Lecture 8
Models of Volatility Dynamics

Lars A. Lochstoer

UCLA Anderson School of Management

Winter 2025

Outline

1. Stylized facts of volatility clustering
2. Realized Variance
3. GARCH models
4. Mixed-frequency Volatility Models

ARMA(p,q)

- so far, we have focused on modeling the conditional mean:

$$r_t = E\left[r_t | r_{t-1}, r_{t-2}, ...; \boldsymbol{\theta}\right] + \varepsilon_t$$

- the benchmark models for the conditional mean are ARMA models

$$r_t = \phi_0 + \phi_1 r_{t-1} + \ldots + \phi_p r_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} + \ldots - \theta_q \varepsilon_{t-q}$$

- given estimates, $\hat{\phi}_0, \hat{\phi}_1, \ldots, \hat{\theta}_{q-1}, \hat{\theta}_q$, we can calculate the residuals $\hat{\varepsilon}_t = v_t(0)$ recursively from the initial condition.

- the key modeling goal is to make sure that the residuals $\{\hat{\varepsilon}_t\}$ are white noise
  - No additional predictable components left

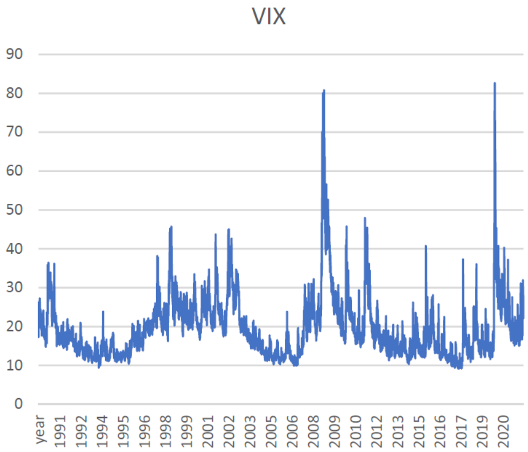
Volatility Clustering

- suppose we have a good model for the conditional mean μ_t (e.g. ARMA(p, q))
- now we look at the squared residuals $\{\hat{\varepsilon}_t^2\}$ to test for conditional heteroskedasticity
- for returns on most financial assets, there is a lot **more autocorrelation** in conditional second moments than in the conditional first moments.
- volatility is predictable!

Volatility?

- In most asset markets, volatility varies dramatically over time.
- episodes of high volatility seem to be clustered.
- we don't simply observe volatility
- how do we measure *vol* (vol is short-hand for volatility)?
  1. implied volatility (back out volatility from option prices)
  2. (non-parametric) realized volatility (e.g., realized volatility of stock returns over one-month using daily data)
  3. (parametric) model volatility

VIX: Option Implied Market Volatility



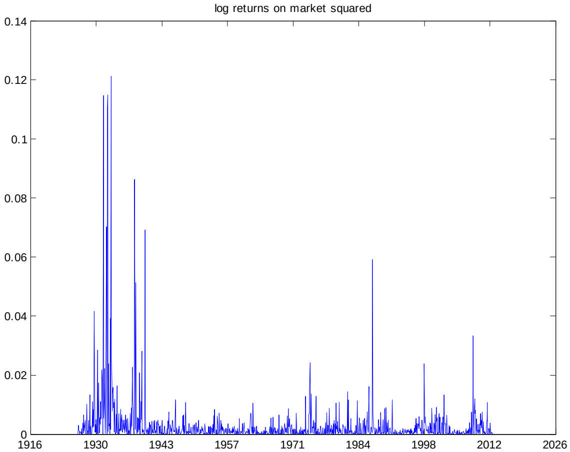
Annualized measure of 30-day vol. VIX, designed to measure the market's expectation of 30-day volatility implied by at-the-money S&P 100 Index (OEX) option prices. Monthly data. 1990-2009. VIX white paper: https://www.cboe.com/micro/vix/vixwhite.pdf.

Why do we care about Volatility?

- why do we care about modeling vol?
  1. portfolio allocation
  2. risk management
  3. option pricing

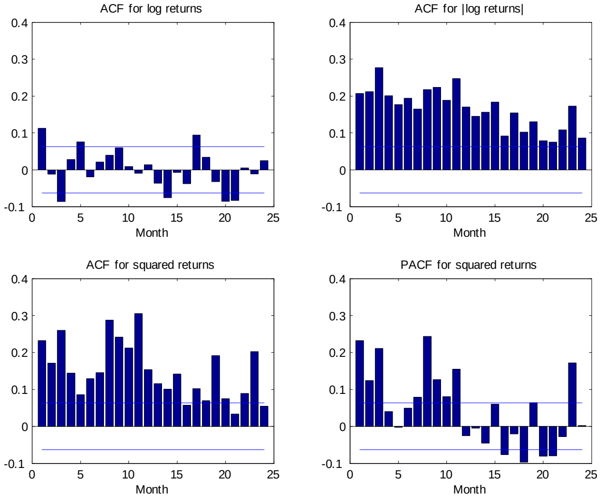
Stylized Facts on Volatility Clustering

Squared Monthly Log Market Returns



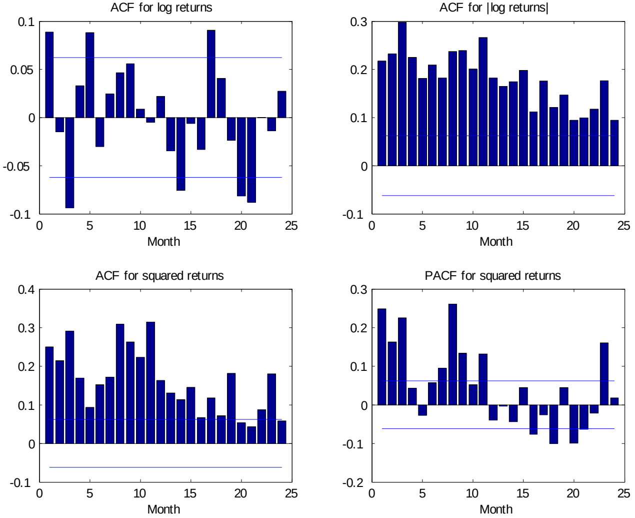
log stock returns on market -squared. Monthly data. 1926-2012.

