

Financial Econometrics

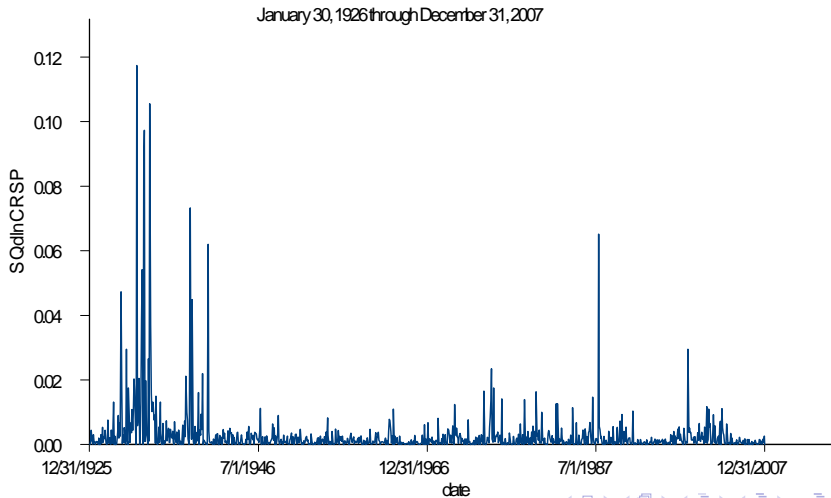
Volatility

Gerald P. Dwyer, Jr.

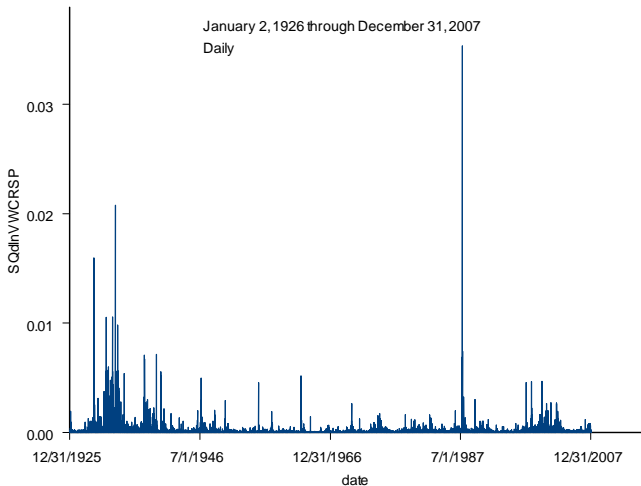
Trinity College, Dublin

February 2009

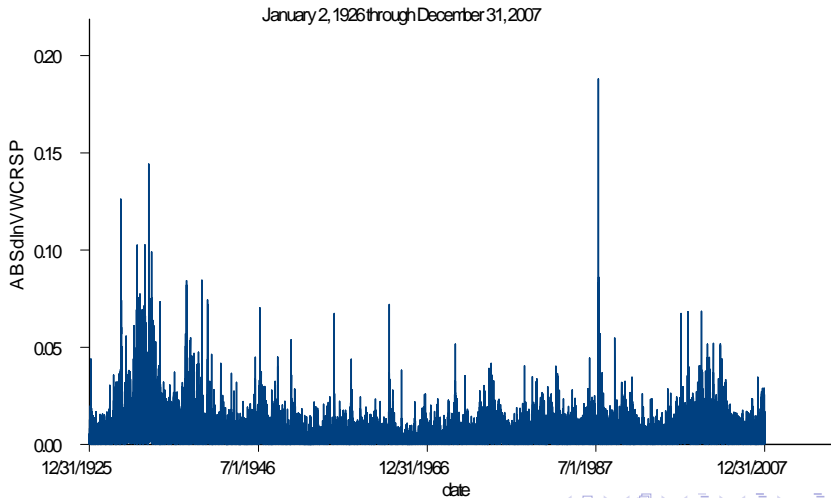
Squared changes in log returns for CRSP



Squared changes in log returns for CRSP daily



Absolute value of log returns for CRSP daily



Heteroskedasticity over time

- These graphs suggest heteroskedasticity over time
 - Time-varying volatility of returns
 - Of interest in itself to characterize returns
 - Matters for prices of options and some other financial instruments
 - Volatility clustering
- These graphs are suggestive but don't tell us too much
 - Using individual observations on squared changes and absolute value to estimate variance and standard deviation as it changes
 - Similar to using each individual observations to estimate means over time

