TIME SERIES MODELING AND FORECASTING

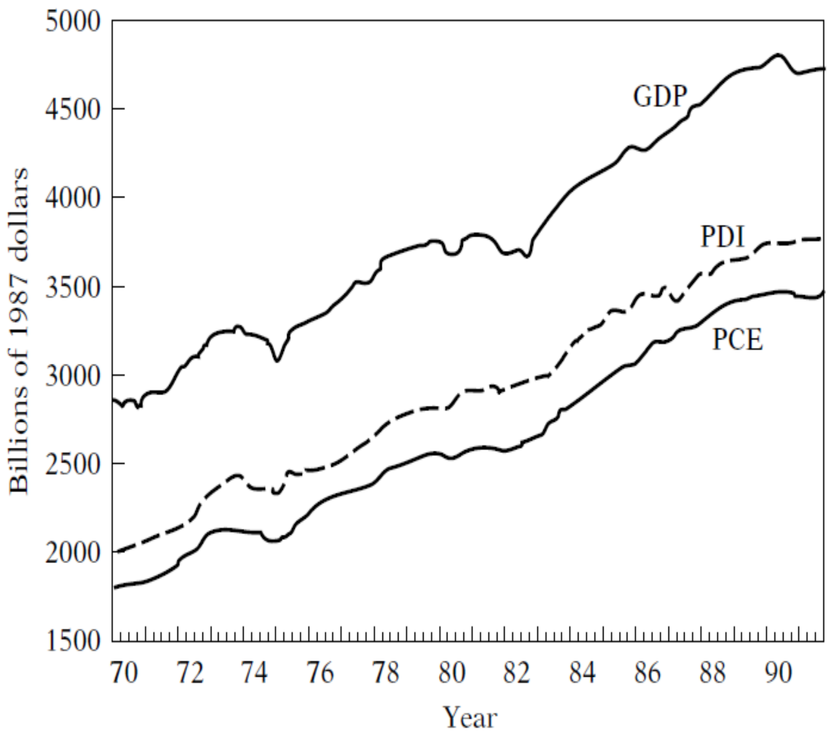
Sunny Kumar Singh

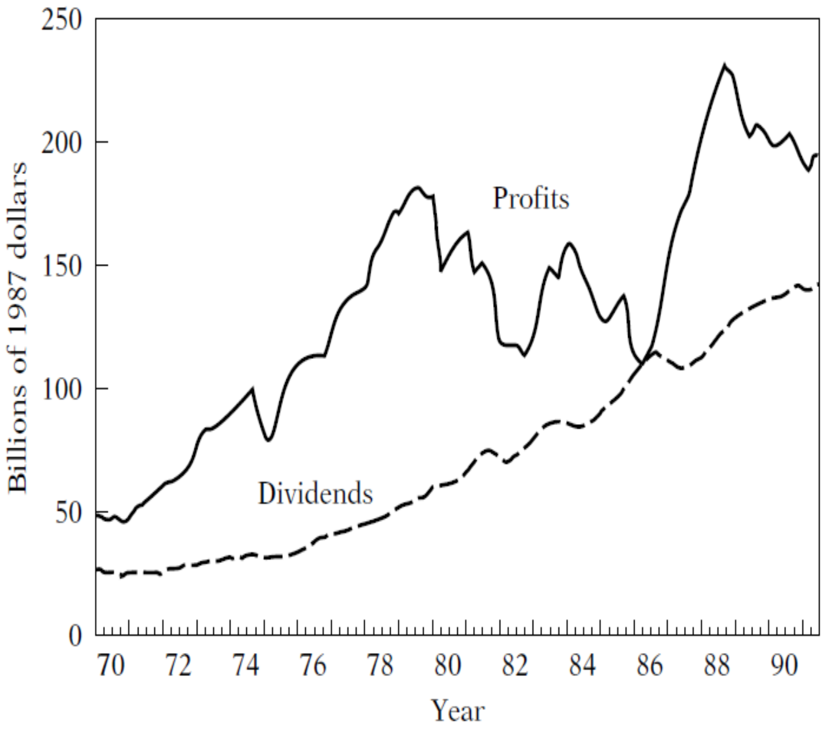
BITS Pilani, Hyderabad

Plan of presentation

- Overview of Time Series Data
- Types of Time Series Models
- Importance of Stationarity in Time Series Analysis
- Univariate Time Series Modelling
  - Modelling Returns: AR, MA, ARIMA
  - Modelling Volatility: Unconditional (Standard Deviation, Extreme Value Estimators) and Conditional Volatility (EWMA, ARCH, GARCH, E-GARCH)

Some example of economic & financial time series data





TIME SERIES ANALYSIS

- Sometimes, we are interested in examining behaviour of a series/ variable (say, y_t), for which it is very difficult to come out with a structural model (e.g., $y_t = b_0 + b_1 * x_t + error_t$) because of:
  - *Measurement of explanatory variables needed to explain this variable is difficult;*
  - *Frequency of such explanatory variables is not in tune with the variable to be explained (y_t). For example, we are interested in studying some variable based on daily data, however, variables needed are available at very low frequency (say, monthly, Quarterly, etc.).*
  - *Further, it is observed that structural models are not very good for 'out-of-sample' forecast.*

Time series analysis

- For example, we are interested in modeling daily stock returns. We would require some variables like ratios and other macroeconomic variables; these are not available on daily basis. Besides, other variables capturing psychological aspects of investors will be very difficult to measure. Therefore, 'structural models' cannot be applied in such situations.

- **Consequently, we move to another class of models, called as TIME SERIES MODELS. Such models utilize past realisations/ values of a series to model and forecast it's behaviour over time.**

- **Financial return & volatility are usually modeled using a class of such time series models, known as *'Univariate Time Series models'*. That is, we use only one series to develop the model, e.g., return & volatility are modeled based on their own past realisations/ values.**

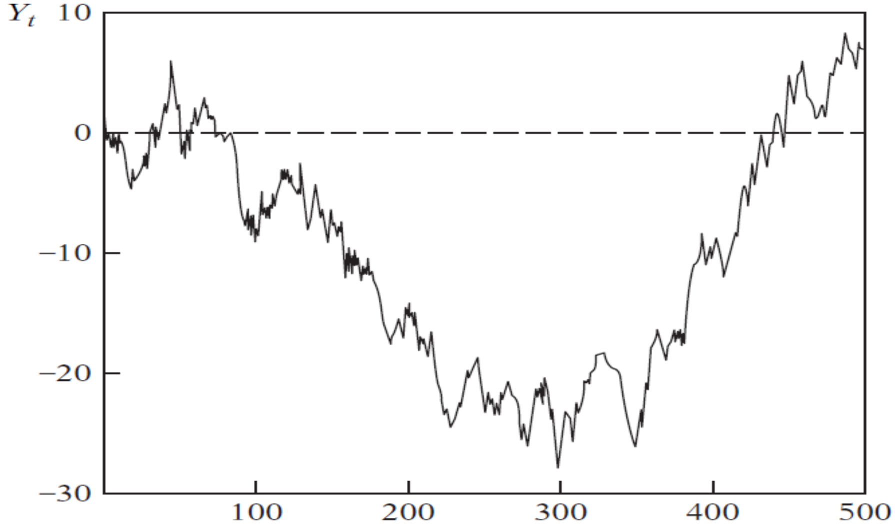
TIME SERIES ANALYSIS

- ***NOTE: Unlike structural models, Time series models, do not necessarily build on some economic reasoning to model the relationship.***

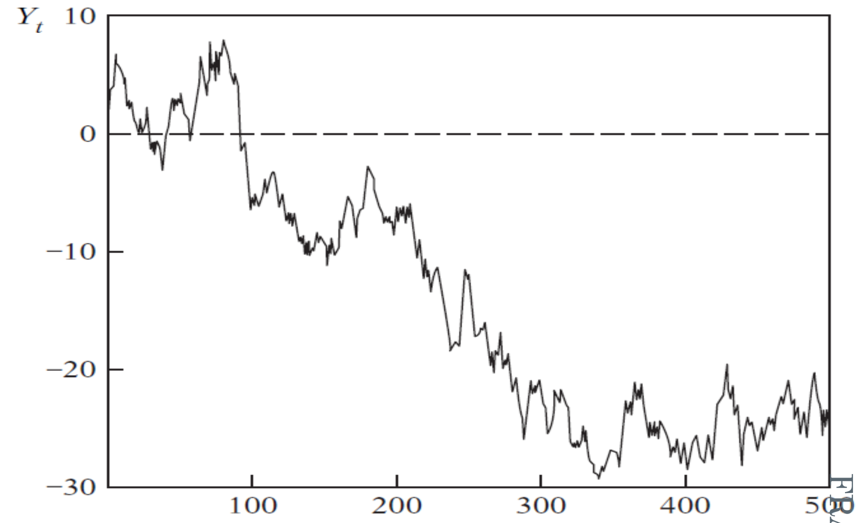
- **Such models builds more on a data-driven approach. That is, in general, these models attempt to utilize characteristics of data to fit the model rather than their economic reasoning.**

Some key concepts

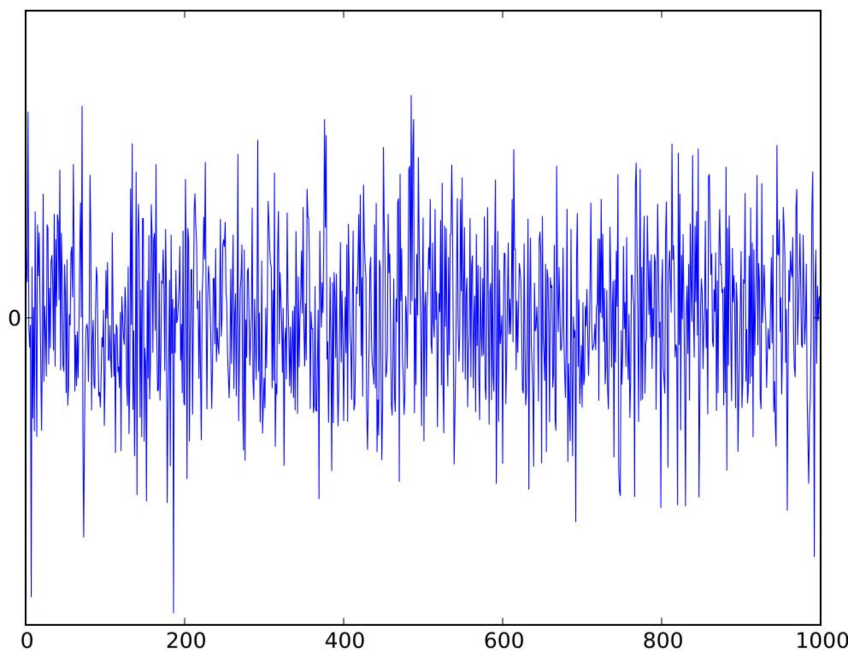
- Stochastic Processes
- Stationary Processes
- Purely Random Processes
- Nonstationary Processes
- Random Walk Models
- Deterministic and Stochastic Trends

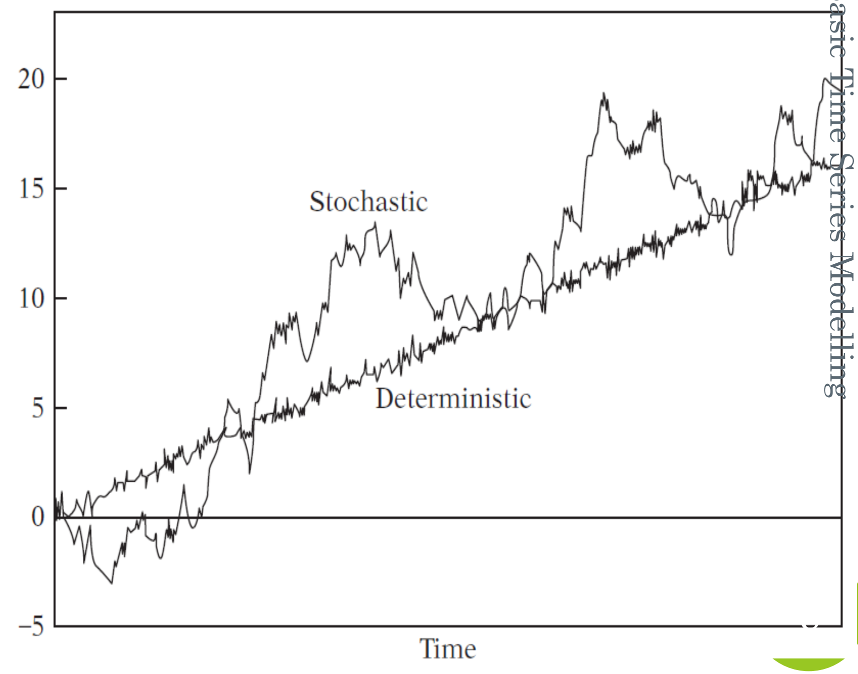


$Y_t = Y_{t-1} + u_t$



$Y_t = 2 + Y_{t-1} + u_t$





Types of Time Series Models

- **Stationary Time Series Models (Modelling Return and Risk)**
  - For example: AR, MA, ARMA, ARIMA, EWMA, ARCH, GARCH, EGARCH, GJR-GARCH, etc.

- **Non-stationary Time Series Models**
  - For Example: COINTEGRATION & VECTOR ERROR CORRECTION MODELS (VECM)

Stationarity condition for Auto-Regressive (AR) models

- **Condition of Stationarity simply means that the series is fairly stable over time and, therefore, the *history is useful for future predictions*** *as, for example,* we attempt to predict returns using only information contained in their past values.

- ***Stationarity condition is needed to ensure STABILITY of analysis.***

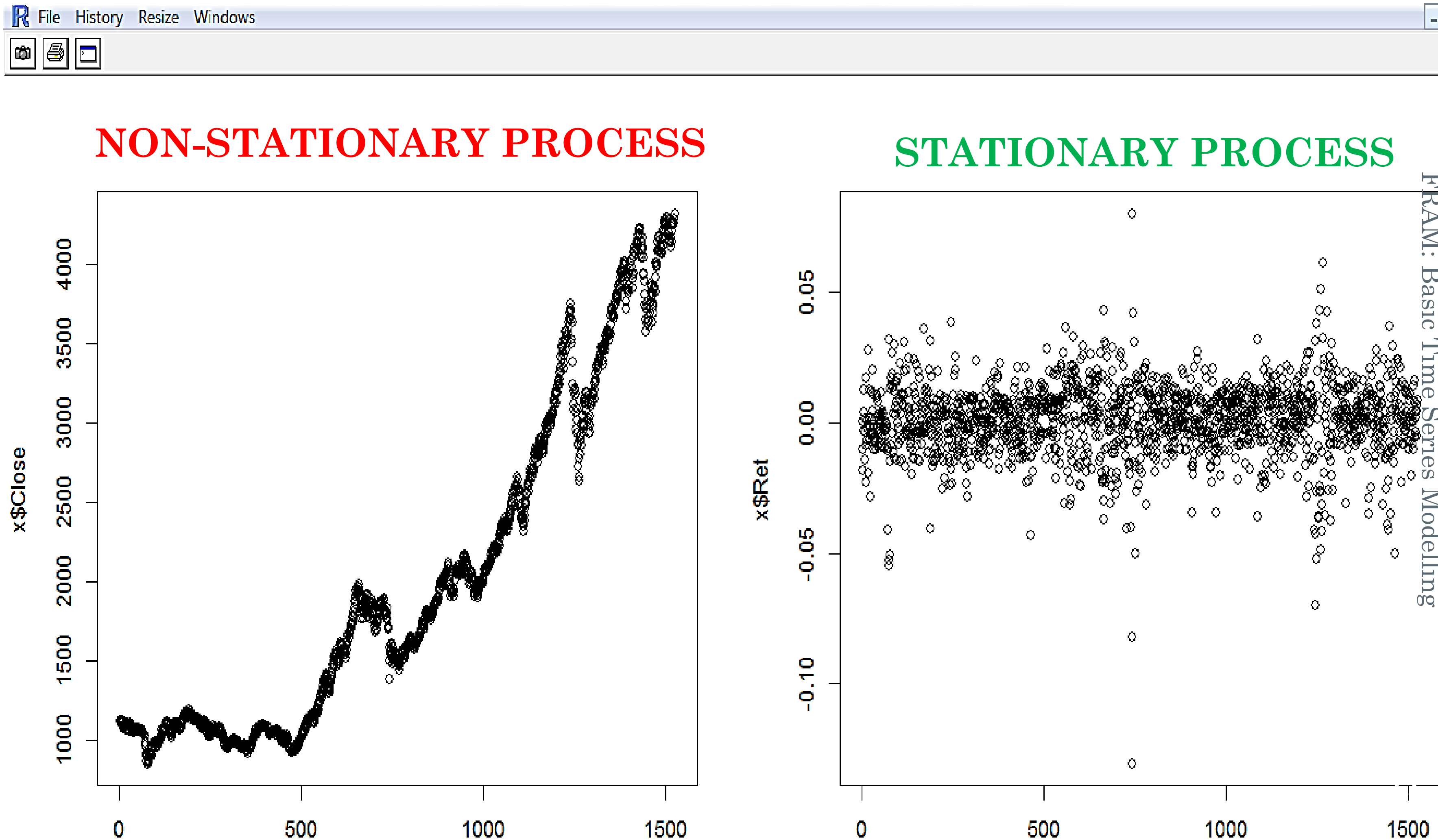
Stationarity Conditions:

(A) Strict Stationary: All the four moments defining distribution of observations are time invariant. That is, the joint probability of observing n values is the same irrespective of time.

$$P\{y_{t_1} \leq b_1, \ldots, y_{t_n} \leq b_n\} = P\{y_{t_1+m} \leq b_1, \ldots, y_{t_n+m} \leq b_n\}$$

(B) Weakly Stationary: Assumes that only first two moments are time invariant. (Practically applicable)

Non-Stationary Vs. Stationary Series



Stationarity of a Series

- **Why do we need stationarity?**
  - **Stationarity in data leads to stable results. A non-stationary data may lead to instable coefficients (especially in univariate-time series analysis).**

- **How do we detect it?**

1. Correlograms (Visual inspection)

- *ACF Plots*
- *PACF Plots*

2. Formal statistical tests to detect stationarity of a series:

- *Dickey-Fuller (DF test)*
- ***Augmented Dickey-Fuller (ADF test)***

3. Roots or Zeros of 'Lag Polynomial' of a series.

AUTO CORRELATION FUNCTION (ACF) & PARTIAL AUTO CORRELATION FUNCTION (PACF)

- **Autocorrelation function** shows relation between autocorrelation and lags. Autocorrelation between Y_t and its lagged values Y_{t-k} (That is, at lag K) is determined as follows:

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\text{covariance at lag } k}{\text{variance}}$$

$$\hat{\gamma}_k = \frac{\sum (Y_t - \bar{Y})(Y_{t+k} - \bar{Y})}{n}$$

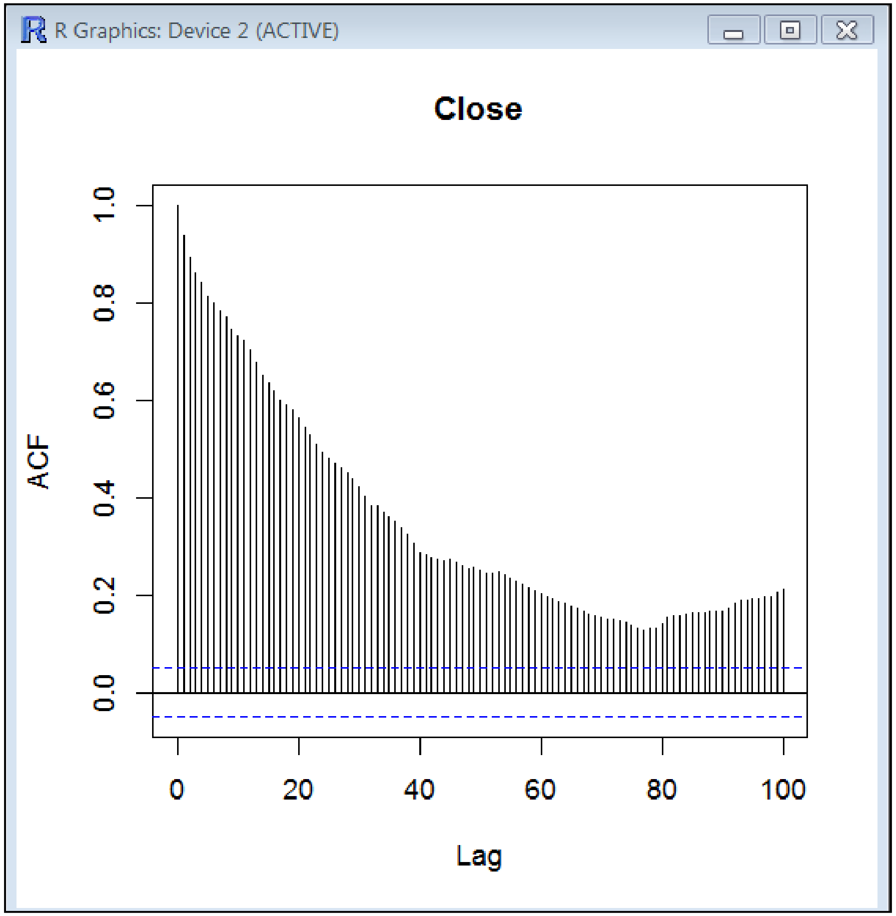
$$\hat{\gamma}_0 = \frac{\sum (Y_t - \bar{Y})^2}{n}$$

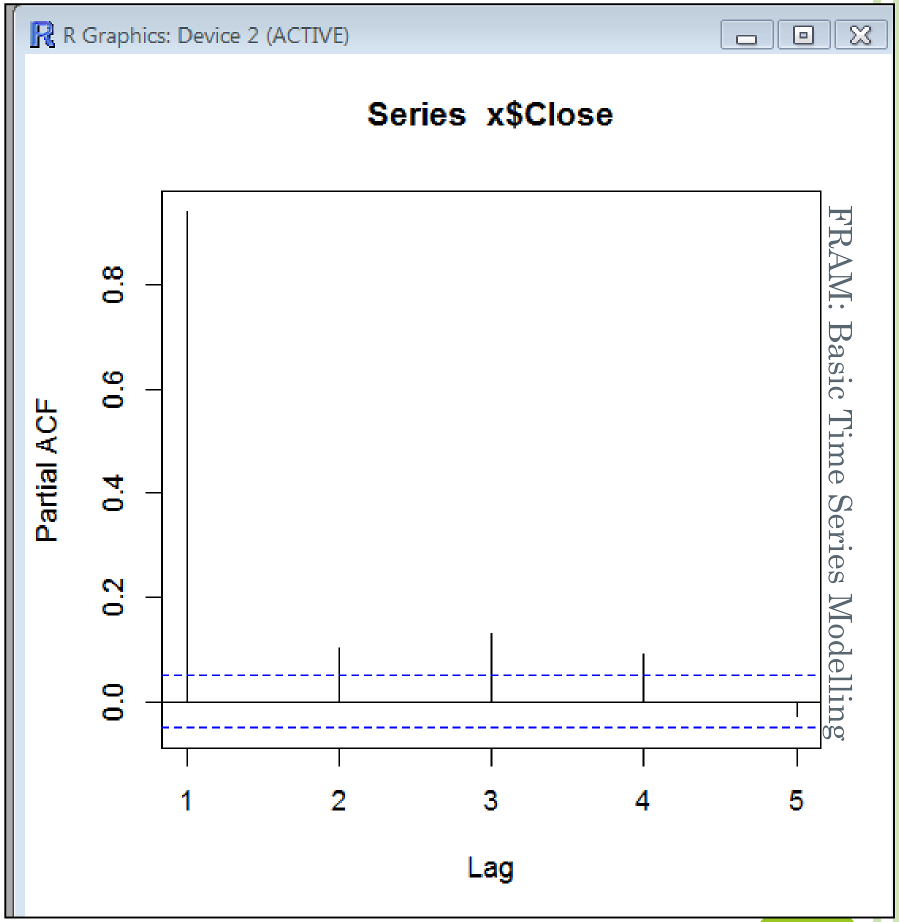
Partial Auto Correlation Function (PACF)

- In time series data, typically, a large proportion of correlation between Y_t & Y_{t-k} happens due to their correlation with intermediate variables.

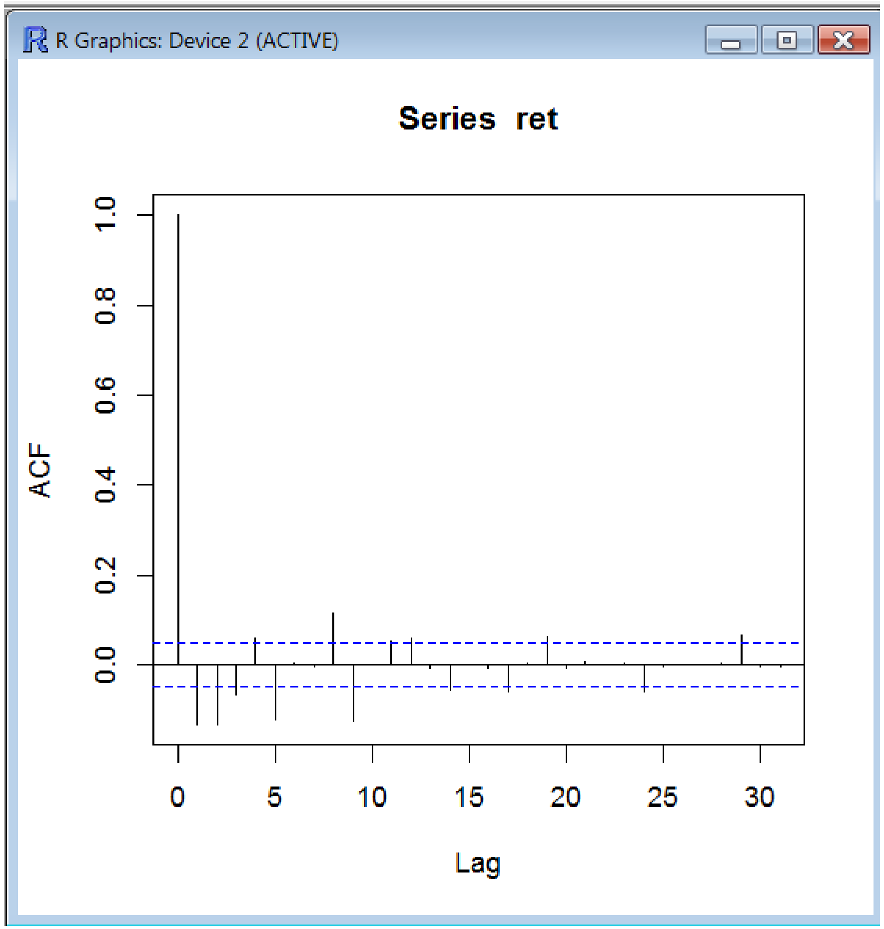
- PACF measures correlation between Y_t and Y_{t-k} after removing the correlation that Y_t could have with the intermediate lags (Y_{t-1}, Y_{t-2}, …, Y_{t-k+1}).

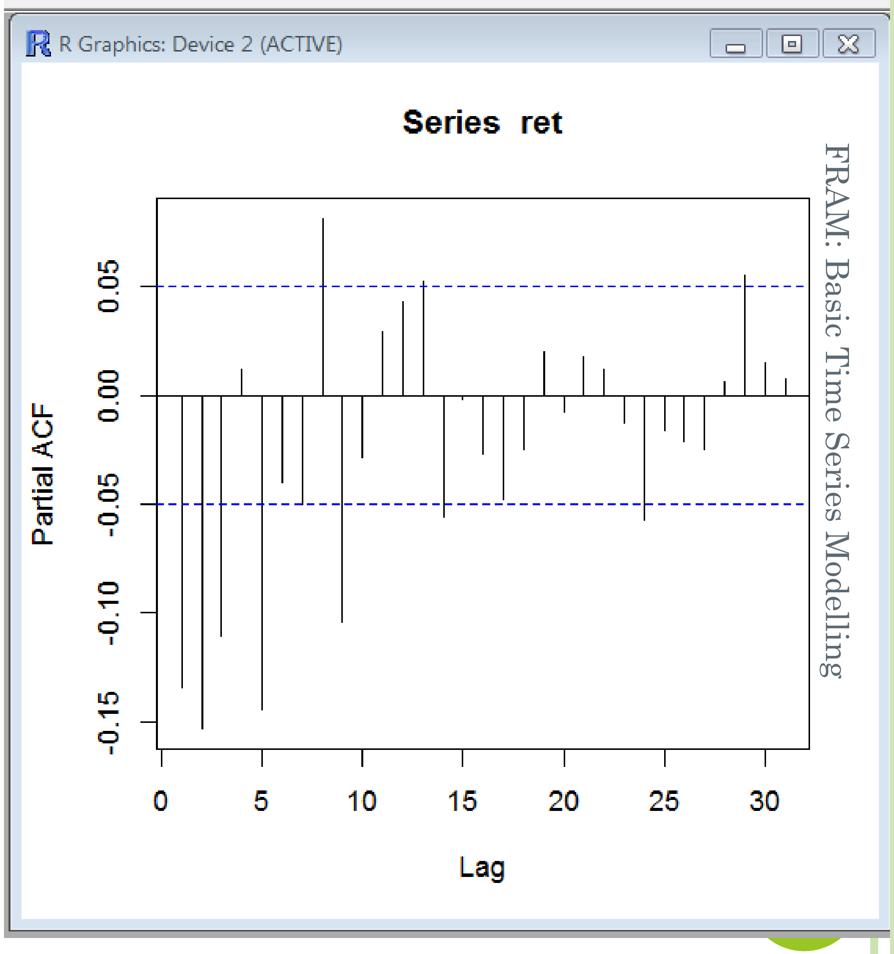
Correlograms (ACF & PACF) of a Non-Stationary Time Series





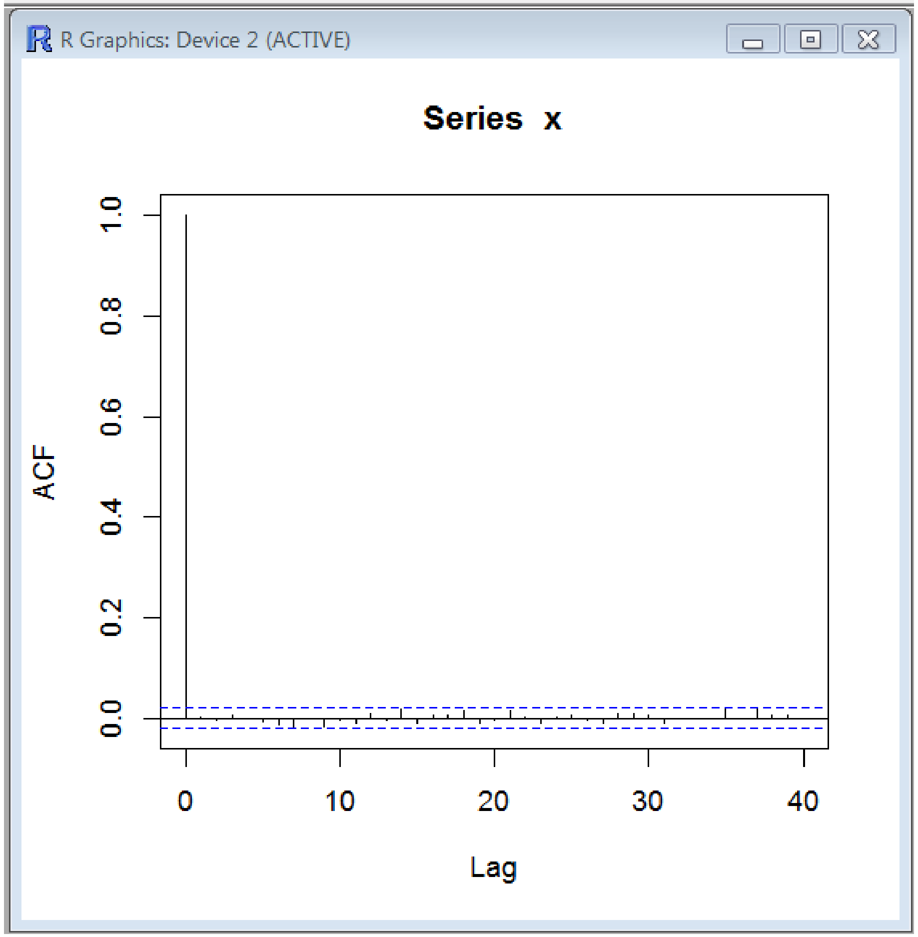
Correlograms (ACF & PACF) of a Stationary Time Series