(P)ACF's of absolute value, squared, and regular market returns; Long sample



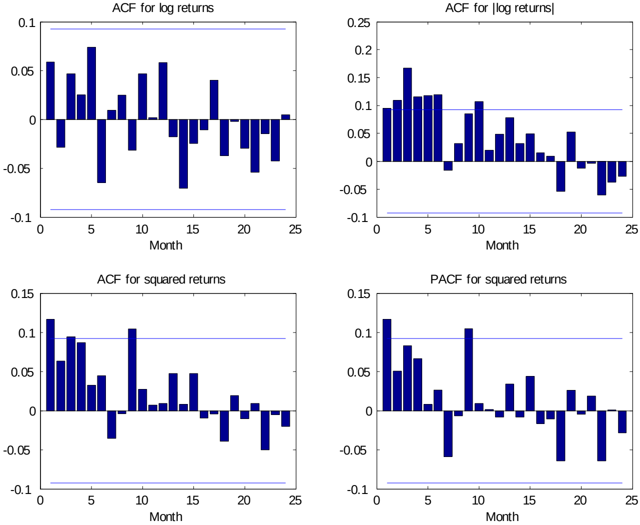
log stock returns on CRSP-VW. Monthly data, 1926-2012.

Dynamics of Variance: Monthly Returns, postwar Sample



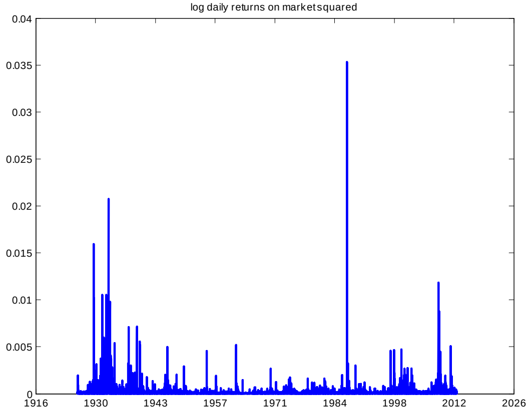
log stock returns on CRSP-VW. Monthly data. 1945-2012.

Dynamics of Variance: Monthly Returns, Short Sample



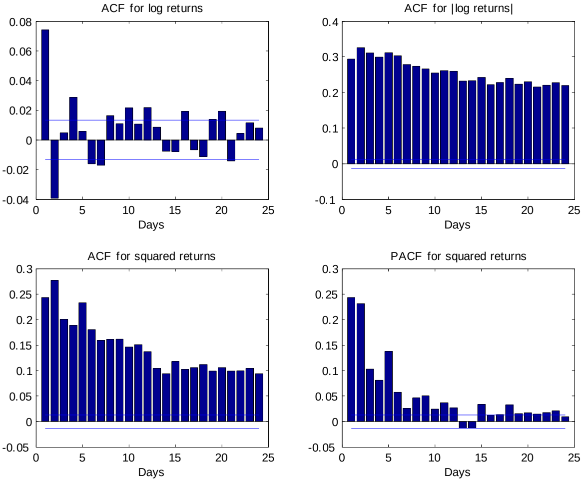
log stock returns on CRSP-VW. Monthly data. 1970-2012.

Daily Market Returns Squared



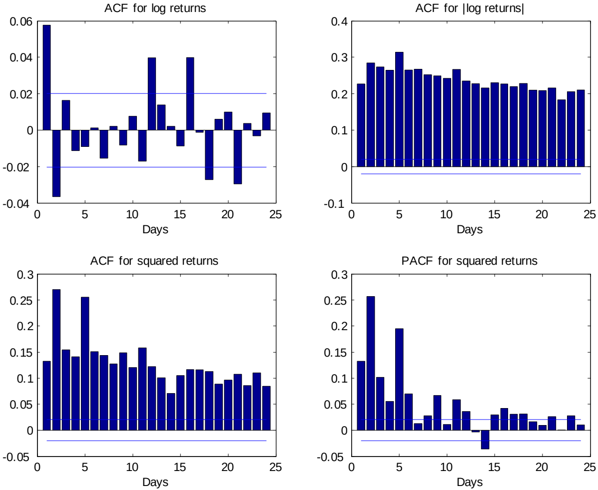
log stock returns on market -squared. Daily data. 1988-2012.

Dynamics of Variance: Daily Returns, Long Sample



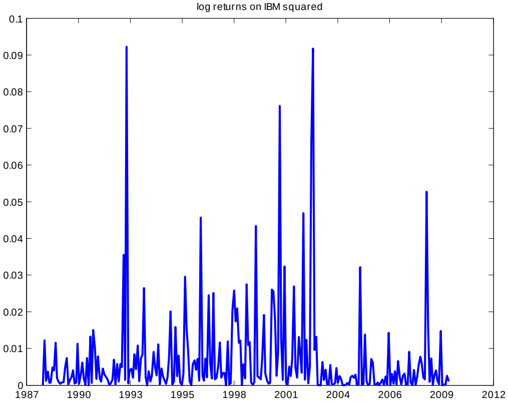
log stock returns on CRSP-VW. Daily data. 1926-2012.

Dynamics of Variance: Daily Returns, Short Sample



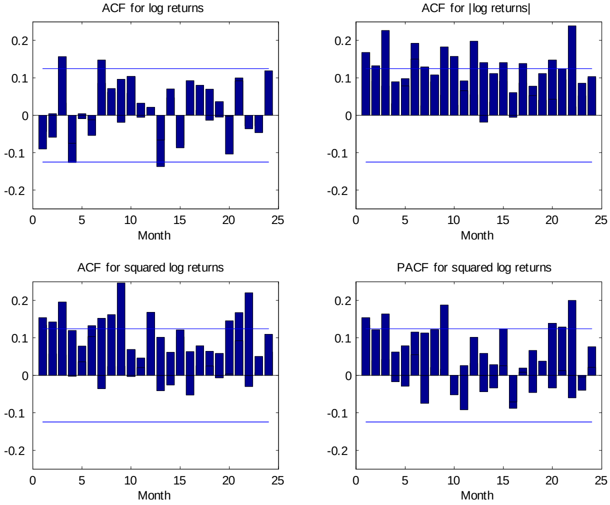
log stock returns on CRSP-VW. Daily data. 1970-2012.

IBM



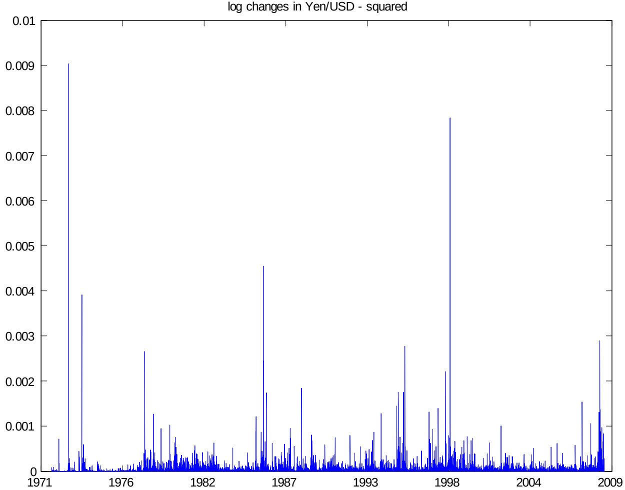
log stock returns on IBM -squared. Monthly data. 1988-2009.

