-No-Leverage: Bivariate Gaussian?

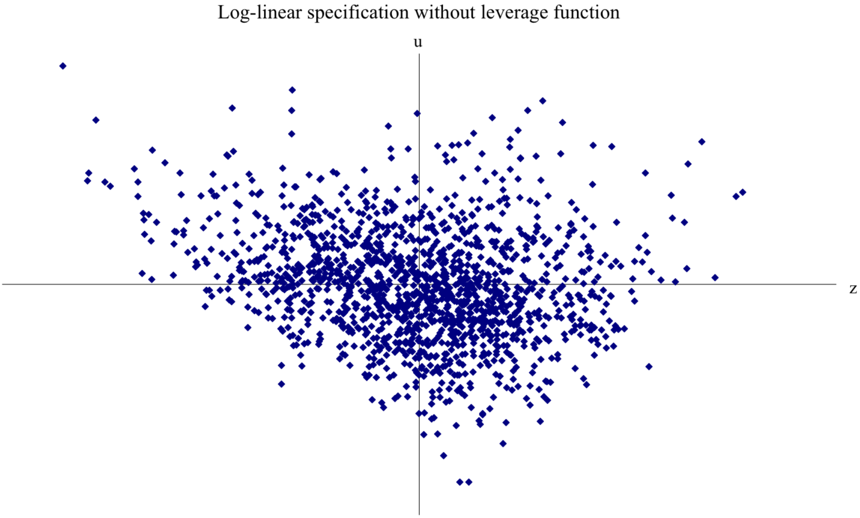


Lin-Leverage: Bivariate Gaussian?

Linear specification with quadratic leverage function

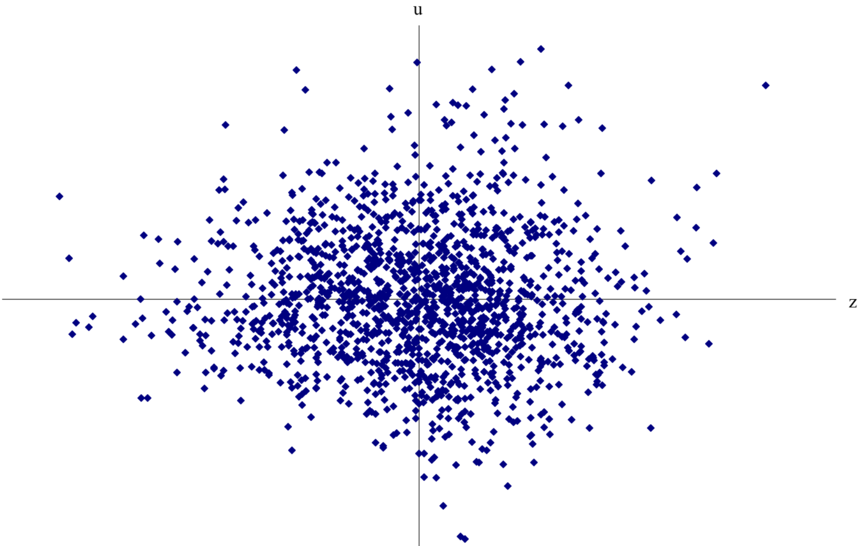


Log-No-Leverage: Bivariate Gaussian?



Log-Leverage: Bivariate Gaussian?

Log-linear specification with quadratic leverage function



Misspecification

Standard Errors for the RealGARCH(1,2) Model

	Linear Model			Log-linear Model		
	$\mathcal{I}^{-1}$	$\mathcal{J}^{-1}$	$\mathcal{I}^{-1}\mathcal{J}\mathcal{I}^{-1}$	$\mathcal{I}^{-1}$	$\mathcal{J}^{-1}$	$\mathcal{I}^{-1}\mathcal{J}\mathcal{I}^{-1}$
ω	0.007	0.004	0.019	0.015	0.015	0.016
β	0.034	0.017	0.125	0.040	0.031	0.053
γ_1	0.053	0.040	0.133	0.030	0.025	0.040
γ_2	0.054	0.032	0.177	0.046	0.036	0.062
ξ	0.038	0.037	0.096	0.044	0.042	0.051
φ	0.080	0.064	0.212	0.044	0.033	0.069
σ_u	0.009	0.002	0.054	0.005	0.005	0.006
τ_1	0.013	0.014	0.016	0.010	0.011	0.011
τ_2	0.008	0.013	0.011	0.006	0.008	0.006

Alternative: MEM/HEAVY ("Parallel GARCH")

- The return equation implies

$$r_t^2 = h_t z_t^2$$

- MEM (Engle & Gallo, 2006) and HEAVY (Shephard & Sheppard, 2010)

$$x_t = h_{x,t} z_{x,t}^2,$$

- where $h_{x,t}$ is an additional latent variable $z_{x,t} \sim (0, 1)$.
- Parallel GARCH structure:

$$h_{x,t} = \omega_x + \alpha_x r_{t-1}^2 + \beta_x h_{x,t-1} + \gamma_x x_{t-1}.$$

