Advanced Risk Management

Week 2

Univariate Volatility Models

UNIVERSITY OF AMSTERDAM

AMSTERDAM BUSINESS SCHOOL Business

▸ Copyright statement

Outline

- The GARCH model
- Variance forecasting
- (Quasi-)Maximum likelihood estimation
- News impact curve
- Model checking
- Realized variance
- Volatility forecasting using realized variance

Volatility modelling

- Volatility $=$ standard deviation, $\sqrt{\text{variance}}$.
- Recall the "generic" RiskMetrics volatility model:

$$R_{t+1} = \sigma_{t+1} z_{t+1}, \qquad \sigma^2_{t+1} = \lambda \sigma^2_t + (1-\lambda) R^2_t.$$

  Motivation: exploits volatility clustering ($\text{Corr}(R^2_t, R^2_{t-j}) > 0$).
- Solution is exponentially weighted moving average of squared returns:

$$\sigma^2_{t+1} = (1-\lambda) \sum_{j=0}^{\infty} \lambda^j R^2_{t-j};$$

  more reasonable than equally weighted moving average $\sigma^2_{t+1} = m^{-1} \sum_{j=0}^{m-1} R^2_{t-j}$.
- GARCH models provide a generalization of the RiskMetrics model, allowing for more flexibility.

The GARCH model

- GARCH $=$ "Generalized AutoRegressive-Conditional Heteroskedasticity".
- Developed by Rob Engle (2003 Nobel laureate) and Tim Bollerslev.
- Returns R_{t+1} satisfy $E_t(R_{t+1}) = 0$ and $\text{Var}_t(R_{t+1}) = \sigma^2_{t+1}$, where

$$\sigma^2_{t+1} = \omega + \alpha R_t^2 + \beta \sigma_t^2, \qquad \omega, \alpha, \beta > 0, \qquad \alpha + \beta < 1.$$

- Three coefficients ω, α, β need to be positive so that σ^2_{t+1} is always positive.
- Like RiskMetrics, variance can be written as weighted sum of squared returns plus constant:

$$\sigma^2_{t+1} = \frac{\omega}{1-\beta} + \alpha \sum_{j=0}^{\infty} \beta^j R^2_{t-j},$$

but weights do not sum to 1! RiskMetrics is obtained by $\omega = 0$ and $\beta = 1 - \alpha = \lambda$.

The GARCH model

- σ^2_{t+1} is the time-varying conditional variance of R_{t+1}; is there a constant unconditional variance?
- If so, it should satisfy

$$\sigma^2 = \mathrm{Var}(R_{t+1}) = E(R^2_{t+1}) = E(E_t[R^2_{t+1}]) = E(\sigma^2_{t+1}),$$

so taking expectations in the GARCH equation gives

$$\begin{aligned} \sigma^2 = E(\sigma^2_{t+1}) &= \omega + \alpha E(R^2_t) + \beta E(\sigma^2_t) \\ &= \omega + \alpha\sigma^2 + \beta\sigma^2, \end{aligned}$$

which is solved by

$$\sigma^2 = \frac{\omega}{1 - \alpha - \beta}.$$

- Therefore, $\alpha + \beta < 1$ is needed for the unconditional variance to exist. In that case,

$$\sigma^2_{t+1} = \sigma^2 + \alpha(R^2_t - \sigma^2) + \beta(\sigma^2_t - \sigma^2).$$

Variance forecasting

- In risk management applications, we will be interested in

$$\mathrm{Var}_t(R_{t+1:t+K}) = \mathrm{Var}_t\left(\sum_{k=1}^{K} R_{t+k}\right) = \sum_{k=1}^{K} E_t(\sigma_{t+k}^2).$$

  Covariance terms are zero because $E_t(R_{t+1}) = 0 \Rightarrow$ no autocorrelation in returns.