IBM



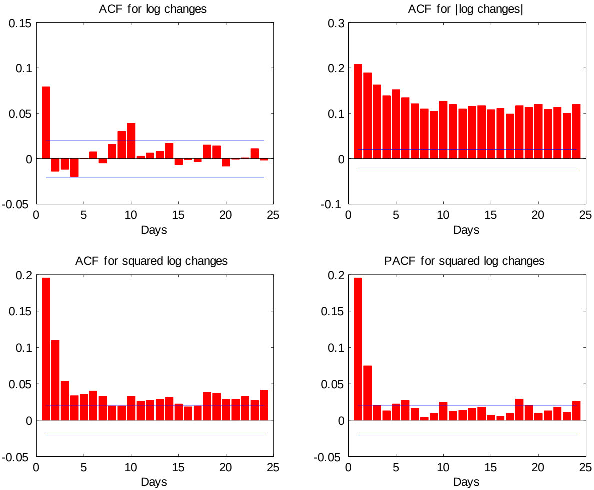
log stock returns on IBM. Monthly data. 1988-2009.

Yen/USD Squared Log Changes: Daily Data



log changes in Yen/USD squared. Daily data. 1971-2009.

Yen/USD (P)ACF's



log changes in Yen/USD. Daily data. 1971-2009.

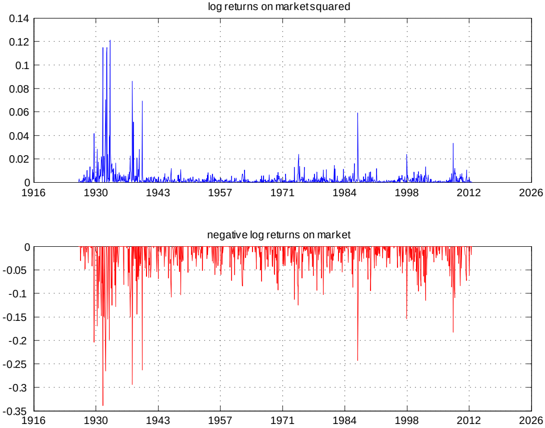
Testing for ARCH

- we can test for autoregressive conditional heteroscedasticity using a **Ljung-Box** Q-test:
- suppose the model for the conditional mean is ARMA(p, q)
- estimate the model
- run a Q-test on the estimated squared residuals $\{\hat{\varepsilon}_t^2\}$
  - test the null that:

$$H_0 : \rho_1 = \rho_2 = \ldots = \rho_m = 0$$

    where ρ_i denotes the autocorrelation of the *squared* residuals
  - if you cannot reject the null, no need for ARCH machinery

Leverage Effects: Monthly Returns on Market Squared



log stock returns on market -squared. Monthly data. 1926-2012.

Volatility Summary

- volatility clusters
- volatility is stationary
- leverage effect (asymmetry)
  - negative returns seem to be followed by larger increases in volatility than positive returns
  - first emphasized by Black (1976)
  - Black (1976) suggested that negative returns (decreases in price) change a company's debt/equity ratio, increasing their leverage.

Realized Variance

High-Frequency Data

Let r_t denote log of gross returns over a period of time t

- t may denote 1 day, 1 week, 1 month.

Suppose over the time t, we observe log-returns $\{r_{t,i}\}_{i=1}^{N_t}$ at a higher frequency.

- Example: t is monthly and i are days in the month

$$r_t = \sum_{i=1}^{N_t} r_{t,i}$$

for log returns, the conditional variance of the return over time t is given by:

$$V\left(r_t \mid F_{t-1}\right) = \sum_{i=1}^{N_t} V\left(r_{t,i} \mid F_{t-1}\right) + 2\sum_{i<j} \text{Cov}\left(r_{t,i}, r_{t,j} \middle| F_{t-1}\right)$$

Realized Variance (Volatility)

Realized variance in month t using daily frequency within-month data i is typically:

$$RV_t = \sum_{i=1}^{n} r_{t,i}^2$$

- In practice, t is often 1 day or 1 month.

Notice two implicit assumptions:

1. The conditional mean of high-frequency returns (e.g., daily) is assumed to be equal to zero
2. High-frequency returns are assumed to be uncorrelated across time (no covariance terms)

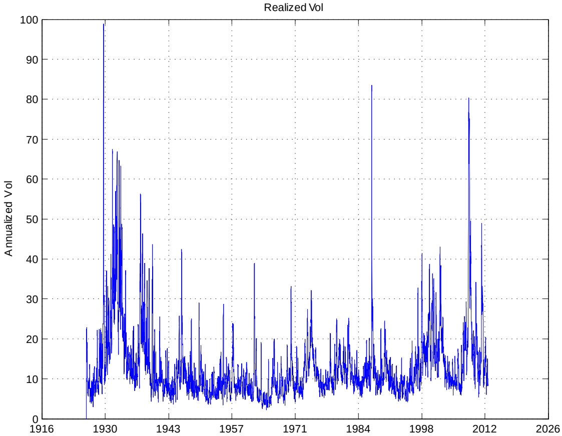
These assumptions are pretty good for financial market returns

- Estimating sample mean is noisy so may actually worsen variance estimate in MSE sense
- Markets are competitive so autocorrelations small, again estimating these may bring in more noise than it is worth

Realized Volatility

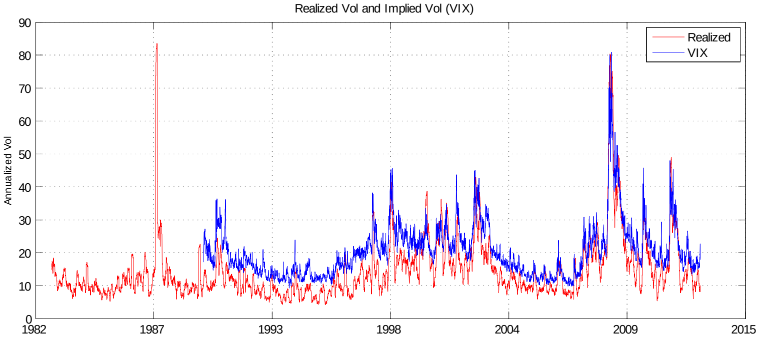
Realized Volatility: square root of $(Periods\_of\_t\_in\_year \times RV)$

- Annualized measure: monthly RV means Realized Vol $= \sqrt{12 \times RV}$



Annualized (Monthly) Realized Vol. Daily data. 1926-2012.

