■ **Mixed ARMA model:** A compact form for flexible models. As if things weren't complicated enough, a process can be a mixture of AR and MA components. That is, there is a more general class of process called ARMA(p,q) models that consist of (a) an autoregressive component with p lags, and (b) a moving average component with q lags.

Focus on the ARMA(1,1) model for

1. simplicity
2. useful for understanding GARCH models in Ch. 3 for volatility modeling.

ARMA(1,1) Model:

- Form: $r_t = \phi_1 r_{t-1} + \phi_0 + a_t - \theta a_{t-1}$ or

$$(1 - \phi_1 B) r_t = \phi_0 + (1 - \theta B) a_t$$

  where $\{a_t\}$ is assumed to be white noise with mean 0 and variance σ_a^2. This is a combination of an AR(1) on the LHS and an MA(1) on the RHS.
- Stationarity: same as AR(1)
- Invertibility: same as MA(1)
- (Unconditional) mean: same as AR(1), i.e. $E[r_t] = \mu = \frac{\phi_0}{1-\phi_1}$
- Variance: $Var[r_t] = \gamma_0 = \frac{1+\theta_1^2 - 2\phi_1\theta_1}{1-\phi_1^2}\sigma_a^2$. If $|\phi_1| < 1$, then $\gamma_0 > 0$ and finite.
- $\gamma_1 = \phi_1\gamma_0 - \theta_1\sigma_a^2$.
- $\gamma_\ell = \phi_1\gamma_{\ell-1},\ \text{for } \ell \geq 2.$
- ACF: Satisfies $\rho_\ell = \phi_1\rho_{\ell-1}$, for $\ell > 1$, but

$$\rho_1 = \phi_1 - \frac{\theta_1\sigma_a^2}{\gamma_0} \neq \phi_1.$$

  This is the difference between AR(1) and ARMA(1,1) models.
- PACF: does not cut off at finite lags.
- Note that the ACF and PACF are not informative in determining the order of the ARMA(p,q) model.

Building an ARMA(1,1) model

- Specification: use extended ACF (EACF) or AIC
- Estimation: cond. or exact likelihood method
- Model checking: as before
- Forecast: MA(1) affects the 1-step ahead forecast. others are similar to those of AR(1) models.

Three model representations:

1. ARMA form: compact, useful in estimation and forecasting

2. AR representation:
$$r_t = \phi_0 + \pi_1 r_{t-1} + \pi_2 r_{t-2} + ...$$
It tells how r_t depends on its past values.
3. MA representation:
$$r_t = \mu + a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + ...$$
It tells how r_t depends on the past shocks.
For a stationary series, $\psi_i = \phi_1^i(\phi_1 - \theta_1)$ converges to zero as $i \longrightarrow \infty$. Thus, the effect of any shock is transitory.
The MA representation is particularly useful in computing variances of forecast errors.

For an ℓ-step ahead forecast,

$$r_n(\ell) = E[r_{n+\ell}|r_{1:n}] = E[\mu + a_{n+\ell} + \psi_1 a_{n+\ell-1} + \psi_2 a_{n+\ell-2} + ...|a_{1:n}] = \mu + 0 + ... + 0 + \psi_\ell a_n + \psi_{\ell-1} a_{n-1} + ..$$

Therefore, $r_n(\ell) \to \mu$ as $\ell \to \infty$.The speed by which $r_n(\ell)$ approaches μ determines the speed of mean reversion. The forecast error is

$$e_n(\ell) = a_{n+\ell} + \psi_1 a_{n+\ell-1} + ... + \psi_{\ell-1} a_{n+1}$$

The variance of forecast error is

$$Var[e_n(\ell)] = (1 + \psi_1^2 + ... + \psi_{\ell-1}^2)\sigma_a^2.$$

and $Var[e_n(\ell)] \to Var[r_t]$ as $\ell \to \infty$.