MEM and HEAVY Model

	Latent Variable	Observables
MEM	$\begin{aligned} h_t &= \omega + \alpha r_{t-1}^2 + \beta h_{t-1} + \delta r_{t-1} + \varphi R_{t-1}^2 \\ h_{R,t} &= \omega_R + \alpha_R R_{t-1}^2 + \beta_R h_{R,t-1} + \delta_R r_{t-1} \\ h_{RV,t} &= \omega_{RV} + \alpha_{RV} RV_{t-1} + \beta_{RV} h_{RV,t-1} \\ &\quad + \delta_{RV} r_{t-1} + \vartheta_{RV} RV_{t-1} 1_{(r_{t-1}<0)} + \varphi_{RV} r_{t-1}^2 \end{aligned}$	$\begin{aligned} r_t^2 &= h_t z_t^2 \\ R_t^2 &= h_{R,t} z_{R,t}^2 \\ RV_t &= h_{RV,t} z_{RV,t}^2 \end{aligned}$
HEAVY	$\begin{aligned} h_t &= \omega + \beta h_{t-1} + \gamma x_{t-1} \\ \mu_t &= \omega_R + \alpha_R x_{t-1} + \beta_R \mu_{t-1} \end{aligned}$	$\begin{aligned} r_t &= \sqrt{h_t} z_t \\ x_t &= \mu_t z_{RK,t}^2 \end{aligned}$
Realized GARCH	$h_t = \omega + \beta h_{t-1} + \gamma x_{t-1}$	$\begin{aligned} r_t &= \sqrt{h_t} z_t \\ x_t &= \xi + \varphi h_t + \tau(z_t) + u_t \end{aligned}$

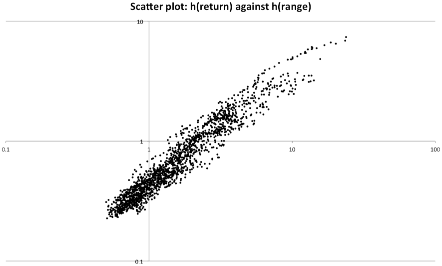
MEM/HEAVY Structure

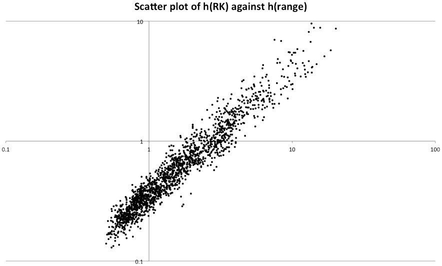
- Estimate latent volatility for: Return, Range, and Realized Kernel, separately.

$$
\begin{aligned}
r_t &= \sqrt{h_t} z_t & h_t &= \omega + \alpha r_{t-1}^2 + \beta h_{t-1} \\
R_t^2 &= h_{Rt} z_{R,t}^2 & h_{R,t} &= \omega_R + \alpha_R R_{R,t-1}^2 + \beta_R h_{R,t-1} \\
RK_t &= h_{RK,t} z_{RK,t} & h_{RK,t} &= \omega_{RK} + \alpha_{RK} R_{RK,t-1}^2 + \beta_{RK} h_{RK,t-1}
\end{aligned}
$$

- Do we need 3 volatility factors?

Single versus Multiple Latent Volatility Variable



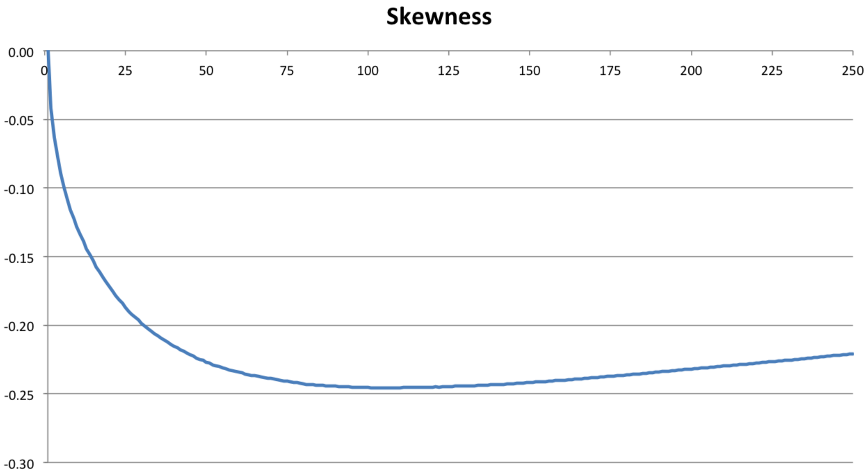


Moments, Forecasting, and More

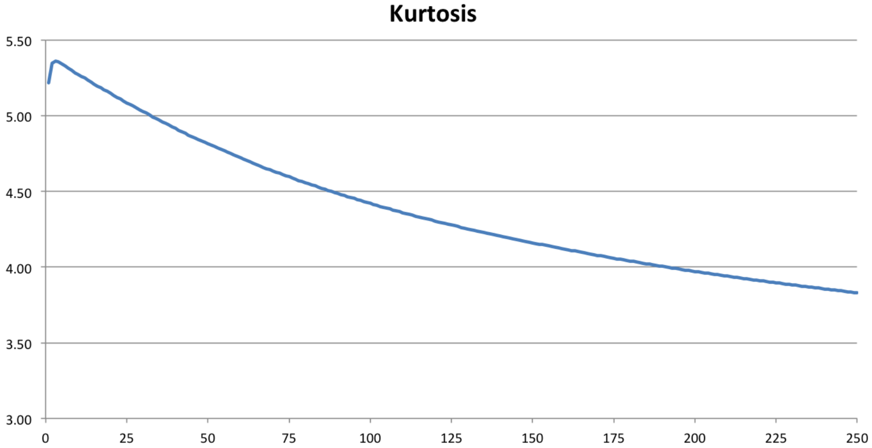
Skewness and Kurtosis of Cumulative Returns

- No skewness for r_t, unless z_t is skewed.
- "Leverage effect" induces skewness for multiperiod returns.
- Simulation using log-linear Realized GARCH(1,2) parameters for SPY.