Realized Volatility and Implied Vol



Annualized (Monthly) Realized Vol and the VIX. Daily data. 1983-2012.

GARCH models

ARCH(1)

Definition

The ARCH(1) process is given by:

$$\varepsilon_t = \sigma_t \eta_t, \qquad \sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2$$

where η_t is i.i.d., mean zero, has variance of one , $\alpha_0 > 0$ and $\alpha_i \geq 0$ for $i > 0$

Notice: this is an AR(1) in ε_t^2

- See next slide

ARCH(1) as AR(1)

First, note that generally:

$$E_{t-1}\left(\varepsilon_t^2\right) = \sigma_t^2.$$

- I would have preferred to call this σ_{t-1}^2 as it is the conditional variance of ε_t, which is known at time t
- But, the GARCH literature decided to call it σ_t^2, so we will be consistent with this literature when discussing these models

AR(1) in ε_t^2:

$$\varepsilon_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \nu_t,$$

where ν_t is a mean-zero error term (it's not Normal as ε^2 can't go negative).

- Expected variance is

$$\begin{aligned} E_{t-1}\left(\varepsilon_t^2\right) &= \sigma_t^2 \\ &= \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 \end{aligned}$$

Unconditional moments

The unconditional mean, variance, and autocorrelations are found the same way as for the AR model

- So, not point in going through it again!
  - e.g., we need $|\alpha_1| < 1$ for stationarity
  - Forecasting works the same way as for an AR(1)

- But, additional requirement: positive expected variance
  - So, we also need $\alpha_0, \alpha_1 > 0$

Fourth Moment: fat tails through heteroskedasticity!

- the unconditional fourth moment is :

$$E\left[\varepsilon_t^4\right] = E(E\left[\varepsilon_t^4 | F_{t-1}\right]) = E\left(3\sigma_t^4\right) = 3\left[E(\alpha_0 + \alpha_1 \varepsilon_{t-1}^2)\right]^2$$

- using stationarity, this implies

$$m_4 = \frac{3\alpha_0^2(1+\alpha_1)}{(1-\alpha_1)(1-3\alpha_1^2)}$$

which means $0 \leq \alpha_1^2 < 1/3$

- the unconditional kurtosis is given by:

$$\frac{E\left[\varepsilon_t^4\right]}{V\left[\varepsilon_t\right]^2} = 3\frac{1-\alpha_1^2}{1-3\alpha_1^2} > 3$$

- **positive kurtosis** even though innovations are conditionally Gaussian

Normality

The residual ε_t is conditionall Normal if $\eta_t \sim$N$(0,1)$:

$$\varepsilon_t = \sigma_t \eta_t.$$

- ARCH(1) process is not unconditionally normal
  - time variation in the variance generates fat tails unconditionally

Building an ARCH model

- we can pick the order of an ARCH model by looking at *PACF* at $\{\varepsilon_t^2\}$; you might need a large number of lags
- for a well-specified ARCH model, the standardized residual:

$$\hat{\eta}_t = \frac{\hat{\varepsilon}_t}{\hat{\sigma}_t}$$

  should be white noise
- this can be tested by examining $\hat{\eta}_t$
  1. check the volatility spec by doing a Q-test on $\hat{\eta}_t^2$
  2. check the mean spec by doing a Q-test on $\hat{\eta}_t$

Weakness of ARCH Models

1. symmetry in effects of positive and negative shocks on vol
2. ARCH model: mechanical description of volatility (no economics)
3. sometimes many lags are needed to describe vol dynamics

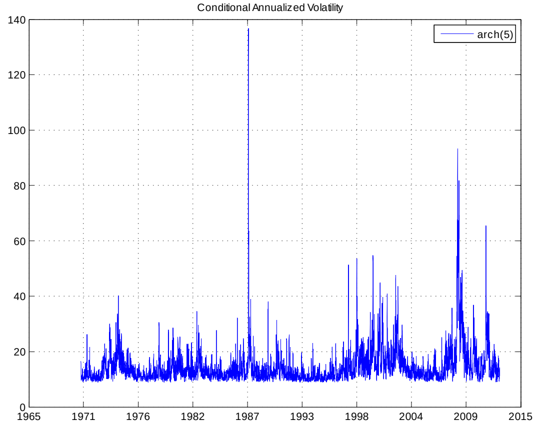
ARCH(5)

Table: ARCH(5)

Parameter	Value	s.e.	t-stat
Constant	3.09E-05	6.65E-07	46.472
ARCH1	0.099307	0.005071	19.5838
ARCH2	0.144208	0.008707	16.5623
ARCH3	0.141901	0.009809	14.4668
ARCH4	0.168116	0.01012	16.6119
ARCH5	0.150002	0.010459	14.3424

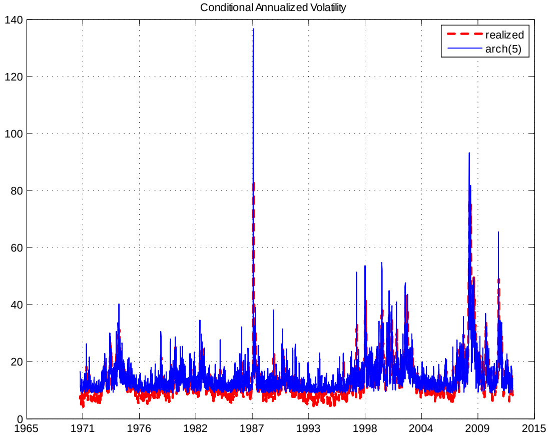
ARCH(5) estimated on daily CRSP-VW stock returns 1971-2012.

Parametric Volatility σ_t



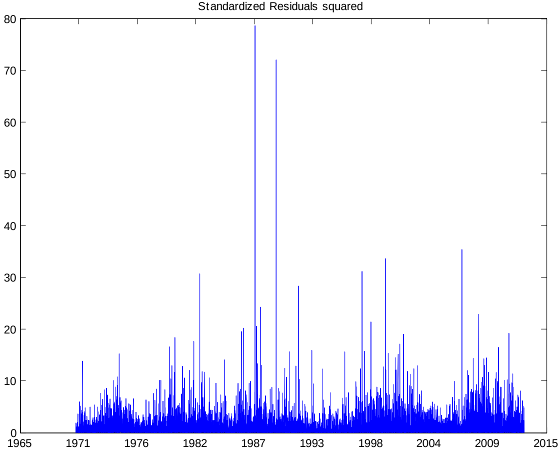
ARCH(5). Daily data. CRSP-VW.1971-2012.