Unit-root Non-stationarity:

Roots that lie on the unit circle are right at the threshold that marks the transition from stationarity. The problem with unit-root processes is that they look stationary in small samples. But treating them as stationary leads to very misleading results. Moreover, regressing one non-stationary process on another, leads many "false positives" where two variables seem related when they are not.

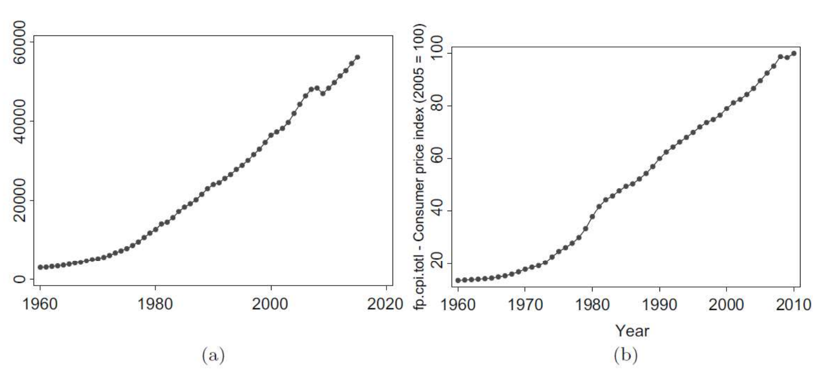

Figure 29: Two non-stationary economic time series. (a) Nominal GDP per capita. (b) Consumer price index.

Random Walk

- Random walk is the best known unit-root non-stationary series.
- Form $p_t = p_{t-1} + a_t$
- Unit root? It is an AR(1) model with coefficient $\phi_1 = 1$.
- Repeated substitution shows

$$p_t = \sum_{i}^{\infty} a_{t-i} = \sum_{i=0}^{\infty} \psi_i a_{t-i}$$

  where $\psi_i = 1$ for all i. Thus, ψ_i does not converge to zero as i increases. The effect of any shock is permanent.
- Strong memory: sample ACF approaches 1 for any finite lag.
- Nonstationary: Because the variance of r_t diverges to infinity as t increases.

Random Walk with Drift

- Form $p_t = \mu + p_{t-1} + a_t$, $\mu \neq 0$
- Has a unit root.
- Nonstationary
- Strong memory
- Has a time trend with slope μ. Why?

  Remark: There are many similarities between (a) random walks with drift and (b) deterministic trend processes. They are both non-stationary, but the source of the non-stationarity is different.

a) Random walks with drift	b) Trend-stationary series
$X_t = \mu + X_{t-1} + a_t = \mu t + \sum_{i=1}^{t} a_i$	$X_t = \mu t + a_t$
$E[X_t] = \mu t$	$E[X_t] = \mu t$
$Var[X_t] = \sigma_a^2 t$	$Var[X_t] = \sigma_a^2$

  Both models have means which increase linearly over time. This makes it very difficult to visually identify which process generated the data. The variance of the random walk with drift, however, grows over time, while the variance of the deterministic trend model does not.

  First differencing (discussed later) the deterministic trend model introduces an MA unit root in the error terms. Never take first differences to remove a deterministic trend. Rather, regress X on time, and then work with the residuals. These residuals now represent X that has been linearly detrended. Such a process is called "trend stationary."

■ Meaning of the constant term in a model

- MA model: mean

- AR model: related to mean
- Random walk with drift. time trend

Practical implication in financial time series:

Example: Monthly log returns of General Electrics (GE) from 1926 to 1999 (74 years) Sample mean: 1.04%, std($\hat{\mu}$) = 0.26. Very significant! is about 12.45% a year

$1 investment in the beginning of 1926 is worth

- annual compounded payment: $5,907
- quarterly compounded payment: $8,720
- monthly compounded payment: $9,570
- Continuously compounded? \$ $e^{0.1245\times 73}$

Differencing: If we have a time series whose mean is increasing, we can apply the difference operator enough times to render the series stationary. If a series needs to be differenced once in order to make it stationary, we say that the series is "integrated of order one" or "I(1)." A series that needs to be differenced twice is "integrated of order two" and is "I(2)." In general, if a series needs to be differenced d times, it is said to be "integrated of order d" and is "I(d)."

