# ARCH/GARCH Volatility Models

Junye Li

Finance Department
ESSEC Business School

# Variance

- For a random variable $x$, its expected value is $E[x]$,
- and $x - E[x]$ defines the unexpected part of $x$;
- $V[x]$ is the variance of $x$
  - $V[x] = E[(x - E[x])^2]$
  - Standard deviation is the square-root of variance
- In Financial Time Series, *volatility* is the standard deviation of asset returns.

# Time-Varying Volatility

- Volatility determines the behaviour of the unpredictable part of asset returns;
- Suppose that we can model asset returns as being conditionally normal. If mean and variance are time-varying, then

$$f(r_t \mid F_{t-1}) = N(\mu_t, \sigma_t^2)$$

- We have already investigated models for $\mu_t$, now we turn to models for the conditional variance $\sigma_t^2$.

- Consider an AR(1) model for asset returns:

$$r_t = \beta_0 + \beta_1 r_{t-1} + \varepsilon_t \ .$$

- Then the conditional mean is $\mu_t = \beta_0 + \beta_0 r_{t-1}$ .
- We define the conditional variance as:

$$\sigma_t^2 = E[(r_t - \mu_t)^2 \mid F_{t-1}] = E[\varepsilon_t^2 \mid F_{t-1}].$$

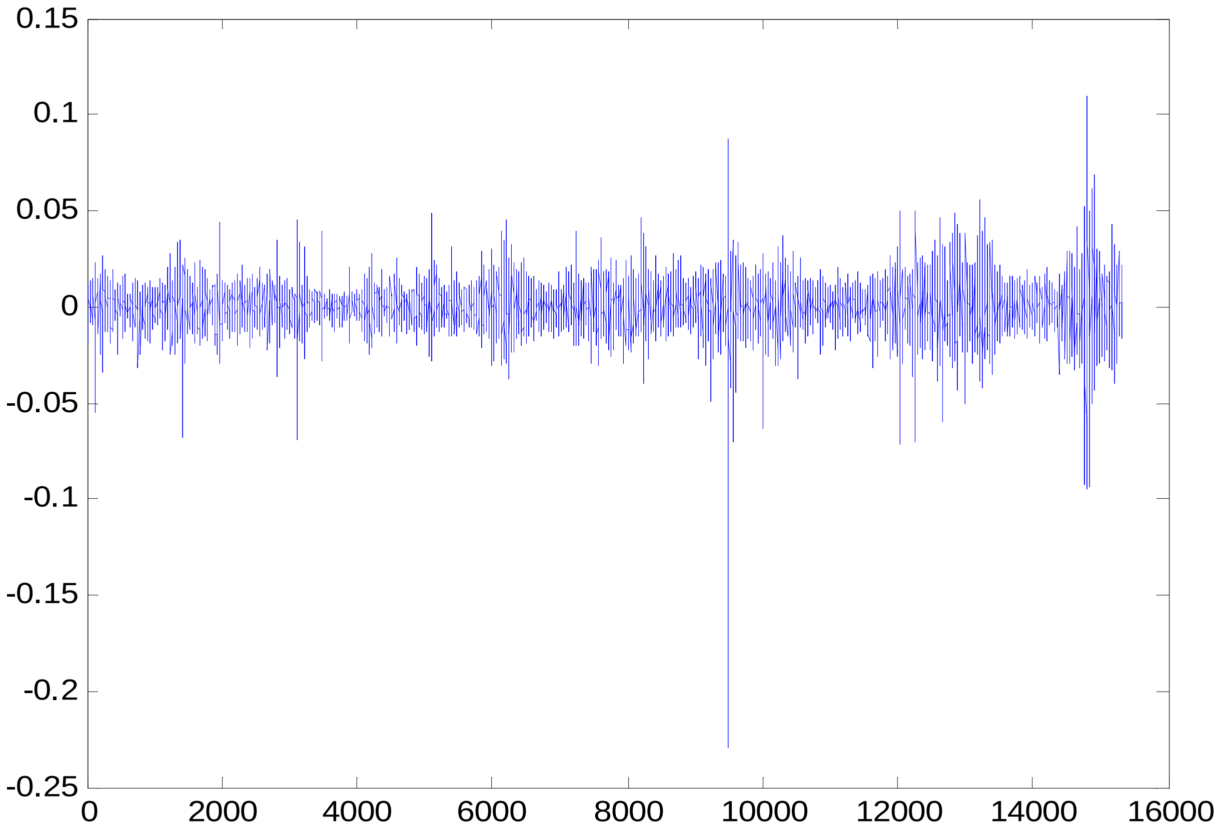
- Since variance is time-varying, it is allowed to vary as the information at time t-1 varies;

- If markets are nearly efficient, returns are not predictable. If the conditional mean is very small, the conditional variance can be calculated by

  $$\sigma_t^2 = E[(r_t - \mu_t)^2 \mid F_{t-1}] = E[r_t^2 \mid F_{t-1}].$$

- Hence, the variance over multiple days is the sum of the variances over each of days;
- If the variance were the same every day of the year, the annualized variance is simply given by: Total number of days x daily variance.

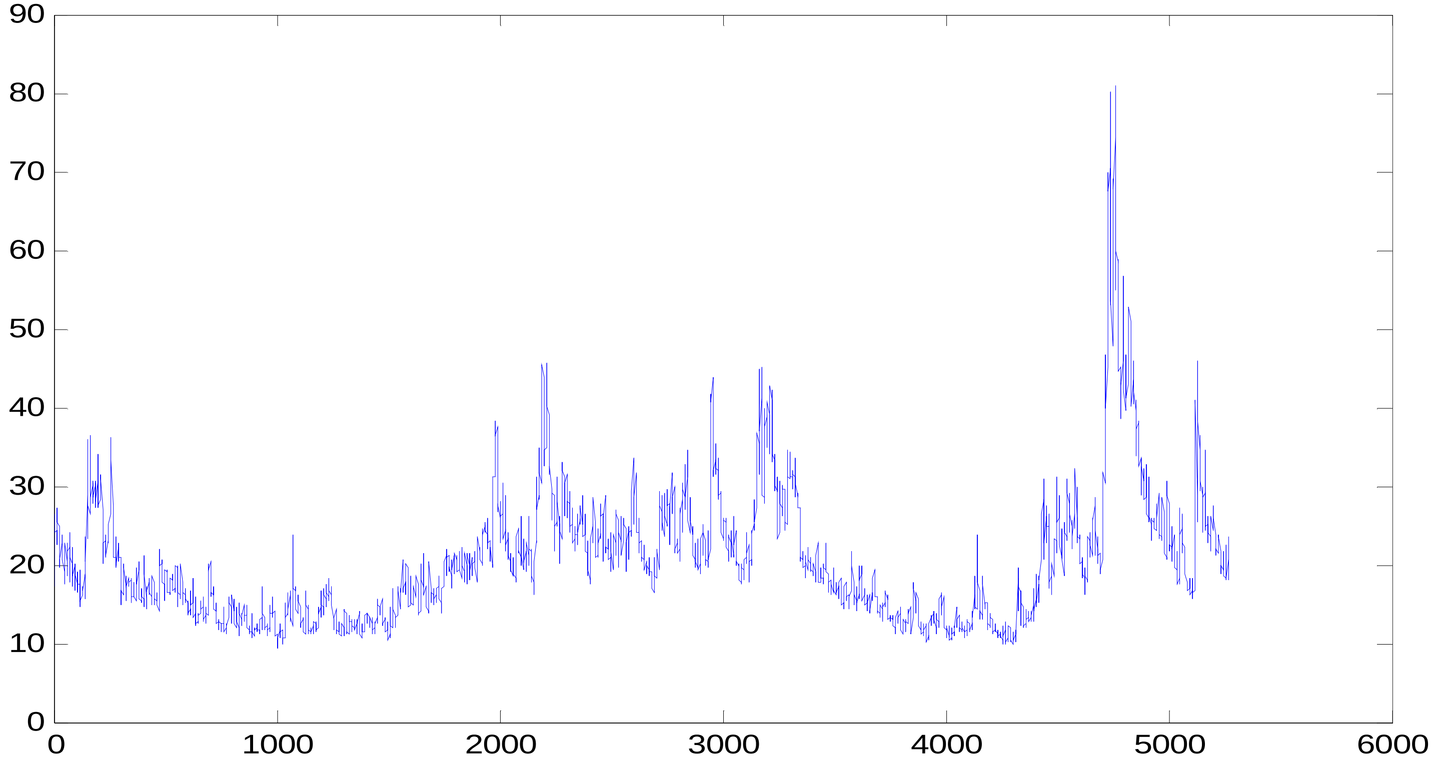
# Review of Stylized Facts



# Review of Stylized Facts

- Non-normal distribution:
  - Left-skewed;
  - Fat-tailed.
- Periods of high and low volatilities
  - Volatility clustering
- Mean-reverting volatility
- VIX as a Market Volatility Index

# VIX Index



# Why Do We Need Volatility?

- Volatility is a measure of risk
  - Knowledge of risk allows us to avoid it;
  - Risk-return trade-off.
- Markowitz Portfolio Theory:
  - What risk must we take to achieve a satisfactory return;
  - 1990 Nobel Prize
- CAPM:
  - Systematic risk and idiosyncratic risk;
  - 1990 Nobel Prize
- Option Pricing
  - Volatility is a key input
  - 1997 Nobel Prize

# How to Estimate Volatility

- A special feature of stock volatility is that it is not directly observable.
- Historical Volatility:

$$\hat{\sigma} = \sqrt{252 \sum_{j=T-K}^{T} r_j^2 / K}$$

  - Choose K large so that the estimate is more accurate;

# How to Estimate Volatility

- Exponential Smoothing

$$\sigma_t^2 = \lambda\ \sigma_{t-1}^2 + (1- \lambda)\ r_{t-1}^2$$

  - This model is used by RISKMETRICS
  - Volatility is time-varying
  - Current volatility depends on the previous volatility and squared return
  - How to choose $\lambda$?
  - No mean reversion

# Modeling Volatility

- Volatility models can be classified into two general categories.
  - The first uses an exact function to govern the evolution of asset variance.
  - The second uses a stochastic equation to describe asset variance.
- The ARCH/GARCH models belong to the first category, whereas the stochastic volatility models are in the second category.