Skewness of Cumulative Returns



Kurtosis of Cumulative Returns



Multi-Period Forecast

- Multi-period ahead predictions with the Realized GARCH model is straightforward.
- When $p = q = 1.$ we obtain VARMA(1,1) structure

$$\begin{bmatrix} \tilde{h}_t \\ \tilde{x}_t \end{bmatrix} = \begin{bmatrix} \beta & \gamma \\ \varphi\beta & \varphi\gamma \end{bmatrix} \begin{bmatrix} \tilde{h}_{t-1} \\ \tilde{x}_{t-1} \end{bmatrix} + \begin{bmatrix} \omega \\ \xi + \varphi\omega \end{bmatrix} + \begin{bmatrix} 0 \\ \tau(z_t) + u_t \end{bmatrix},$$

so we can write, $Y_t = AY_{t-1} + \mu + \epsilon_t.$

Realized Exponential GARCH
An Improved Variant

Realized Exponential GARCH

- $$\begin{aligned} r_t &= \mu + \sqrt{h_t} z_t \\ \log h_t &= \omega + \beta \log h_{t-1} + \tau(z_t) + \gamma' u_{t-1} \\ \log x_{k,t} &= \xi_k + \varphi_k \log h_t + \delta_{(k)}(z_t) + u_{k,t} \quad k = 1, \ldots, K. \end{aligned}$$

- Leverage functions (a particular choice)

$$\begin{aligned} \tau(z_t) &= \tau' a_t \\ \delta_{(k)}(z_t) &= \delta_k' a_t, \end{aligned} \qquad a_t = a(z_t) = \begin{pmatrix} z_t \\ z_t^2 - 1 \\ \vdots \end{pmatrix}.$$

(hermite polynomials).

Estimates for SPY with RK

- Open-to-Close

$$\log h_t = \underset{(0.004)}{-0.017} + \underset{(0.005)}{0.970} \log h_{t-1} - \underset{(0.009)}{0.102} z_{t-1} + \underset{(0.005)}{0.051}(z_{t-1}^2 - 1) + \underset{(0.024)}{0.272} u_{t-1}$$

$$\log x_t = \underset{(0.042)}{-0.1643} + \underset{(0.046)}{1.099} \log h_t - \underset{(0.009)}{0.072} z_t + \underset{(0.009)}{0.073}(z_t^2 - 1) + u_t,$$

with $\hat{\sigma}_u^2 = 0.132$.

- Close-to-close

$$\log h_t = \underset{(0.004)}{-0.008} + \underset{(0.005)}{0.970} \log h_{t-1} - \underset{(0.009)}{0.133} z_{t-1} + \underset{(0.005)}{0.03}(z_{t-1}^2 - 1) + \underset{(0.024)}{0.273} u_{t-1}$$

$$\log x_t = \underset{(0.042)}{-0.410} + \underset{(0.046)}{1.054} \log h_t - \underset{(0.009)}{0.011} z_t + \underset{(0.009)}{0.038}(z_t^2 - 1) + u_t,$$

with $\hat{\sigma}_u^2 = 0.133$.

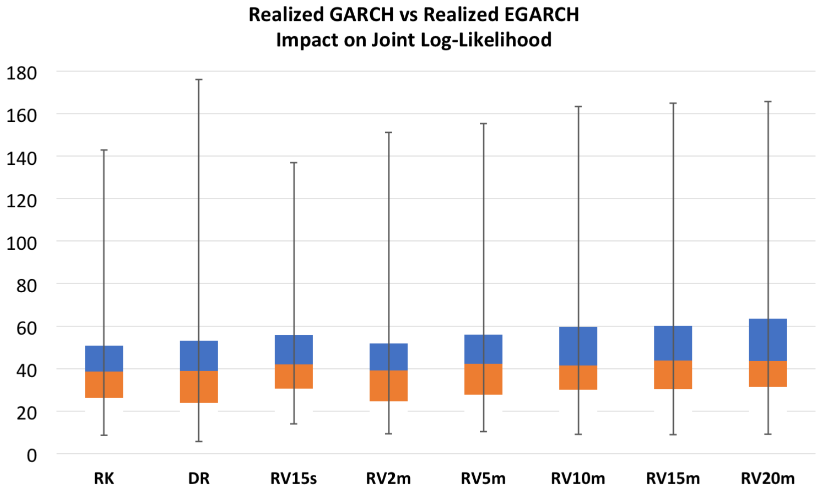
Multiple Measures (Partial Log-Likelihood)

	In-sample	Out-of-sample
EGARCH	-6.84	-19.75
RK	-1.31	-0.27
RG	-2.16	-3.22
RV15s	-7.64	-4.49
RV2m	-2.82	-0.70
RV5m	-1.49	-0.68
RV15m	-2.23	-1.15
RV20m	-2.27	-1.51
RK&RG	-0.76	0
RK&RV15s	-9.28	-5.21
RK&RV5m	-0.89	-0.13
RV15s&RV5m	-8.44	-4.71
RV5m&RV20m	-1.21	-0.47
RG&RV5m	-0.67	-0.34
RG&RV20m	-1.06	-1.25
RK&RG&RV15s	-8.13	-4.91
RK&RG&RV5m	-0.30	-0.01
RK&RV2m&RV5m&RV20m	-2.31	-0.54
RK&RG&RV5m&RV20m	0	-0.01

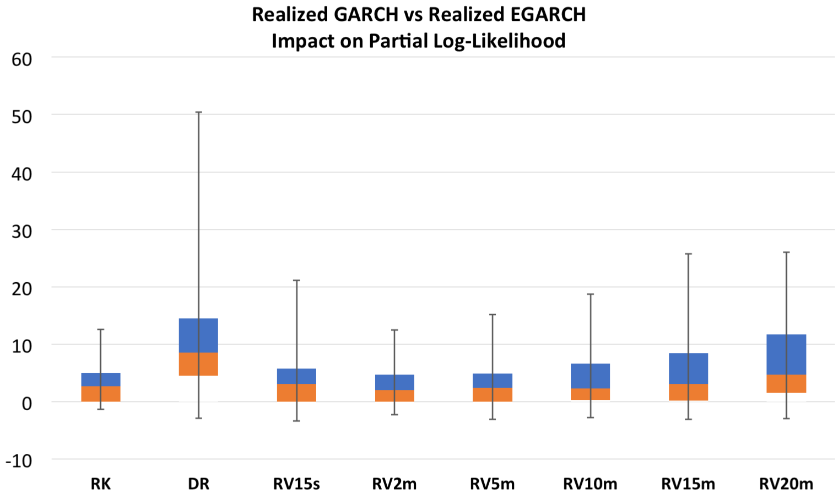
Top-3: Partial Out-of-Sample Log-Likelihood