- 1st difference: $r_t = p_t - p_{t-1}$
  If p_t is the log price, then the 1st difference is simply the log return. Typically, 1st difference means the "change" or "increment" of the original series.
- **ARIMA(p,d,q):** If the dth difference of the series x_t follows an ARMA(p,q), the original series (x_t) follows an ARIMA(p,d,q) model.

  For example, if $c_t = x_t - x_{t-1}$ follows an ARMA(p,q), then x_t follows an ARMA(p,1,q).

Unit-root Tests

Let p_t be the log price of an asset. To test that p_t has a unit root (i.e. is not predictable), the following model is employed:

$$p_t = \phi_0 + \phi_1 p_{t-1} + e_t$$

The hypothesis of interest is $H_0 : \phi_1 = 1$ vs $H_a : \phi_1 < 1$.

Dickey-Fuller test is the usual t-ratio of the OLS estimate of ϕ_1 being 1. This is the DF unit-root test. The t-ratio, however, has a nonstandard limiting distribution (because under H_0 the series in both sides of the equation are non-stationary).

$$t \text{ ratio} = \frac{\hat{\phi}_1 - 1}{\text{std}(\hat{\phi}_1)}$$

Let $\Delta p_t = p_t - p_{t-1}$. Then, the *augmented DF unit-root test* for an AR(p) model is based on

$$\Delta p_t = c_t + \beta p_{t-1} + \sum_{i=1}^{p-1} \phi_i \Delta p_{t-i} + e_t$$

The t-ratio of the OLS estimate of β is the ADF (Augmented Dickey-Fuller) unit-root test statistic. Again, the statistic has a non-standard limiting distribution.

$$t \text{ ratio} = \frac{\hat{\beta} - 0}{\text{std}(\hat{\beta})}$$

Example: Consider the log series of U.S. quarterly real GDP series from 1947.I to 2009.IV. (data from Federal Reserve Bank of St. Louis). See *q-gdpc96.txt* on the course web.

R Demonstration:

```
> library(fUnitRoots)
> help(UnitrootTests) % See the tests available
>da=read.table(‘‘q-gdpc96.txt’’,header=T)
>gdp=log(da[,4])
> adfTest(gdp,lag=4,type=c("c")) #Assume an AR(4) model for the series.
Title:
Augmented Dickey-Fuller Test
Test Results:
PARAMETER:
Lag Order: 4
STATISTIC:
Dickey-Fuller: -1.7433
P VALUE:
0.4076 # cannot reject the null hypothesis of a unit root.
*** A more careful analysis
> x=diff(gdp)
> ord=ar(x) # identify an AR model for the differenced series.
> ord
Call:
ar(x = x)
Coefficients:
1 2 3
0.3429 0.1238 -0.1226
Order selected 3 sigma^2 estimated as 8.522e-05
# An AR(3) for the differenced data is confirmed.
# Our previous analysis is justified.
```

Discussion: The command *arima* on R.

1. Dealing with the constant term. If there is any differencing, no constant is used. The subcommand include.mean=T in the *arima* command.

2. Fixing some parameters. Use subcommand *fixed* in *arima*. Use unemployment rate series as an example.

Really, there is little new here regarding estimation. We simply need to find out whether differencing a variable a small number of times renders a variable stationary. We then proceed with the ARMA(p,q) portion of the analysis.

■ **Seasonal Time Series:**

TS with periodic patterns and useful in

- predicting quarterly earnings
- pricing weather-related derivatives
- analysis of transactions data (high-frequency data), e.g., U-shaped pattern in intraday trading intensity, volatility, etc.

Of course, when we say "seasonality" here, we simply mean any sort of periodicity. A weekly recurring pattern is seasonal, but at a weekly frequency.