- Final equation on previous slide implies

$$E_t(\sigma_{t+k}^2 - \sigma^2) = \alpha E_t(R_{t+k-1}^2 - \sigma^2) + \beta E_t(\sigma_{t+k-1}^2 - \sigma^2) = (\alpha+\beta) E_t(\sigma_{t+k-1}^2 - \sigma^2),$$

  which can be further solved into

$$E_t(\sigma_{t+k}^2 - \sigma^2) = (\alpha + \beta)^{k-1}(\sigma_{t+1}^2 - \sigma^2).$$

- So GARCH with $\alpha + \beta < 1$ leads to mean-reversion in volatility:

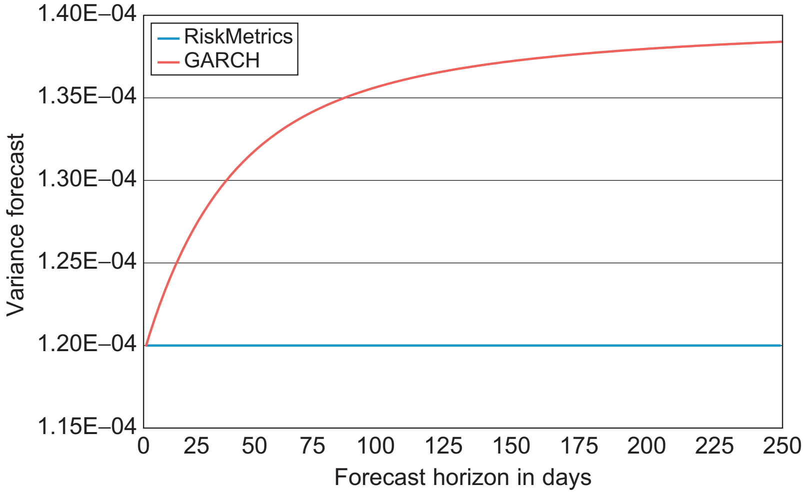
$$\lim_{k\to\infty} \mathrm{Var}_t(R_{t+k}) = \sigma^2.$$

Variance forecasting

- Combining $\text{Var}_t(R_{t+1:t+K}) = \sum_{k=1}^{K} E_t(\sigma^2_{t+k})$ with
$$E_t(\sigma^2_{t+k}) = \sigma^2 + (\alpha + \beta)^{k-1}(\sigma^2_{t+1} - \sigma^2),$$
we find
$$\text{Var}_t(R_{t+1:t+K}) = K\sigma^2 + \sum_{k=1}^{K}(\alpha + \beta)^{k-1}(\sigma^2_{t+1} - \sigma^2).$$
- The may be contrasted with RiskMetrics:
$$\text{Var}_t(R_{t+1:t+K}) = K\sigma^2_{t+1}$$
- So rule of thumb: "K-day VaR = $\sqrt{K} \times$ 1-day VaR", does not apply with GARCH.

Variance forecasting – illustration

Figure 4.2 Variance forecast for 1–250 days cumulative returns.



Maximum likelihood estimation

- To apply model in practice, parameters ω, α, β need to be estimated from data $R_1, \ldots, R_T$.
- Model cannot be estimated by OLS. Instead, use maximum likelihood estimation.
- If $z_t = R_t/\sigma_t$ is i.i.d. $N(0,1)$, then conditional density of R_t given information at $t-1$ is

$$\ell_t = \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(-\frac{R_t^2}{2\sigma_t^2}\right), \qquad \sigma_t^2 = \omega + \alpha R_{t-1}^2 + \beta\sigma_{t-1}^2.$$

- Maximum likelihood estimators $\hat{\omega}$, $\hat{\alpha}$ and $\hat{\beta}$ are solutions to

$$\max_{\omega,\alpha,\beta} \ln L = \sum_{t=1}^{T} \ln \ell_t \qquad \Leftrightarrow \qquad \min_{\omega,\alpha,\beta} \sum_{t=1}^{T} \left(\ln \sigma_t^2 + \frac{R_t^2}{\sigma_t^2}\right).$$

Need to make assumption about σ_1^2, e.g., $\sigma_1^2 = T^{-1} \sum_{t=1}^{T} R_t^2$.