# Model Building

- Specify a mean equation by testing serial dependence in the data (ARMA models);
- Use the residuals to test for ARCH effects.
- Specify a volatility model if ARCH effects are statistically significant.
- Perform a joint estimation of mean and volatility models
- Check the goodness-of-fit.

# Testing for ARCH Effect

- Let $e_t = r_t - \mu_t$ be the residuals of the mean equation.
- The squared residuals $e^2_t$ is then used to check for conditional heteroskedasticity, which is also known as the ARCH effect.
- Two tests are available:
  – Ljung-Box Q-test for the series of $e^2_t$;
  – Engle's Lagrange multiplier test (Engle, 1982)

# Lagrange Multiplier Test

- The test is equivalent to the usual F-test. For a linear model

$$e_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2 + \ldots + \alpha_m e_{t-m}^2 + \varepsilon_t$$

- The null hypothesis H0: α = … α = 0. Denote

$$SSR_0 = \sum_{t=m+1}^{T} (e_t^2 - \overline{e}); SSR_1 = \sum_{t=m+1}^{T} \hat{\varepsilon}_t^2$$

- The F-test is

$$F = \frac{(SSR_0 - SSR_1)/m}{SSR_1/(T-2m-1)} \sim \chi^2(m)$$

# ARCH/GARCH

- (Generalized) Autoregressive Conditional Heteroskedasticity:
  - Volatility is predictable (conditional)
  - Uncertainty (Heteroskedasticity)
  - Time-varying (Autoregressive)
- ARCH idea: use a weighted average of the volatility over a long period with higher weights on the recent past and small weights on the distant past.

# ARCH(q)

- The ARCH(q) model is:

$$r_t = \mu_t + \varepsilon_{t,} \qquad \varepsilon_t \sim (0, h_t)$$

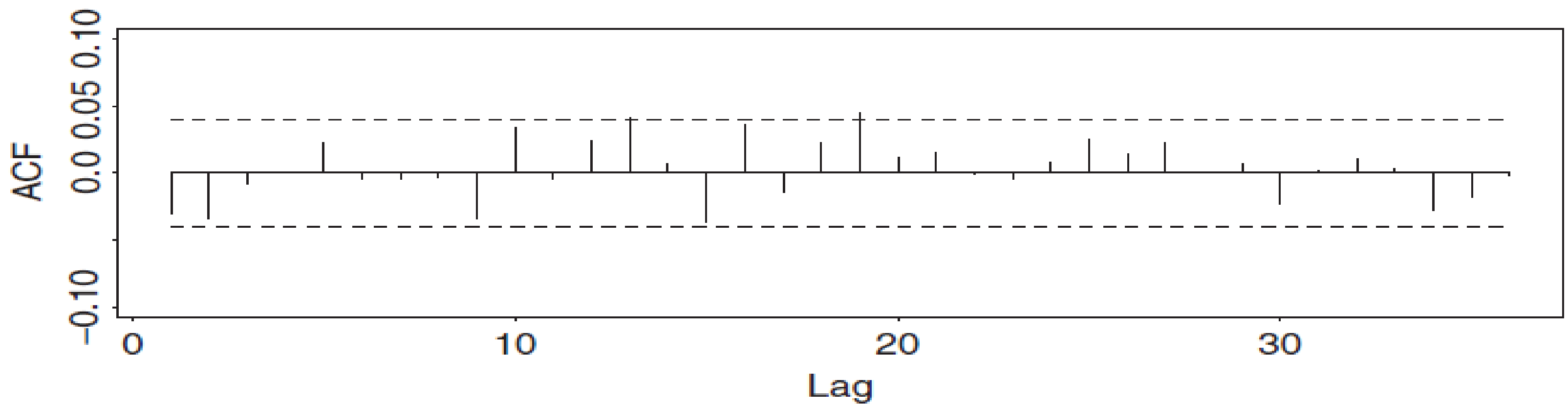
$$h_t = \omega + \alpha_1 \varepsilon_{t-1}^2 + \ldots + \alpha_q \varepsilon_{t-q}^2$$

- $h_t$ is the conditional variance

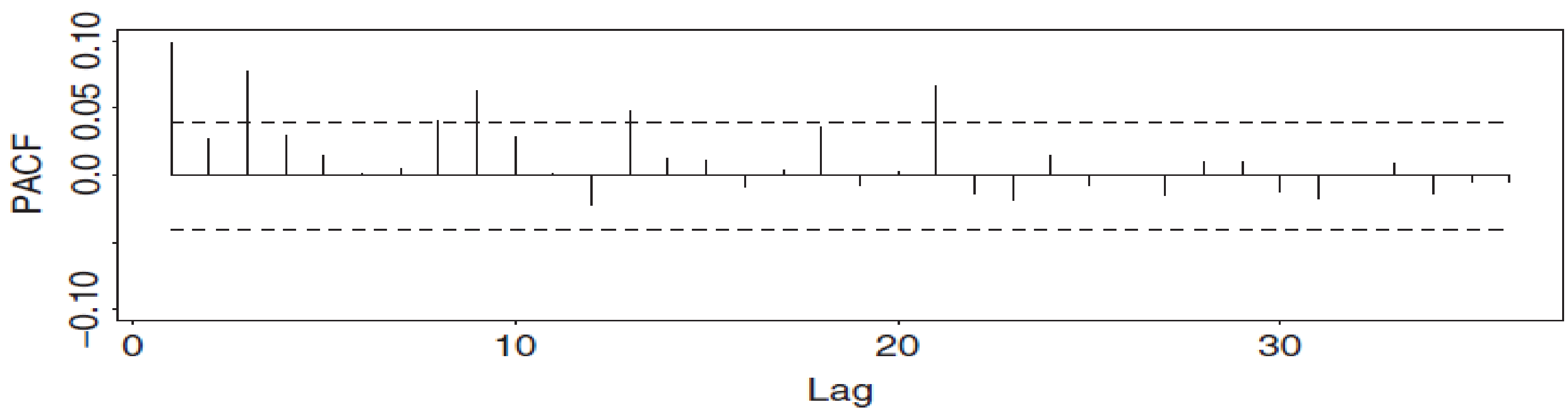
$$\sigma_t^2 = E[(r_t - \mu_t)^2 \mid F_{t-1}] = E[\varepsilon_t^2 \mid F_{t-1}].$$

- Large shocks tend to be followed by another large shock – volatility clustering.
- In practice, we need long lags of q.

# An Example: Exchange Rate



(a)



(b)

# Properties of ARCH Models

- We focus on ARCH(1) model:

$$r_t = \mu + \varepsilon_{t,} \qquad \varepsilon_t \text{ -- } N(0, h_t)$$

$$h_t = \omega + \alpha_1 \varepsilon_{t-1}^2$$

- Unconditional mean of $r_t$ : E[$r_t$ ] = $\mu$;
- Unconditional variance of $r_t$ :

$$Var\,(r_t) = E[\varepsilon_t^2] = E[E[h_t \mid F_{t-1}]]$$

$$= E[\omega + \alpha_1 \varepsilon_{t-1}^2] = \omega + \alpha_1 E[\varepsilon_{t-1}^2]$$

$$\Rightarrow Var\,(r_t) = \omega / (1 - \alpha_1)$$

- The skewness: zero
- The fourth moment:

$$E[(r_t - \mu)^4] = E[\varepsilon_t^4] = E[E[\varepsilon_t^4 \mid F_{t-1}]]$$

$$= 3E[(E[\varepsilon_t^2 \mid F_{t-1}])^2] = 3E[\omega^2 + 2\omega\alpha_1\varepsilon_{t-1}^2 + \alpha_1^2\varepsilon_{t-1}^4]$$

$$\Rightarrow m_4 = 3(\omega^2 + 2\omega\alpha_1 E[\varepsilon_{t-1}^2] + \alpha_1^2 m_4]$$

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

- The kurtosis

$$kurt = \frac{E[(r_t - \mu)^4]}{[Var(r_t)]^2} = 3\frac{1-\alpha_1^2}{1-3\alpha_1^2} > 3$$

# Weakness of ARCH Models

- The model assumes that positive and negative shocks have the same effects on volatility;
- The model is rather restrictive;
- It gives no indication about what causes such behavior to occur;
- In practice, we need large q;
- ARCH models are likely to overpredict the volatility.

# GARCH(p, q) Model

- Generalized ARCH models are most important extensions of ARCH models (Bollerslev, 1986);
- Tomorrow’s variance is predicted to be a weighted average of the
  - Long-run mean of variance
  - Today’s variance forecast
  - The news/shocks (today’s squared error)
- The simplest but very powerful model is GARCH(1, 1)

# GARCH(1, 1) Model

$r_t = \mu + \varepsilon_{t,} \quad \varepsilon_t$ is $N(0, h_t)$

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

- Generalization of Exponential Smoothing model
- Generalization of ARCH model
- Generalization of Constant volatility
- Parameters: *ω, α*, and *β.* What roles do they play?