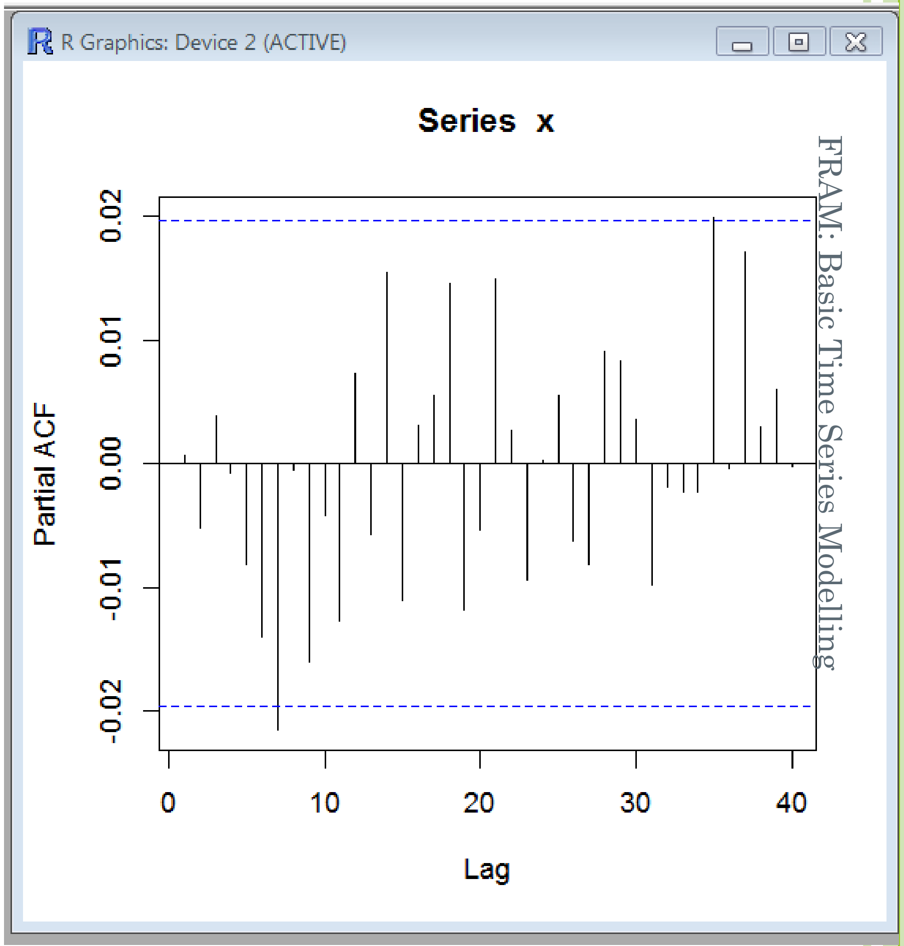




CORRELOGRAMS (ACF & PACF) OF A WHITE NOISE PROCESS

(A STATIONARY PROCESS WITH NO DISCERNIBLE STRUCTURE)





FORMAL TESTS FOR STATIONARITY
DICKEY FULLER & AUGMENTED DICKEY FULLER TESTS

- **AR models & condition for stationarity**
- **AR(1) model** $y_t = \mu + \phi_1 y_{t-1} + u_t$

$$|\phi_1| < 1$$

- **Stationarity Condition**
- **Testing the unit root hypothesis using Dickey-Fuller (DF) test:** δ

Null Hypothesis (H_0: = 0) ***Alternate Hypothesis (H_1: δ < 0)***

Three different models for testing unit root hypothesis

Y_t **is a random walk:** $\Delta Y_t = \delta Y_{t-1} + u_t$

Y_t **is a random walk with drift:** $\Delta Y_t = \beta_1 + \delta Y_{t-1} + u_t$

Y_t **is a random walk with drift around a stochastic trend:** $\Delta Y_t = \beta_1 + \beta_2 t + \delta Y_{t-1} + u_t$

TESTING UNIT ROOT IN AR(P) MODEL
AUGMENTED DICKEY-FULLER TEST

- ***AR(p) model***

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_p y_{t-p} + u_t$$

- ***Augmented Dickey Fuller (ADF) Test***

$$\Delta Y_t = \beta_1 + \beta_2 t + \delta Y_{t-1} + \sum_{i=1}^{m} \alpha_i \Delta Y_{t-i} + \varepsilon_t$$

Remedy for Non-Stationarity
Transforming the series

Methods Applied

(1) Differencing a series: *(when a series contains a stochastic trend, such a series is also called as Integrated series)*

- ***First-order differencing*** *(Empirically, majority of series become stationary after first order differencing; in other words, majority of series are said to contain one unit-root).*
  - *Original series:* Y_t ***(Non-stationary series)***
  - *First-order differencing:*
    Differenced series denoted as $\Delta Y_t = (Y_t - Y_{t-1})$ [test ΔY_t for stationarity]
- ***Second-order & higher-order differencing*** *(in case the series is integrated of higher order)*

(2) De-trending the series; *(when a series contains a deterministic trend)*

A WHITE NOISE PROCESS

- A white noise process is one with (virtually) no discernible structure (a pure random process). A definition of a white noise process is:

$$E(y_t) = \mu$$
$$Var(y_t) = \sigma^2$$
$$E(y_t y_{t-r}) = \text{cov}ariance \;=\; \gamma_{t-r} = \begin{cases} \sigma^2 & if \quad t = r \\ 0 & \quad otherwise \end{cases}$$

- Thus the autocorrelation function will be zero apart from a single peak of 1 at s = 0. $\tau_s \sim$ approximately N(0,1/T) where T = sample size

- We can use this to do significance tests for the autocorrelation coefficients by constructing a confidence interval.

- For example, a 95% confidence interval would be given by $\pm .196 \times \frac{1}{\sqrt{T}}$. If the sample autocorrelation coefficient, $\hat{\tau}_s$, falls outside this region for any value of s, then we reject the null hypothesis that the true value of the coefficient at lag s is zero.

Modeling Time Series Data

Some popular models include:

- Auto-Regressive (AR) models
- Moving Average (MA) models
- ARMA models
- Auto-regressive Integrated Moving Average (ARIMA) models

Steps in Building Time Series Models

- *Identification*
- *Estimation*
- *Diagnostics*

Auto-Regressive (AR) Models

AUTO-REGRESSIVE (AR) MODELS

- AR(1) model

$$y_t = \mu + \phi_1 y_{t-1} + u_t$$

- AR(p) model

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_p y_{t-p} + u_t$$

- **Identifying Structure of the Model:**

a. Correlograms (ACF & PACF)

b. AIC & SBIC or SC criteria to select the best model

SAMPLE AR PROBLEM

- Consider the following simple AR(1) model

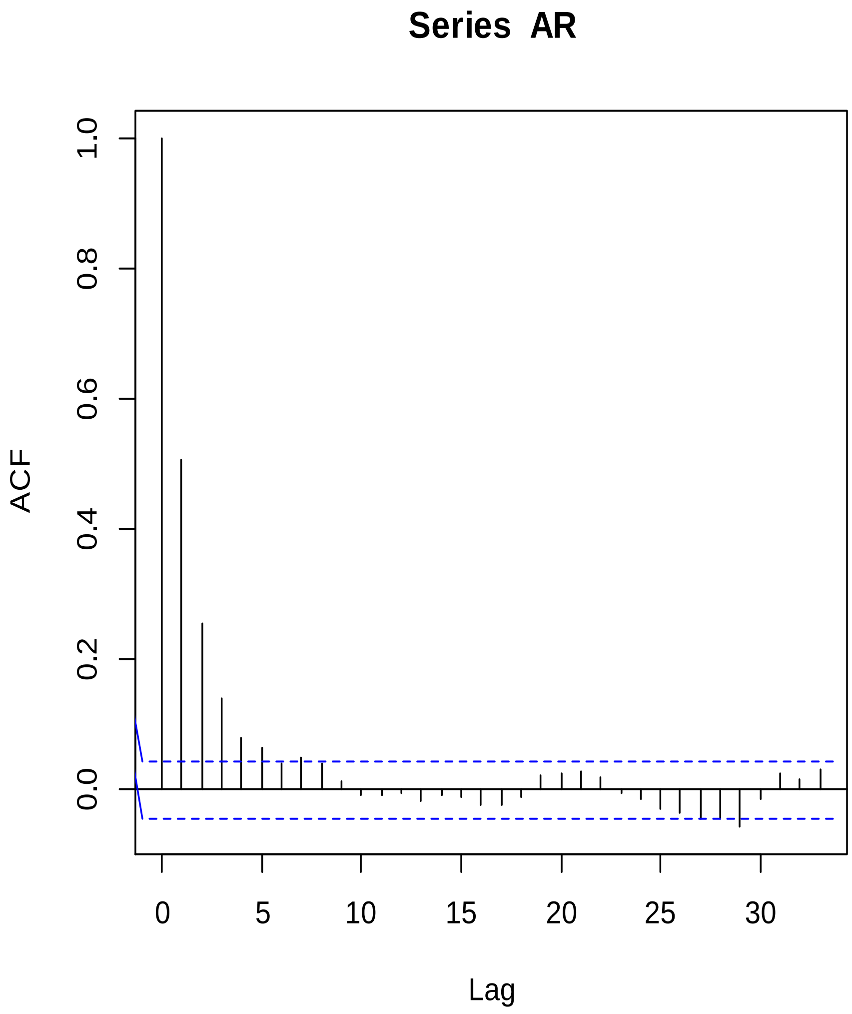
$$y_t = \mu + \phi_1 y_{t-1} + u_t$$

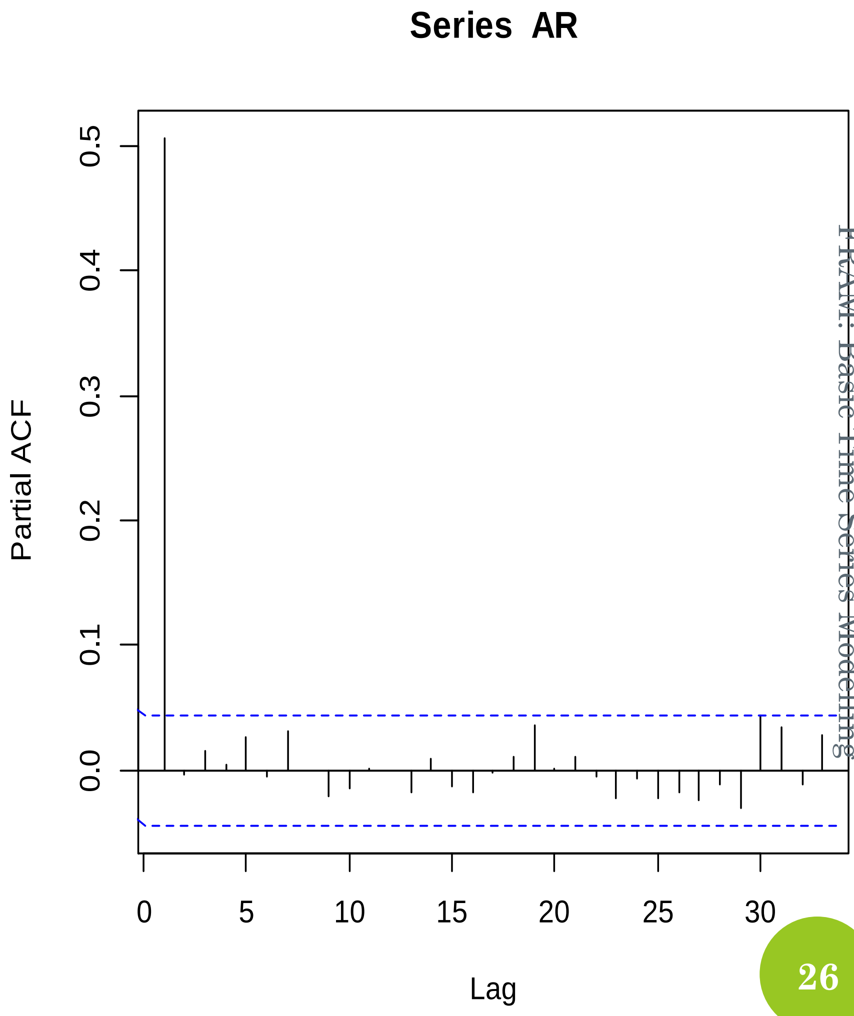
(i) Calculate the (unconditional) mean of y_t. For the remainder of the question, set μ=0 for simplicity.

(ii) Calculate the (unconditional) variance of y_t.

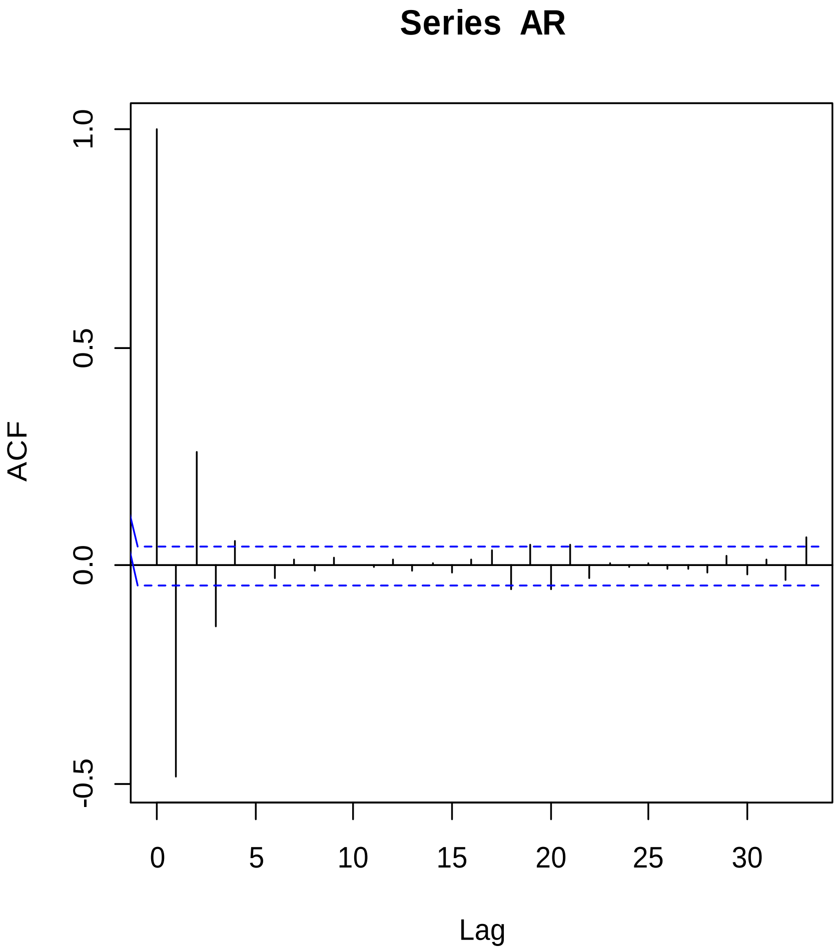
(iii) Derive the autocorrelation function for y_t.

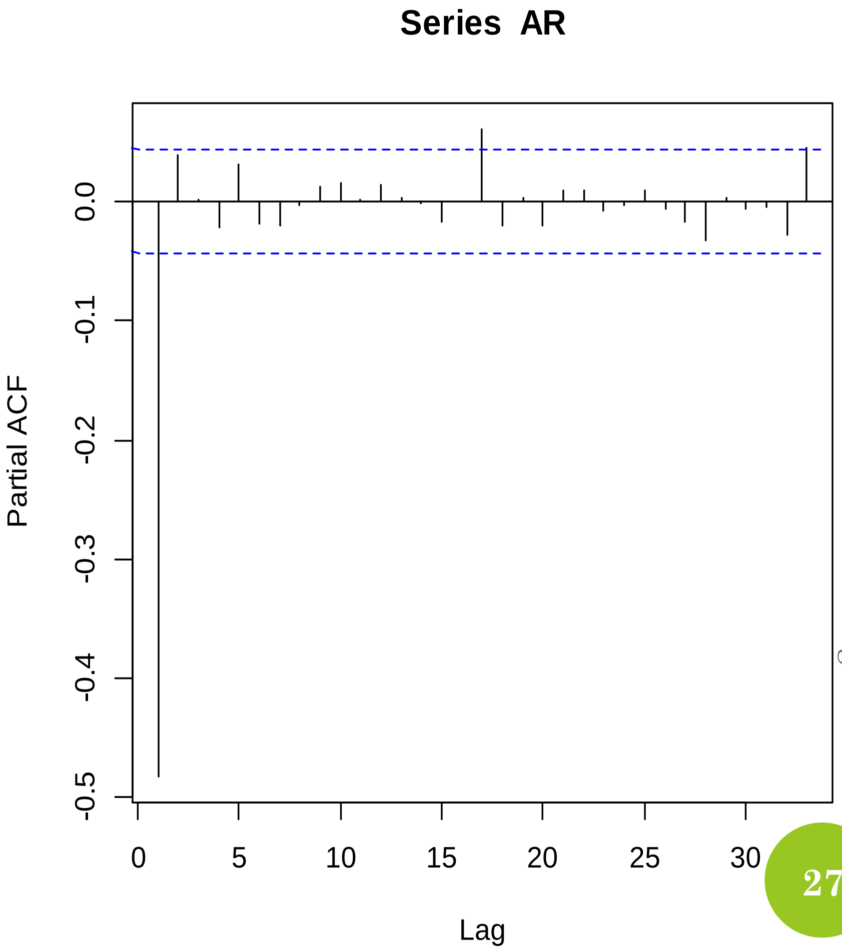
AR(1) model with Rho or Phi =+0.5





AR(1) MODEL WITH RHO OR PHI = -0.5





CRITERIA FOR CHOOSING THE BEST MODEL

- ***Akaike Information Criteria (AIC)***

$$AIC = \ln(\hat{\sigma}^2) + 2k / T$$

- OR

$$\text{AIC} = -2l/T + 2k/T$$

- ***Schwarz Bayesian Information Criteria (SBIC, also known as SIC or SC)***

$$SBIC = \ln(\hat{\sigma}^2) + (k \ln T) / T$$

- OR

$$SBIC = -2l + (k \ln T) / T$$

where $k = p + q + 1$, T = sample size, and $\hat{\sigma}^2$ is the variance of residual terms. And, in the other formula of AIC and SBIC, l represents value of log-likelihood.