Quasi-maximum likelihood estimation

- In practice, distribution of standardized returns z_t does not look like $N(0,1)$ (longer tails).
- We still get good estimates by maximizing the normal likelihood, but need to adjust the standard errors of the estimates. This is known as QMLE.
- Sometimes "variance targeting" is useful:
  - first estimate $\hat{\sigma}^2 = T^{-1} \sum_{t=1}^{T} R_t^2$;
  - then estimate α and β by maximizing $\ln L$, using

$$\begin{aligned} \sigma_t^2 &= \hat{\sigma}^2 + \alpha(R_{t-1}^2 - \hat{\sigma}^2) + \beta(\sigma_{t-1}^2 - \hat{\sigma}^2) \\ &= \hat{\sigma}^2(1 - \alpha - \beta) + \alpha R_{t-1}^2 + \beta \sigma_{t-1}^2, \end{aligned}$$

  implying an estimate $\hat{\omega} = \hat{\sigma}^2(1 - \hat{\alpha} - \hat{\beta})$.

Implementation in Python

- Install arch package:

  `conda install -c bashtage arch`

- Code:

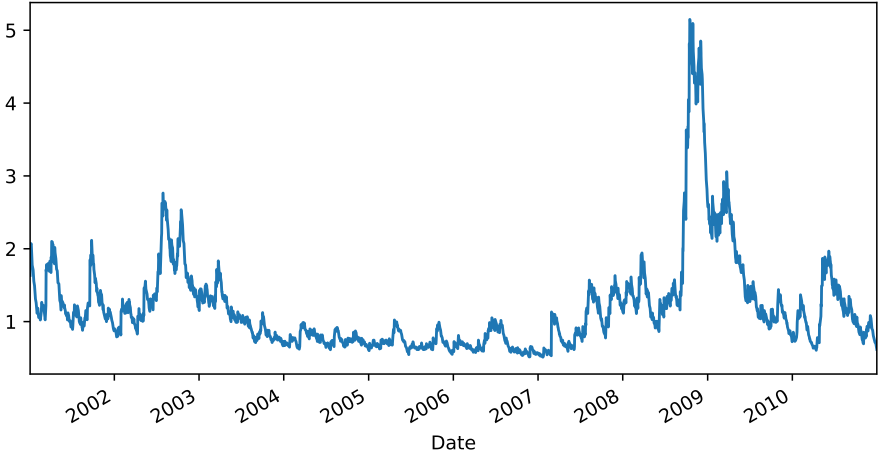
```
from arch import arch_model
am = arch_model(R)
res = am.fit()
print(res.summary())
sigma = res.conditional_volatility
sigma.plot()
```

- To avoid numerical problems, `R` has been multiplied by 100 (percentage returns).

Results – estimation output

```
                     Constant Mean - GARCH Model Results
==============================================================================
Dep. Variable:                      R   R-squared:                      -0.001
Mean Model:             Constant Mean   Adj. R-squared:                 -0.001
Vol Model:                      GARCH   Log-Likelihood:               -3719.07
Distribution:                  Normal   AIC:                           7446.15
Method:            Maximum Likelihood   BIC:                           7469.47
                                        No. Observations:                 2514
Date:                Wed, Feb 07 2018   Df Residuals:                     2510
Time:                        22:40:22   Df Model:                            4
                                 Mean Model
============================================================================
                 coef    std err          t      P>|t|      95.0% Conf. Int.
----------------------------------------------------------------------------
mu             0.0420  1.759e-02      2.390  1.685e-02 [7.561e-03,7.650e-02]
                              Volatility Model
============================================================================
                 coef    std err          t      P>|t|      95.0% Conf. Int.
----------------------------------------------------------------------------
omega          0.0125  5.468e-03      2.279  2.267e-02 [1.744e-03,2.318e-02]
alpha[1]       0.0791  1.154e-02      6.858  6.986e-12   [5.652e-02,  0.102]
beta[1]        0.9122  1.214e-02     75.162      0.000     [  0.888,  0.936]
============================================================================
```

Results – volatility plot



Implementation in Stata

- Code:

```
import excel "C:\Work\ARM2020\Week2\R.xlsx", sheet("R") firstrow
replace R = 100*R
bcal create busdates, from(Date)
bcal load busdates
replace Date = bofd("busdates",Date)
format %tbbusdates Date
tsset Date
save "C:\Work\ARM2020\Week2\R.dta"

arch R if Date>252, arch(1) garch(1) nolog
```

- The `bcal` statements (“business calendar”) are to communicate to the program that business holidays should not be interpreted as “gaps” in the data set.