Seasonality can have different lengths. Retail sales vary seasonally with the holiday shopping season, but they also have "seasonality" at a weekly frequency: weekend sales are higher than weekday sales.

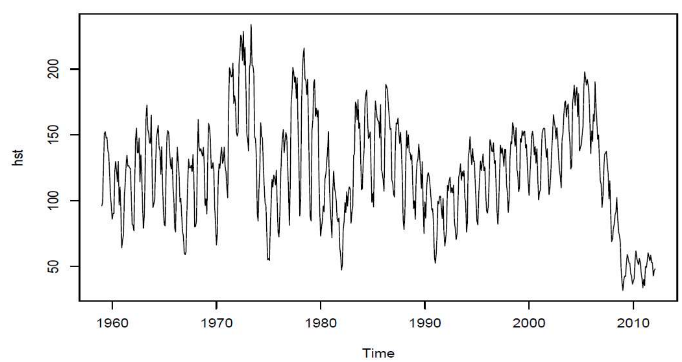

Figure 30: Time plot of monthly U.S. housing starts: 1959.1-2012.2. Data obtained from US Bureau of the Census. The data are in thousand units.

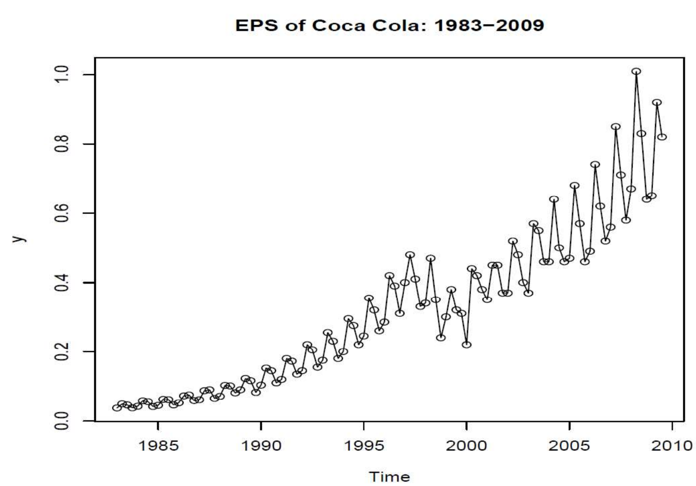

Figure 31: Time plot of quarterly earnings per share of KO (Coca Cola) from 1983 to 2009.