Parametric Volatility against Realized Vol.



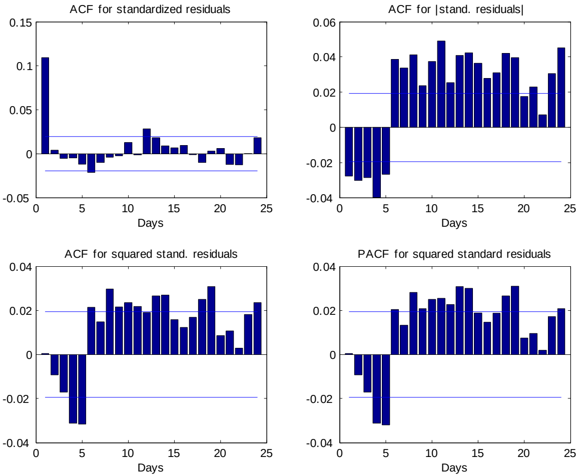
ARCH(5). Daily data. CRSP-VW. 1971-2012.

Standardized Residuals $(\varepsilon_t / \sigma_t)^2$



ARCH(5). Daily data. CRSP-VW.1971-2012.

Standardized Residuals ε_t / σ_t



ARCH(5). Daily data. CRSP-VW.1971-2012.

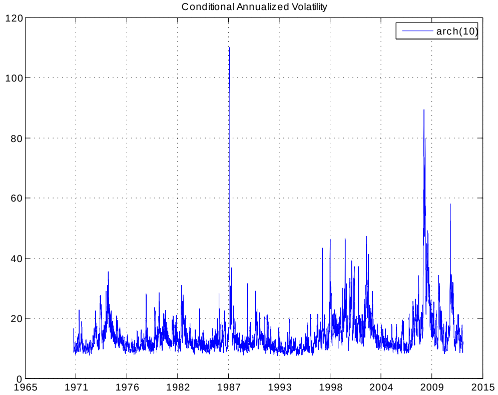
ARCH(10)

Table: ARCH(10)

Parameter	Value	standard error	t-stat
Constant	2.04E-05	7.56E-07	26.927
ARCH1	0.083715	0.004496	18.6207
ARCH2	0.113967	0.008698	13.1022
ARCH3	0.080502	0.009147	8.80109
ARCH4	0.095266	0.00937	10.1667
ARCH5	0.107659	0.011581	9.29659
ARCH6	0.075962	0.009292	8.1747
ARCH7	0.052205	0.008305	6.28631
ARCH8	0.083885	0.00877	9.56551
ARCH9	0.073778	0.008694	8.4864
ARCH10	0.049606	0.008884	5.58371

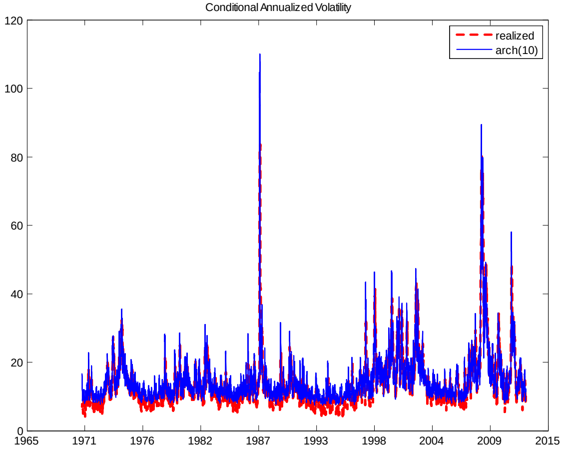
ARCH(10) estimated on daily CRSP-VW stock returns 1971-2012.

Parametric Volatility σ_t



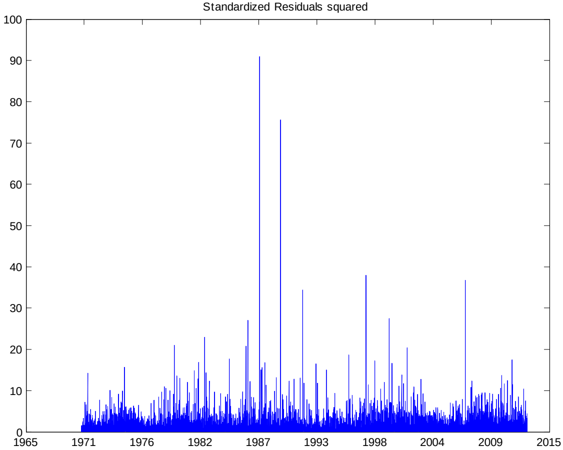
ARCH(10). Daily data. CRSP-VW.1971-2012.

Parametric Volatility against Realized Variance.



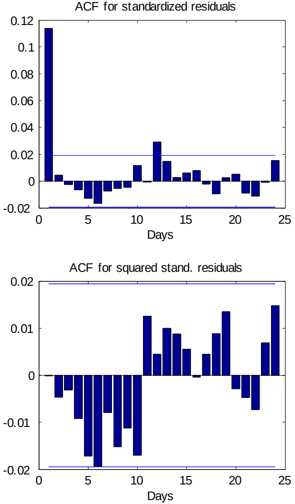
ARCH(10). Daily data. CRSP-VW.1971-2012.

Standardized Residuals $(\varepsilon_t/\sigma_t)^2$



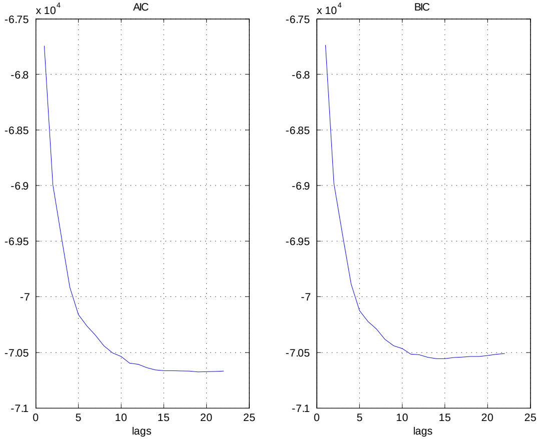
ARCH(10). Daily data. CRSP-VW.1971-2012.

Standardized Residuals ε_t/σ_t



ARCH(10). Daily data. CRSP-VW.1971-2012.

AIC/BIC



ARCH(lags). Daily data. CRSP-VW.1971-2012.