Estimation output in Stata

```
ARCH family regression

Sample: 03jan2001 - 31dec2010                   Number of obs     =      2,514
Distribution: Gaussian                          Wald chi2(.)      =          .
Log likelihood = -3719.971                      Prob > chi2       =          .

------------------------------------------------------------------------------
             |                 OPG
           R |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
R            |
       _cons |   .0421904    .018227     2.31   0.021     .0064661    .0779147
-------------+----------------------------------------------------------------
ARCH         |
        arch |
         L1. |   .0786345   .0080383     9.78   0.000     .0628798    .0943891
             |
       garch |
         L1. |   .9128176   .0085148   107.20   0.000     .8961289    .9295063
             |
       _cons |   .0123212   .0020622     5.97   0.000     .0082794     .016363
------------------------------------------------------------------------------
```

News impact curve

- Standardized return $z_t = R_t/\sigma_t$ may be interpreted as today's "shock to return", or "news".
- News impact curve (function) is effect of z_t on σ^2_{t+1}, keeping σ^2_t fixed.
- Standard GARCH model:

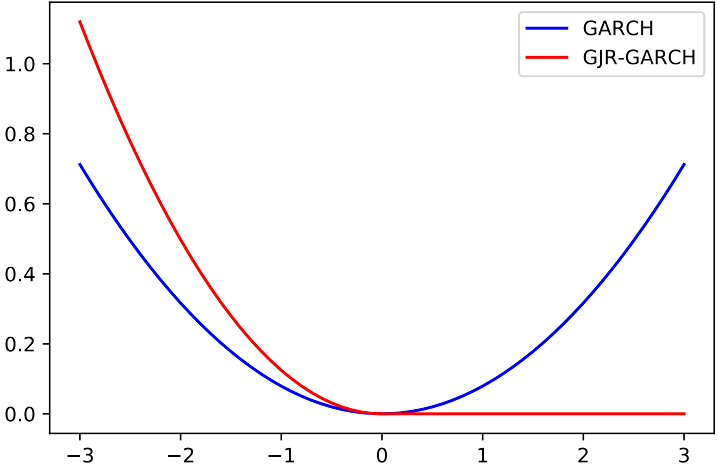
$$\sigma^2_{t+1} = \omega + \alpha R_t^2 + \beta\sigma_t^2 = \omega + \alpha\sigma_t^2 z_t^2 + \beta\sigma_t^2,$$

  so NIC is quadratic. So positive shock $z_t = c$ has same effect as negative shock $z_t = -c$.
- Leverage effect implies for equity returns: $z_t = -c$ has bigger impact than $z_t = c$.
- Extension that allows for this is GJR-GARCH:

$$\sigma^2_{t+1} = \omega + \alpha R_t^2 + \gamma I_t R_t^2 + \beta\sigma_t^2, \qquad I_t = \mathbb{I}(R_t < 0),$$

  where $\mathbb{I}(A) = 1$ if A is true, and zero otherwise.
- Model can be estimated by QMLE; in practice, $\hat{\gamma}$ is much larger than $\hat{\alpha} \approx 0$.
- Python implementation: `am = arch_model(R,p=1,o=1,q=1)`