- **Decision Rule:** Select the model with least AIC and SBIC (or BIC) values. In case of conflict, go for SBIC as it assigns very stringent penalty for higher lags compared to AIC as we choose the model with lower AIC OR SBIC values (as evident from formula itself).

Estimating AR models

Example

Results
AR model using R

x=read.csv('nifty.csv') # nifty.csv contains two columns/ variables, namely, Close and Ret#
m=arima(x$Ret, order=c(4,0,0))
m

Call:
arima(m= x$Ret, order = c(4, 0, 0))

Coefficients:

	ar1	ar2	ar3	ar4	intercept
Coefficients	0.0940	-0.0948	0.0318	0.0672	9e-04
S.E.	0.0256	0.0257	0.0256	0.0255	4e-04

sigma^2 estimated as 0.0001908; ***log likelihood = 4366.5; aic = -8721***

Diagnostics of the Fitted Model

Joint Hypothesis Tests

- If the model is fitted correctly, the error term left should become a white noise process;
- It can be verified using ACF & PACF.
- OR, We can also test the joint hypothesis that all m of the τ_k correlation coefficients are simultaneously equal to zero using the Q-statistic, called the Ljung-Box statistic:

$$Q^* = T(T+2)\sum_{k=1}^{m}\frac{\tau_k^2}{T-k} \sim \chi_m^2$$

where T = sample size, m = maximum lag length

In the formula, τ_k denotes the autocorrelation of Y_t with Y_{t-k}. It is also denoted as ρ_k or Φ_k

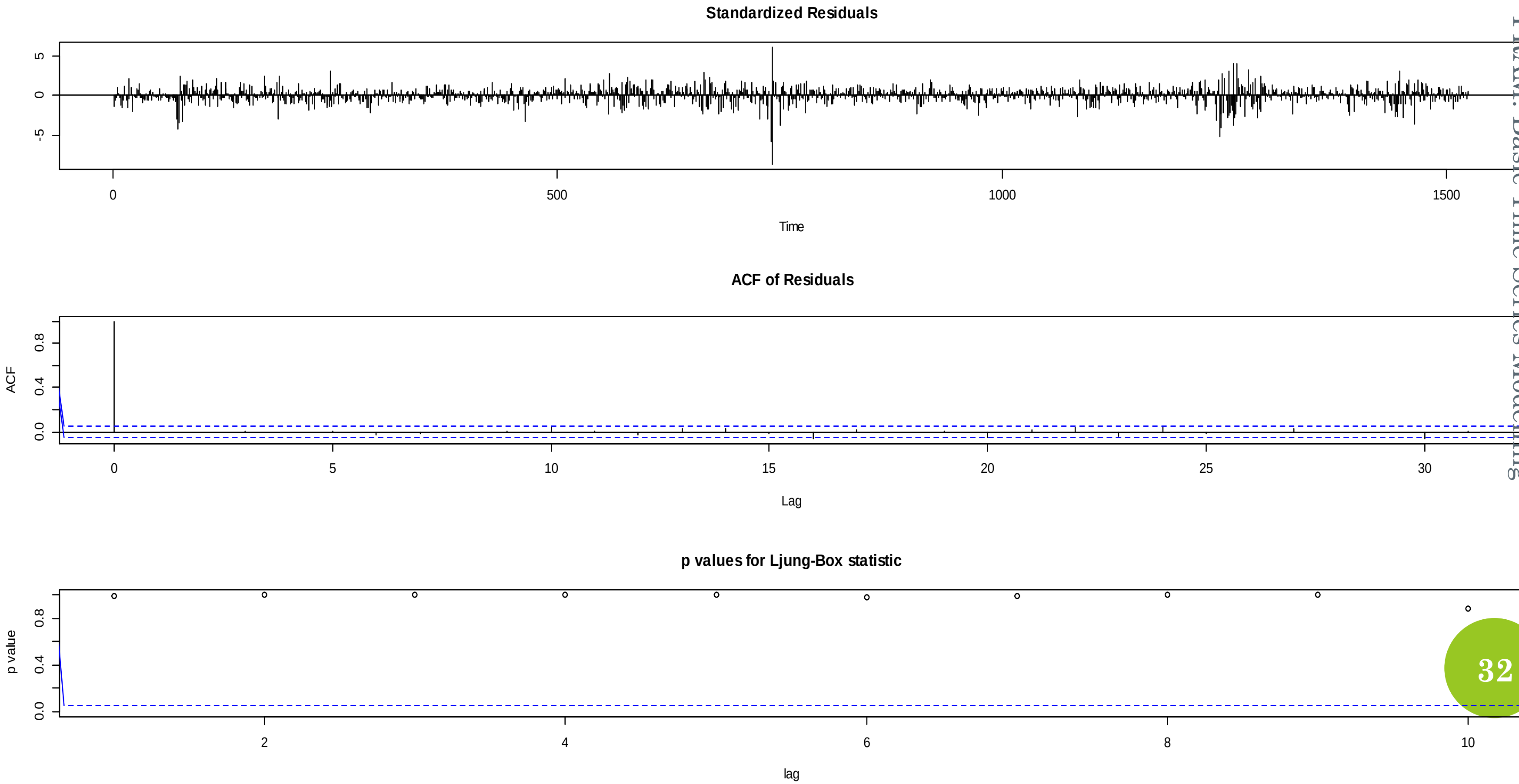
NULL HYPOTHESIS TESTED IS:

$$H_0 : \tau_1 = \tau_2 = \tau_3 \ldots = \tau_m = 0$$

- This statistic is very useful as a portmanteau (general) test of linear dependence in time series.

Diagnosis Using R

- For the diagnosis of any Time Series model in R:
- tsdiag(name of the fitted model)
- For example, we fitted the AR(1, 2 & 4) model in the class and named it as 'm'.
- m=arima(x$Ret, order=c(4,0,0))
- *tsdiag(m)* will yield the following results:



Moving Average (MA) Models

MA MODELS

- Let u_t (t=1,2,3,...) be a sequence of independently and identically distributed (iid) random variables with E(u_t)=0 and Var(u_t)= σ_ε^2, then

$$y_t = \mu + u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2} + \ldots + \theta_q u_{t-q}$$

is a q^{th} order moving average model, denoted as **MA(q).**

- **Its properties are**

E(y_t)=μ;

Var(y_t) = γ_0 = $(1+ \theta_1^2 + \theta_2^2 + \ldots + \theta_q^2)\sigma^2$

Covariances

$$\gamma_s = \begin{cases} (\theta_s + \theta_{s+1}\theta_1 + \theta_{s+2}\theta_2 + \ldots + \theta_q \theta_{q-s})\sigma^2 & for \quad s = 1,2,\ldots,q \\ 0 \quad for \quad s > q \end{cases}$$

MA MODELS

Example 1

Consider the following MA(2) process:

$$X_t = u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2}$$

where u_t is a zero mean white noise process with variance.

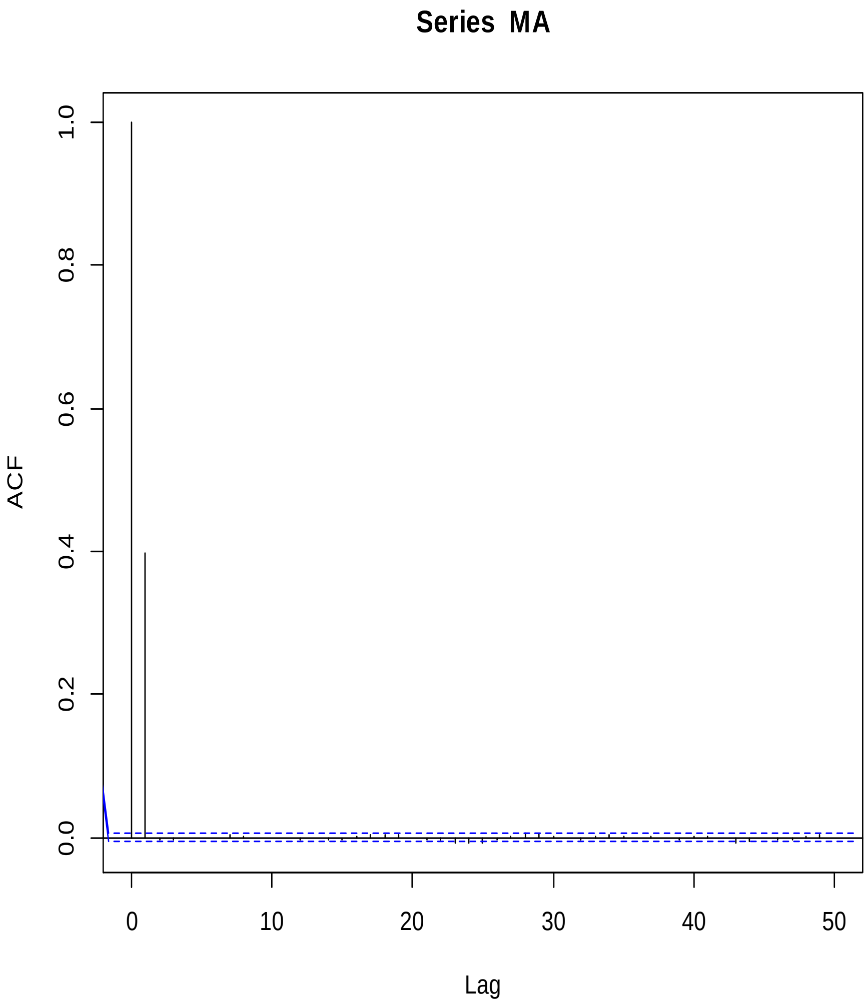
(i) Calculate the mean and variance of X_t

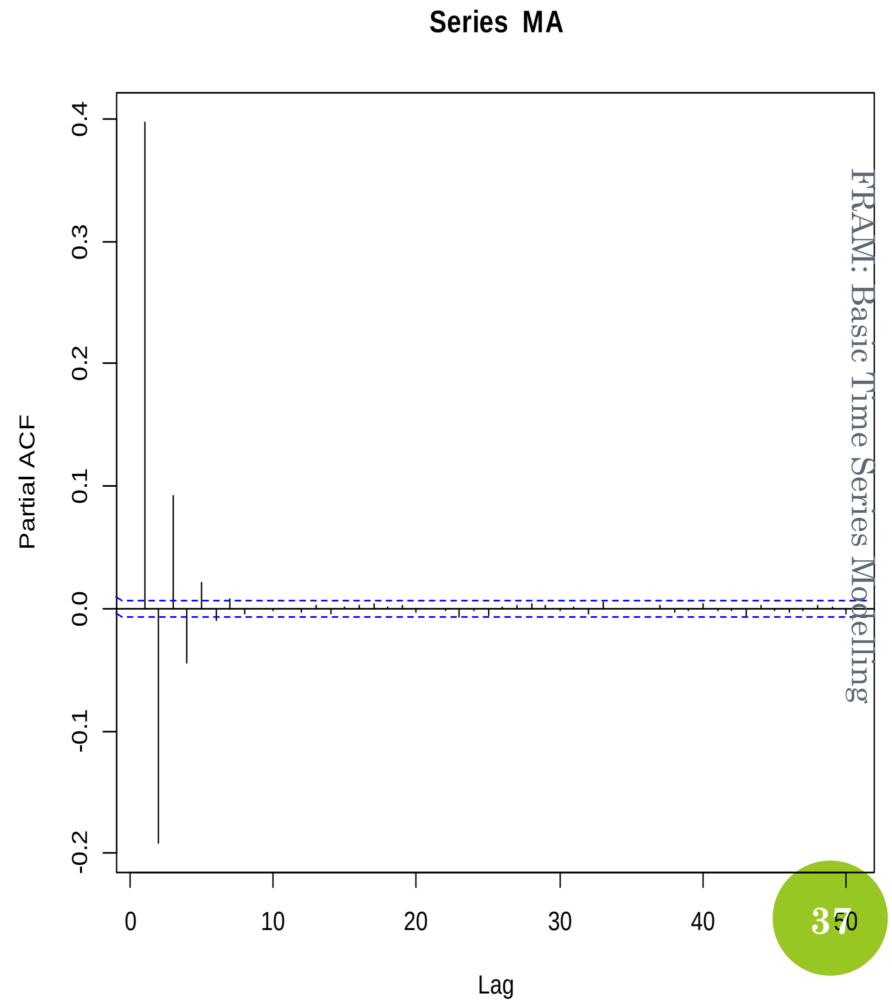
(ii) Derive the autocorrelation function for this process (i.e. express the autocorrelations, τ_1, τ_2, ... as functions of the parameters θ_1 and θ_2).

(iii) If θ_1 = -0.5 and θ_2 = 0.25, sketch the acf of X_t.

Identifying a Moving Average Process

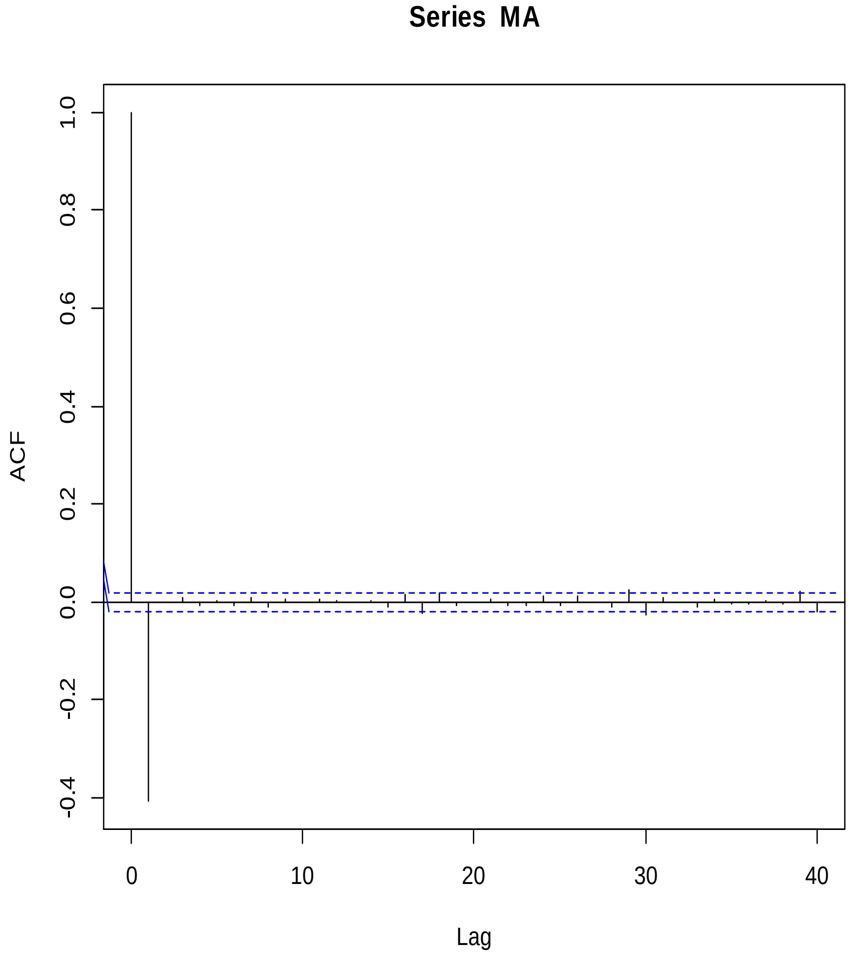
MA(1) MODEL WITH THETA = 0.5

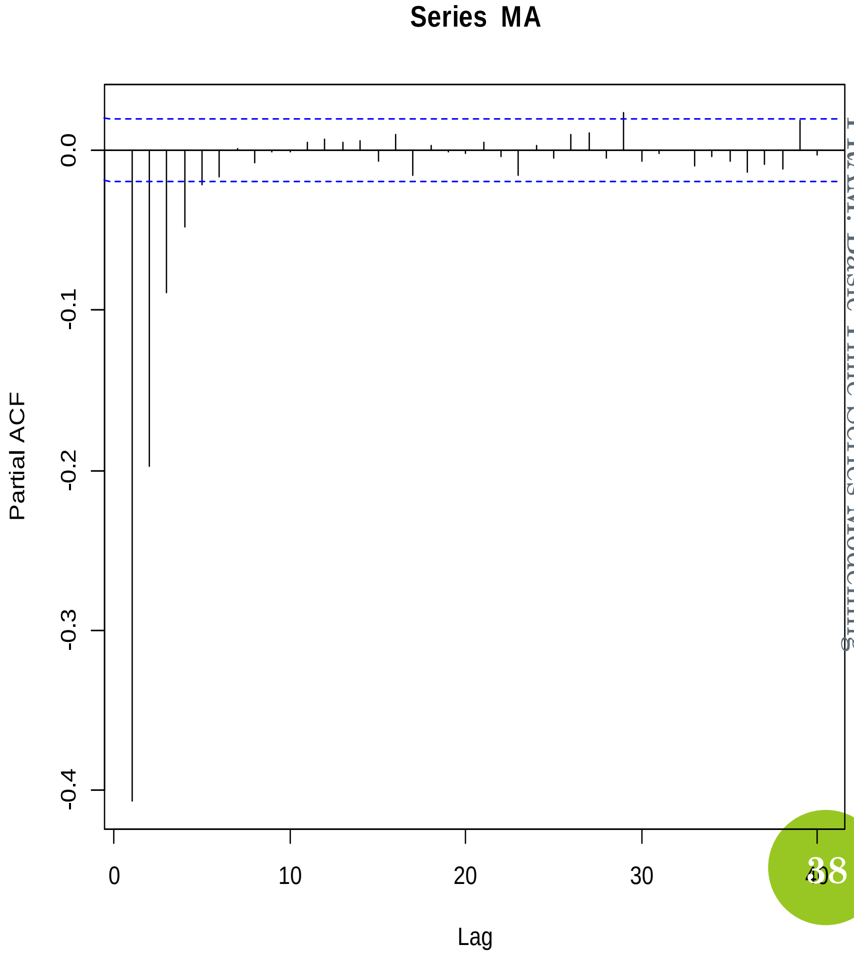




FRAM: Basic Time Series Modelling

MA(1) model with Theta = -0.5





Estimation & Diagnosis of an MA Model is done on the same lines as we did for AR models
(Please try it on your own)