# Properties of GARCH(1, 1)

- Using the fact that

$$z_t = (r_t - \mu)/h_t^{1/2}$$

is a standard normal (0, 1);

- The GARCH(1, 1) can be written as

$r_t = \mu + h_t^{1/2} z_{t,} \quad z_t$ i.i.d $N(0, 1)$

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

- Stationarity:

$$h_t = \omega + \alpha\, (h_{t-1}^{1/2} z_{t-1})^2 + \beta\, h_{t-1}$$
$$= \omega + (\alpha\, z_{t-1}^2 + \beta)\, h_{t-1}$$

$$E[h_t] = \omega + E(\alpha z_{t-1}^2 + \beta)E[h_{t-1}]$$
$$= \omega + (\alpha + \beta)E[h_{t-1}]$$
$$\Rightarrow \sigma^2 = \omega/(1 - (\alpha + \beta))$$

- The process is covariance stationary if and only if $\alpha + \beta < 1$;
- The volatility process can also be written as

$$h_t = (1 - (\alpha + \beta))\sigma^2 + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1}$$

  – Weighted average of three component:
    - The unconditional variance
    - Yesterday's forecast of variance
    - Yesterday's shocks/news

# Return Moments

- The unconditional mean of returns:

$$E[r_t] = \mu + E[h_t^{1/2}]\, E[z_t] = \mu$$

- The unconditional variance of returns:

$$\gamma_0 = Var(r_t) = E(r_t - \mu)^2 = E[h_t]\, E[z_t^2]$$
$$= \omega/(1 - (\alpha + \beta)) = \sigma^2$$

- The unconditional third moment:

$$E(r_t - \mu)^3 = E[h_t^{3/2}]\, E[z_t^3] = 0$$

- Kurtosis of returns:

$$E[(r_t - \mu_t)^4] = E[h_t^2]E[z_t^4] = 3E[h_t^2]$$

$$\textit{Kurtosis} \quad = 3E[h_t^2]/\sigma^4$$

$$= 3[(1-(\alpha+\beta))/(1-2\alpha^2 - (\alpha+\beta)^2)]$$

--- when $2\alpha^2 + (\alpha + \beta)^2 < 1$, kurtosis $> 3$

- Autocorrelations of returns

$$\gamma_j = \text{cov}(r_t, r_{t-j}) = E[h_t^{1/2} z_t h_{t-j}^{1/2} z_{t-j}]$$

$$= E[h_t^{1/2} z_t h_{t-j}^{1/2}]\, E[z_{t-j}]$$

$$= 0$$

# What Stylized Facts Can Be Explained

- Time-varying volatility
  - Mean reverting (long-run mean $\sigma^2$)
  - Volatility clustering (determined by $\alpha + \beta$)
- Return non-normality
  - Fat-tailed: yes, as kurtosis is larger than 3
  - Skewed: no, as skewness is still zero

# Estimating GARCH Models

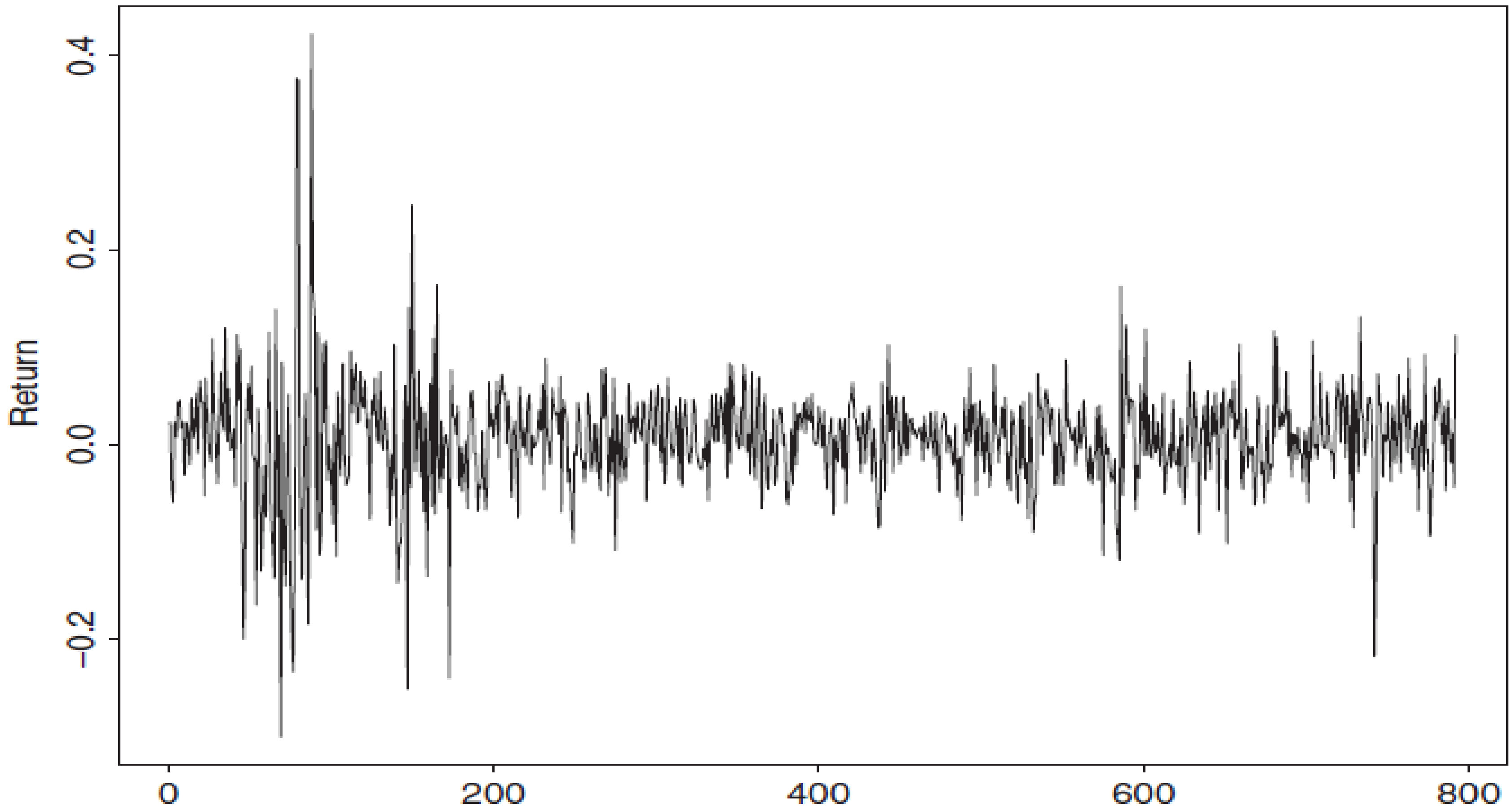
- MLE
- Find the likelihood function using an iterative method.
- This estimation is optimal for large sample if the errors are really normal.
- It is still good without normality: Quasi-Maximum Likelihood Estimation
- Robust Variance Covariance Matrix

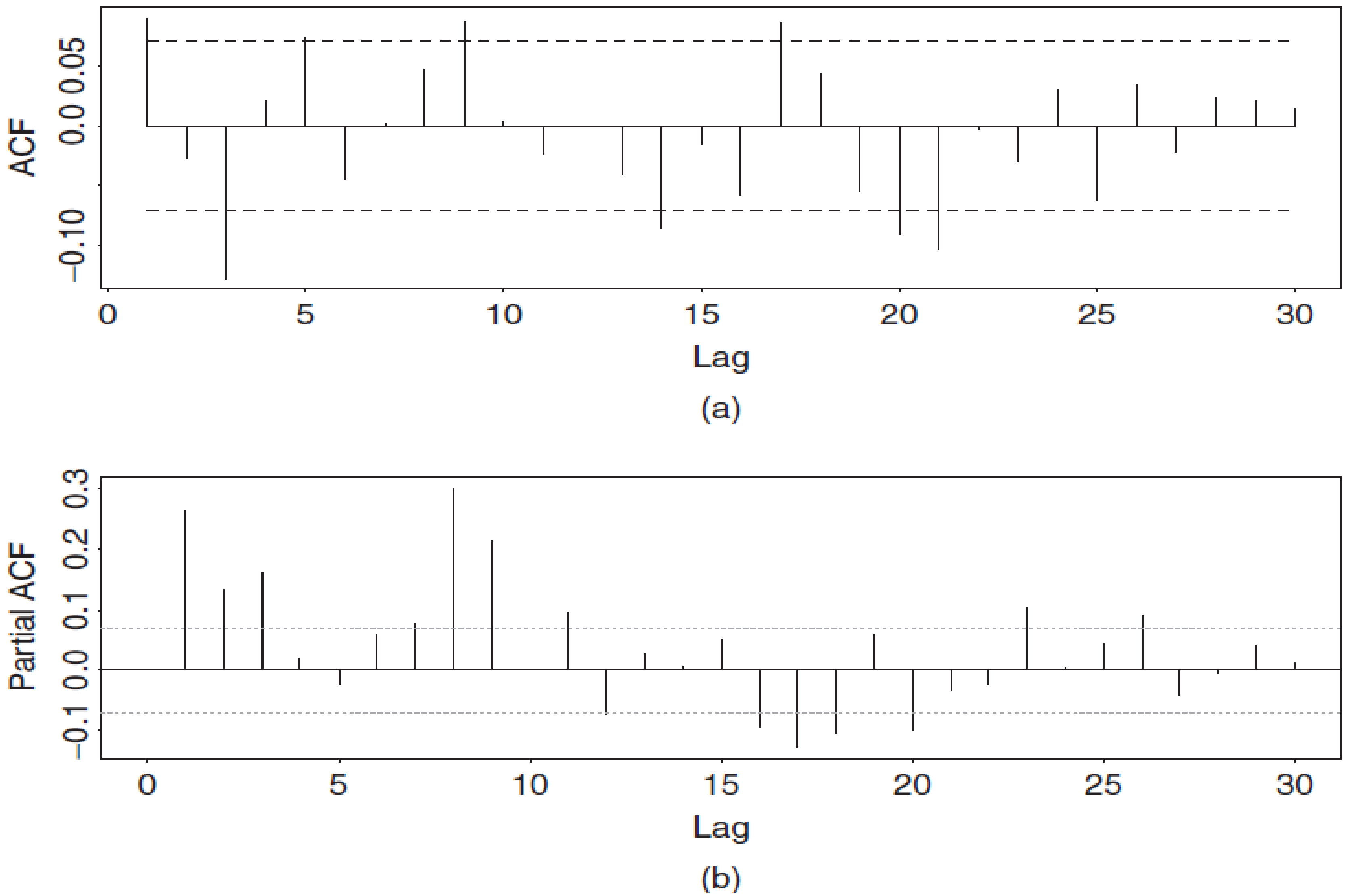
$$V = [I_{2D}(\vartheta)\, I_{OP}^{-1}(\vartheta)\, I_{2D}(\vartheta)]^{-1}$$

# Estimating GARCH Models

- All three parameters in GARCH(1, 1) should be positive;
- The sum of alpha and beta should be less than 1 in order to make sure that volatility process is stationary. But it is very close to one, indicating the volatility is very persistent.
- The estimated unconditional variance should be close to the data variance.