Example – NIC of GARCH and GJR-GARCH, S&P 500 returns



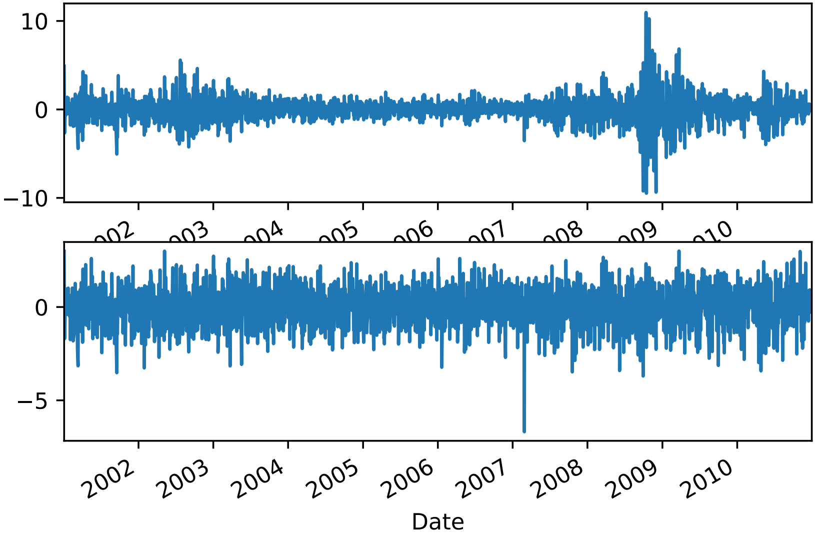
Model checking

- Assumption that $\sigma^2_{t+1} = \text{Var}_t(R_{t+1})$ satisfies (GJR-)GARCH specification may be violated:
  - we might need another form of NIC;
  - we might need (more) lags of R^2_t and σ^2_t in the equation $\Rightarrow$ GARCH(p, q);
  - perhaps it's better to formulate a linear model for $\ln \sigma^2_{t+1} \Rightarrow$ EGARCH.
- When model 0 is a special case of model 1 ("nested") then test of 0 against 1 is

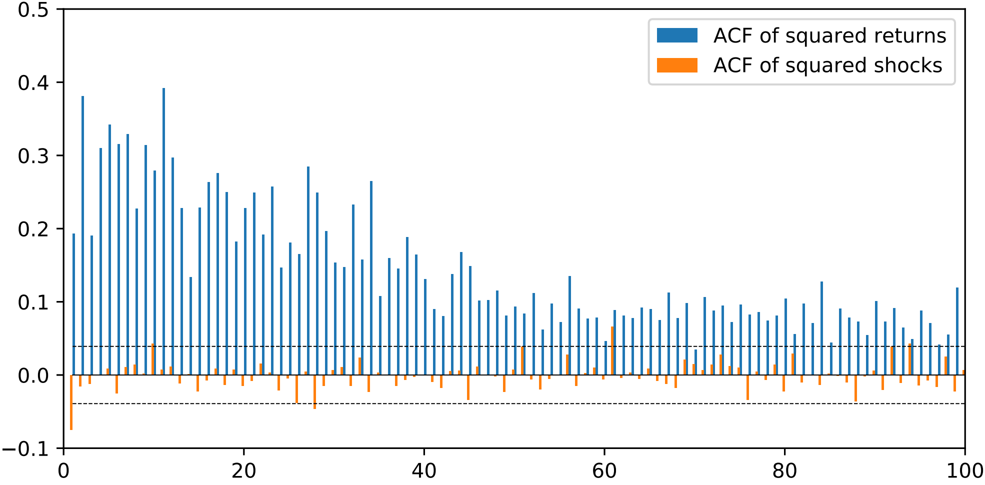
$$LR = 2\left(\ln(L_1) - \ln(L_0)\right) \overset{H_0}{\sim} \chi^2_m,$$

  where m is number of extra parameters in model 1, and L_i is likelihood of model $i = 0, 1$.
- We can also check whether $\text{Var}_t(z_{t+1}) = 1$, as predicted by model, using estimated shocks $\hat{z}_t$:
  - plot of $\hat{z}_t$ against time should not display volatility clustering;
  - autocorrelations of $\hat{z}^2_t$ should not be significantly different from zero.

S&P 500 returns and shocks

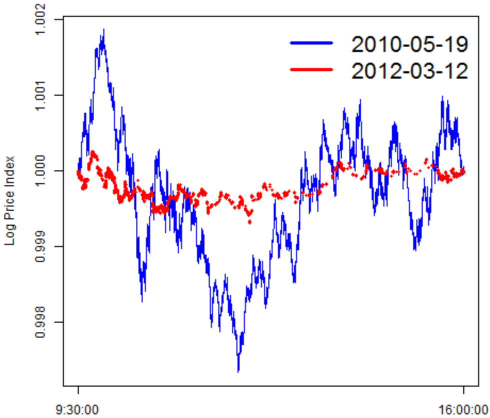


Autocorrelations of squared S&P 500 returns and shocks



High-frequency data

- GARCH & RiskMetrics applied to daily returns: log-difference in closing prices.
- Behaviour of prices during the day should be even more informative about risk than daily return.
- Example: Chart of SPY (S&P 500 tracking ETF) on two days with returns ≈ 0:



Realized variance – motivation

- Suppose we observe price every minute during 24 hours in interval $[t-1, t]$ (realistic for forex).
- This leads to $m = 24 \times 60 = 1440$ log-returns during the day:

$$R_{t,j} = \ln(S_{t-1+j/m}) - \ln(S_{t-1+(j-1)/m}), \qquad j = 1, \ldots, m.$$

  The daily return then satisfies $R_t = \ln(S_t) - \ln(S_{t-1}) = \sum_{j=1}^{m} R_{t,j}$.
- For estimation of mean daily return, these intra-day returns do not help:

$$\hat{\mu} = \frac{1}{T}\sum_{t=1}^{T}\sum_{j=1}^{m} R_{t,j} = \frac{1}{T}\sum_{t=1}^{T} R_t = \frac{1}{T}\left(\ln(S_T) - \ln(S_0)\right).$$

- For estimation of daily return variance, it can be shown that

$$\hat{\sigma}_m^2 = \frac{1}{T}\sum_{t=1}^{T}\sum_{j=1}^{m} R_{t,j}^2$$

  is a factor $1/m$ more precise than $\hat{\sigma}^2 = \frac{1}{T}\sum_{t=1}^{T} R_t^2$.

Realized variance

- Definition (notation slightly different from Christoffersen):

$$RV_t^m = \sum_{j=1}^{m} R_{t,j}^2.$$

- Ex-post estimate of variance during day t; can only be calculated at end of day;
  - do not divide by m: we want estimate of daily return variance, not 1-minute return variance.
- Theoretical result: if log-asset price follows Ito process $d\ln(S_t) = \mu_t dt + \sigma_t dW_t$, then

$$\operatorname*{plim}_{m\to\infty} RV_t^m = \int_{t-1}^{t} \sigma_s^2 ds.$$

  So in the continuous-time limit ($m \to \infty$), we estimate average variance without error.

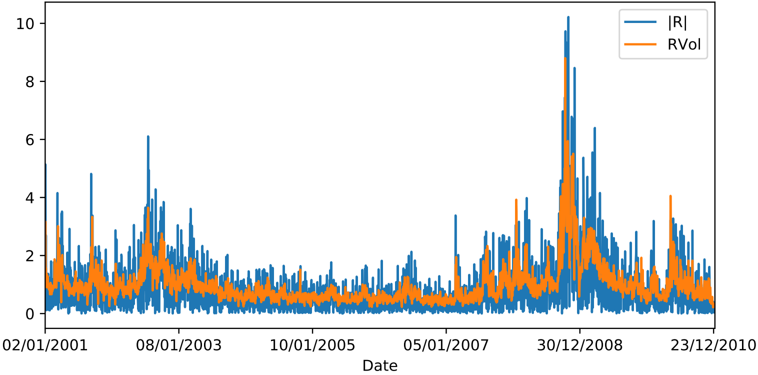
Realized variance in practice

- When taking m very large, so very small time intervals, then RV is affected by (market microstructure) noise.
- Therefore, in practice it is recommended to use 5-minute returns.
- Stock markets are open 9.30–16.00. Then RV_t^m estimates open-to-close return variance;
  - close-to-close return variance requires analysis overnight return $\ln(S_t^{open}) - \ln(S_{t-1}^{close})$.
- Unlike daily financial data, historical intra-day data are not available free of charge.
- Daily RV time series of some common stock indices are available from

