

Financial Econometrics

Volatility

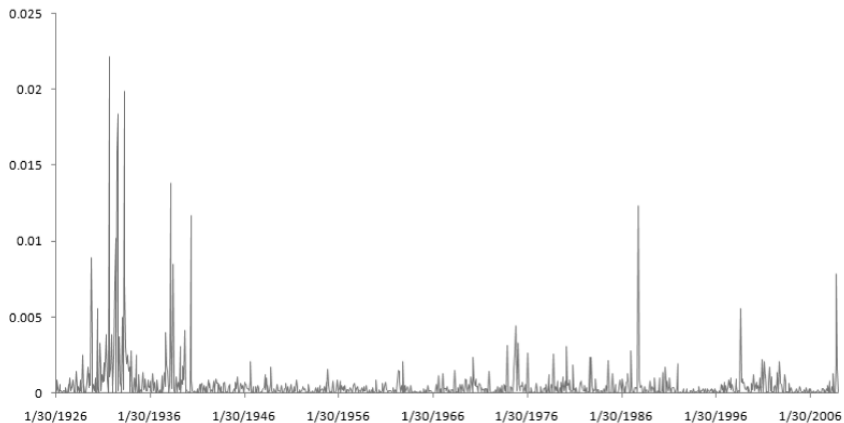
Gerald P. Dwyer

Trinity College, Dublin

January 2011

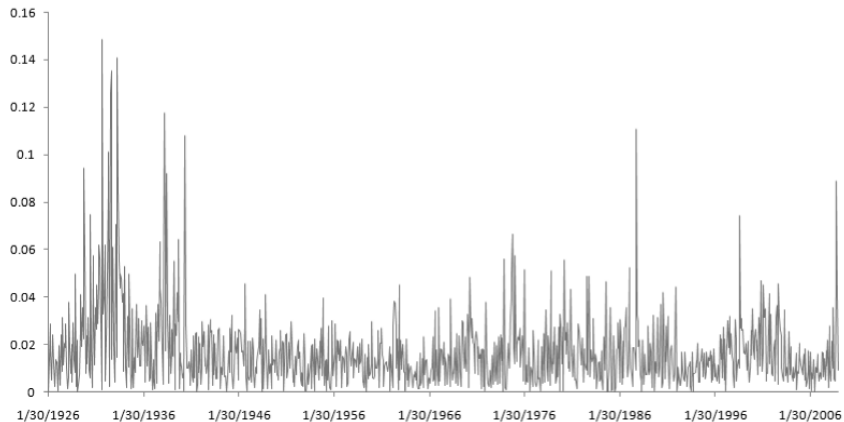
Squared changes in log returns for CRSP daily

**Squared changes in log of CRSP index
January 31, 1926 to December 31, 2008**



Absolute value of log returns for CRSP daily

**Absolute value of changes in log of CRSP index
January 31, 1926 to December 31, 2008**



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 mean as it changes

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

Correlogram of RETD_SQ

Date: 01/25/10 Time: 09:00
Sample: 1925:12 2008:12
Included observations: 997

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 0.294	0.294	95.723	0.000
		2 0.168	0.089	115.13	0.000
		3 0.224	0.168	165.35	0.000
		4 0.124	0.012	180.77	0.000
		5 0.063	-0.010	184.78	0.000
		6 0.104	0.052	195.65	0.000
		7 0.110	0.054	207.73	0.000
		8 0.329	0.308	316.65	0.000
		9 0.333	0.192	428.45	0.000
		10 0.205	0.035	470.98	0.000
		11 0.268	0.107	543.65	0.000
		12 0.131	-0.081	561.10	0.000
		13 0.112	0.042	573.71	0.000
		14 0.098	0.000	583.40	0.000
		15 0.125	0.080	599.16	0.000
		16 0.057	-0.094	602.49	0.000
		17 0.084	-0.101	609.65	0.000
		18 0.067	-0.094	614.22	0.000
		19 0.163	0.040	641.43	0.000
		20 0.053	-0.081	644.29	0.000
		21 0.025	-0.045	644.91	0.000
		22 0.076	-0.021	650.78	0.000
		23 0.153	0.102	674.65	0.000
		24 0.052	-0.026	677.47	0.000
		25 0.053	0.046	680.38	0.000
		26 0.093	0.060	689.34	0.000
		27 0.022	-0.037	689.84	0.000
		28 0.015	-0.006	690.07	0.000
		29 0.005	0.022	690.09	0.000
		30 0.018	0.013	690.44	0.000
		31 0.034	0.009	691.64	0.000
		32 0.038	-0.030	693.10	0.000
		33 0.067	0.049	697.75	0.000