Serial correlation of change in logarithm of value-weighted CRSP index

- [acf_dlnrcrsp.pdf](#)

Serial correlation of absolute value of change in logarithm of value-weighted CRSP index

- `abs_dlncrsp.pdf`

Autoregressive conditional heteroskedasticity (ARCH)

is intended to deal with this

- Simple ARCH model

$$\begin{aligned}r_t &= \mu_t + a_t & \mathbb{E} a_t &= 0, \mathbb{E} a_t^2 = \sigma_t^2 \\ \sigma_t^2 &= \alpha_0 + \alpha_1 a_{t-1}^2\end{aligned}$$

where σ_t^2 is the variance of a_t conditional on past values of the squared innovations, a_{t-1}^2

- $r_t = \mu_t + a_t$ is the mean equation for r_t
- $\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2$ is the volatility equation for r_t
- a_t is the innovation in r_t
- A slightly different version of the same equations is

$$\begin{aligned}r_t &= \mu_t + \sigma_t \varepsilon_t & \mathbb{E} \varepsilon_t &= 0, \mathbb{E} \varepsilon_t^2 = 1 \\ \sigma_t^2 &= \alpha_0 + \alpha_1 a_{t-1}^2\end{aligned}$$

where $a_t = \sigma_t \varepsilon_t$

Estimating an ARCH model

$$\begin{aligned}r_t &= \mu_t + a_t & \mathbb{E} a_t &= 0, \mathbb{E} a_t^2 = \sigma_t^2 \\ \sigma_t^2 &= \alpha_0 + \alpha_1 a_{t-1}^2\end{aligned}$$

- Steps in estimating an ARCH model
 - 1 Estimate a model for the mean equation
 - 2 Use the residuals of the mean equation to test for ARCH effects
 - 3 Specify a volatility model with ARCH effects if it seems warranted
 - 4 Check the fitted model and refine as suggested by diagnostic statistics

Mean equation

- In general, there is no reason the mean equation can't be as complicated as we like

$$r_t = \mu_t + a_t$$

- r_t can be a complicated ARMA(p,q)
 - Generally need not worry about unit roots in returns
- May be mis-specified if ignore conditional heteroskedasticity of a_t

- Simple model

$$\begin{aligned}r_t &= \mu_t + \sigma_t \varepsilon_t & \mathbb{E} \varepsilon_t &= 0, \mathbb{E} \varepsilon_t^2 = 1 \\ \sigma_t^2 &= \alpha_0 + \alpha_1 a_{t-1}^2\end{aligned}$$

- Let a_t be the actual residuals from the mean equation so

$$a_t = r_t - \mu_t$$

Two tests for ARCH

- Box-Ljung test applied to squared residuals, a_t^2 , for some pre-specified number of lags k
- Engle test based on a regression for the squared residuals

$$a_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \alpha_2 a_{t-2}^2 + \dots + \alpha_k a_{t-k}^2 + e_t$$

where e_t is the error term in the regression

- Test whether $\alpha_1 = \alpha_2 = \dots = \alpha_k$ using an F-ratio

$$F = \frac{(SSR_0 - SSR_1) / k}{SSR_1 / (T - 2k - 1)}$$

where $SSR_0 = \sum_{t=1}^T (a_t^2 - \bar{a}^2) / T$, \bar{a}^2 is the mean of a_t^2 , and $SSR_1 = \sum_{t=k+1}^T e_t^2 / (T - k + 1)$ and $F \sim^A \chi_k^2$

Properties of ARCH models

Mean

- Simple model

$$\begin{aligned}r_t &= \sigma_t \varepsilon_t \quad \mathbb{E} \varepsilon_t = 0, \mathbb{E} \varepsilon_t^2 = 1 \\ \sigma_t^2 &= \alpha_0 + \alpha_1 a_{t-1}^2\end{aligned}$$

where $\alpha_0 > 0$ and $\alpha_1 \geq 0$. Why?