GARCH models

GARCH model of Bollerslev (1986)

Definition

Consider $\varepsilon_t = r_t - \mu_t$. The ε_t follows a GARCH(m, s) model if

$$\varepsilon_t = \sigma_t \eta_t, \qquad \sigma_t^2 = \alpha_0 + \sum_{i=1}^{m} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2$$

where η_t is i.i.d., mean zero, has variance of one , $\alpha_0 > 0$ and $\alpha_i \geq 0$ for $i > 0$ and $\beta_j \geq 0$ for $j > 0$, and

$$\sum_{i=1}^{max(m,s)} (\alpha_i + \beta_i) < 1$$

GARCH is ARMA for vol

- define:

$$\kappa_t = \varepsilon_t^2 - \sigma_t^2$$

- hence

$$\sigma_t^2 = \varepsilon_t^2 - \kappa_t$$

- this delivers an ARMA-like representation:

$$\varepsilon_t^2 = \alpha_0 + \sum_{i=1}^{max(m,s)} (\alpha_i + \beta_i)\varepsilon_{t-i}^2 - \sum_{j=1}^{s} \beta_j \kappa_{t-j} + \kappa_t$$

where $E(\kappa_t) = 0$ and $E(\kappa_t \kappa_{t-j}) = 0$ but is not white noise

- the unconditional variance is simply:

$$E\left(\varepsilon_t^2\right) = \frac{\alpha_0}{1 - \sum_{i=1}^{max(m,s)} (\alpha_i + \beta_i)}$$

GARCH(1,1) of Bollerslev (1986)

- consider GARCH(1,1) model:

$$\begin{aligned} \varepsilon_t &= \sigma_t \eta_t \qquad\qquad \eta_t \sim \mathrm{N}(0,1) \\ \sigma_{t+1}^2 &= \alpha_0 + \alpha_1 \varepsilon_t^2 + \beta_1 \sigma_t^2 \end{aligned}$$

  - large $\varepsilon_t^2, \sigma_t^2$ leads to large σ_{t+1}^2

- the unconditional kurtosis is given by:

$$\frac{E\left[\varepsilon_t^4\right]}{V\left[\varepsilon_t\right]^2} = 3\frac{1-(\alpha_1+\beta_1)^2}{1-(\alpha_1+\beta_1)^2-2\alpha_1^2} > 3$$

- positive kurtosis even though innovations η_t are Gaussian

- Forecasting, unconditional mean, variance, and autocorrelations same procedure as for ARMA(1,1) in ε_t^2

Picking Order and Checking Model

- for a well-specified GARCH model, the standardized residual:

$$\hat{\eta}_t = \frac{\hat{\varepsilon}_t}{\hat{\sigma}_t}$$

  should be white noise

- this can be tested by examining $\hat{\eta}_t$
  1. check the volatility specification by doing a Q-test on $\hat{\eta}_t^2$
  2. check the mean spec. by doing a Q-test on $\hat{\eta}_t$

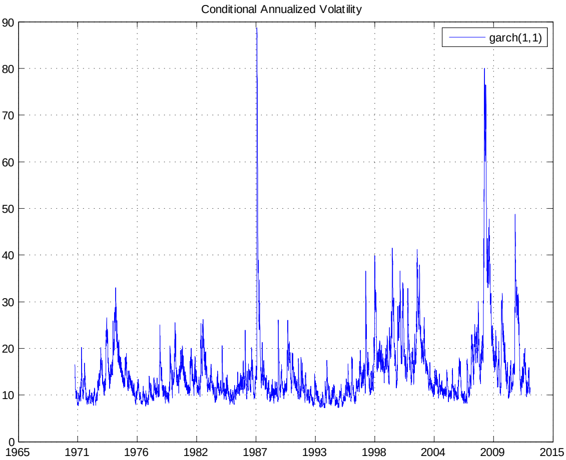
GARCH(1,1)

Table: GARCH(1,1)

Parameter	Value	standard error	t-stat
Constant	1.16E-06	2.23E-07	5.18121
GARCH1	0.909294	0.002906	312.933
ARCH1	0.07998	0.001725	46.3601

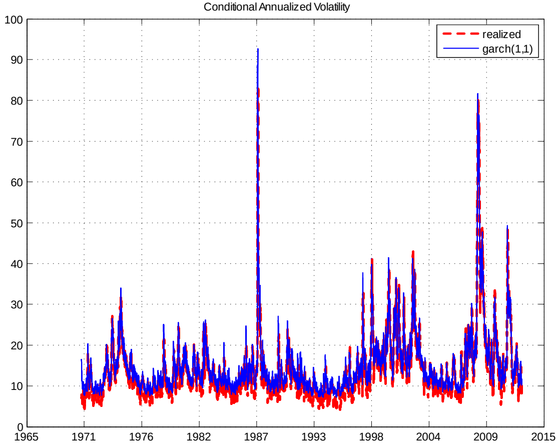
GARCH(1,1). Daily data. CRSP-VW.1971-2012.

Parametric Volatility σ_t



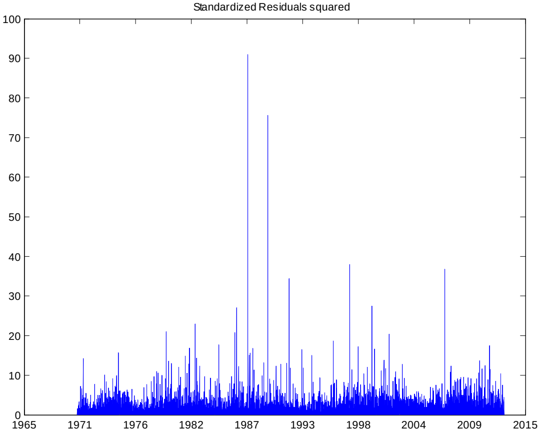
GARCH(1,1). Daily data. CRSP-VW.1971-2012.

Parametric Volatility against Realized Vol.



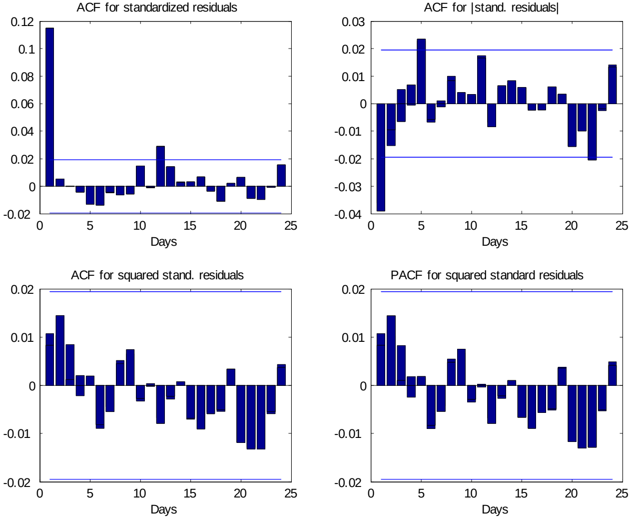
GARCH(1,1). Daily data. CRSP-VW.1971-2012.