Example

Summary of the Behaviour of Correlograms for AR and MA Processes

An autoregressive process has

- *a geometrically decaying acf*
- ***number of spikes of pacf = AR order***

A moving average process has

- ***Number of spikes of acf = MA order***
- *a geometrically decaying pacf*

IMPORTANT RESULTS

- An stationary AR (1) model can be shown as MA(∞) process. (Wold's decomposition);

- Similarly, a invertible MA(1) process can be shown as AR(∞) process;

- ***The implication of this result is that:***

  - ***A model with long AR structure can be fairly approximated with a parsimonious MA(1) model;***

  - ***similarly, a model with long MA structure can be estimated by an parsimonious AR(1) model.***

Auto-Regressive Moving Average (ARMA) Models

ARMA PROCESSES

- By combining the AR(p) and MA(q) models, we can obtain an ARMA(p,q) model:

$$\phi(L)y_t = \mu + \theta(L)u_t$$

where

$$\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \ldots - \phi_p L^p$$

and

$$\theta(L) = 1 + \theta_1 L + \theta_2 L^2 + \ldots + \theta_q L^q$$

or

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_p y_{t-p} + \theta_1 u_{t-1} + \theta_2 u_{t-2} + \ldots + \theta_q u_{t-q} + u_t$$

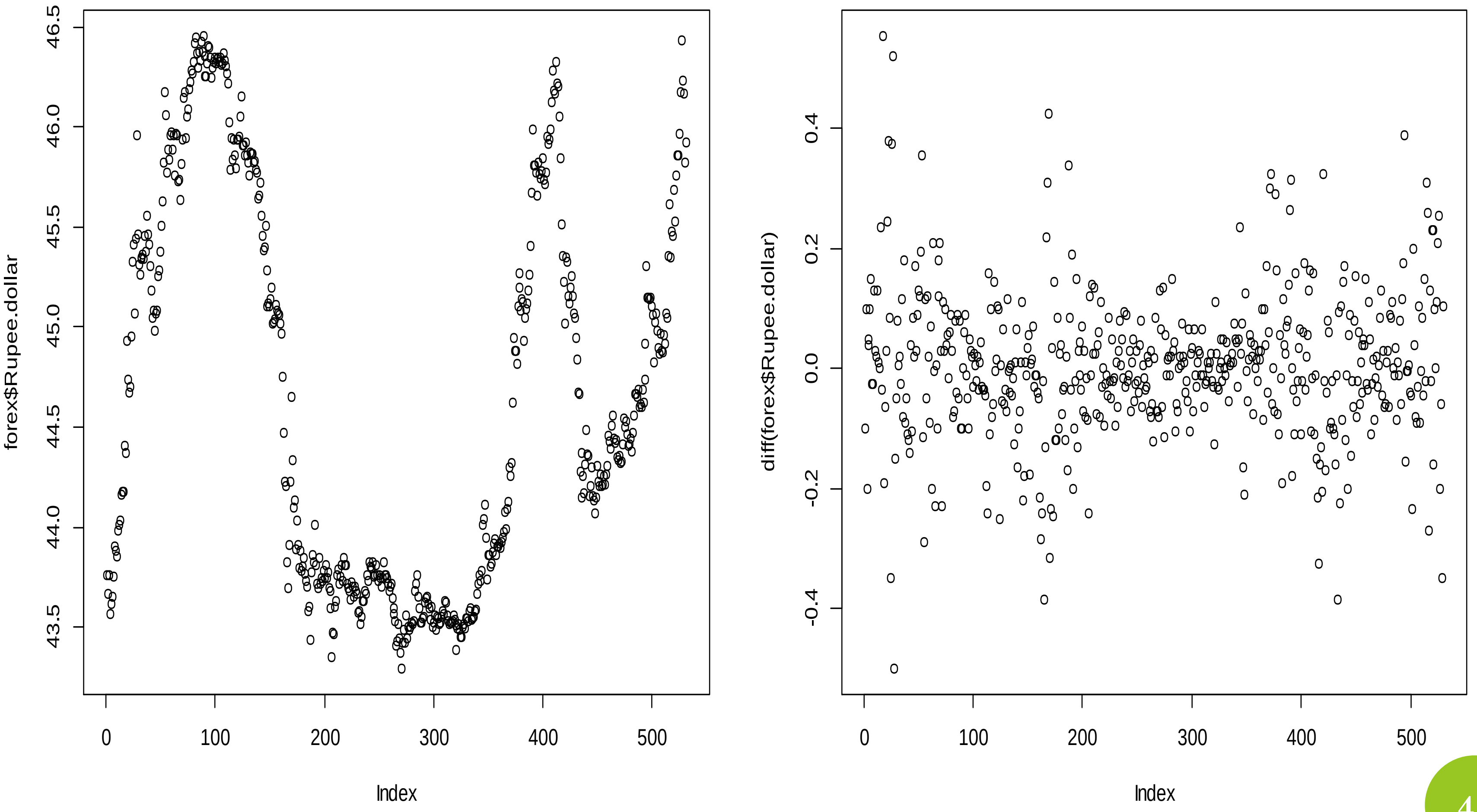
with

$$E(u_t) = 0;\ E(u_t^2) = \sigma^2;\ E(u_t u_s) = 0,\ t \neq s$$

ARMA MODELS

- ARMA models are capable of capturing very complex patters of temporal correlation.
- Thus, they are a useful and interesting class of models.
- In fact, they can capture any valid autocorrelation!

Forecasting Exchange Rate with ARMA(1,1)



Forecasting Exchange Rate with ARMA(1,1)
Stationarity of the data

> adf.test(forex$Rupee.dollar)

Augmented Dickey-Fuller Test

data: forex$Rupee.dollar
Dickey-Fuller = -1.3648, Lag order = 8, p-value = 0.8472
alternative hypothesis: stationary

> adf.test(diff(forex$Rupee.dollar))

Augmented Dickey-Fuller Test

data: diff(forex$Rupee.dollar)
Dickey-Fuller = -7.2883, Lag order = 8, p-value = 0.01
alternative hypothesis: stationary

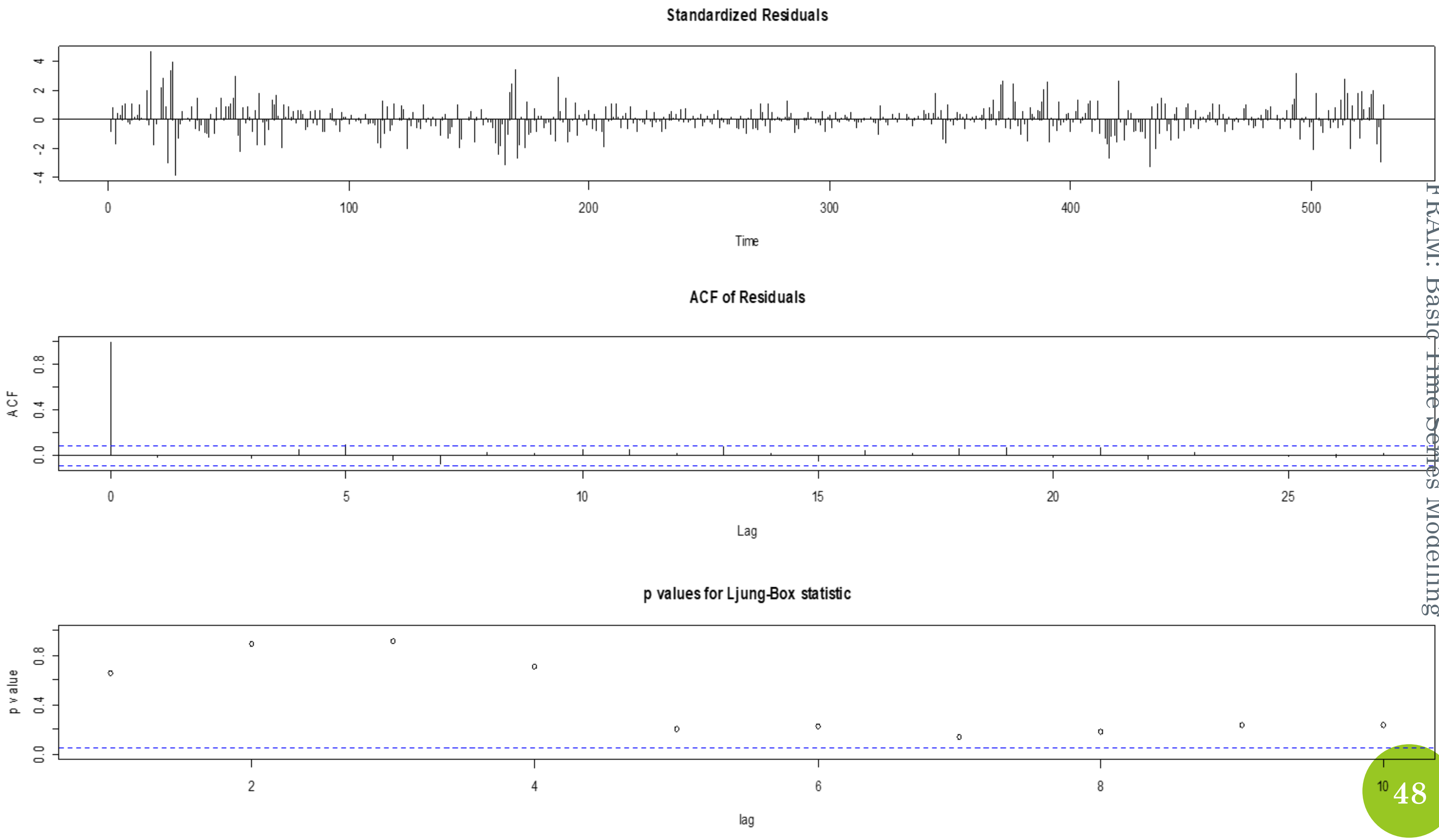
Forecasting Exchange Rate with ARMA(1,1)

- Exchange rates ARMA output

```
> forex=read.csv('Exchange rates.csv')
> head(forex)
  Rupee.dollar
1       43.765
2       43.665
3       43.765
4       43.565
5       43.615
6       43.655

> ret=diff(forex$Rupee.dollar)
> m=arima(ret, order=c(1,0,1))
> m
Call: arima(x = ret, order = c(1, 0, 1))

Coefficients:
          ar1     ma1  intercept
      -0.9420  0.9728     0.0041
s.e.   0.0286  0.0182     0.0054

sigma^2 estimated as 0.01486:  log likelihood = 363.25,  aic = -718.5
```

DIAGNOSTICS OF THE FITTED MODEL

- tsdiag(m)



Identification
ACF & PACF for Different Processes

Process	ACF	PACF
White noise	No significant coefficients	No significant coefficients
AR(2)	Geometrically declining or damped sinusoid acf	First 2 pacf coefficients significant, all others insignificant
MA(1)	First acf coefficient significant, all others insignificant	Geometrically declining or damped sinusoid pacf
ARMA(2,1)	Geometrically declining or damped sinusoid acf	Geometrically declining or damped sinusoid pacf

Estimation & Forecasting Financial Volatility

Volatility: An Introduction

- *Volatility refers to the spread of all likely outcomes of an uncertain variable (say, returns).*

- *Typically, in financial markets, we are often concerned with the spread of asset returns.*

- It is a key input to many important finance applications such as:

  - Investment & portfolio construction,
  - option pricing,
  - hedging, and risk management (VaR model).
  - Trading Strategies

- One of the most studied areas in finances equally in academia & industry given its importance;

Volatility Estimation & Forecasting
Select Models

- **Different models have been developed over time.**

- Choice of a model depends on the horizon one is looking for, i.e., ***short-term, medium-term, long-term.***

Historical Volatility Estimators

- Standard Deviation
- Extreme Value Estimators
- ***Conditional Volatility Models***: ARCH models, e.g., EWMA, GARCH, EGARCH, GJR, etc.
- **Realised Volatility (based on High-frequency data)**

Historical Volatility
Using Standard Deviation

- Statistically, volatility, in it's simplest form, is often measured as the sample standard deviation:

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(r_i - \bar{r})^2$$

- *Where, r_i is the return on day i, $\bar{r}$ is the average return, and N= number of observation in the sample.*

- Inherently, it assigns equal weightage to all the observations in the sample.

- Fails to capture short-term dynamics of the returns (dependence);

- Therefore, empirically, it has not been appreciated as a good estimator of volatility, especially for short-term volatility as it fails to capture important characteristics revealed by a financial time series in short term;

- Can be used a fair measure of long-run volatility.

Extreme Value Estimators of Volatility

- ***Builds on the logic that price range contains more information about the volatility than close prices.***

- Instead of Close value, these estimators build on 'Open', 'High', 'Low', & 'Close' (OHLC) prices.

- Better captures variation in the data;

- Requires much less data compared to Traditional Measure of Volatility;

- Empirically, such estimators have been found to be very good in estimation & forecasting volatility in Indian context as well as in different markets across the globe.

Select Extreme Value Estimators

- Parkinson's Formula (only on High Low prices)

$$\hat{\sigma}_{pk}^2 = \left(\frac{1}{4\ln 2}\right)\frac{1}{n}\sum_{t=1}^{n}(H_t - L_t)^2$$

- Garman & Klass

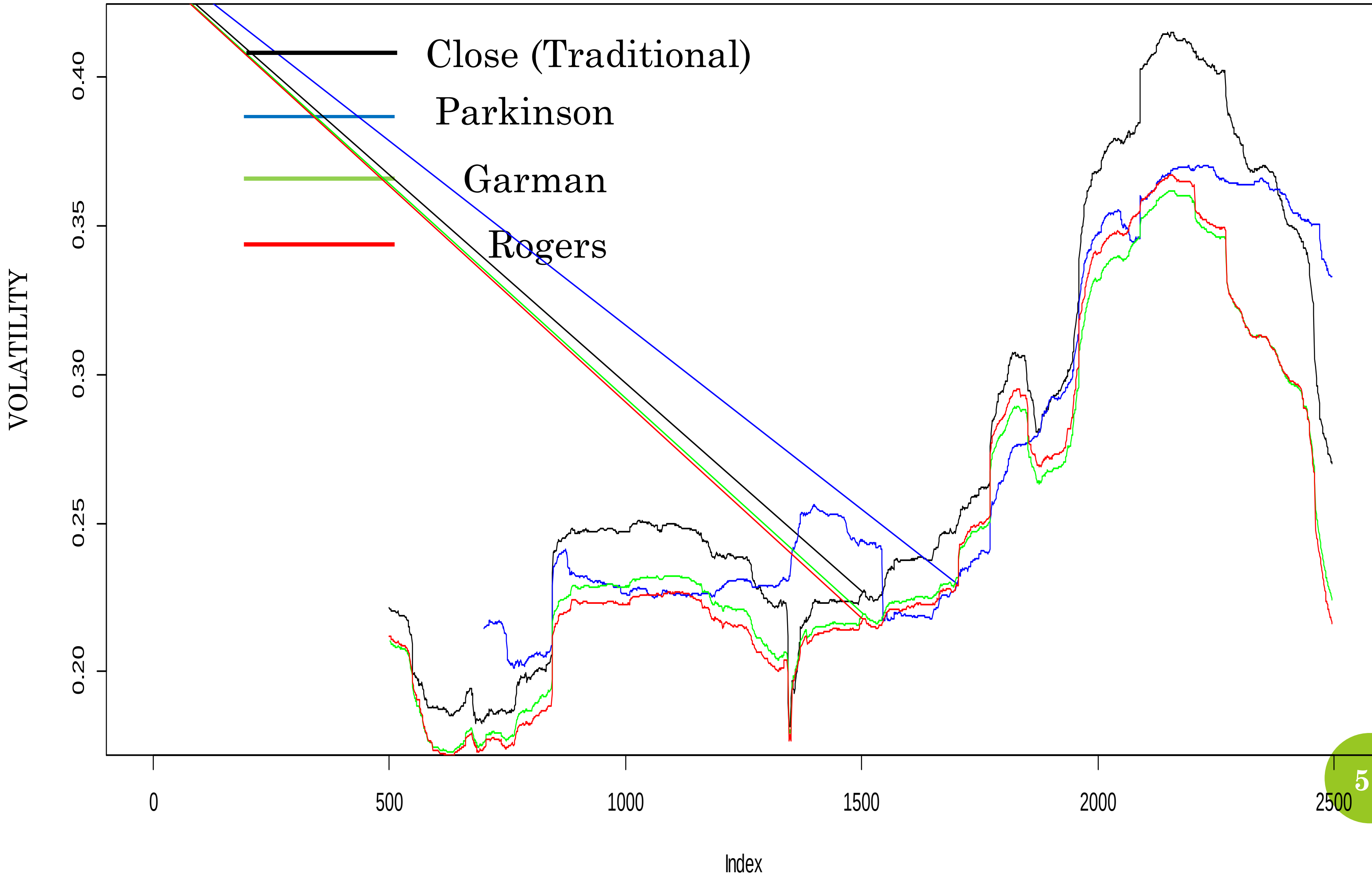
$$\hat{\sigma}_{gk}^2 = \frac{1}{n}\sum_{t=1}^{n}[0.511(H_t - L_t)^2 - .019\{(C_t - O_t)(H_t + L_t - 2O_t) - 2(H_t - O_t)(L_t - O_t)\} - 0.383(C_t - O_t)^2]$$