# An Example

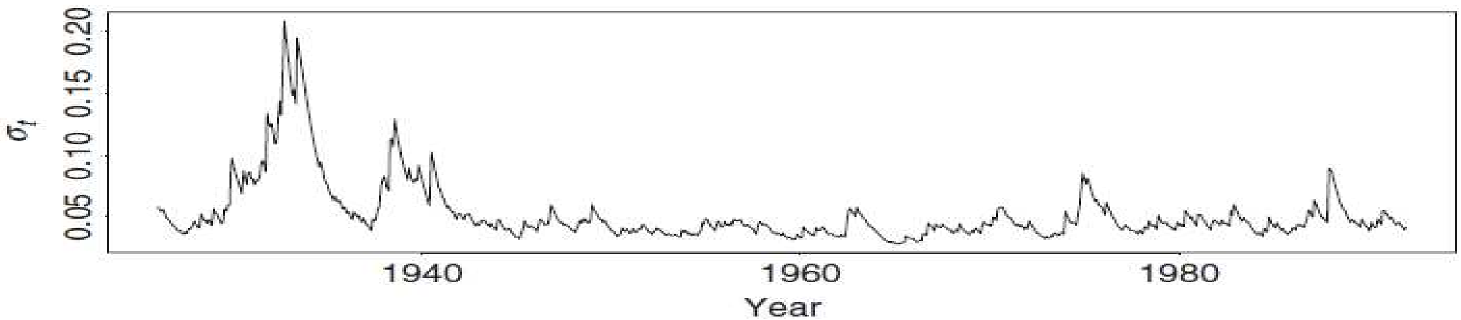



ACF
−0.10
0.0
0.05
0
5
10
15
20
25
30
Lag
(a)
Partial ACF
−0.1
0.0
0.1
0.2
0.3
0
5
10
15
20
25
30
Lag
(b)

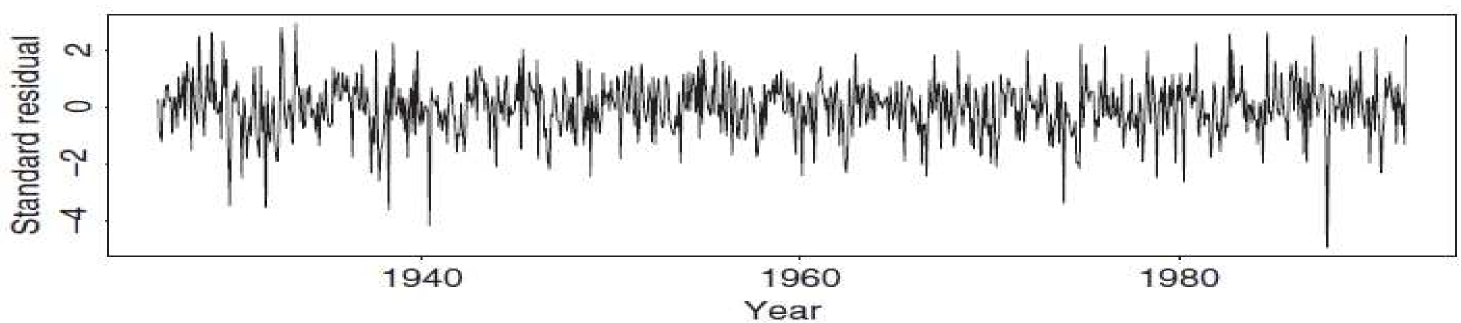
- We use a joint AR(3)-GARCH(1, 1) model

$$r_t = 0.0078 + 0.032 r_{t-1} - 0.029 r_{t-2} - 0.008 r_{t-3} + a_t,$$

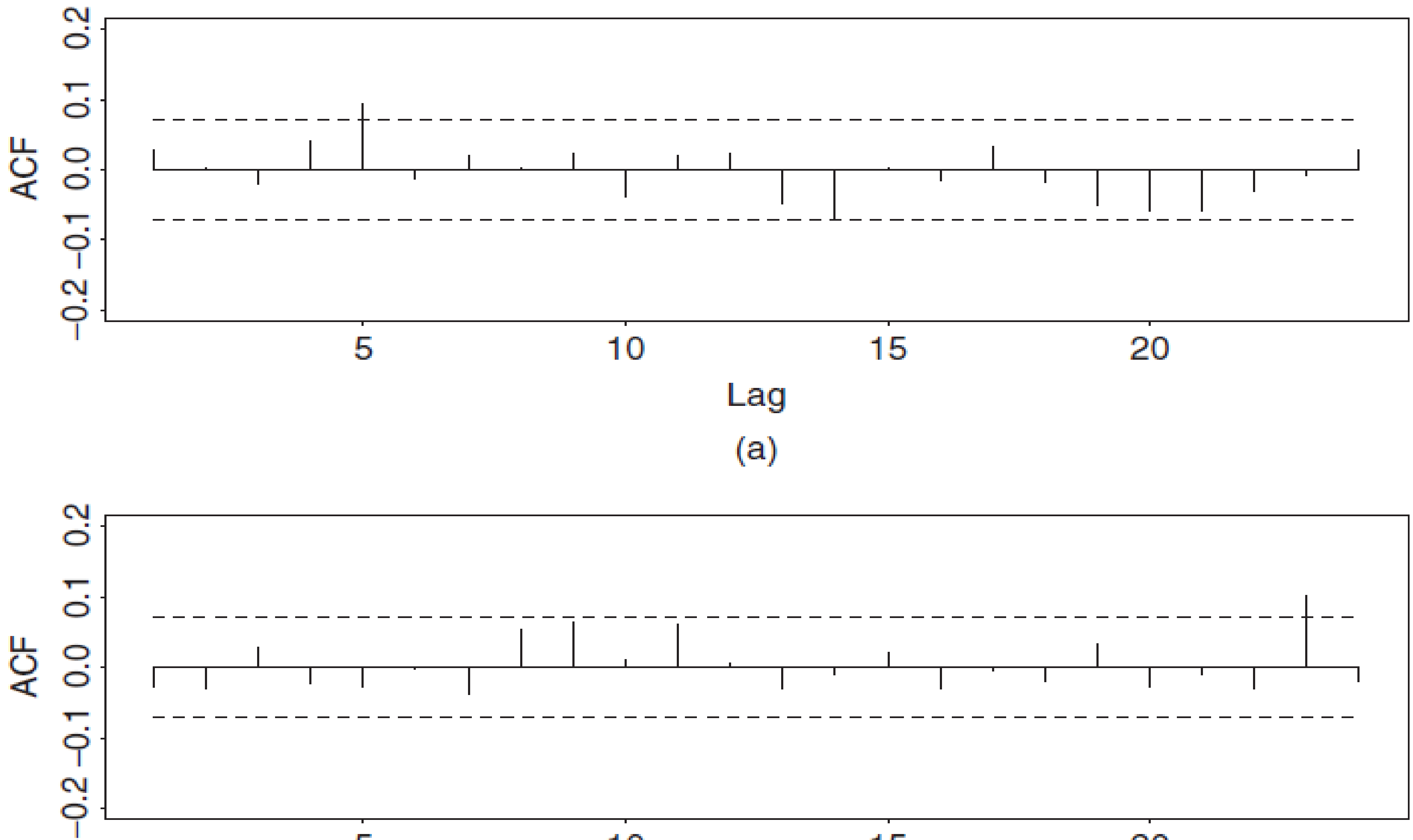
$$\sigma_t^2 = 0.000084 + 0.1213 a_{t-1}^2 + 0.8523 \sigma_{t-1}^2.$$




(a)




(b)




(a)

(b)