- Let F_{t-1} denote the set of all information available in $t - 1$ and earlier, especially $r_{t-1}, r_{t-2}, \dots, a_{t-1}, a_{t-2}, \dots$

$$\begin{aligned}\mathbb{E}[a_t] &= \mathbb{E}[\mathbb{E}(a_t | F_{t-1})] \quad (\text{application of law of iterated expectations}) \\ &= \mathbb{E}[\mathbb{E}(\sigma_t \varepsilon_t | F_{t-1})] \\ &= \mathbb{E}[\mathbb{E}(\sigma_t | F_{t-1}) \mathbb{E}(\varepsilon_t | F_{t-1})] \\ &= \mathbb{E}[\mathbb{E}(\sigma_t | F_{t-1}) \cdot 0] \\ &= 0\end{aligned}$$

Properties of ARCH models

Variance

- Simple model

$$r_t = \sigma_t \varepsilon_t \quad \mathbb{E} \varepsilon_t = 0, \mathbb{E} \varepsilon_t^2 = 1$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2$$

$$\alpha_0 > 0, \alpha_1 \geq 0$$

$$\mathbb{E}[a_t] = 0$$

$$\begin{aligned} \text{Var}[a_t] &= \mathbb{E}[a_t^2] = \mathbb{E}[\mathbb{E}(a_t^2 | F_{t-1})] \\ &= \mathbb{E}[\mathbb{E}(\alpha_0 + \alpha_1 a_{t-1}^2 | F_{t-1})] \\ &= \mathbb{E}[\alpha_0 + \alpha_1 a_{t-1}^2] = \alpha_0 + \alpha_1 \mathbb{E} a_{t-1}^2 \\ &= \alpha_0 + \alpha_1 \mathbb{E} a_t^2 \end{aligned}$$

- Therefore, if $0 \leq \alpha_1 < 1$,

$$\text{Var}[a_t] = \frac{\alpha_0}{1 - \alpha_1}$$

Properties of ARCH models

Kurtosis – fourth moment

- Simple model

$$\begin{aligned}r_t &= \sigma_t \varepsilon_t & \mathbb{E} \varepsilon_t &= 0, \mathbb{E} \varepsilon_t^2 = 1 \\ \sigma_t^2 &= \alpha_0 + \alpha_1 a_{t-1}^2 \\ \alpha_0 &> 0, \alpha_1 &\geq 0 \\ \mathbb{E}[a_t] &= 0 \\ \text{Var}[a_t] &= \frac{\alpha_0}{1 - \alpha_1}\end{aligned}$$

- Tail behavior

- Assume ε_t is normally distributed
- Do we get fatter tails than from the normal distribution?

Fourth moment, tail behavior

$$E a_t^4 = \frac{3\alpha_0^2 (1 + \alpha_1)}{(1 - \alpha_1) (1 - 3\alpha_1^2)}$$

- $E a_t^4 > 0$ obviously and therefore α_1 must satisfy $(1 - 3\alpha_1^2) > 0$ and therefore $0 \leq \alpha_1^2 \leq \frac{1}{3}$
- The unconditional kurtosis of a_t with normally distributed ε_t is

$$\frac{E a_t^4}{\text{Var}[a_t]^2} = 3 \frac{1 - \alpha_1^2}{1 - 3\alpha_1^2}$$

- Therefore,

$$\frac{E a_t^4}{\text{Var}[a_t]^2} > 3$$

- This implies fatter tails than for a normal distribution

Properties of ARCH models

Restrictions on estimated volatility equation in practice

- Simple model

$$r_t = \sigma_t \varepsilon_t \quad \mathbb{E} \varepsilon_t = 0, \mathbb{E} \varepsilon_t^2 = 1$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2$$

$$\alpha_0 > 0, \alpha_1 \geq 0$$

$$\mathbb{E}[a_t] = 0$$

$$\text{Var}[a_t] = \frac{\alpha_0}{1 - \alpha_1}$$

$$\frac{\mathbb{E} a_t^4}{\text{Var}[a_t]^2} = 3 \frac{1 - \alpha_1^2}{1 - 3\alpha_1^2} > 3, \quad 0 \leq \alpha_1^2 \leq \frac{1}{3}$$

- α_i need not all be positive when more than one lag in the volatility equation

- Sufficient to make sure all the estimated conditional volatilities $\widehat{\sigma}_t^2 > 0$
- If one $\widehat{\sigma}_t^2$ is negative, the estimated results make no sense

Limitations of ARCH models

- 1 Symmetric effects of shocks. This is too restrictive for stock returns, where positive shocks have a smaller effect on future volatility than negative shocks
- 2 Returns tend to have some clusters of high and low volatility, whereas ARCH models tend predict slow decay to mean from any current volatility
- 3 Restrictive parameterization, e.g. $0 \leq \alpha_1^2 \leq \frac{1}{3}$ for kurtosis to be well defined for ARCH(1)
- 4 Deterministic equation for volatility; no error term in
$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2$$
- 5 Provides no evidence on source of changes in volatility