- Rogers & Satchell

$$\hat{\sigma}_{rs}^2 = \frac{1}{n}\sum_{t=1}^{n}[(H_t - C_t)(H_t - O_t) + (L_t - C_t)(L_t - O_t)]$$

- **Note: The first two measures assume that the security follows drift-less GBM. Roger & Satchel's model relaxes this assumption.**

Parkinson, Garman & Rogers: A comparison



EXTREME VALUE ESTIMATORS VS. TRADITIONAL MEASURE OF VOLATILITY

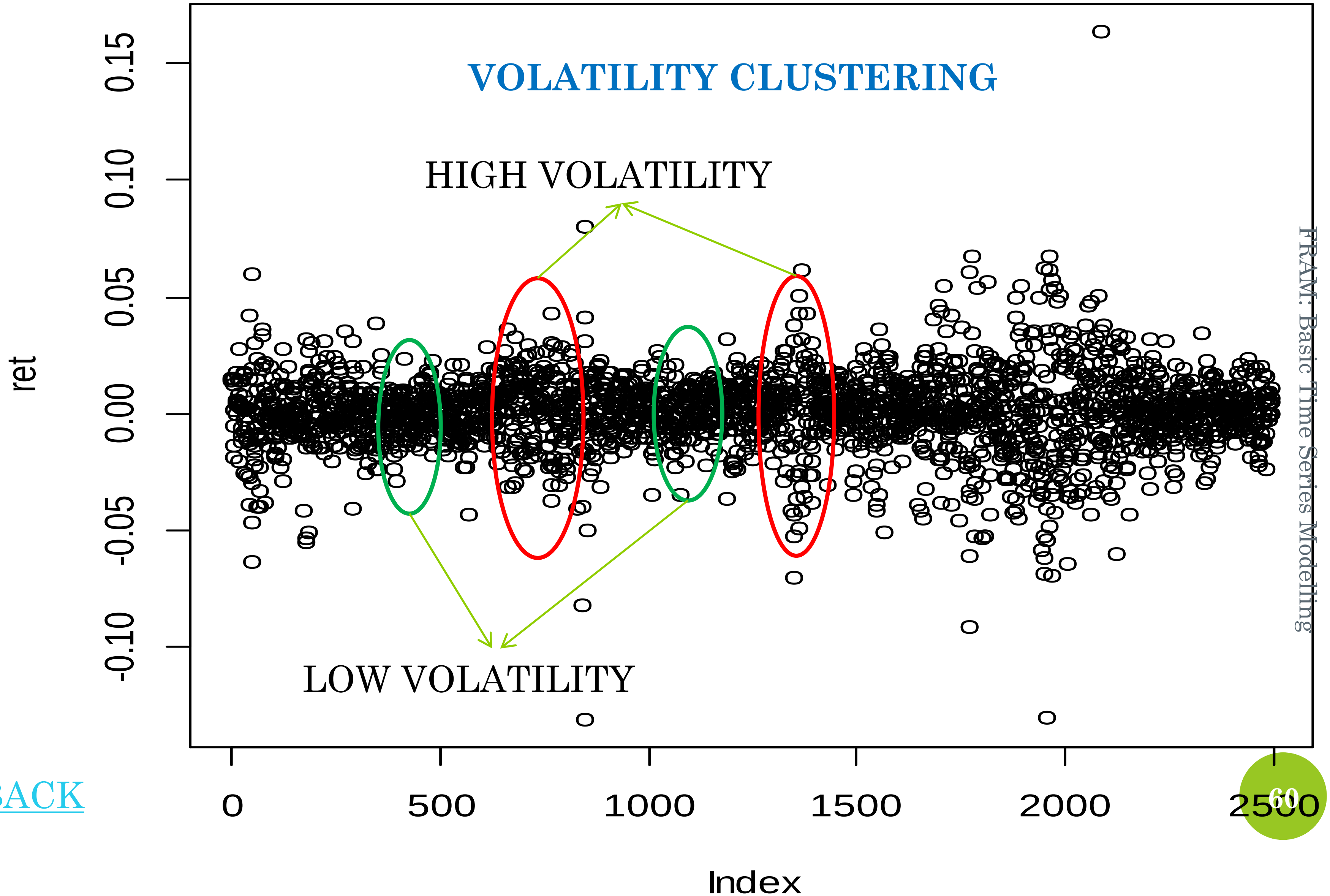
- Extreme Value estimators of volatility (viz., Parkinson, Garman-Klass & Rogers-Satchell) are empirically found to understate the volatility vis-à-vis Traditional measures;
- However, they are found to be many times more efficient (5 to 7 times) compared to Traditional estimator of volatility.
- It is important to note that the downward bias in Extreme value estimators is with respect to Traditional measures; naturally, question arises is it (Traditional measure) a true measure of volatility as it itself might be biased (overstated).
- Therefore, a correct measure of benchmark volatility needs to be in place to make any comment on performance of other measures. For the purpose, literature supports ***'Realised Volatility Measures'***

CONDITIONAL VOLATILITY MODELS

Volatility of current Period is conditioned on it's previous values

ARCH MODELS

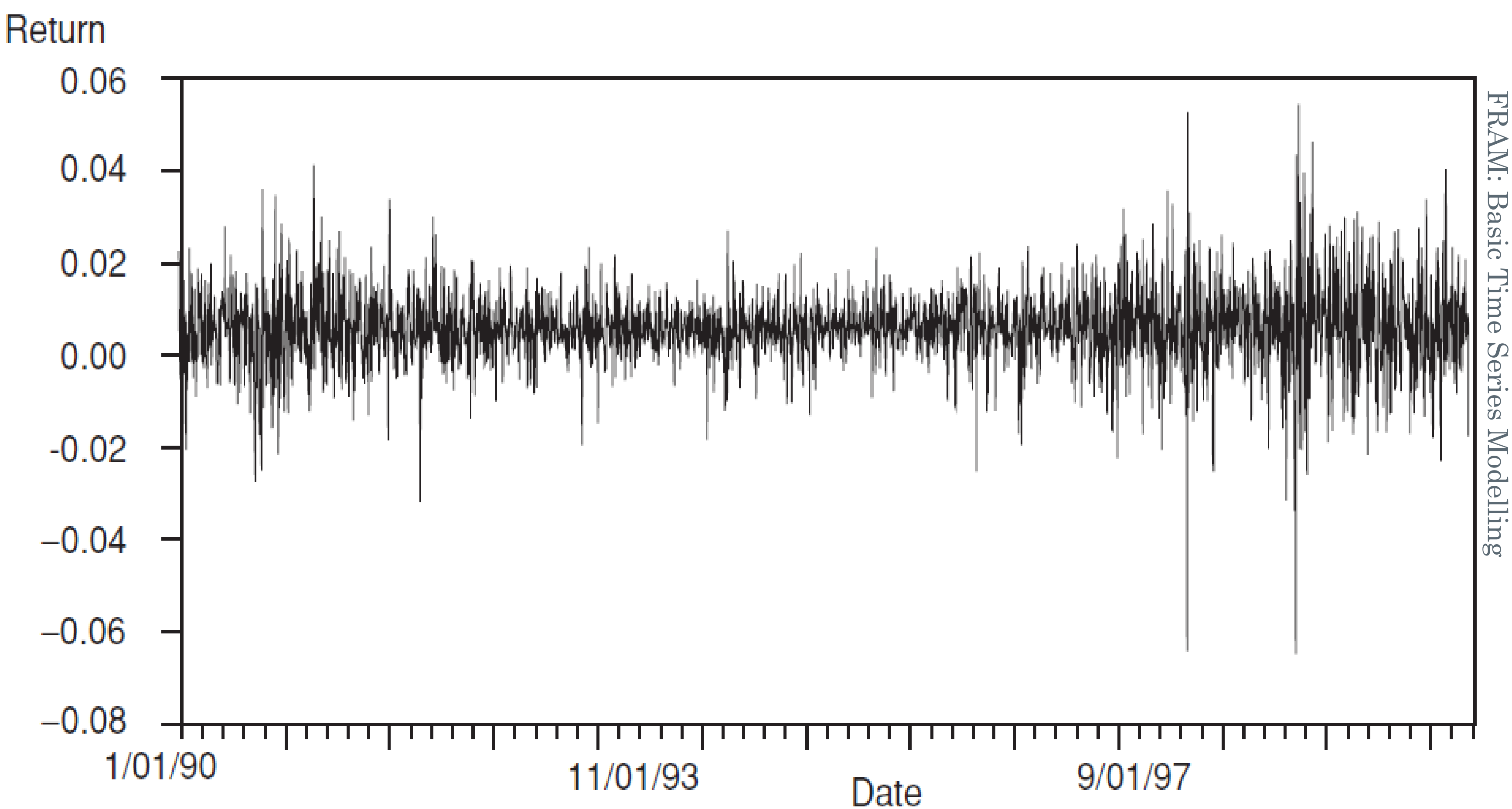
Some Stylized Facts about the Returns
Daily Returns 2001-2009
Stylized Fact 1 : Clustering of Volatility



BACK

Some Stylized Facts about the Returns
Daily Returns 2001-2009
Stylized Fact 1 : Clustering of Volatility – S&P 500

VOLATILITY CLUSTERING



Testing the ARCH effect
Identification of Model

- ARCH effect

- Testing of ARCH Effect:
  - ACF & PACF of Squared-residuals (e^2)
  - **ARCH LM Test**

Testing the "ARCH Effects" ARCH LM test

1. First, run any postulated linear regression of the form given in the equation above, e.g. $y_t = \beta_1 + \beta_2 x_{2t} + ... + \beta_k x_{kt} + u_t$
   saving the residuals, $\hat{u}_t$.

2. Then square the residuals, and regress them on q own lags to test for ARCH of order q, i.e. run the regression
   $$\hat{u}_t^2 = \gamma_0 + \gamma_1 \hat{u}_{t-1}^2 + \gamma_2 \hat{u}_{t-2}^2 + ... + \gamma_q \hat{u}_{t-q}^2 + v_t$$
   where v_t is iid.
   Obtain R^2 from this regression

3. The test statistic is defined as TR^2 (the number of observations multiplied by the coefficient of multiple correlation) from the last regression, and is distributed as a $\chi^2(q)$.

4. F statistics can also be used to test equality of variance of error terms to establish ARCH effect.

TESTING THE "ARCH EFFECTS" (CONT'D)

4. The null and alternative hypotheses are

$H_0 : \gamma_1 = 0$ and $\gamma_2 = 0$ and $\gamma_3 = 0$ and ... and $\gamma_q = 0$

$H_1 : \gamma_1 \neq 0$ or $\gamma_2 \neq 0$ or $\gamma_3 \neq 0$ or ... or $\gamma_q \neq 0$.

If the value of the test statistic is greater than the critical value from the χ^2 distribution, reject the null hypothesis.

Autoregressive Conditionally Heteroscedastic (ARCH) models for Estimating & forecasting Volatility

- What could the current value of the variance of the errors plausibly depend upon?
  - Previous squared error terms.

- This leads to the autoregressive conditionally heteroscedastic model for the variance of the errors:

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

This is known as an ARCH(1) model.

Autoregressive Conditionally Heteroscedastic (ARCH) Models (cont'd)

- The full model would be
  $y_t = \beta_1 + \beta_2 x_{2t} + \ldots + \beta_k x_{kt} + u_t, \ u_t \sim N(0, \sigma_t^2)$
  where $\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2$

- We can easily extend this to the general case where the error variance depends on q lags of squared errors:

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \ldots + \alpha_q u_{t-q}^2$$

$$s.t. \quad \alpha_i > 0 \ \forall \ i = 1, 2, \ldots, q$$

- This is an ARCH(q) model.

PROBLEMS WITH ARCH(q) MODELS

- How do we decide on q? (One option is likelihood ratio test, however, there is no clearly best approach)

- The required value of q might be very large. (affect the parsimony of the model)

- Non-negativity constraints might be violated. (because of large number of parameters)
  - When we estimate an ARCH model, we require $\alpha_i > 0 \; \forall \; i=1,2,...,q$ (since variance cannot be negative)

- A natural extension of an ARCH(q) model which gets around some of these problems is a GARCH model.

GENERALISED ARCH (GARCH MODELS)

- The ARCH model, generalised by Bollerslev, is popularly known as GARCH model.

- A GARCH Model
  - AR(1)-GARCH(1,1) model

Distributional assumption

$$y_t = \mu + \phi y_{t\text{-}1} + u_t \quad , u_t \sim \mathrm{N}(0, \sigma_t^2)$$ **Mean equation**

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \beta \sigma_{t\text{-}1}^2$$ **Variance equation**

*In conditional volatility measures, it is popularly dented as **ht**, to differentiate from traditional measures.*

- A GARCH model requires three specifications to be made, viz., (i) Mean equation, (ii) Variance Equation & (iii) Distributional assumption.

Another way of writing GARCH (1,1) model

$$R_t = c + \epsilon_t$$

$$\sigma_t^2 = \omega + \alpha \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2$$

THE UNCONDITIONAL (LONG-RUN) VARIANCE UNDER THE GARCH SPECIFICATION

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

- where α_1 and $\beta > 0$ (NON-NEGATIVITY CONSTRAINT)
- The unconditional/ long-run variance of u_t is given by

$$\text{Var}(u_t) = \frac{\alpha_0}{1-(\alpha_1+\beta)}$$

  when $\alpha_1 + \beta < 1$

- $\alpha_1 + \beta \geq 1$ is termed “non-stationarity” in variance
- $\alpha_1 + \beta = 1$ is termed intergrated GARCH
- For non-stationarity in variance, the conditional variance forecasts will not converge on their unconditional value as the horizon increases.