	In-sample	Out-of-sample
EGARCH	-6.84	-19.75
RK	-1.31	-0.27
RG	-2.16	-3.22
RV15s	-7.64	-4.49
RV2m	-2.82	-0.70
RV5m	-1.49	-0.68
RV15m	-2.23	-1.15
RV20m	-2.27	-1.51
RK&RG	-0.76	0
RK&RV15s	-9.28	-5.21
RK&RV5m	-0.89	-0.13
RV15s&RV5m	-8.44	-4.71
RV5m&RV20m	-1.21	-0.47
RG&RV5m	-0.67	-0.34
RG&RV20m	-1.06	-1.25
RK&RG&RV15s	-8.13	-4.91
RK&RG&RV5m	-0.30	-0.01
RK&RV2m&RV5m&RV20m	-2.31	-0.54
RK&RG&RV5m&RV20m	0	-0.01

Realized EGARCH vs Realized GARCH (Full)



Realized EGARCH vs Realized GARCH (Partial)



Realized GARCH Volatility during the Global Financial Crisis

Realized EGARCH: Global Financial Crisis

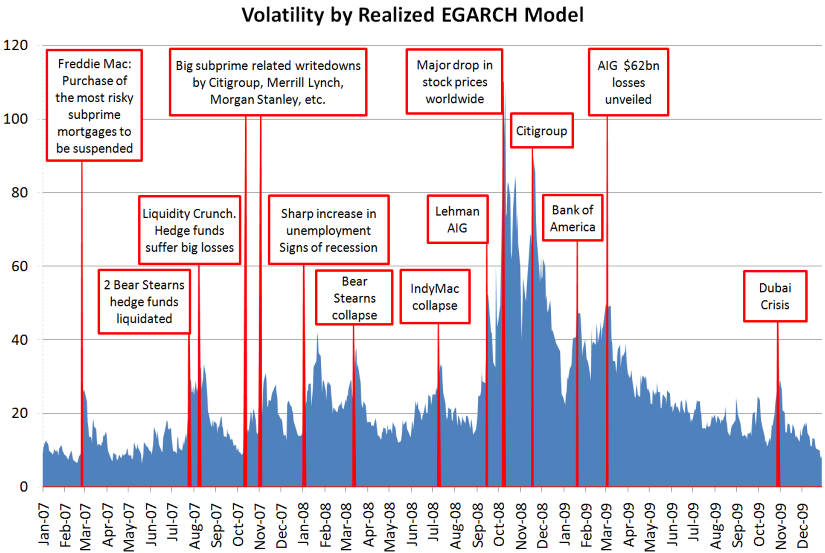
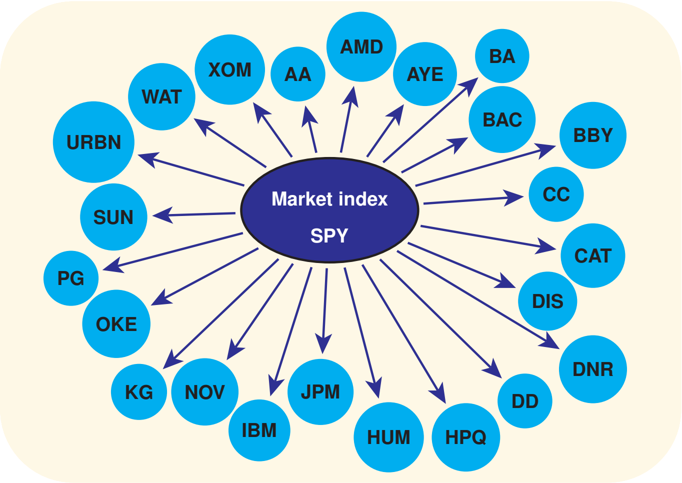


Figure: Conditional volatility during the global financial crisis with some of the

Realized Beta GARCH

- Multivariate GARCH model that utilizes realized measures of volatility and covolatility
- In essence: Bivariate GARCH model (+model for realized measures)
  - With a particular structure: Market return + individual return
  - Easy to extend to large number of "individual returns"
  - One-factor structure... can extract "betas".

We assume a simple one-factor structure: Where assets are taged on one-by-one



Multivariate Realized GARCH

- Bivariate return:
  - r_{0t} Market return (SPX returns used as proxy)
  - r_{1t} "Individual asset" return
  - x_{0t} Realized measure of market volatility
  - x_{1t} Realized volatility measure for r_{1t}.
  - y_{1t} realized correlation measure

$$y_{1t} = \frac{x_{01t}}{\sqrt{x_{0t}x_{1t}}},$$

  where x_{01t} is a realized covariance measure..