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

Correlogram of RETD_ABS

Date: 01/25/10 Time: 09:09
Sample: 1925:12 2008:12
Included observations: 997

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 0.217	0.217	47.053	0.000
		2 0.212	0.173	92.085	0.000
		3 0.262	0.202	161.08	0.000
		4 0.189	0.089	197.08	0.000
		5 0.158	0.047	222.03	0.000
		6 0.177	0.068	253.55	0.000
		7 0.151	0.043	276.64	0.000
		8 0.231	0.142	330.50	0.000
		9 0.256	0.148	396.31	0.000
		10 0.187	0.050	431.57	0.000
		11 0.242	0.099	490.63	0.000
		12 0.164	-0.008	517.94	0.000
		13 0.151	0.003	540.93	0.000
		14 0.159	0.015	566.52	0.000
		15 0.186	0.067	601.50	0.000
		16 0.096	-0.045	610.83	0.000
		17 0.145	0.003	632.12	0.000
		18 0.103	-0.047	642.96	0.000
		19 0.126	0.001	659.20	0.000
		20 0.072	-0.065	664.47	0.000
		21 0.072	-0.033	669.76	0.000
		22 0.101	0.006	680.27	0.000
		23 0.166	0.090	708.56	0.000
		24 0.095	0.000	717.86	0.000
		25 0.100	0.006	728.18	0.000
		26 0.110	0.001	740.67	0.000
		27 0.069	-0.011	745.59	0.000
		28 0.033	-0.050	746.70	0.000
		29 0.037	-0.014	748.11	0.000
		30 0.044	-0.004	750.14	0.000
		31 0.049	0.009	752.62	0.000
		32 0.072	0.013	758.01	0.000
		33 0.067	0.014	762.63	0.000

Autoregressive conditional heteroskedasticity (ARCH)

is intended to deal with this

- Simple ARCH model for returns

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

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

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

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

where $u_t = \sigma_t \varepsilon_t$

Estimating an ARCH model

$$\begin{aligned}y_t &= \mu_t + u_t & \mathbb{E} u_t &= 0, \mathbb{E} u_t^2 = \sigma_t^2 \\ \sigma_t^2 &= \alpha_0 + \alpha_1 u_{t-1}^2\end{aligned}$$

- 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 variance 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

$$y_t = \mu_t + u_t$$

- y_t can be a complicated ARMA(p,q) or can have variables included
 - y_t is stationary in mean
 - y_t may be first difference of original series
 - for example $r_t = p_t - p_{t-1}$
- May be mis-specified if ignore conditional heteroskedasticity of u_t

- Simple model

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

- Let u_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, u_t^2 , for some pre-specified number of lags k
- Engle test based on a regression for the squared residuals

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

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 $SSR_0 = \sum_{t=1}^T (u_t^2 - \overline{u^2}) / T$, $\overline{u^2}$ is the mean of u_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

$$\begin{aligned}y_t &= u_t = \sigma_t \varepsilon_t & \mathbb{E} \varepsilon_t &= 0, \mathbb{E} \varepsilon_t^2 = 1 \\ \sigma_t^2 &= \alpha_0 + \alpha_1 u_{t-1}^2\end{aligned}$$

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 $y_{t-1}, y_{t-2}, \dots, u_{t-1}, u_{t-2}, \dots$

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

Properties of ARCH models

Variance

- Simple model

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

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

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

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

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

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

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

Properties of ARCH models

Kurtosis – fourth moment

- Simple model

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

- Tail behavior

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

Fourth moment, tail behavior

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

- $E a_t^4 > 0$ 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

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

- Therefore,

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

- This implies fatter tails than for a normal distribution

Properties of ARCH models

Restrictions on estimated variance equation in practice

- Simple model

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

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

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

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

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

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

- α_i need not all be positive when more than one lag in the variance 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 variance than negative shocks
- 2 Returns tend to have some clusters of high and low variance, whereas ARCH models tend predict slow decay to mean from any current variance
- 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 variance; no error term in
$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2$$
- 5 Provides no evidence on source of changes in variance