Steps in Estimating a GARCH Model

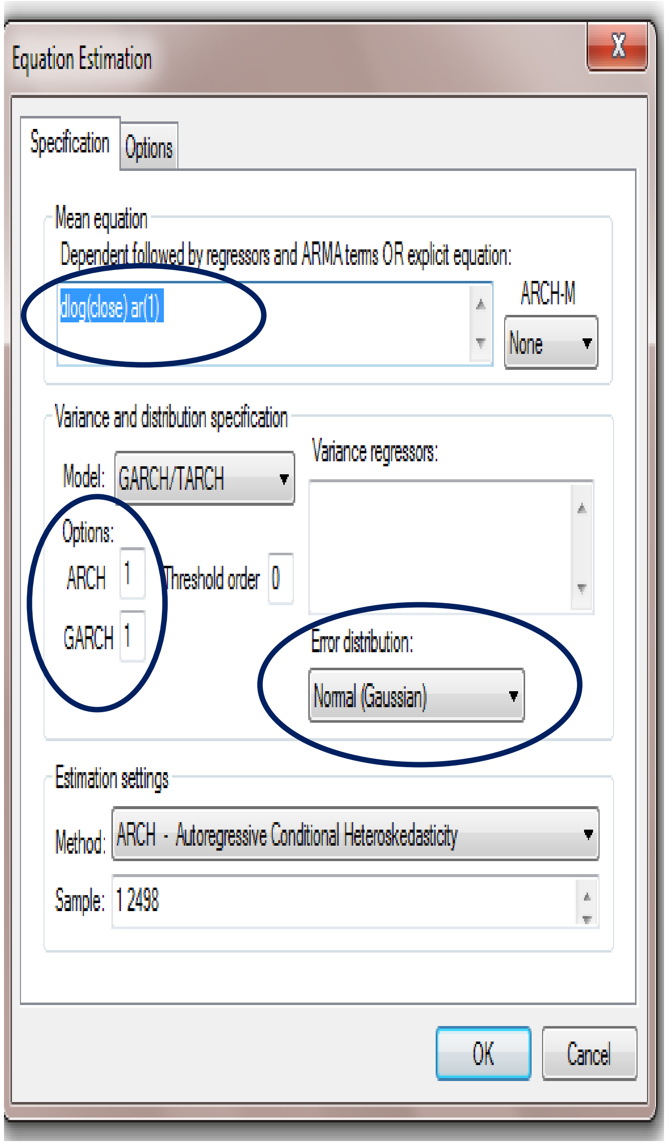
- ***Identification & Specification* (That is, to ascertain the ARCH Effect and specifying correct model);**
  - ***By examining the squared residuals. Or, by formally conducting the ARCH LM test.***
  - ***Specify mean equation, variance equation and conditional distribution of errors. IT IS IMPORATNT TO NOTE THAT SPECIFYING A CORRECT MEAN EQUATION IS EQUALLY IMPORTANT AS THE VARIANCE IS MEASURED AROUND THE MEAN.***

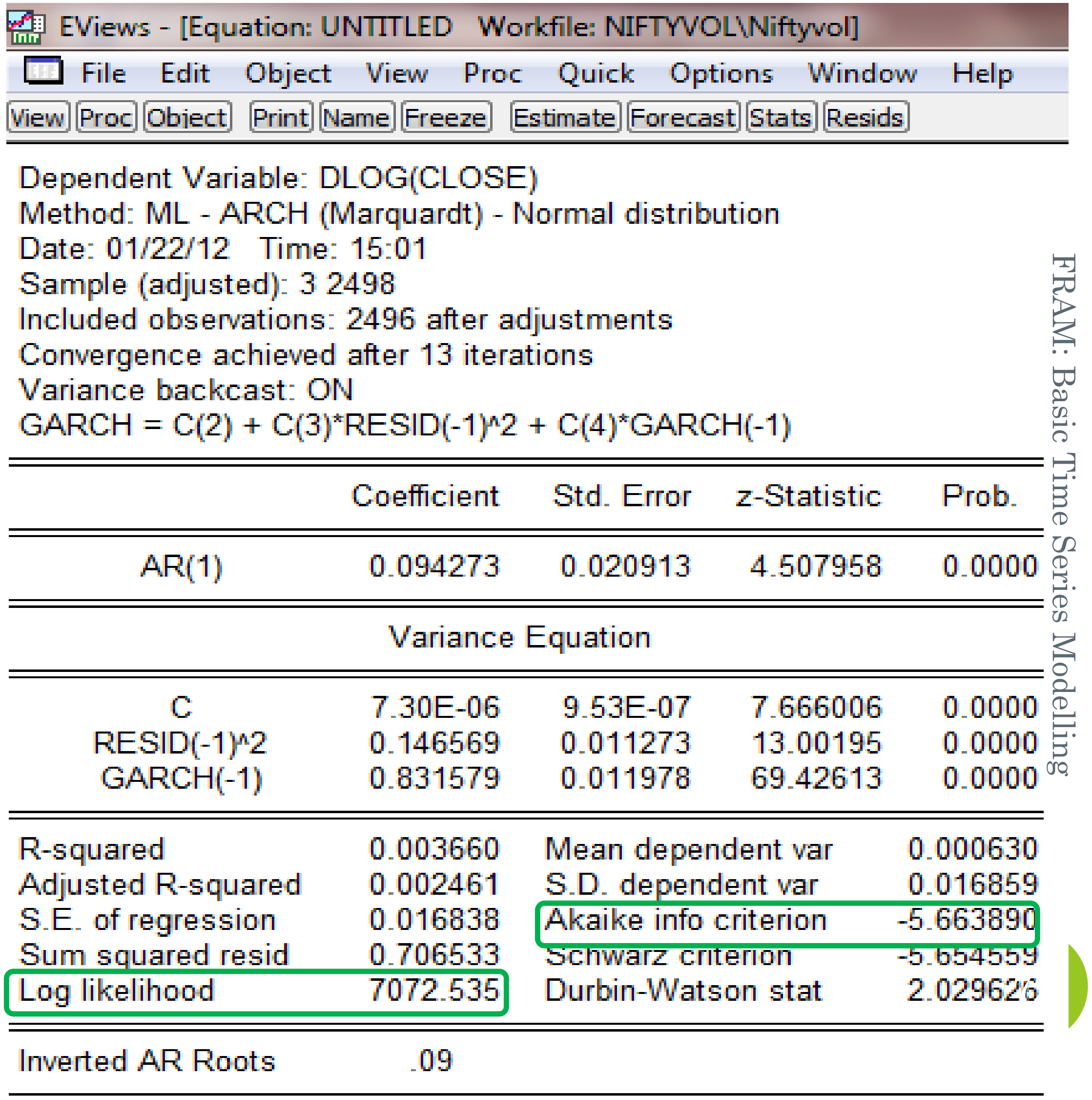
- ***Estimation of the model* (Requires estimation of a mean & variance equation simultaneously);**
  - ***Based on distributional assumption (Normal or Non-normal), we define the Log-Likelihood Function (LLF) to be maximized. The unknown parameters are estimated by maximizing the LLF with respect to each unknown parameter.***

- ***Diagnosis of the Fitted Model;***
  - ***Finally, the diagnosis of the fitted model is carried out to ensure that all ARCH effects present in the data have been captured by the specified model. For the purpose, a diagnosis of squared-residuals is carried out.***

GARCH MODEL: AN EXAMPLE (DATA)

AR(1)-GARCH (1,1) MODEL WITH NORMALLY DISTRIBUTED ERRORS





EViews - [Equation: UNTITLED Workfile: NIFTYVOL\Niftyvol]

File Edit Object View Proc Quick Options Window Help

View Proc Object Print Name Freeze Estimate Forecast Stats Resids

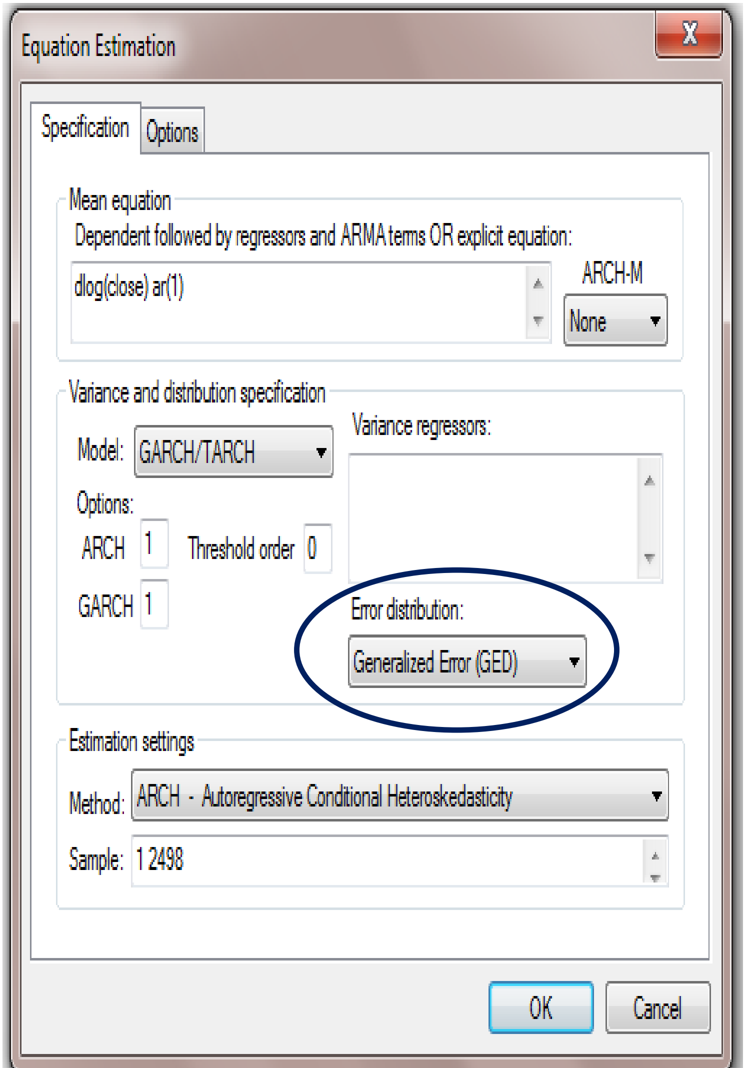
Dependent Variable: DLOG(CLOSE)
Method: ML - ARCH (Marquardt) - Normal distribution
Date: 01/22/12 Time: 15:01
Sample (adjusted): 3 2498
Included observations: 2496 after adjustments
Convergence achieved after 13 iterations
Variance backcast: ON
GARCH = C(2) + C(3)*RESID(-1)^2 + C(4)*GARCH(-1)

	Coefficient	Std. Error	z-Statistic	Prob.
AR(1)	0.094273	0.020913	4.507958	0.0000
	Variance Equation			
C	7.30E-06	9.53E-07	7.666006	0.0000
RESID(-1)^2	0.146569	0.011273	13.00195	0.0000
GARCH(-1)	0.831579	0.011978	69.42613	0.0000

R-squared	0.003660	Mean dependent var	0.000630
Adjusted R-squared	0.002461	S.D. dependent var	0.016859
S.E. of regression	0.016838	Akaike info criterion	-5.663890
Sum squared resid	0.706533	Schwarz criterion	-5.654559
Log likelihood	7072.535	Durbin-Watson stat	2.029626
Inverted AR Roots	.09		

AR(1)-GARCH (1,1) MODEL
WITH (GENERALISED ERROR DISTRIBUTION) GED ERRORS

GED is used to model high kurtosis and fat-tails which are typically observed in financial time series data (specially in returns).



EViews - [Equation: UNTITLED Workfile: NIFTYVOL\Niftyvol]

File Edit Object View Proc Quick Options Window Help

View Proc Object Print Name Freeze Estimate Forecast Stats Resids

Dependent Variable: DLOG(CLOSE)
Method: ML - ARCH (Marquardt) - Generalized error distribution (GED)
Date: 01/22/12 Time: 15:15
Sample (adjusted): 3 2498
Included observations: 2496 after adjustments
Convergence achieved after 16 iterations
Variance backcast: ON
GARCH = C(2) + C(3)*RESID(-1)^2 + C(4)*GARCH(-1)

	Coefficient	Std. Error	z-Statistic	Prob.
AR(1)	0.091018	0.020916	4.351671	0.0000
Variance Equation				
C	6.90E-06	1.45E-06	4.772059	0.0000
RESID(-1)^2	0.136307	0.016095	8.468879	0.0000
GARCH(-1)	0.840781	0.017142	49.04909	0.0000
GED PARAMETER	1.450219	0.042049	34.48876	0.0000

R-squared	0.003782	Mean dependent var	0.000630
Adjusted R-squared	0.002183	S.D. dependent var	0.016859
S.E. of regression	0.016840	Akaike info criterion	-5.695346
Sum squared resid	0.706446	Schwarz criterion	-5.683683
Log likelihood	7112.792	Durbin-Watson stat	2.023526

Inverted AR Roots .09

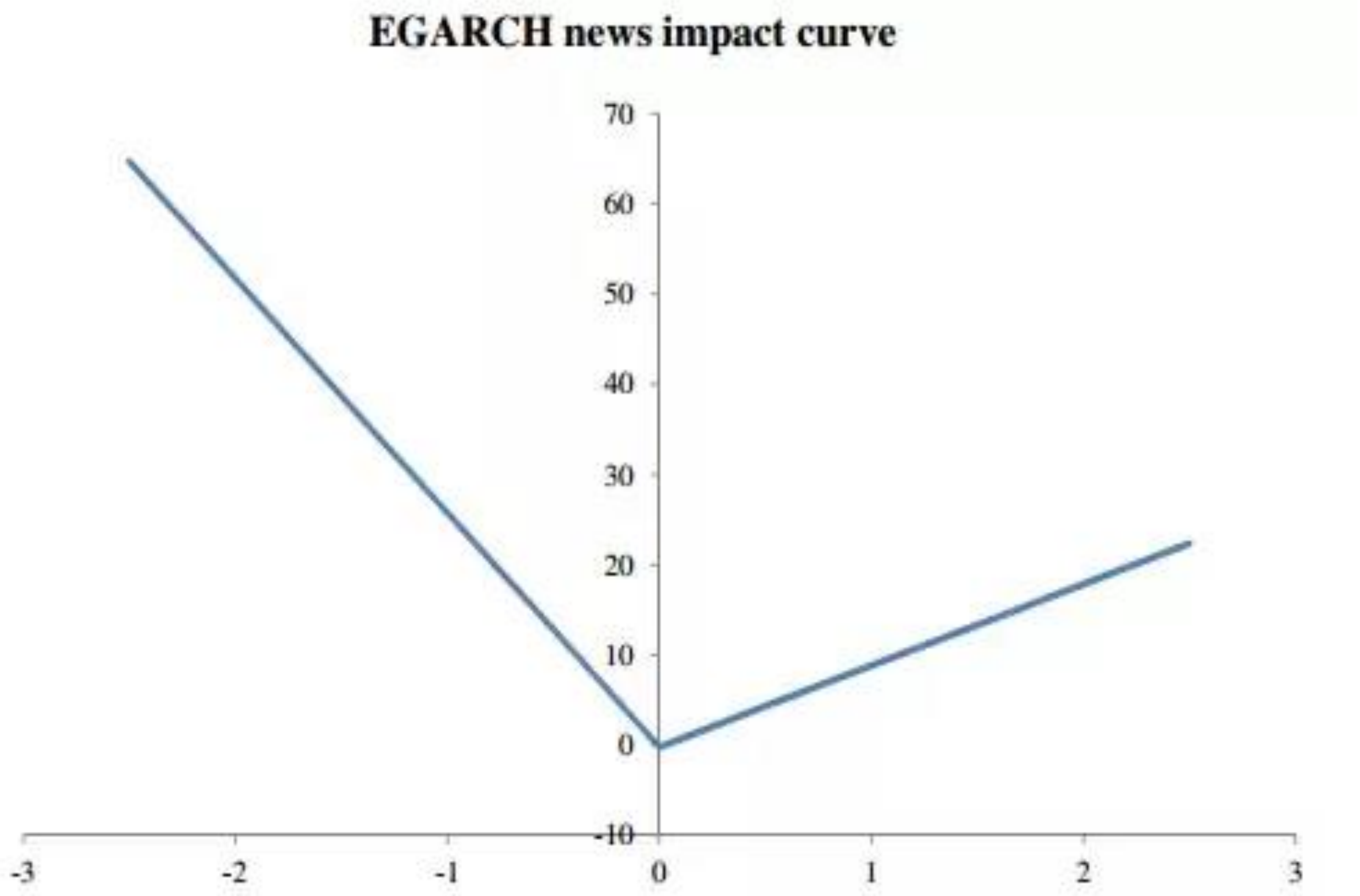
Capturing asymmetric Volatility

responding to the another stylized fact that

returns and volatility are negatively correlated in financial markets

(also known as Leverage Effect)

Asymmetric volatility



LIMITATION OF GARCH & STYLIZED FACT 2: ASYMMETRIC RESPONSE TO NEGATIVE & POSITIVE SHOCKS

- *Such models capture the empirical phenomenon that assets returns and volatility are negatively correlated*

- *Two such models are GJR-GARCH and EGARCH models.*

- *GJR (*Glosten, Jaganathan and Runkle*) or TGARCH (p,q,r) Model:*

$$\sigma_t^2 = \omega + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 + \sum_{i=1}^{p} \alpha_i \epsilon_{t-i}^2 + \sum_{k=1}^{r} \gamma_k \epsilon_{t-k}^2 I_{t-k}^-$$

GJR GARCH(1,1,1) model

$$\sigma_t^2 = \omega + \beta * \sigma_{t-1}^2 + \alpha * \varepsilon_{t-1}^2 + \gamma * \varepsilon_{t-1}^2 * \mathrm{I}_{t-1}^-$$

where $I_t^- = 1$ if $\epsilon_t < 0$ and 0 otherwise.

In this model, good news, $\epsilon_{t-i} > 0$, and bad news. $\epsilon_{t-i} < 0$, have differential effects on the conditional variance; good news has an impact of α_i, while bad news has an impact of $\alpha_i + \gamma_i$. If $\gamma_i > 0$, bad news increases volatility, and we say that there is a *leverage effect* for the i-th order. If $\gamma_i \neq 0$, the news impact is asymmetric.