Notation and Key Variables

- Filtration:

$$\mathcal{F}_t = \sigma(r_{0s}, r_{1s}, x_{0s}, x_{1s}, y_{1s} | s \leq t).$$

- Conditional Correlation

$$\rho_t = \frac{\mathrm{cov}(r_{0t}, r_{1t} | \mathcal{F}_{t-1})}{\sqrt{\mathrm{var}(r_{0t} | \mathcal{F}_{t-1}) \mathrm{var}(r_{1t} | \mathcal{F}_{t-1})}},$$

- Dynamic Beta

$$\beta_t = \frac{\mathrm{cov}(r_{0t}, r_{1t} | \mathcal{F}_{t-1})}{\mathrm{var}(r_{0t} | \mathcal{F}_{t-1})}.$$

Decomposing the Problem

- Seek model for
$$f(r_{0t}, x_{0t}, r_{1t}, x_{1t}, y_{1t} | \mathcal{F}_{t-1}).$$
- Decompose into:
$$f(r_{0t}, x_{0t} | \mathcal{F}_{t-1}) f(r_{1t}, x_{1t}, y_{1t} | r_{0t}, x_{0t}, \mathcal{F}_{t-1}).$$

Realized EGARCH For Market Returns

- $$\begin{aligned} r_{0t} &= \mu_0 + \sqrt{h_{0t}} z_{0t}, & z_{0t} &\sim \mathrm{iid}\, N(0,1), \\ \log h_{0t} &= a_0 + b_0 \log h_{0t-1} + c_0 \log x_{0t-1} + \tau_{(0)}(z_{0t-1}) \\ \log x_{0t} &= \xi_0 + \varphi_0 \log h_{0t} + \delta_{(0)}(z_{0t}) + u_{0t}, & u_{0t} &\sim \mathrm{iid}\, N(0, \sigma_u^2). \end{aligned}$$

- With $z_{0t} \perp\!\!\!\perp u_{0t}$. This specifies

$$f(r_{0t}, x_t | \mathcal{F}_{t-1}).$$

Now the Conditional Model

- $$f(r_{1t}, x_{1t}, y_{1t} | r_{0t}, x_{0t}, \mathcal{F}_{t-1}).$$

- (another) decomposition into...

$$f(r_{1t} | r_{0t}, x_{0t}, \mathcal{F}_{t-1}) f(x_{1t}, y_{1t} | r_{1t}, r_{0t}, x_{0t}, \mathcal{F}_{t-1}).$$

Conditional Model for Individual Return

- $$\begin{aligned} r_{1t} &= \mu_1 + \sqrt{h_{1t}} z_{1t}, \\ \log h_{1t} &= a_1 + b_1 \log h_{1t-1} + c_1 \log x_{0t-1} + d_1 \log h_{0t-1} + \tau_{(1)}(z_{1,t-1}) \end{aligned}$$

- With $z_{1t} | r_{0t}, x_{0t}, \mathcal{F}_{t-1} = z_{1t} | z_{0t} \sim N(\rho_t z_{0t}, 1 - \rho_t^2)$ i.e.
$$\begin{pmatrix} z_{0t} \\ z_{1t} \end{pmatrix} \sim N \begin{pmatrix} 1 & \rho_t \\ \rho_t & 1 \end{pmatrix}.$$

- This specifies
$$f(r_{1t} | r_{0t}, x_{0t}, \mathcal{F}_{t-1})$$

Conditional Model for Individual RM

- Finally $f(x_{1t}, y_{1t}|r_{1t}, r_{0t}, x_{0t}, \mathcal{F}_{t-1})$ defined by

$$
\begin{aligned}
\log x_{1t} &= \xi_1 + \varphi_1 \log h_{1t} + \delta_{(1)}(z_{1t}) + u_{1t}, \\
F(y_{1t}) &= \xi_{10} + \varphi_{10} F(\rho_{1t}) + v_{1t},
\end{aligned}
$$

- where

$$F(\rho) = \frac{1}{2} \log \frac{1+\rho}{1-\rho}, \qquad \text{(Fisher transform)}$$

- and

$$\begin{pmatrix} u_{0t} \\ u_{1t} \\ v_{1t} \end{pmatrix} \sim \mathrm{iid} N(0, \begin{pmatrix} \sigma^2_{u_0} & \sigma^2_{u_0,u_1} & \sigma^2_{u_0,v_1} \\ \bullet & \sigma^2_{u_1} & \sigma^2_{u_1,v_1} \\ \bullet & \bullet & \sigma^2_{v_1} \end{pmatrix}$$

Estimation
Log-likelihood Components

- We formulate a quasi likelihood function of the joint model, by adopting a Gaussian specification.
- So $(z_{0,t}, z_{1,t})$ are independent of $(u_{0,t}, u_{1,t}, v_{1,t})$ which implies that the log-likelihood of the model is given by

$$\ell(r_0, x_0, r_1, x_1, y_1) = -\frac{1}{2}\left(\ell_{z_0} + \ell_{u_0} + \ell_{z_1|z_0} + \ell_{u_1,v_1|u_0}\right)$$

- The likelihood contributions given a normal distribution are straightforward to derive (and available in the paper)

Results and Discussion
The Data Set

- In estimating our proposed model, we use high-frequency prices of 594 assets.
- SPY is an exchange-traded fund that holds all of the S&P 500 Index stocks and has enormous liquidity. We use this as a proxy for the market.
- The data source is the collection of trades and quotes recorded on the NYSE, taken from the TAQ database through the Wharton Research Data Services (WRDS) system.
- The sample period runs from January 3, 2002 to the end of 2009, delivering (for most of the stocks) 2008 distinct trading days.
- We match TAQ data to CRSP data and eventually use CRSP's permnos to identify stocks and to adjust for splits and dividends.