Standardized Residuals $(\varepsilon_t/\sigma_t)^2$



GARCH(1,1). Daily data. CRSP-VW.1971-2012.

Standardized Residuals ε_t / σ_t



GARCH(1,1). Daily data. CRSP-VW.1971-2012.

AIC/BIC

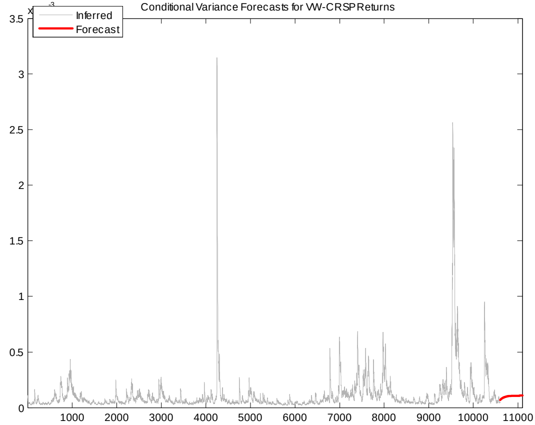
- AIC/BIC for GARCH(1,1)

$$\begin{aligned} aic &= -7.0738e+04 \\ bic &= -7.0665e+04 \end{aligned}$$

- GARCH(1,1) has smaller AIC and BIC than ARCH(10)

- you could use the BIC/AIC criterion to find the best GARCH(p,q)

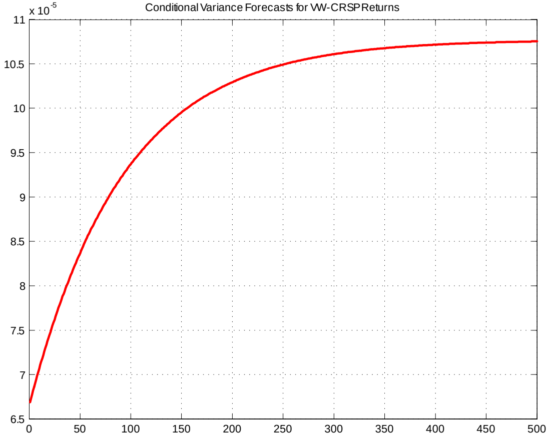
Predicted one-period ahead variance from GARCH(1,1)



GARCH(1,1). Daily data. CRSP-VW.1971-2012.

Prediction variance

- x-axis gives number of periods ahead forecast is



GARCH(1,1). Daily data. CRSP-VW.1971-2012.

I-GARCH, EGARCH, GARCH-M, GJR models

(lots more GARCH models exist)

I-GARCH

Definition

Consider $\varepsilon_t = r_t - \mu_t$. The ε_t follows a *I*-GARCH$(1,1)$ model if

$$\varepsilon_t = \sigma_t \eta_t, \qquad \sigma_t^2 = \alpha_0 + (1 - \beta_1)\varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

where η_t is i.i.d., mean zero, has variance of one. $0 < \beta < 1$

- the unconditional variance is not defined!
- hard to justify for returns

Forecasting with I-GARCH(1,1)

- in general for GARCH(1,1):

$$\sigma_t^2(h) = \alpha_0 + (\alpha_1 + \beta_1)[\sigma_t^2(h-1)]$$

- now set $\alpha_1 + \beta_1 = 1$
- repeated substitution yields:

$$\sigma_t^2(h) = (h-1)\alpha_0 + [\sigma_t^2(1)]$$

  - effects on future vols are persistent!

GARCH(1,1)-M

Definition

Consider

$$\varepsilon_t = r_t - \mu_t - \gamma\sigma_t^2.$$

The ε_t follows a GARCH$(1,1)$-M model if

$$\varepsilon_t = \sigma_t\eta_t, \qquad \sigma_t^2 = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \beta_1\sigma_{t-1}^2$$

where η_t is i.i.d., mean zero, has variance of one. $0 < \beta < 1$.

M stands for GARCH in the mean.

- the parameter γ measures a variance risk premium
- the risk premium increases when volatility increases

E-GARCH model of Nelson (1991)

Definition

Consider $\varepsilon_t = r_t - \mu_t$. The ε_t follows a E-GARCH$(1,1)$ model if

$$\varepsilon_t = \sigma_t \eta_t, \qquad \ln \sigma_t^2 = (1-\alpha)\alpha_0 + \alpha \ln \sigma_{t-1}^2 + g(\eta_{t-1})$$

where η_t is i.i.d., mean zero, has variance of one and the shock-term is

$$g(\eta_t) = \delta \eta_t + \gamma \left(|\eta_t| - E(|\eta_t|) \right)$$

Built-in asymmetry to capture leverage effect

- we expect $\delta < 0$

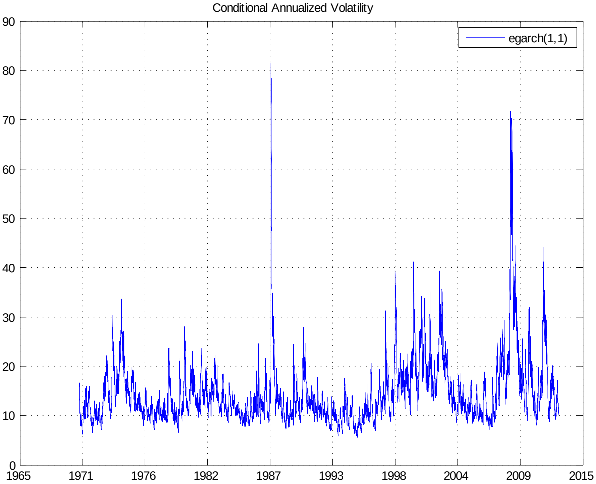
E-GARCH(1,1)

Table: E-GARCH(1,1)

Parameter	Value	standard error	t-stat
———	———	———	———
Constant	-0.17256	0.011062	-15.599
GARCH1	0.980926	0.001135	864.546
ARCH1	0.13937	0.005217	26.7121
Leverage1	-0.07601	0.003161	-24.0473

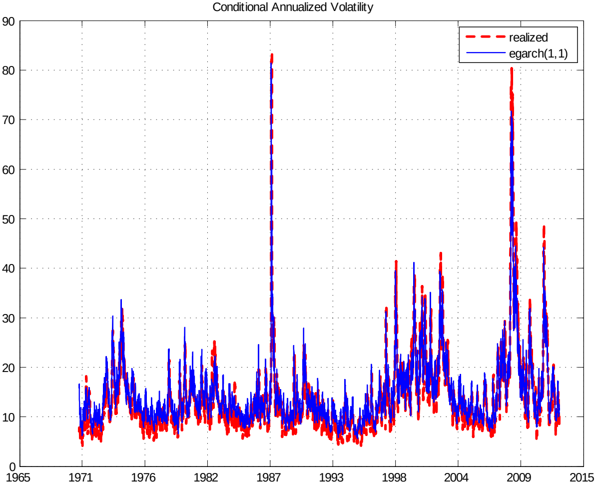
EGARCH(1,1). Daily data. CRSP-VW.1971-2012.