# Forecasting for GARCH Models

- One-step ahead forecast

$$h_{t+1} = \omega + \alpha\, \varepsilon_t^2 + \beta\, h_t$$

- Two-step ahead forecast

$$h_{t+2} = \omega + \alpha\, \varepsilon_{t+1}^2 + \beta\, h_{t+1}$$

$$E[h_{t+2} | F_t] = \omega + (\alpha + \beta)\, h_{t+1}$$

$$= \sigma^2 + (\alpha + \beta)\, (h_{t+1} - \sigma^2)$$

- Multi-step forecast

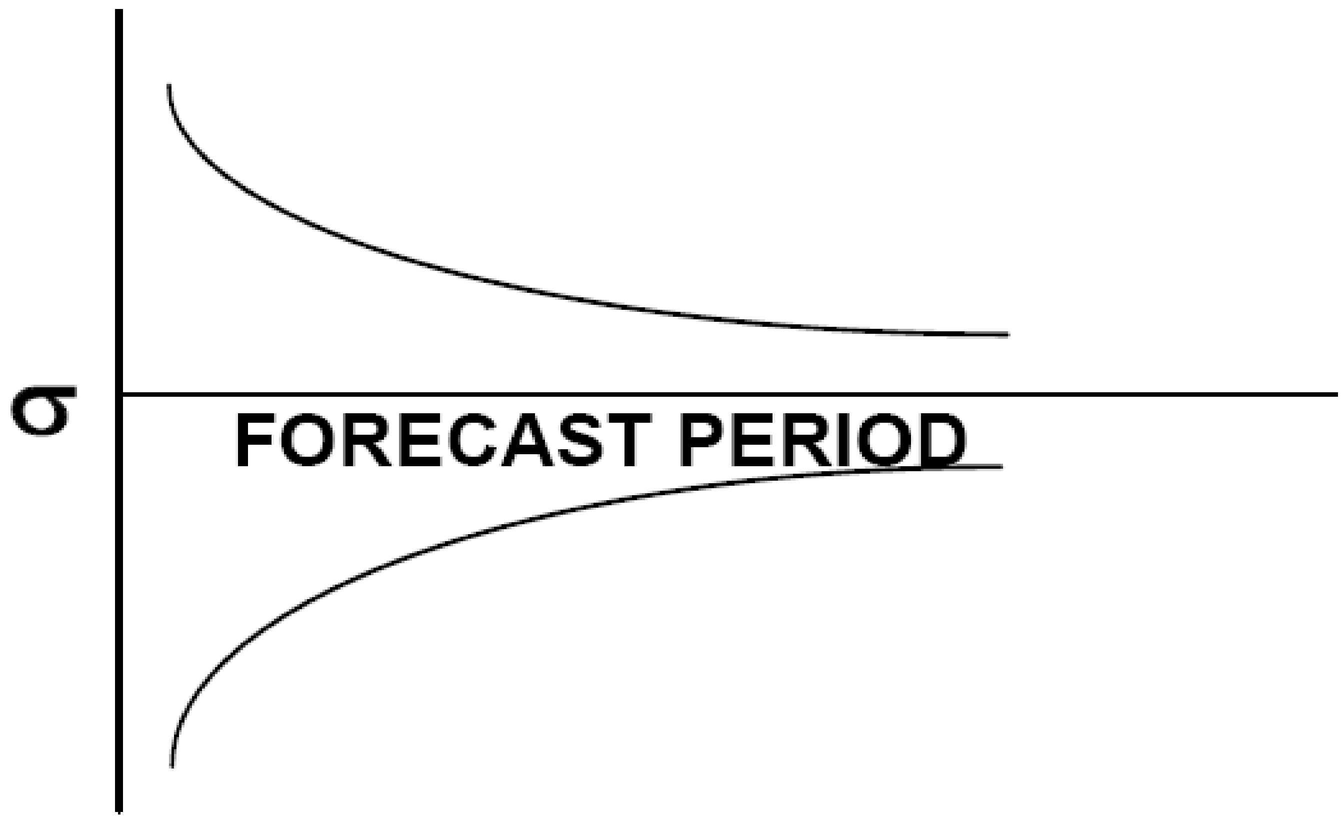
$$
\begin{aligned}
E[h_{t+k} | F_t] &= \omega + (\alpha + \beta)\, E[h_{t+k-1} | F_t] \\
&= \sigma^2 + (\alpha + \beta)\, (E[h_{t+k-1} | F_t] - \sigma^2) \\
&= \sigma^2 + (\alpha + \beta)^{k-1} (h_{t+1} - \sigma^2)
\end{aligned}
$$

- Forecasts converge to the same value no matter what the current volatility is

$$
\begin{aligned}
E[h_{t+k} | F_t] &= \sigma^2 + (\alpha + \beta)^{k-1} (h_{t+1} - \sigma^2) \\
&\rightarrow \sigma^2 \quad \text{if } \alpha + \beta < 1
\end{aligned}
$$

- Little or no updating for Long-horizon volatility

# Term Structure of Volatility

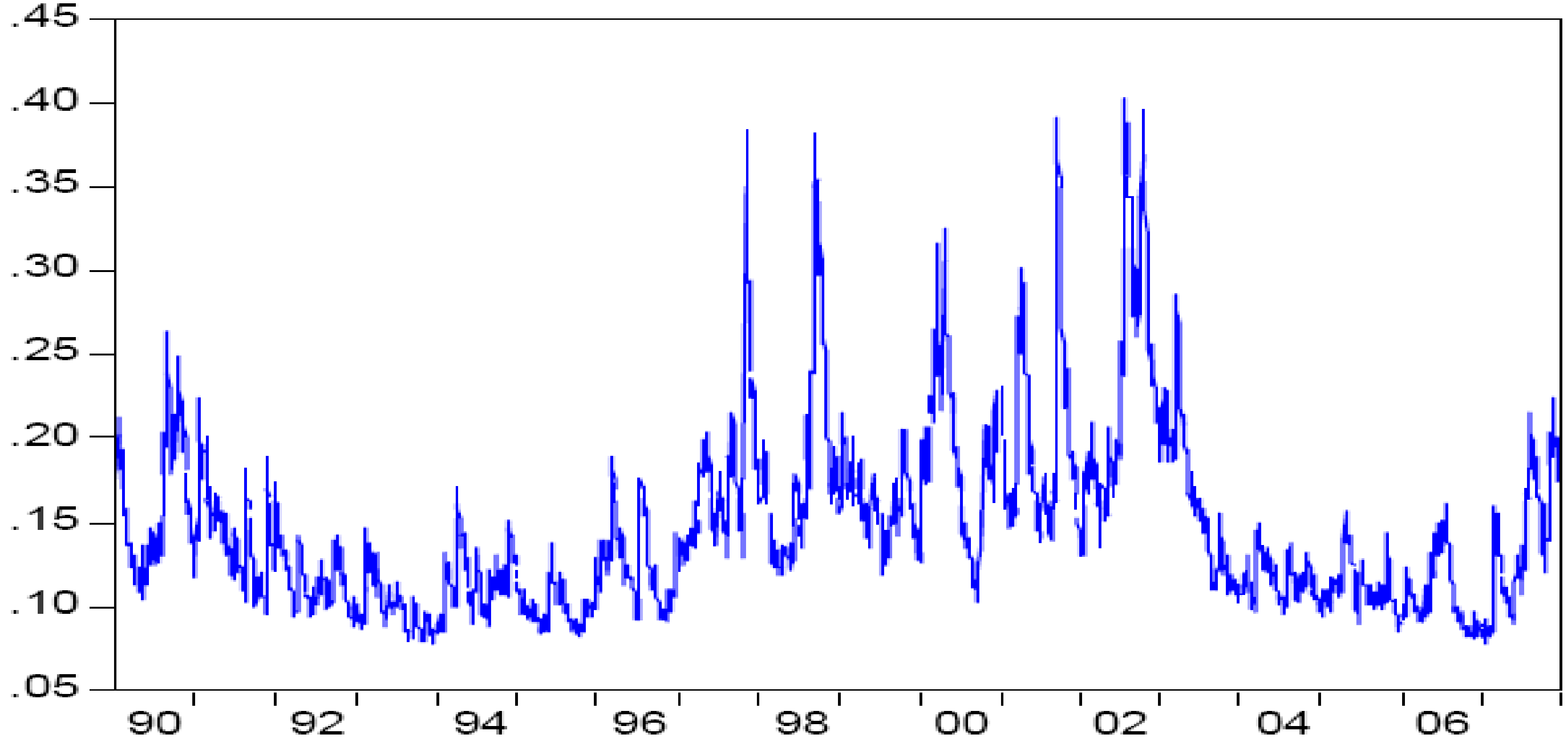


# An Example: Dow-Jones 1990-2008

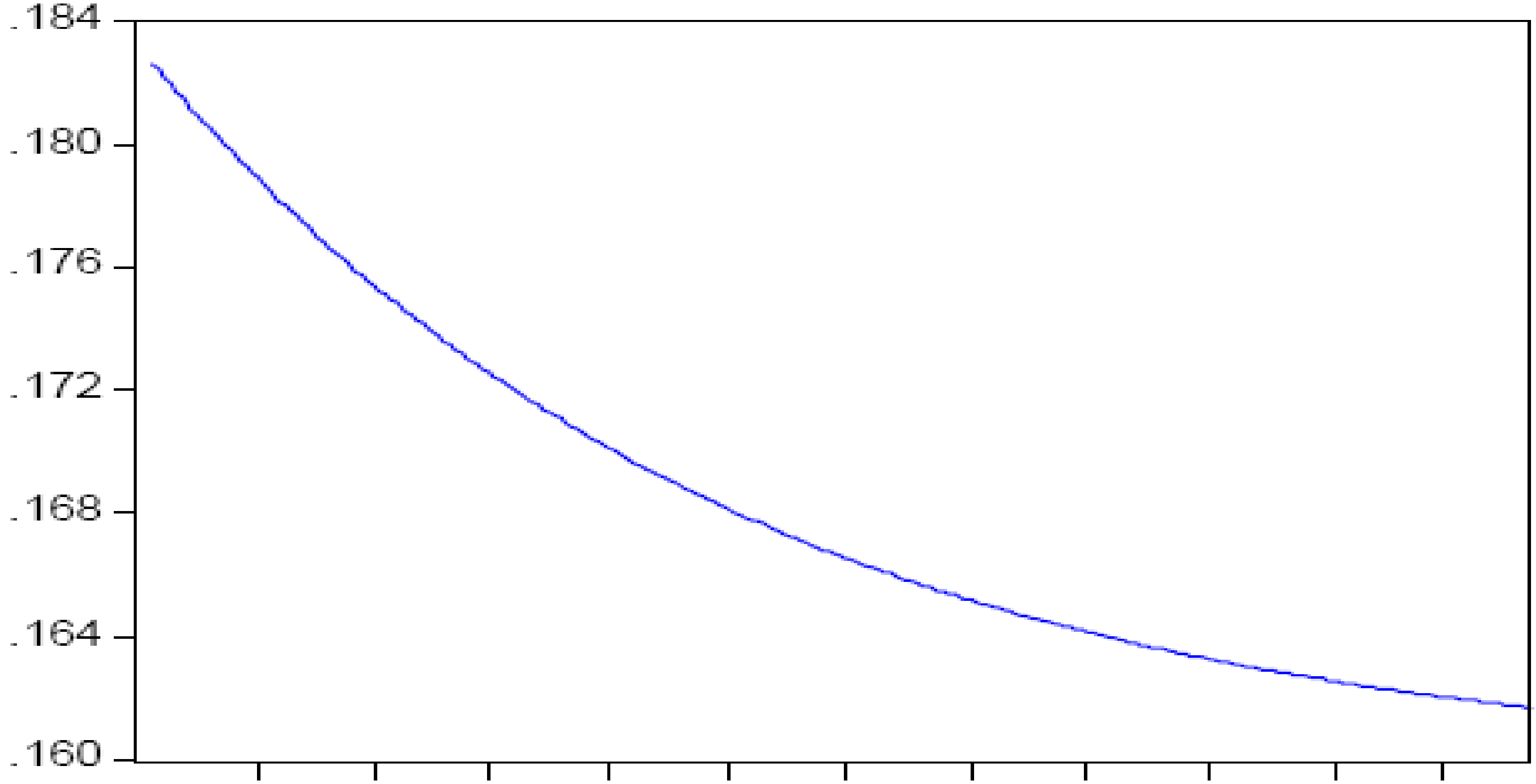
Date: 01/10/08 Time: 13:42
Sample: 1/02/1990 1/04/2008
Included observations: 4541
Convergence achieved after 15 iterations
GARCH = C(2) + C(3)*RESID(-1)^2 + C(4)*GARCH(-1)