Results and Discussion
The Data Set

- We start with a ticker list of all S&P 500 constituents as of 31.12.2009. We match these to CRSP permnos.
- The ticker history of each permno is collected and HF data for tickers that were detected in the second step and were not in the initial ticker symbol list is extracted.
- Finally, the daily data from the CRSP is matched to the high-frequency based realized measures using the PERMNOs as a company identifier resulting in a total of 743 assets.
- Tickers as company identifiers are risky:
  - Over the sample period we consider, around 10% of the companies had 2 or more ticker symbols
  - A particular ticker symbol can be used over time for more than one company: e.g. ”T” (AT&T) has two permnos (i.e. CRSP identifies two distinct companies for this ticker)

Estimated Market Model

- Return eq.

$$r_{0t} = 0.01 + \sqrt{h_{0t}} z_{0t}, \qquad (h_{01} = 1.33).$$

- GARCH eq.

$$\log h_{0t} = 0.175 + 0.601 \log h_{0t-1} + 0.377 x_{0t-1} - 0.092 z_{0t-1} + 0.020(z_{0t-1}^2 - 1),$$

- Measurement eq.

$$\log x_{0t} = -0.469 + \log h_{0t} - 0.104 \times z_{0t} + 0.018 \times (z_{0t}^2 - 1) + u_{0t}.$$

Estimated Model (Individual Assets)

$$\begin{aligned}\log h_{it} &= a_i + b_i \log h_{it-1} + c_i x_{it-1} + d_i h_{0t} + \tau_{i1} z_{it-1} + \tau_{i2}(z_{it-1}^2 - 1),\\ \log x_{it} &= \xi_i + \varphi_i \log h_{it} + \delta_{i1} z_{it} + \delta_{i2}(z_{it}^2 - 1) + u_{it},\end{aligned}$$

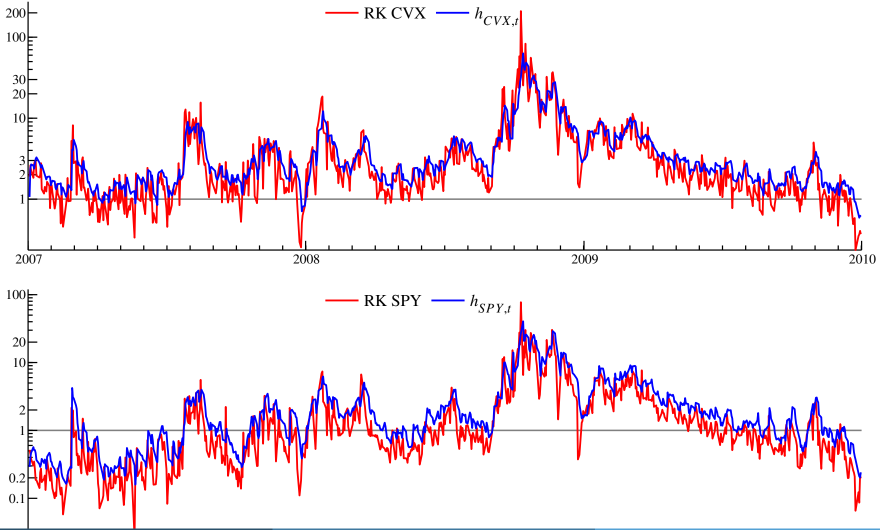
	Volatility parameters								
	$h_{i,1}$	μ_i	a_i	b_i	c_i	d_i	τ_{i1}	τ_{i2}	ξ_i
Mean	8.157	0.023	0.212	0.584	0.326	0.055	-0.037	0.011	-0.332
Med.	3.461	0.019	0.197	0.588	0.329	0.042	-0.038	0.010	-0.339
Min	0.025	-0.132	-0.023	0.291	0.152	-0.021	-0.086	-0.022	-0.895
1%	0.252	-0.070	0.022	0.422	0.213	-0.007	-0.072	-0.009	-0.737
5%	0.738	-0.044	0.064	0.461	0.251	0.002	-0.059	-0.004	-0.558
95%	27.54	0.102	0.419	0.701	0.396	0.144	-0.012	0.028	-0.080
99%	58.54	0.148	0.521	0.752	0.425	0.183	-0.000	0.036	-0.005
Max	323.9	0.230	0.679	0.805	0.473	0.345	0.014	0.048	0.086

Estimated Model (Correlation Parameters)

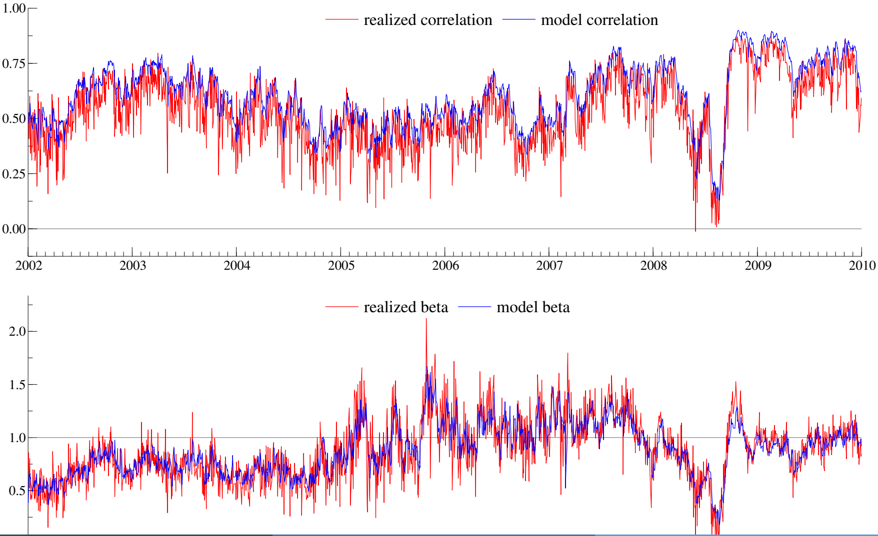
$$
\begin{aligned}
F(\rho_{it}) &= a_{i0} + b_{i0} F(\rho_{it-1}) + c_{i0} F(y_{it-1}) \\
F(y_{it}) &= \xi_{i0} + \varphi_{i0} F(\rho_{it}) + v_{it},
\end{aligned}
$$