Estimation of ARCH model

- 1 Use partial autocorrelation function of a_t^2 to determine order of ARCH specified
- 2 Maximize the likelihood of distribution of ε_t
 - Distributions
 - Normal distribution
 - t-distribution with degrees of freedom ν
 - generalized error distribution
- 3 a_t/σ_t is a sequence of IID variables if correctly specified
 - Use $\hat{a}_t/\hat{\sigma}_t$ to examine whether the serial correlation is adequately estimated
 - Partial autocorrelation function of $\hat{a}_t/\hat{\sigma}_t$
 - Partial autocorrelation function of $(\hat{a}_t/\hat{\sigma}_t)^2$ and Engle regression test on $(\hat{a}_t/\hat{\sigma}_t)^2$
 - Compare distribution to the one assumed using Kolmogorov-Smirnoff tests
 - These tests are asymptotically correct but have nontrivial estimation error and the reported p-values are only a guide to decisions

- ARCH models can require many lags
 - Reduce lags in mean equations by using ARMA models
 - MA terms can substitute for several AR terms
 - Maybe including something like MA terms in ARCH equation can reduce number of lags

- GARCH (Generalized ARCH) model

- ARCH Model

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \dots + \alpha_k a_{t-m}^2$$

- Instead, try GARCH, here a GARCH(m,s) (order not consistent between Tsay and Eviews even)

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \dots + \alpha_k a_{t-m}^2 + \beta_1 \sigma_{t-1}^2 + \dots + \beta_s \sigma_{t-s}^2$$

- Lag lengths are m for the part analogous to the moving average and s for the part analogous to an autoregression
- May be able to reduce lag length substantially by having both sets of terms

Properties of GARCH models

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \dots + \alpha_k a_{t-m}^2 + \beta_1 \sigma_{t-1}^2 + \dots + \beta_k \sigma_{t-s}^2$$

- Restrictions on parameters

$$\alpha_i > 0, \beta_i > 0, \sum_{i=1}^{\max(m,s)} (\alpha_i + \beta_i) < 1$$

- Properties of estimates and relation to parameters

$$E[a_t^2] = \frac{\alpha_0}{1 - \sum_{i=1}^{\max(m,s)} (\alpha_i + \beta_i)}$$

- For GARCH(1,1), $\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2$, with $1 - (\alpha_1 + \beta_1)^2 - 2\alpha_1^2 > 0$

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

- Generally speaking, it is hard to estimate more than a few lags

- What if the volatility is very persistent?

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \dots + \alpha_k a_{t-m}^2 + \beta_1 \sigma_{t-1}^2 + \dots + \beta_k \sigma_{t-s}^2$$

- with $\sum_{i=1}^{\max(m,s)} (\alpha_i + \beta_i) = 1$, indicating a unit root in the volatility process
- Actually pretty common with returns
 - Change in logarithm of value-weighted CRSP index 1/2/1926 to 12/31/1007
 - $d\ln vwc_{crsp} = 0.000632 + \hat{a}_t$
 - $\hat{\sigma}_t^2 = 1.22 \cdot 10^{-6} + 0.09947 \hat{a}_{t-1}^2 + 0.8897 \hat{\sigma}_{t-1}^2$
 - standard errors of coefficients are $1.22 \cdot 10^{-8}$, 0.0019 and 0.0023
 - sum of coefficients is 0.9892
- IGARCH model, IGARCH(1,1)

$$\begin{aligned} a_t &= \sigma_t \varepsilon_t \\ \sigma_t^2 &= \alpha_0 + \beta_1 \sigma_{t-1}^2 + (1 - \beta_1) a_{t-1}^2 \\ 0 &< \beta_1 < 1 \end{aligned}$$

Properties of IGARCH model

- IGARCH(1,1)

$$\begin{aligned}a_t &= \sigma_t \varepsilon_t \\ \sigma_t^2 &= \alpha_0 + \beta_1 \sigma_{t-1}^2 + (1 - \beta_1) a_{t-1}^2 \\ 0 &< \beta_1 < 1\end{aligned}$$

- Unconditional volatility is undefined
- Constant term is similar to constant term for a random walk – a trend
 - A nonzero constant term suggests a trend in variance
 - Why? One-step-ahead forecast at h is forecast of σ_{h+1}^2
 - Suppose that estimate of σ_h^2 and a_h^2 are available and $\sigma_h^2(1)$ is the forecast made of σ^2 made at h for one step ahead

$$\begin{aligned}\sigma_h^2(1) &= \alpha_0 + \beta_1 \sigma_h^2 + (1 - \beta_1) a_h^2 \\ \sigma_h^2(2) &= \alpha_0 + \beta_1 \sigma_{h+1}^2 + (1 - \beta_1) a_{h+1}^2\end{aligned}$$