| | Coefficient | Std. Error | z-Statistic | Prob. |
|---|---|---|---|---|
| C | 0.000527 | 0.000119 | 4.414772 | 0.0000 |
| Variance Equation | | | | |
| C | 1.00E-06 | 1.37E-07 | 7.290125 | 0.0000 |
| RESID(-1)^2 | 0.064459 | 0.004082 | 15.79053 | 0.0000 |
| GARCH(-1) | 0.925645 | 0.005025 | 184.2160 | 0.0000 |

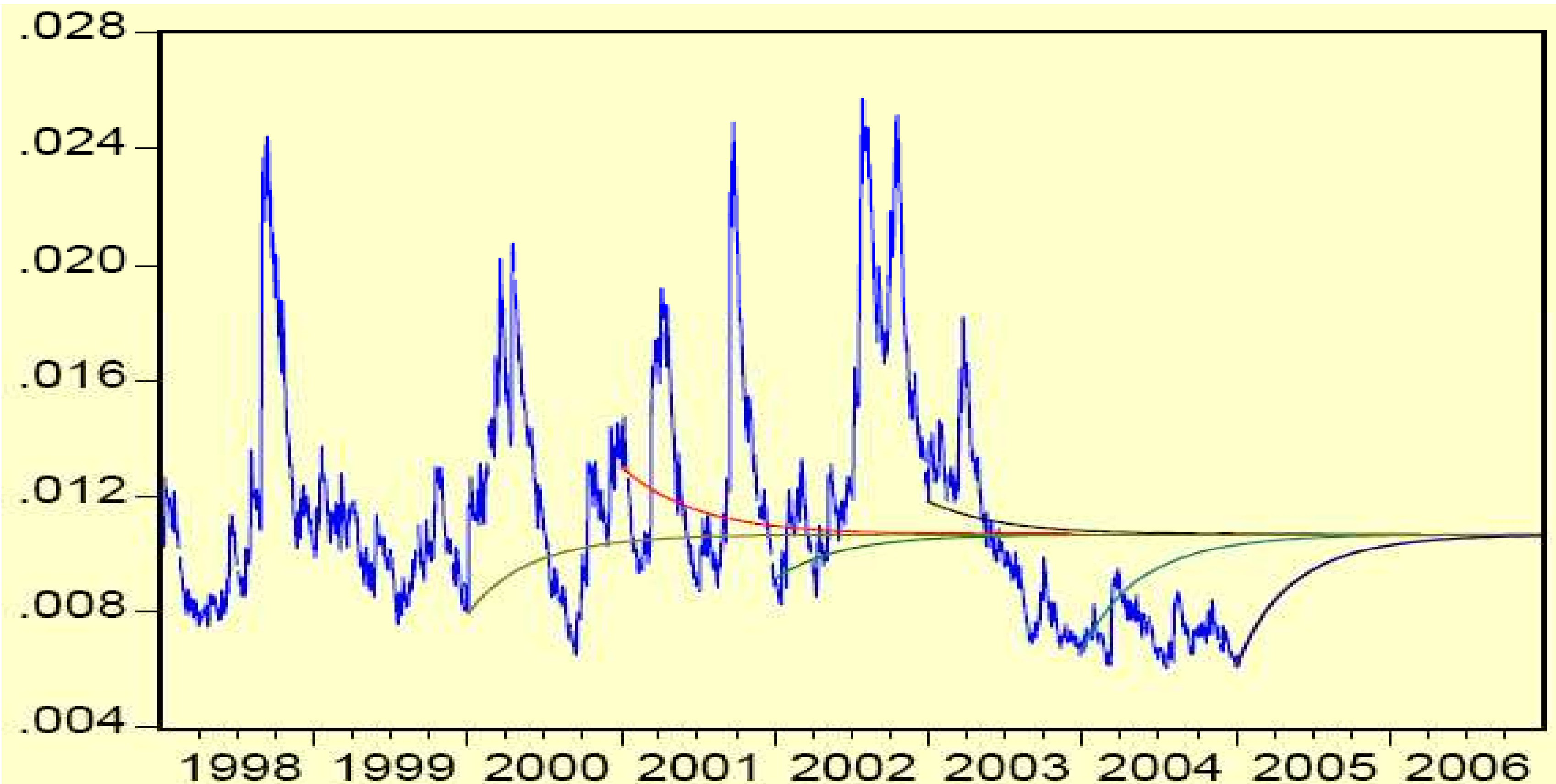
# Volatility Estimate



# Volatility Forecast



# Volatility Forecast



# Another Look at GARCH Model

- Define the forecast error as

  $v_t = \varepsilon_t^2 – h_t = h_t(z_t^2 - 1)$
- $v_t$ is a white noise:
  - Mean: $E[v_t] = 0$
  - Covariance: $Cov(v_t , v_s) = 0$, for $t \neq s$
- $h_t = (r_t - \mu)^2 - v_t$
- $(r_t - \mu)^2 = \omega + (\alpha + \beta) (r_{t-1} - \mu)^2 + v_t – \beta v_{t-1}$
- ARMA(1, 1): $(\alpha + \beta) < 1$ for stationarity.

# The GARCH-M Model

- In finance, the return of a security may depend on its volatility.
- To model such a phenomenon, we may consider the GARCH-M model, where M stands for GARCH *in the mean*.
- The GARCH(1,1)-M model is

$r_t = \mu + \gamma h_t + \varepsilon_{t,} \quad \varepsilon_t$ is $N(0, h_t)$

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

- The parameter γ is called the risk premium parameter.
- The existence of risk premium is, therefore, another reason that some historical stock returns have serial correlations.

$$r_t = 0.0055 + 1.09\sigma_t^2 + a_t, \qquad \sigma_t^2 = 8.76 \times 10^{-5} + 0.123a_{t-1}^2 + 0.849\sigma_{t-1}^2,$$

| $g(\sigma_t)$ | Command |
|---|---|
| $\sigma_t^2$ | var.in.mean |
| $\sigma_t$ | sd.in.mean |
| $\ln(\sigma_t^2)$ | logvar.in.mean |

# Non-Normal Distributions

- Student *t* distribution:

  $f(z) = c(v)[\ 1 + z^2/(v-2)]^{-(v+1)/2}$

  Where

  $c(v) = \Gamma(0.5(v+1))/[\Gamma(0.5v)\mathrm{sqrt}(\pi(v-2))]$

  – $v > 2$ is the degree of freedom parameter
  – The condition for a finite moment of order n is $n < v$. In particular, the kurtosis is finite when $v > 4$;
  – As $v \to \infty$, it converges to the standard normal

- The generalized error distribution (GED)

$$f(z) = C(\eta)\exp\left(-0.5\left|\frac{z}{\lambda(\eta)}\right|^{\eta}\right),$$

*where*

$$C(\eta) = 2^{-1/\eta}\left[\frac{\Gamma(\eta^{-1})}{\Gamma(3\eta^{-1})}\right]^{1/2}, \lambda(\eta) = \frac{\eta}{2}\left[\frac{\Gamma(3\eta^{-1})}{\Gamma(\eta^{-1})^{3}}\right]^{1/2}$$

- The parameter η is positive. It becomes standard normal when η = 2.
- It has fatter tails than the normal when η < 2.

# What is the Best Model?

- The most reliable and robust model is GARCH(1, 1);
- A student-t error assumption gives better estimates;
- For equities, asymmetry is always important. However,
  - Both normal and student-t are symmetric, and our models can not generate skewed distribution.

# Asymmetric Volatility

- Often negative return shocks have a bigger effect on volatility than positive return shocks.
- GJR-GARCH Model

  $h_t = \omega + \alpha\ \varepsilon_{t-1}^2 + \gamma S_{t-1}\ \varepsilon_{t-1}^2 + \beta\ h_{t-1}$

  Where $S_{t-1} = 1$ if $\varepsilon_{t-1} < 0$, otherwise, $S_{t-1} = 0$.