	Correlation Parameters					
	$\rho_{i,1}$	a_{i0}	b_{i0}	c_{i0}	ξ_{i0}	ϕ_{i0}
Mean	0.355	0.036	0.709	0.285	-0.098	0.974
Median	0.353	0.035	0.704	0.277	-0.064	0.912
Min	-0.196	-0.112	0.533	0.029	-1.697	0.469
1%	0.012	-0.084	0.563	0.072	-0.740	0.515
5%	0.078	-0.030	0.602	0.121	-0.426	0.599
95%	0.653	0.107	0.834	0.477	0.145	1.481
99%	0.752	0.136	0.890	0.601	0.225	2.460
Max	0.857	0.204	0.964	0.670	0.255	3.620

Volatities: CVX & SPY

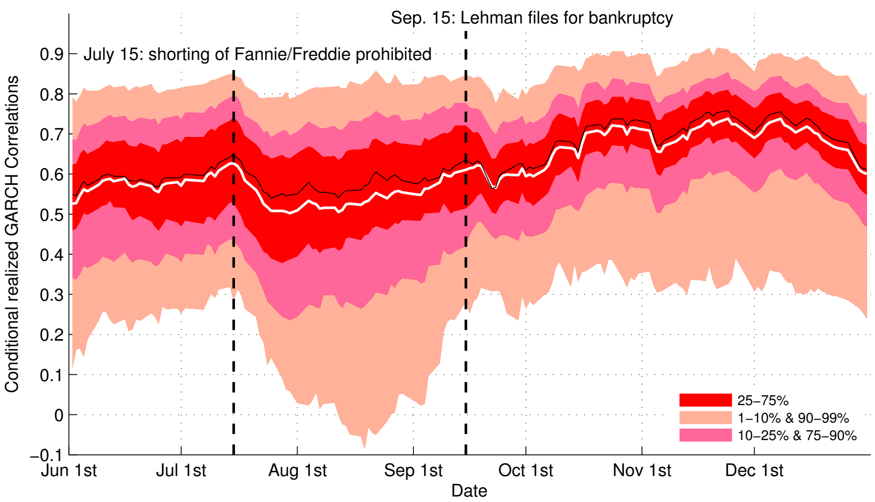


Correlation & Beta (CVX,SPY)