- Best forecast of σ_{h+1}^2 is $\sigma_h^2(1)$ and best forecast of a_{h+1}^2 from $a_{h+1}^2 = \sigma_{h+1}^2 \varepsilon_t$ is $\sigma_h^2(1)$, so

$$\sigma_h^2(2) = \alpha_0 + \sigma_h^2(1)$$

- GARCH-M is GARCH in mean – simple one is

$$\begin{aligned}r_t &= \mu + c\sigma_t^2 + a_t, & a_t &= \sigma_t\varepsilon_t \\ \sigma_t^2 &= \alpha_0 + \alpha_1 a_t^2 + \beta_1 \sigma_{t-1}^2\end{aligned}$$

- c is called a risk premium parameter
- Could use σ_t or $\ln \sigma_t^2$ instead of σ_t^2
- Question: When will the variance an asset's return reflect its risk?

- EGARCH is exponential GARCH

$$r_t = \mu + a_t, \quad a_t = \sigma_t \varepsilon_t$$

model for $\ln \sigma_t^2$

- An EGARCH model starts from the function $g(\varepsilon_t)$

$$g(\varepsilon_t) = \theta \varepsilon_t + \gamma (|\varepsilon_t| - \mathbf{E}|\varepsilon_t|)$$
$$\mathbf{E} g(\varepsilon_t) = 0$$

- which can be rewritten as

$$g(\varepsilon_t) = \begin{cases} (\theta + \gamma) \varepsilon_t - \gamma \mathbf{E}|\varepsilon_t| & \text{if } \varepsilon_t \geq 0 \\ (\theta - \gamma) \varepsilon_t - \gamma \mathbf{E}|\varepsilon_t| & \text{if } \varepsilon_t < 0 \end{cases}$$

- If θ is negative, then $g(\varepsilon_t)$ is larger for $\varepsilon_t < 0$ than for $\varepsilon_t \geq 0$

- Define the lag operator L such that $Lx_t = x_{t-1}$ and $L^i x_t = x_{t-i}$
 - B is used by statisticians to mean the same thing
 - EGARCH model, EGARCH(m,s)

$$\begin{aligned}r_t &= \mu + a_t, \quad a_t = \sigma_t \varepsilon_t \\g(\varepsilon_t) &= \theta \varepsilon_t + \gamma (|\varepsilon_t| - E|\varepsilon_t|) \\ \ln \sigma_t^2 &= \alpha_0 + \frac{1 + \beta_1 L + \dots + \beta_{s-1} L^{s-1}}{1 - \alpha_1 L - \dots - \alpha_m L^m} g(\varepsilon_{t-1})\end{aligned}$$

- Take an EGARCH(1,1) with ε_t iid and normally distributed

$$\begin{aligned}a_t &= \sigma_t \varepsilon_t \\(1 - \alpha_L) \ln \sigma_t^2 &= \left\{ \begin{array}{l} \alpha_* + (\gamma + \theta) \frac{a_{t-1}}{\sigma_{t-1}} \text{ if } a_{t-1} \geq 0 \\ \alpha_* + (\gamma - \theta) \frac{|a_{t-1}|}{\sigma_{t-1}} \text{ if } a_{t-1} < 0 \end{array} \right\} \\ \alpha_* &= (1 - \alpha_1) \alpha_0 - \sqrt{2\pi} \gamma\end{aligned}$$

- Threshold GARCH –

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^s (\alpha_i + \gamma_i N_{t-i}) a_{t-i}^2 + \sum_{j=1}^m \beta_j \sigma_{t-j}^2$$

$$N_{t-i} = 1 \text{ if } a_{t-i} < 0$$

$$N_{t-i} = 0 \text{ if } a_{t-i} \geq 0$$

$$\alpha_i, \gamma_i, \beta_j \geq 0$$

- This also allows for bigger effects of negative shocks

- Autocorrelation function of squared residuals or absolute value of returns often decays slowly
 - For a covariance stationary process, autocorrelations decline at an exponential rate
 - For a long-memory process, autocorrelations decline at hyperbolic rate (slower)
- Stochastic volatility
 - The evolution of volatility is not a deterministic function of past volatility and shocks to the mean equation