Asymmetric Volatility & EGARCH Model

Exponential GARCH (EGARCH(p,q,r) Models

$$\log(\sigma_t^2) = \omega + \sum_{j=1}^{q} \beta_j \log(\sigma_{t-j}^2) + \sum_{i=1}^{p} \alpha_i \left| \frac{\epsilon_{t-i}}{\sigma_{t-i}} \right| + \sum_{k=1}^{r} \gamma_k \frac{\epsilon_{t-k}}{\sigma_{t-k}}$$

Exponential GARCH (EGARCH(1,1,1) Models

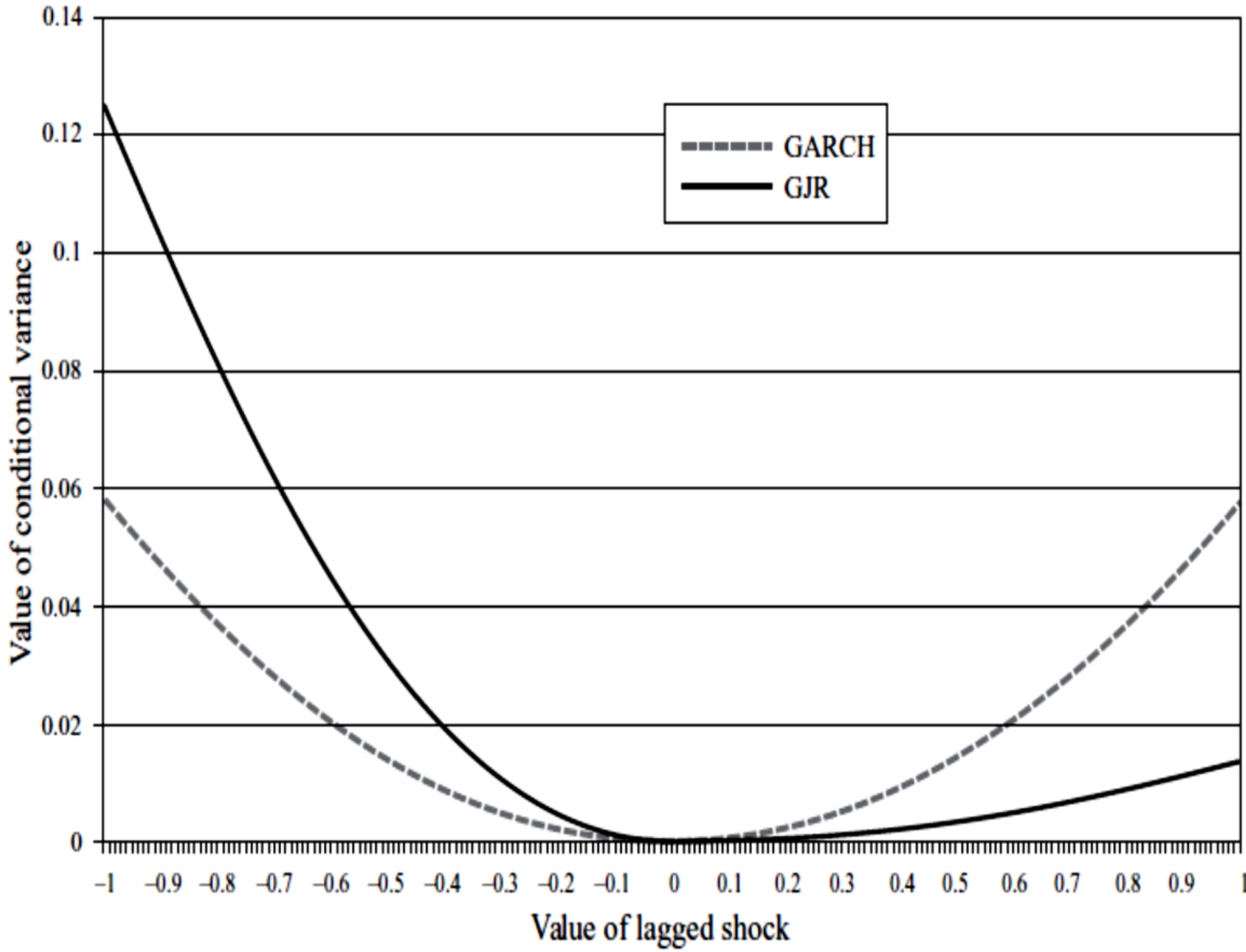
$$\log(\sigma_t^2) = \omega + \beta * \log(\sigma_{t-1}^2) + \alpha * \left| \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \right| + \gamma * \frac{\varepsilon_{t-1}}{\sigma_{t-1}}$$

Note that the left-hand side is the *log* of the conditional variance. This implies that the leverage effect is exponential, rather than quadratic, and that forecasts of the conditional variance are guaranteed to be nonnegative. The presence of leverage effects can be tested by the hypothesis that $\gamma_i < 0$. The impact is asymmetric if $\gamma_i \neq 0$.

Asymmetric Response of Shocks to Volatility

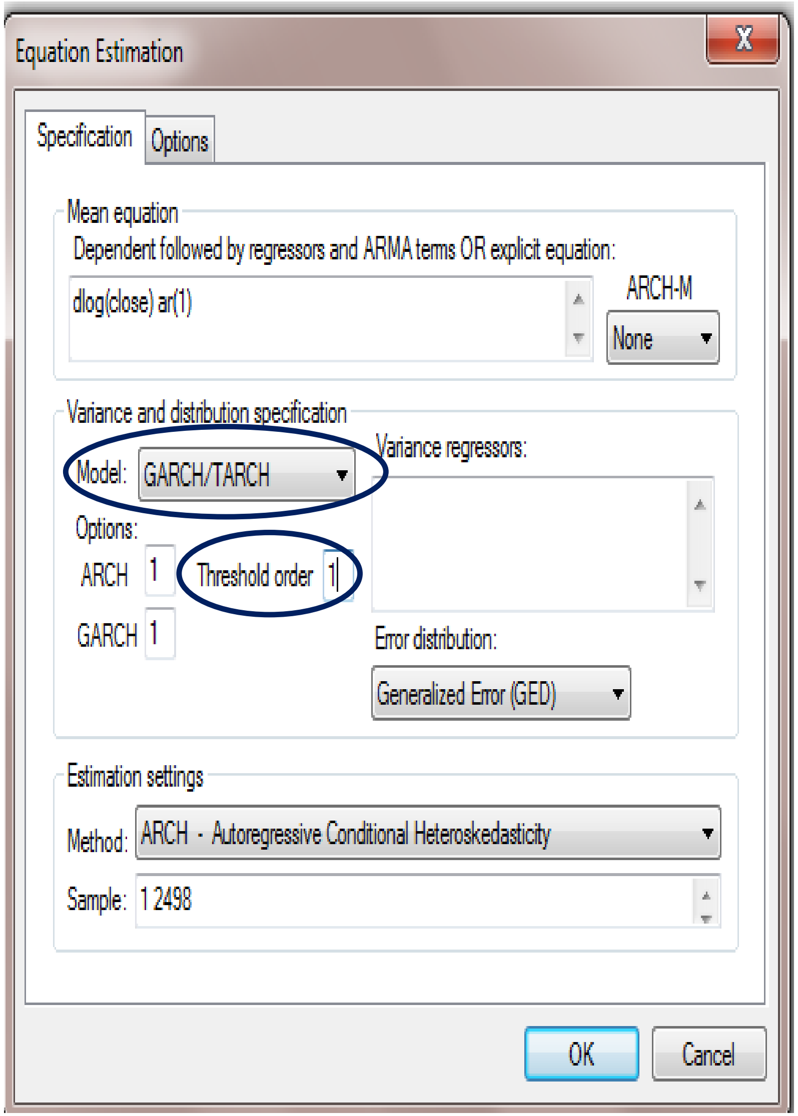
Figure 8.3

News impact curves for S&P500 returns using coefficients implied from GARCH and GJR model estimates



Asymmetric GARCH models: An Example

AR(1)-GJR_GARCH(1,1) MODEL WITH GED ERRORS



EViews - [Equation: UNTITLED Workfile: NIFTYVOL\Niftyvol]

File Edit Object View Proc Quick Options Window Help

View Proc Object Print Name Freeze Estimate Forecast Stats Resids

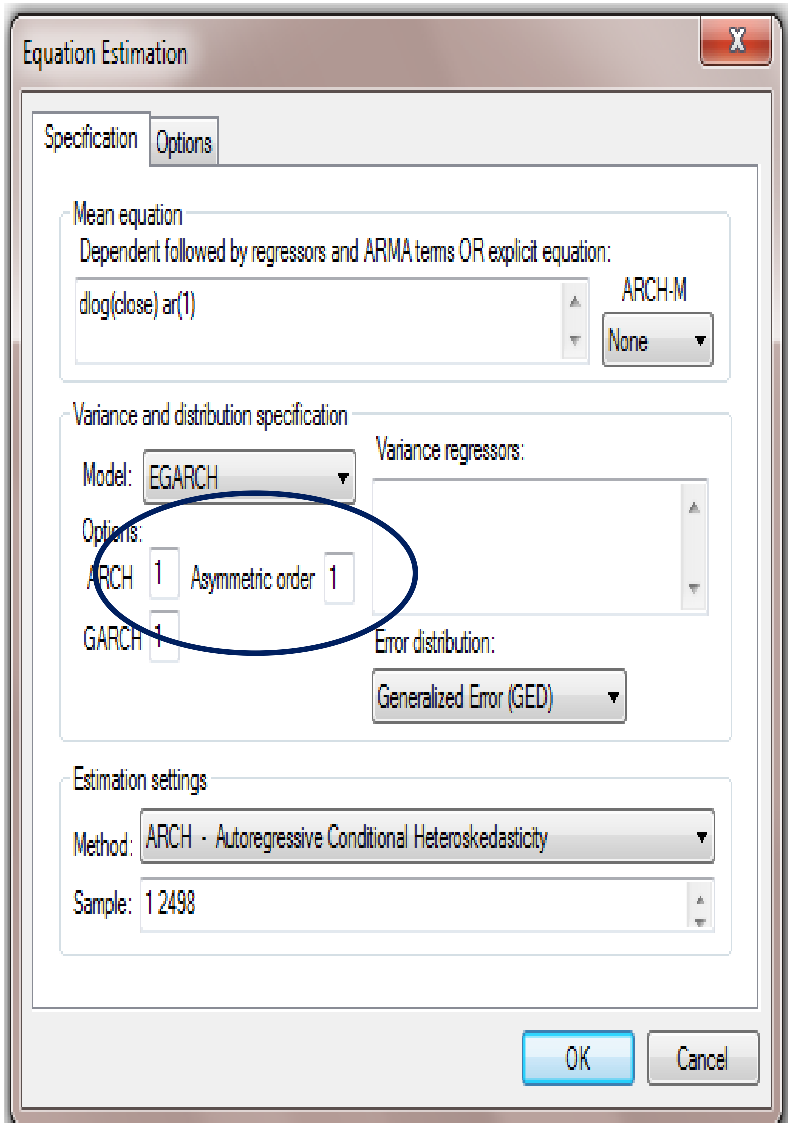
Dependent Variable: DLOG(CLOSE)
Method: ML - ARCH (Marquardt) - Generalized error distribution (GED)
Date: 01/22/12 Time: 15:46
Sample (adjusted): 3 2498
Included observations: 2496 after adjustments
Convergence achieved after 22 iterations
Variance backcast: ON
GARCH = C(2) + C(3)*RESID(-1)^2 + C(4)*RESID(-1)^2*(RESID(-1)<0)
+ C(5)*GARCH(-1)

	Coefficient	Std. Error	z-Statistic	Prob.
AR(1)	0.100087	0.021381	4.681051	0.0000
Variance Equation				
C	9.96E-06	1.59E-06	6.277029	0.0000
RESID(-1)^2	0.044531	0.016170	2.753928	0.0059
RESID(-1)^2*(RESID(-1)<0)	0.205353	0.030748	6.678608	0.0000
GARCH(-1)	0.822246	0.018572	44.27281	0.0000
GED PARAMETER	1.471119	0.039625	37.12588	0.0000

R-squared	0.003389	Mean dependent var	0.000630
Adjusted R-squared	0.001388	S.D. dependent var	0.016859
S.E. of regression	0.016847	Akaike info criterion	-5.715505
Sum squared resid	0.706725	Schwarz criterion	-5.701509
Log likelihood	7138.950	Durbin-Watson stat	2.040307

Inverted AR Roots .10

AR(1)-EGARCH(1,1) MODEL WITH GED ERRORS



EViews - [Equation: UNTITLED Workfile: NIFTYVOL\Niftyvol]

File Edit Object View Proc Quick Options Window Help

View Proc Object Print Name Freeze Estimate Forecast Stats Resids

Dependent Variable: DLOG(CLOSE)
Method: ML - ARCH (Marquardt) - Generalized error distribution (GED)
Date: 01/22/12 Time: 15:37
Sample (adjusted): 3 2498
Included observations: 2496 after adjustments
Convergence achieved after 24 iterations
Variance backcast: ON
LOG(GARCH) = C(2) + C(3)*ABS(RESID(-1)/@SQRT(GARCH(-1))) + C(4)*RESID(-1)/@SQRT(GARCH(-1)) + C(5)*LOG(GARCH(-1))

	Coefficient	Std. Error	z-Statistic	Prob.
AR(1)	0.106917	0.021065	5.075576	0.0000
Variance Equation				
C(2)	-0.639317	0.073509	-8.697153	0.0000
C(3)	0.264911	0.026923	9.839533	0.0000
C(4)	-0.137263	0.016773	-8.183649	0.0000
C(5)	0.948252	0.007460	127.1112	0.0000
GED PARAMETER	1.476599	0.040811	36.18102	0.0000

R-squared	0.002985	Mean dependent var	0.000630
Adjusted R-squared	0.000983	S.D. dependent var	0.016859
S.E. of regression	0.016851	Akaike info criterion	-5.715202
Sum squared resid	0.707012	Schwarz criterion	-5.701205
Log likelihood	7138.572	Durbin-Watson stat	2.053266

Inverted AR Roots	.11

UNCONDITIONAL VOLATILITY IN EGARCH MODEL

$$\sigma_L^2 \, or V_L = \exp\left(\frac{\omega}{1-\beta}\right)$$

UNCONDITIONAL VOLATILITY IN GJR-GARCH MODEL

$$\sigma_L^2 \, or V_L = \frac{\omega}{1-\alpha-\frac{\gamma}{2}-\beta}$$

Can the asymmetry coefficient have opposite sign as well (i.e., Positive for E-GARCH and negative for GJR-GARCH)

GJR-GARCH model on Re/$ Forex Data (Based on Direct Quote)

EViews - [Equation:

File Edit Object View Proc Quick Options Window Help

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: RET_D
Method: ML - ARCH (Marquardt) - Normal distribution
Date: 08/06/13 Time: 11:58
Sample (adjusted): 2 1151
Included observations: 1150 after adjustments
Convergence achieved after 16 iterations
Presample variance: backcast (parameter = 0.7)
GARCH = C(2) + C(3)*RESID(-1)^2 + C(4)*RESID(-1)^2*(RESID(-1)<0)
+ C(5)*GARCH(-1)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	0.000106	0.000157	0.676449	0.4988
		Variance Equation		
C	1.10E-06	3.11E-07	3.521536	0.0004
RESID(-1)^2	0.164144	0.026527	6.187719	0.0000
RESID(-1)^2*(RESID(-1)<0)	-0.068682	0.027820	-2.468753	0.0136
GARCH(-1)	0.843396	0.023192	36.36553	0.0000

R-squared	-0.000282	Mean dependent var	0.000208
Adjusted R-squared	-0.003776	S.D. dependent var	0.006061
S.E. of regression	0.006073	Akaike info criterion	-7.542037
Sum squared resid	0.042225	Schwarz criterion	-7.520092
Log likelihood	4341.671	Hannan-Quinn criter.	-7.533753
Durbin-Watson stat	1.981912		

E-GARCH MODEL ON RE/$ FOREX DATA (BASED ON DIRECT QUOTE)

EViews - [Equati

File Edit Object View Proc Quick Options Window He

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: RET_D
Method: ML - ARCH (Marquardt) - Normal distribution
Date: 08/06/13 Time: 13:21
Sample (adjusted): 2 1151
Included observations: 1150 after adjustments
Convergence achieved after 23 iterations
Presample variance: backcast (parameter = 0.7)
LOG(GARCH) = C(2) + C(3)*ABS(RESID(-1)/@SQRT(GARCH(-1))) +
C(4)*RESID(-1)/@SQRT(GARCH(-1)) + C(5)*LOG(GARCH(-1))

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	0.000143	0.000155	0.920642	0.3572
		Variance Equation		
C(2)	-0.676372	0.137547	-4.917385	0.0000
C(3)	0.264170	0.036800	7.178460	0.0000
C(4)	0.044059	0.017842	2.469344	0.0135
C(5)	0.954828	0.011748	81.27678	0.0000

R-squared	-0.000114	Mean dependent var	0.000208
Adjusted R-squared	-0.003608	S.D. dependent var	0.006061
S.E. of regression	0.006072	Akaike info criterion	-7.541678
Sum squared resid	0.042218	Schwarz criterion	-7.519733
Log likelihood	4341.465	Hannan-Quinn criter.	-7.533394
Durbin-Watson stat	1.982244		

Important link for Time Series analysis

https://onlinecourses.science.psu.edu/stat510/?q=node/41