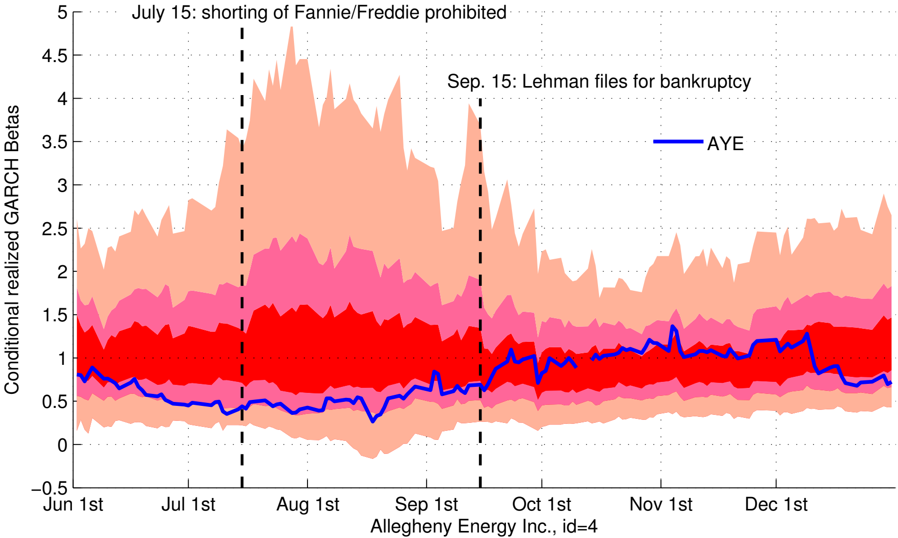


Results and Discussion

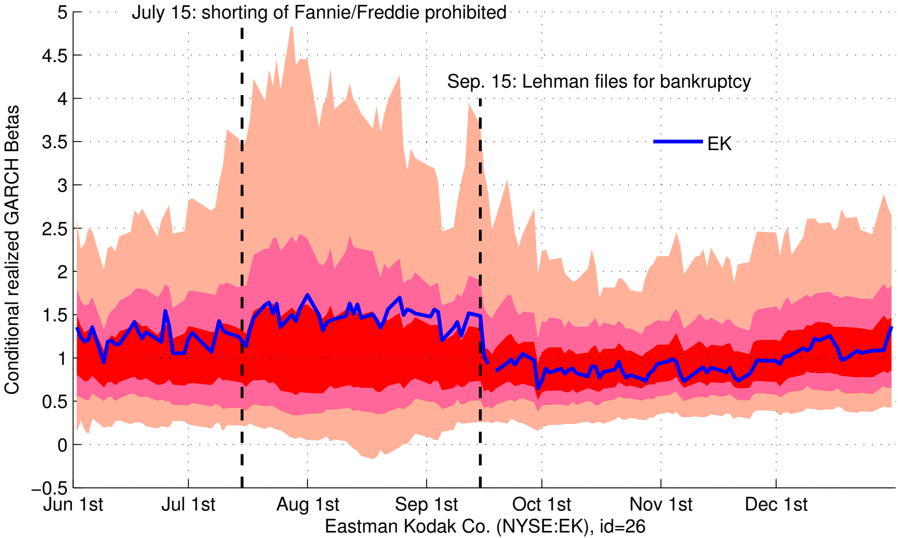
Estimation Results: Correlations

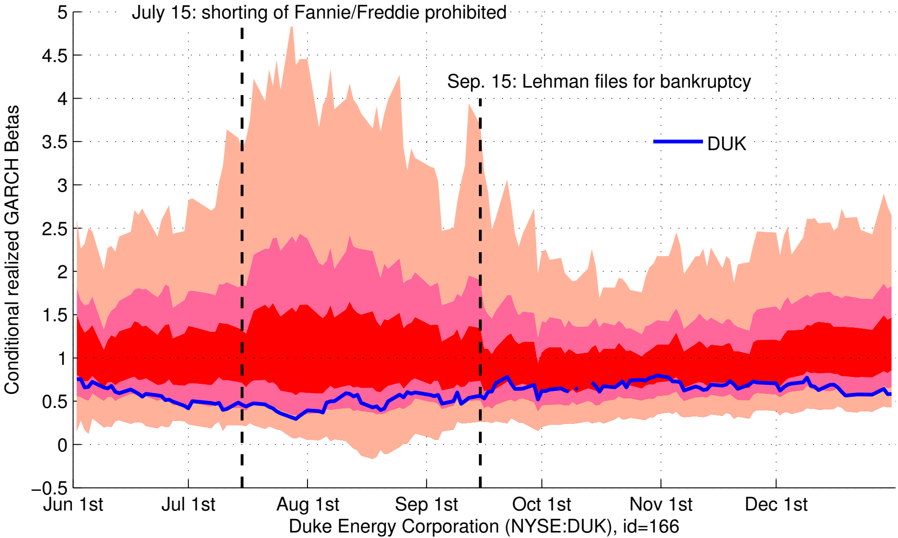


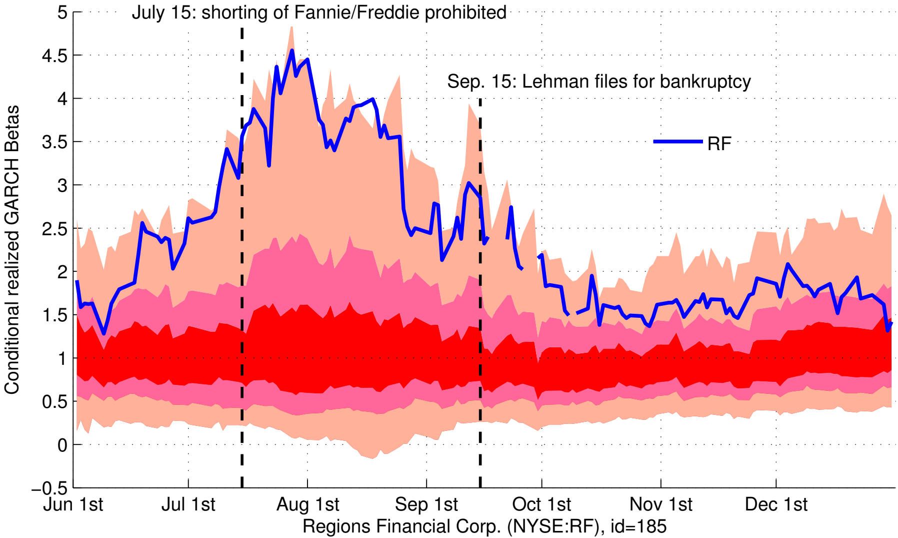
Estimation Results: Betas



Estimation Results: Betas

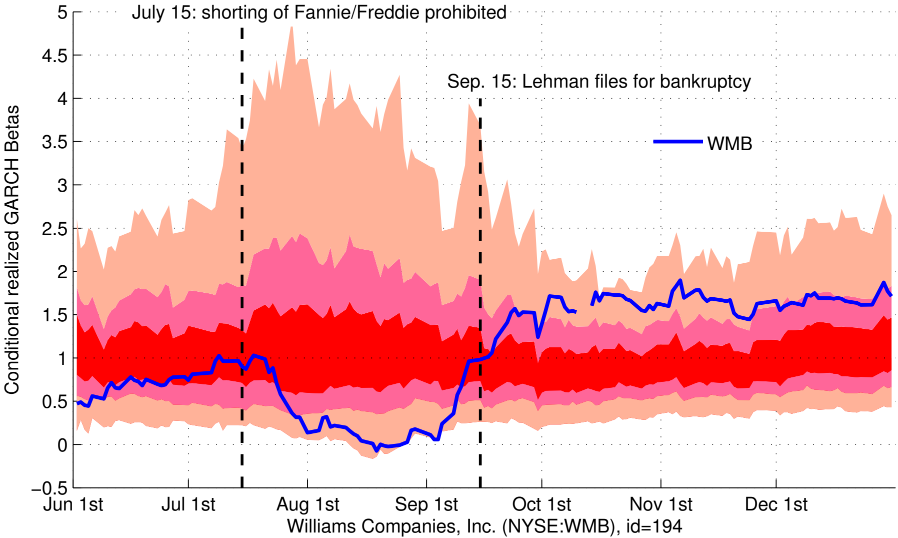



5
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
−0.5
Conditional realized GARCH Betas
July 15: shorting of Fannie/Freddie prohibited
Sep. 15: Lehman files for bankruptcy
DUK
Jun 1st
Jul 1st
Aug 1st
Sep 1st
Oct 1st
Nov 1st
Dec 1st
Duke Energy Corporation (NYSE:DUK), id=166



July 15: shorting of Fannie/Freddie prohibited
Sep. 15: Lehman files for bankruptcy
RF
Conditional realized GARCH Betas
5
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
−0.5
Jun 1st
Jul 1st
Aug 1st
Sep 1st
Oct 1st
Nov 1st
Dec 1st
Regions Financial Corp. (NYSE:RF), id=185

Results and Discussion
Estimation Results: Betas



Summary

- Realized GARCH
  - GARCH that make use of Realized Measures.
- Realized Beta GARCH
  - **Multivariate GARCH Model with Realized Measures**
  - Simplified one-factor structure that is easy to scale
- Key Features
  - **Parsimonious**
  - **Tractable** (quasi maximum likelihood).
  - **Easy** to estimate.
  - Correlations modeled by **Fisher-Transform**
- Empirical Results
  - Great deal **variation** in betas
  - Cross-sectionally and across time.