Estimation of ARCH model

- ① Use partial autocorrelation function of u_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
 - Quasi-maximum likelihood estimation
 - Consistent estimates of parameters
 - Issue of correct standard errors of coefficients
- ③ u_t/σ_t is a sequence of IID variables if correctly specified
- Use $\hat{u}_t/\hat{\sigma}_t$ to examine whether the serial correlation is adequately estimated
 - Partial autocorrelation functions of $\hat{u}_t/\hat{\sigma}_t$ and $(\hat{u}_t/\hat{\sigma}_t)^2$ with Engle regression test on $(\hat{u}_t/\hat{\sigma}_t)^2$
 - Compare distribution to the one assumed using Kolmogorov-Smirnoff tests

- 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

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

- Instead, try GARCH, here a GARCH(m,s) (order of lags often not consistent)

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

- 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

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

- 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

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

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

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

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

- What if the variance is very persistent?

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

- with $\sum_{i=1}^{\max(m,s)} (\alpha_i + \beta_i) = 1$, indicating a unit root in the variance process
- Actually pretty common with returns
 - Change in logarithm of value-weighted CRSP index 1/2/1926 to 12/31/1007
 - $d\ln v w c r s p = 0.000632 + \hat{a}_t$
 - $\hat{\sigma}_t^2 = 1.22 \cdot 10^{-6} + 0.09947 \hat{a}_{t-1}^2 + 0.8897 \hat{\sigma}_{t-1}^2$
 - 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)

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

Properties of IGARCH model

- IGARCH(1,1)

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

- Unconditional variance 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

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

- 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

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

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

- Glosten, Jagannathan and Runkle (1993)
- Represent the asymmetry in returns that negative shocks create more future variance

$$\begin{aligned}y_t &= \mu + c\sigma_t^2 + u_t, & u_t &= \sigma_t \varepsilon_t \\ \sigma_t^2 &= \alpha_0 + \alpha_1 u_{t-1}^2 + \beta_1 \sigma_{t-1}^2 + \gamma u_{t-1}^2 I_{t-1}\end{aligned}$$

- $I_{t-1} = 1$ if $u_{t-1} < 0$ and $I_{t-1} = 0$ if $u_{t-1} \geq 0$
- Greater effect of negative shocks if estimate $\gamma > 0$

- Exponential GARCH – allows for asymmetry
- An EGARCH(1,1) with ε_t iid and normally distributed

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

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

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

- which can be rewritten as

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

- 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 $L x_t = x_{t-1}$ and $L^i x_t = x_{t-i}$
 - EGARCH model, EGARCH(m,s)

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

Test for asymmetric volatility

- Estimate a mean equation and get residuals \hat{u}_t
- Define $I_{t-1} = 1$ if $\hat{u}_{t-1} < 0$ and $I_{t-1} = 0$ if $\hat{u}_{t-1} \geq 0$
- Run regression

$$\hat{u}_t^2 = \phi_0 + \phi_1 I_{t-1} + v_t$$

- If $\phi_1 > 0$, then there is an asymmetric effect of negative shocks

- Threshold GARCH – TGARCH(m,s)

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

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

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

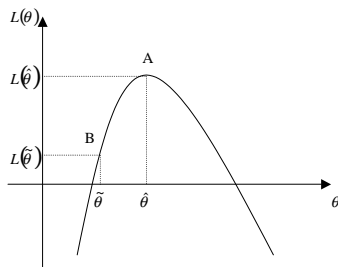
$$\alpha_i, \gamma_i, \beta_j \geq 0$$

- This also allows for bigger effects of negative shocks

Testing restrictions in nonlinear models

- Usual t-tests are estimated for maximum likelihood
 - Inverse of information matrix provides estimator of variance for these tests
- Likelihood ratio, Wald and Lagrange multiplier tests
- Consider estimating a parameter θ
 - Maximum likelihood estimate is $\hat{\theta}$
 - Restricted estimate is $\tilde{\theta}$

Comparison of likelihood ratio, Wald and Lagrange multiplier tests



- Vertical distance from B to A is basis of likelihood ratio tests
- Horizontal distance from B to A is basis of Wald tests (e.g. t-tests)
- Slope of likelihood function at B is basis of Lagrange multiplier test

- 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

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

- non-integer d indicates fractional differencing
 - The long memory comes from the fractional differencing
 - Note that

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

- 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 variance 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 variance

- The estimated monthly variance 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 variance

- Use high, low, opening, and closing prices to estimate variance
 - Can even estimate daily variance 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 variance 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 variance of the log return?
Yes. See next slide

Yang and Zhang estimator

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

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

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

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

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

- Monthly variance 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

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

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

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