Parametric Volatility σ_t



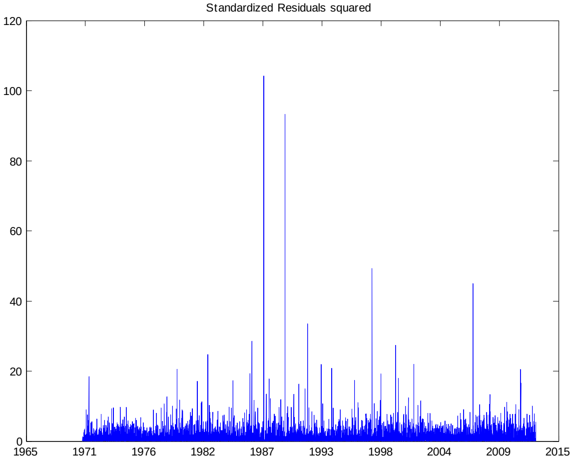
EGARCH(1,1). Daily data. CRSP-VW.1971-2012.

Parametric Volatility against Realized Vol.



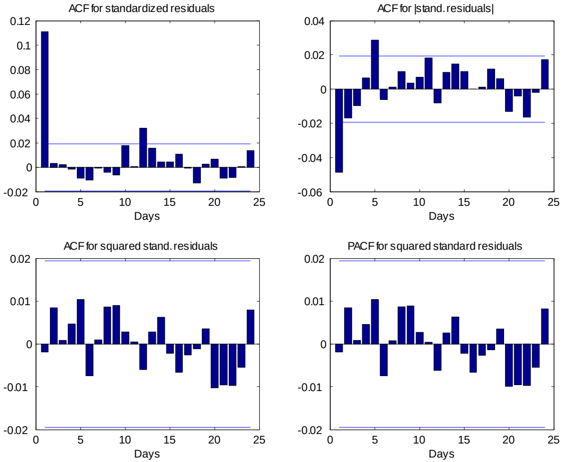
EGARCH(1,1). Daily data. CRSP-VW.1971-2012.

Standardized Residuals $(\varepsilon_t/\sigma_t)^2$



EGARCH(1,1). Daily data. CRSP-VW.1971-2012.

Standardized Residuals ε_t / σ_t



EGARCH(1,1). Daily data. CRSP-VW.1971-2012.

Glosten, Jagannathan, and Runkel (1993)

Definition

Consider $\varepsilon_t = r_t - \mu_t$. The ε_t follows a GJR-GARCH$(1,1)$ model if

$$\varepsilon_t = \sigma_t \eta_t, \qquad \sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 + I_{t-1}\delta$$

where η_t is i.i.d., mean zero, has variance of one.

$$I_t = \begin{cases} 0 & \text{if} \quad \varepsilon_t \geq 0 \\ 1 & \text{if} \quad \varepsilon_t < 0 \end{cases}$$

- simple way to capture leverage effect
- we expect $\delta > 0$

Multi-frequency Volatility Models

Multi-frequency and Realized Variance

As you vividly recall from our earlier discussions and Stochastic Calculus:

- By sampling more frequently, the variance estimate gets more precise
- In the continuous-time limit you effectively observe conditional variance as there is no estimation error

We apply this lesson to our volatility models and make use of the high frequency return data that is available in many markets!

- Chen and Ghysels (2012): models using high-frequency realized variance dominate the standard GARCH class of models
- (This doesn't mean learning GARCH is useless, these are extensions of the same ARMA-type ideas as we shall see)
- (also, sometimes you don't have high-frequency data)

The HAR-RV model

Corsi (2009) proposes the Heterogeneous Autoregressive model of Realized Volatility (HAR-RV)

- Daily realized return variance is given by:

$$RV_{t+1d}^{(d)} = c + \beta^{(d)} RV_t^{(d)} + \beta^{(w)} RV_t^{(w)} + \beta^{(m)} RV_t^{(m)} + \omega_{t+1d}.$$

- $RV_t^{(d)}$ is the daily realized variance based on intraday return observations (e.g., 5 minute returns,or maybe 2-hour returns if not extremely liquid) for the most recent day at time t
- $RV^{(w)}$ and $RV^{(m)}$ are the weekly and monthly realized variances for the most recent week and month at time t
- ω_{t+1d} is the error term

Bekaert and Hoerova (2014) extends this model to include jump-sensitive components as well as information from option markets (the VIX in their case)

High-frequency RV and ARMA models

In general, this style of modeling takes us back to special cases of the ARMA, ARMA-X or VAR frameworks

- ARMA-X is an ARMA that includes exogenous regressors X in the forecasting regression (an $X\beta$ term)
- Notice that the HAR-RV is actually a special case of an AR process in, e.g., 5-minute variance

Using information from option prices is very sensible as markets are forward-looking

- But, markets can make mistakes and/or there is a variance risk premium so market forecast likely not optimal (see, Lochstoer and Muir, 2022)

References

Black, F. (1976). Studies of stock market volatility changes. In Proceedings of the American Statistical Association, Business and Economic Statistics Section, Volume 177, pp. 181.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 31(3), 307-327.

Bekaert and Hoerova (2012). The VIX, the variance premium, and stock market volatility. Journal of Econometrics 183, 181–192.

Corsi, F. (2009). A simple approximate long-memory model of realized volatility. Journal of Financial Econometrics 7, 174–196.

Chen and Ghysels (2012). News – good or bad – and its impact on volatility predictions over multiple horizons. Review of Financial Studies 24(1), 46-81.

Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of the United Kingdom inflation. Econometrica 50(4), 987-1007.

Glosten, L. R., R. Jagannathan, and D. E. Runkle (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. The Journal of Finance 48(5).

Lochstoer, L and T. Muir (2022). Volatility Expectations and Returns. Journal of Finance 77.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: a new approach. Econometrica 59(2), 347-370.