$$\texttt{https}://\texttt{realized.oxford}-\texttt{man.ox.ac.uk}/$$

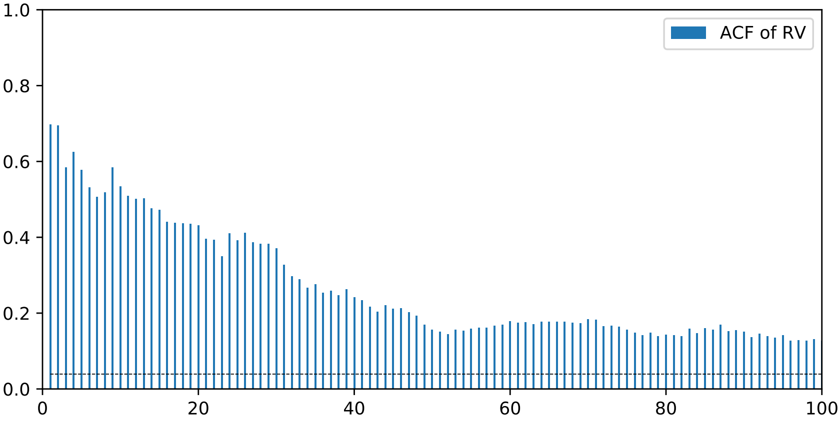
Properties of realized variance

- Realized variance RV_t^m is less noisy measurement than $R_t^2 = RV_t^1$;
  - compare $|R_t| = \sqrt{R_t^2}$ with $RVol_t^m = \sqrt{RV_t^m}$ for S&P 500 index:



Properties of realized variance

- RV is persistent, with positive autocorrelations at long lags ("long memory"):



Volatility forecasting using realized variance

- Realized variance may help in volatility forecasting in two ways:
  1. by using RV_t^m as predictive variable for σ_{t+1}^2 (or RV_{t+1}^m);
  2. by comparing σ_{t+1}^2 from different GARCH models with RV_{t+1}^m.
- For the first part, we may simply extend the GARCH model as

$$\sigma_{t+1}^2 = \omega + \alpha R_t^2 + \beta\sigma_t^2 + \gamma RV_t^m,$$

  which can again be estimated by QMLE;
  - in practice, RV_t^m makes R_t^2 redundant: $\hat{\gamma}$ will be significantly positive, and $\hat{\alpha} \approx 0$.
- For multi-day forecasting $E_t(\sigma_{t+k}^2)$, we need an equation to forecast RV_{t+k-1}^m;
  - Realized GARCH model, see Section 5.3.3.

Volatility forecasting using realized variance

- For forecasting RV_{t+k} directly, models need to replicate the long-memory property.
- In AR(1) model for RV_t or $\ln(RV_t)$, the estimated AR parameter will be close to 1 (random walk).
- Better models are heterogeneous autoregressions (HAR), e.g.

$$RV_{t+1} = \phi_0 + \phi_D RV_{D,t} + \phi_W RV_{W,t} + \phi_M RV_{M,t} + \varepsilon_{t+1},$$

where

$$RV_{D,t} = RV_t, \qquad RV_{W,t} = \frac{1}{5}\sum_{j=0}^{4} RV_{t-j}, \qquad RV_{M,t} = \frac{1}{20}\sum_{j=0}^{19} RV_{t-j}.$$

- Models in logs are preferred; errors are closer to homoskedastic and normally distributed.

Volatility forecast evaluation using realized variance

- Suppose we have estimated a (realized) GARCH model, and wish to evaluate how well the model predicts future variance.
- Assume that model is estimated from sample $\{R_t, t = 1, \ldots, T\}$, and evaluated for out-of-sample period $t = T+1, \ldots, T+n$.
- Measure of forecast quality is mean squared error: traditionally evaluated as

$$\frac{1}{n} \sum_{t=T}^{T+n-1} (R_{t+1}^2 - \sigma_{t+1}^2)^2.$$

- But we know that R_{t+1}^2 is more noisy measurement than RV_{t+1}^m; therefore, we prefer to use

$$MSE = \frac{1}{n} \sum_{t=T}^{T+n-1} (RV_{t+1}^m - \sigma_{t+1}^2)^2.$$

Copyright statement

Figure 4.2 on Slide 8 has been extracted from:

Christoffersen, P. F. (2012), *Elements of Financial Risk Management*, Second Edition. Waltham (MA): Academic Press. ISBN 978-0-12-374448-7.

Back to first slide