  – When the return shock is negative, ARCH parameter becomes $\alpha + \gamma$;

  – When the return shock is positive, ARCH parameter is only $\alpha$;

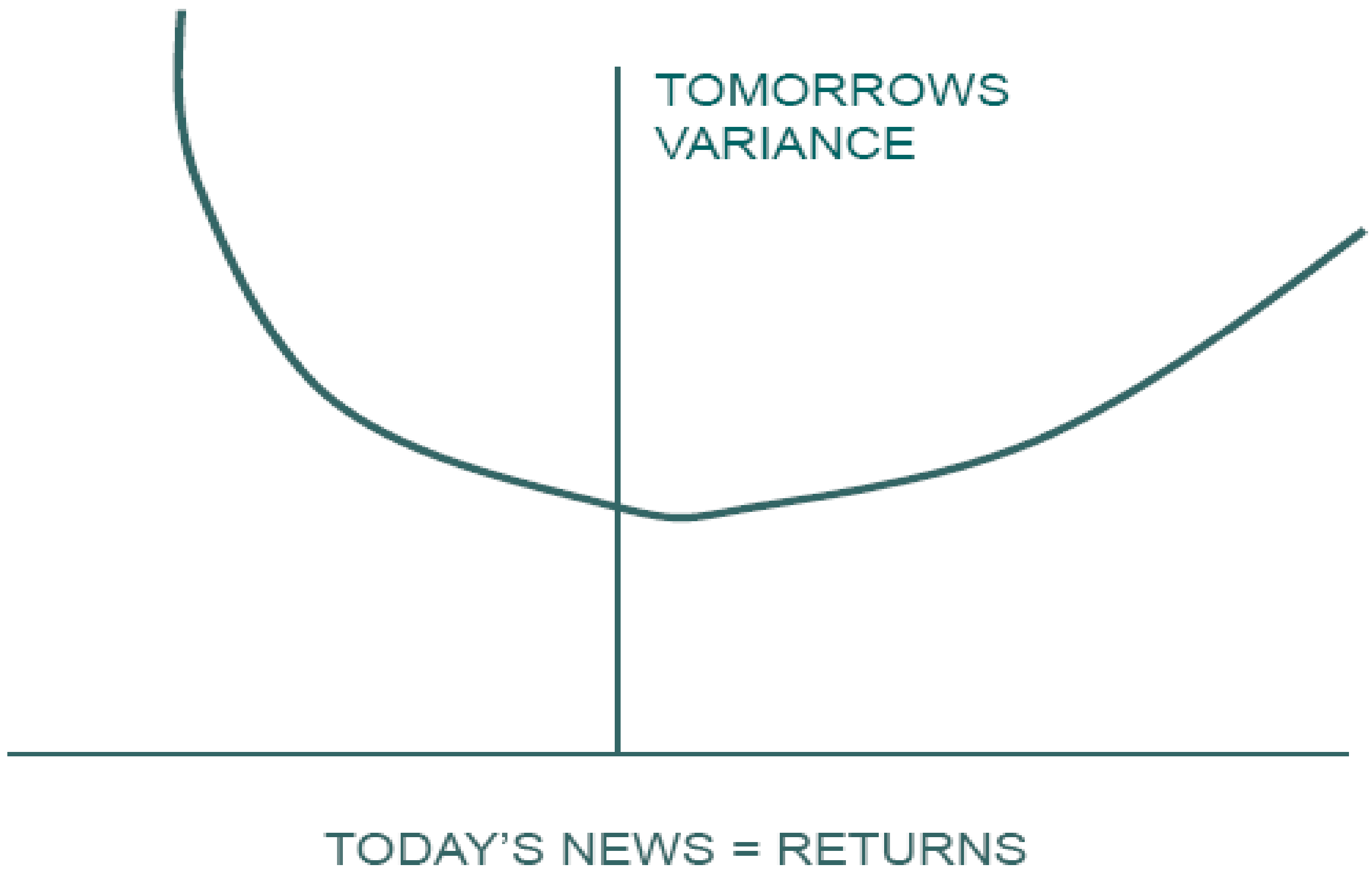
- To obtain theoretical results, we assume that the normalized residuals have symmetric distributions.
  - $E[S_t] = 0.5$;
  - $S_t$ is independent of $z_t$;
- The model can be written

  $$h_t = \omega + [(\alpha + \gamma S_{t-1})z_{t-1}^2 + \beta]\ h_{t-1}$$
- Taking expectation, we have

  $$[h_t] = \omega + [(\alpha + 0.5\gamma) + \beta]\ E[h_{t-1}]$$

  $$\Rightarrow \sigma^2 = \omega/(1 - (\alpha + 0.5\gamma + \beta))$$

# Asymmetric Volatility



# Asymmetric Volatility

- EGARCH (Nelson, 1991)

  $\log(h_t) = \omega + \beta[\log(h_{t-1}) - \omega] + g(z_{t-1})$

  and $g(z_{t-1}) = \alpha z_{t-1} + \gamma[|z_{t-1}| - E([|z_{t-1}|)]$

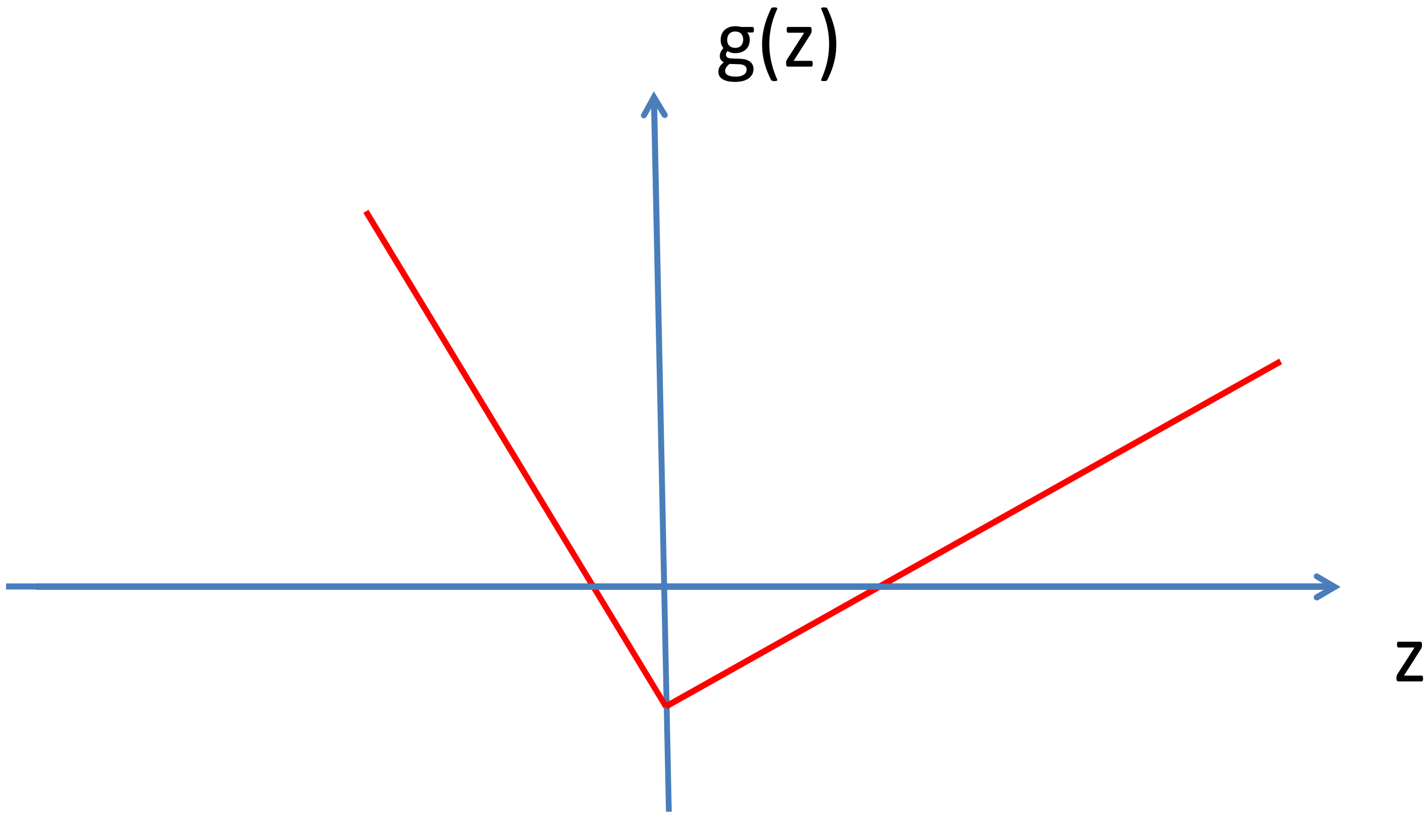
  – For a normal $z_t$, $E(|z_t|) = sqrt(2/\pi)$

  – For a student-*t* $z_t$,

  $E(|z_t|) = 2sqrt(v\text{-}2)\ \Gamma(0.5(v+1))/[sqrt(\pi)\Gamma(0.5v)(v\text{-}1)]$

- *g(z)* is defined by two straight lines that join at *z = 0:*

# The Function of g(z)



# The Function of g(z)

- When $z$ is negative, the function $g(z)$ has a slope $\alpha - \gamma$;
- When $z$ is positive, the function $g(z)$ has a slope $\alpha + \gamma$;
- Empirically, $\gamma$ is negative, indicating that volatility increases more when the market moves downward.