Relatively simple example of long memory

$$\begin{aligned}a_t &= \sigma_t \varepsilon_t \\ \sigma_t &= \sigma \exp(u_t/2), \quad (1-L)^d u_t = \eta_t \\ \varepsilon_t &\sim N(0,1), \quad \eta_t \sim N(0, \sigma_\eta^2) \\ \sigma &> 0, \quad d < \frac{1}{2}\end{aligned}$$

- d not an integer indicates fractional differencing
 - The long memory comes from the fractional differencing
 - Note that

$$\begin{aligned}\ln a_t^2 &= \ln \sigma_t^2 + \ln \varepsilon_t^2 = \ln \sigma^2 + \ln u_t + \ln \varepsilon_t^2 \\ &= \left(\ln \sigma^2 + \mathbb{E} \ln \varepsilon_t^2 \right) + u_t + \left(\ln \varepsilon_t^2 - \mathbb{E} \ln \varepsilon_t^2 \right) \\ &= \mu + u_t + e_t\end{aligned}$$

- where $\mu = \ln \sigma^2 + \mathbb{E} \ln \varepsilon_t^2$ and $e_t = \left(\ln \varepsilon_t^2 - \mathbb{E} \ln \varepsilon_t^2 \right)$

Use higher-frequency returns to estimate variance

- Example: Use daily variance in the month to calculate variance for the month
 - Let r_t^m be the monthly return in month t
 - Let $r_{t,i}$ be the daily return on day i in month t
 - Suppose that daily returns are serially uncorrelated and the daily variance is constant
 - Then

$$r_t^m = \sum_{i=1}^n r_{t,i}$$
$$\text{Var} [r_t^m] = n \text{Var} [r_{t,i}]$$

- and

$$\widehat{\text{Var}} [r_{t,i}] = \frac{\sum_{i=1}^n (r_{t,i} - \bar{r}_t)^2}{n - 1}$$

where \bar{r}_t is the mean of the daily returns

- The estimated monthly volatility thus is

$$\widehat{\sigma}_t^m = n \frac{\sum_{i=1}^n (r_{t,i} - \bar{r}_t)^2}{n - 1}$$

Daily variance to estimate monthly volatility

- The estimated monthly volatility is simple to calculate

$$\sigma_t^m = n \frac{\sum_{i=1}^n (r_{t,i} - \bar{r}_t)^2}{n - 1}$$

- This becomes more complicated if the daily returns are serially correlated, but it's still manageable
- If daily log returns have high excess kurtosis and serial correlations, then this estimator may not be consistent
- Does this make sense from a subject-matter (financial economics) point of view?

Garman-Klass estimator of daily volatility

- Use high, low, opening, and closing prices to estimate volatility
 - Can even estimate daily volatility just knowing opening, high, low and closing prices
 - Assume that price follows a random walk
 - Let C_t be the logarithm of the closing price so $r_t = c_t - c_{t-1}$
 - Conventional estimator is

$$\sigma_t^2 = \text{E} \left[(c_t - c_{t-1})^2 \right]$$

- Using only closing price
- High H_t , low L_t , and open O_t also often are available
- Can estimate daily volatility of price (not log price) from

$$\hat{\sigma}_{GK}^2 = 0.12 \frac{(O_t - C_{t-1})^2}{f} + 0.88 \frac{0.5(H_t - L_t)^2 + 0.386(C_t - O_t)^2}{1 - f}$$

where f is the fraction of the day that the market is open

- Is there something like this to estimate the volatility of the log return?
Yes. See next slide

Yang and Zhang estimator

- Use high, low, opening, and closing prices to estimate volatility of log prices over a longer period
 - Define

$$o_t = \ln O_t - \ln O_{t-1}$$

$$u_t = \ln H_t - \ln O_{t-1}$$

$$d_t = \ln L_t - \ln O_{t-1}$$

$$c_t = \ln C_t - \ln O_{t-1}$$

- Monthly volatility based on n days of trading is

$$\hat{\sigma}_{YZ}^2 = \hat{\sigma}_o + k\hat{\sigma}_c + (1 - k)\hat{\sigma}_{rs}$$

where $\hat{\sigma}_o$ and $\hat{\sigma}_c$ are the estimated variances of o_t and c_t and

$$\hat{\sigma}_{rs}^2 = \frac{1}{n} \sum [u_t (u_t - c_t) + d_t (d_t - c_t)]$$

$$k = \frac{0.34}{1.34 + (n + 1) / (n - 1)}$$

and k is chosen to minimize the variance of the estimator $\hat{\sigma}_{YZ}^2$