# An Example

```
Conditional Variance Equation: ~ egarch(1, 1)
Conditional Distribution:  ged
 with estimated parameter 1.5003 and standard error 0.09912
-------------------------------------------------------------
Estimated Coefficients:
-------------------------------------------------------------
             Value Std.Error t value  Pr(>|t|)
       C  0.01181  0.002012   5.870 3.033e-09
       A -0.55680  0.171602  -3.245 6.088e-04
 ARCH(1)  0.22025  0.052824   4.169 1.669e-05
GARCH(1)  0.92910  0.026743  34.742 0.000e+00
  LEV(1) -0.26400  0.126096  -2.094 1.828e-02
-------------------------------------------------------------
Ljung-Box test for standardized residuals:
-------------------------------------------------------------
Statistic P-value Chi^2-d.f.
     17.87  0.1195         12

Ljung-Box test for squared standardized residuals:
-------------------------------------------------------------
Statistic P-value Chi^2-d.f.
     6.723  0.8754         12
```

# Forecasting Using EGARCH

- In EGARCH, volatility is in log form. We rewrite the model

$$\log(h_t) = \omega(1 - \beta) + \beta\log(h_{t-1}) + g(z_{t-1})$$

- And take exponentials,

$$h_t = h^{\beta}_{t-1}\exp[\omega(1 - \beta)]\exp[g(z_{t-1})]$$

- 1-step ahead forecast:

$$h_{t+1} = h^{\beta}_{t}\exp[\omega(1 - \beta)]\exp[g(z_t)]$$

- 2-step ahead forecast:

$$h_{t+2} = h^{\beta}_{t+1}\exp[\omega(1 - \beta)]\exp[g(z_{t+1})]$$

$$\Rightarrow E[h_{t+2} \mid F_t] = h^{\beta}_{t+1}\exp[\omega(1 - \beta)]E[\exp(g(z_{t+1})) \mid F_t]$$

$$E[e^{g(z)}] = \int_{-\infty}^{+\infty} \exp[\alpha z + \gamma(|z| - \sqrt{2/\pi})] f(z)dz$$

$$= \exp(-\gamma\sqrt{2/\pi})\left[\int_0^{+\infty} e^{(\alpha+\gamma)z} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz\right]$$

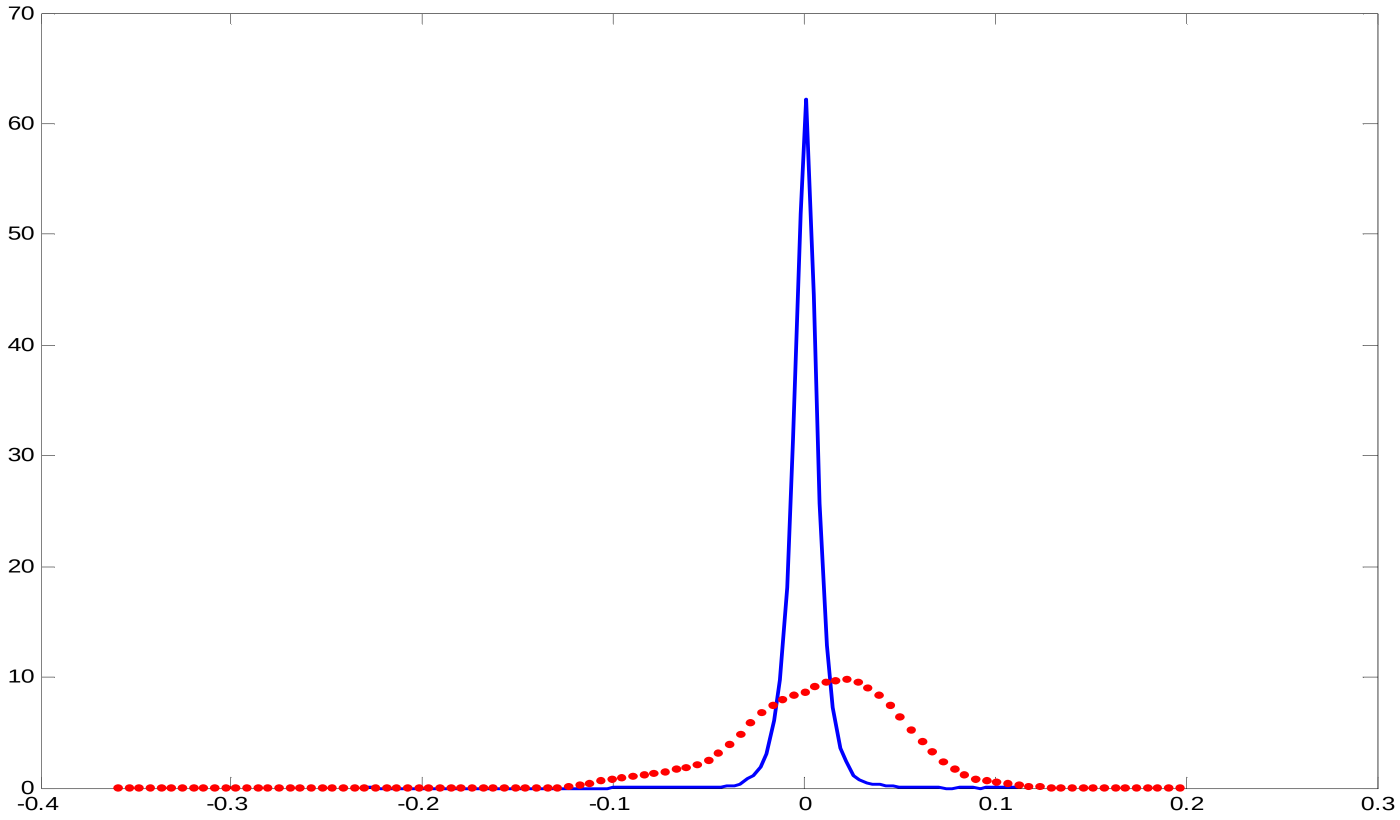
$$+ \exp(-\gamma\sqrt{2/\pi})\left[\int_{-\infty}^{0} e^{(\alpha-\gamma)z} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz\right]$$

$$= \exp(-\gamma\sqrt{2/\pi})\left[e^{(\alpha+\gamma)^2/2}\Phi(\alpha+\gamma) + e^{(\alpha-\gamma)^2/2}\Phi(\alpha-\gamma)\right]$$

# Skewed Return Distribution

- With asymmetric volatility, the return distribution is asymmetric, and empirically has a longer left tail.
- For long horizons, the central limit theorem will reduce this effect and returns will be approximately normal.
- With different data frequency, you may choose different models.

# Skewed Return Distribution



# Where is Asymmetric Volatility From

- Leverage effect:
  - As equity prices decrease, the leverage of a firm increases so that the next shock has higher volatility on stock prices
- Risk Aversion:
  - News of a future volatility event will lead to stock sell and price declining. Since events are clustered, any news event will predict higher volatility in the future.

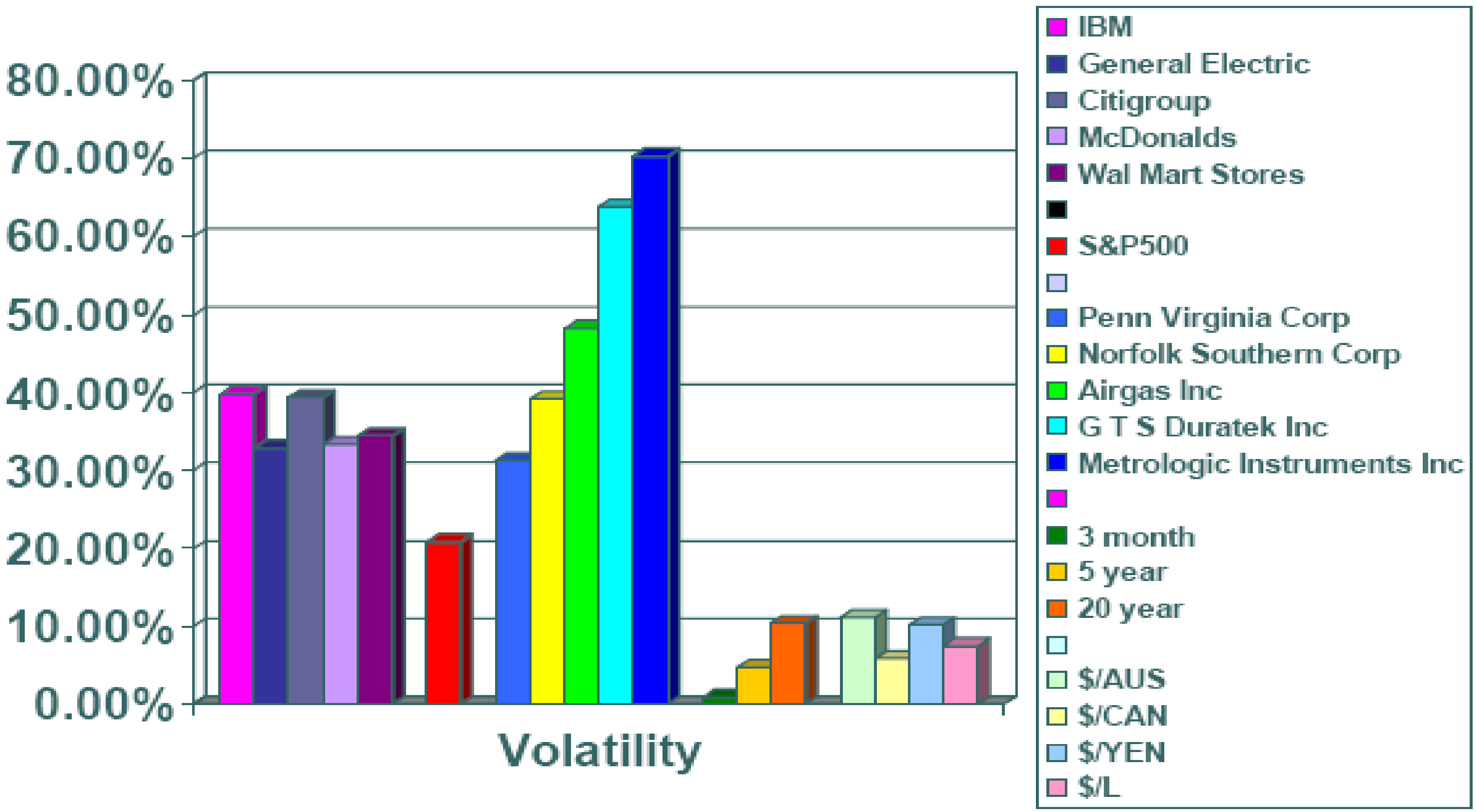
# Why Makes Prices and Volatility Move

- New information on future values moves prices:
  - Volatility is high when there is a lot of new information.
- Trading can move prices, but mostly because it reveals information known to the traders
  - Trading volume
- Volatility reflects the frequency and importance of the news:
  - More important for small stocks than large stocks
  - More important for individual stocks than indices

# What Makes Financial Market Volatility High

- The flow of new information on the macro-economy:
  - High inflation
  - Slow output growth and recession
  - High volatility of short term interest rates
  - High volatility of output growth
  - High volatility of inflation

# Volatility by Asset Class



# Empirical Application

- S&P 500 index returns