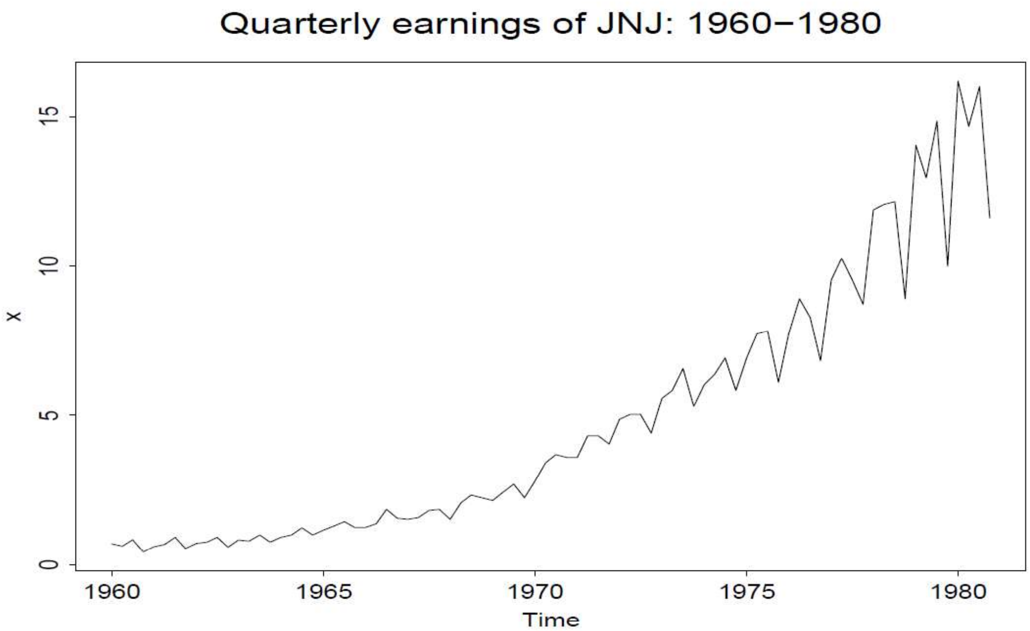

Figure 32: Time plot of quarterly earnings of Johnson and Johnson: 1960-1980.

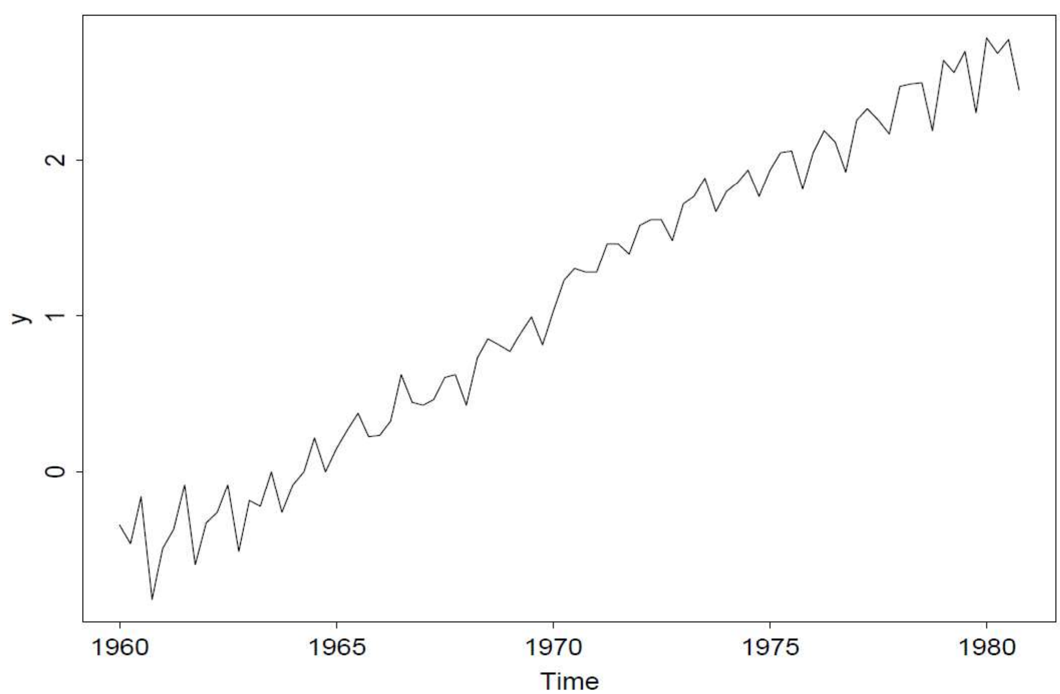

Figure 33: Time plot of quarterly logged earnings of Johnson and Johnson: 1960-1980.

Seasonality can be deterministic or stochastic:

(1) The seasonal differences can vary by the same amount, or by the same percent, each year. Such deterministic seasonality is best captured with the use of seasonal dummy variables. If the dependent variable is in levels, then the dummies capture level shifts; if the dependent variable is logged, then they capture equal percentage changes. For example, if Christmas shopping shows deterministic seasonality, then retail sales might show a spike of approximately 20% growth every winter, for example.

(2) On the other hand, seasonal differences may vary over time. In this case, Christmas shopping might be higher than other seasons, but some years the spike is much larger than 20%, other years much lower, with the differences evolving slowly over time. The best way to capture such evolution is to think of the seasonal spike as being a unit root process, a random walk. Christmas sales might be unusually strong, followed by sequences of Christmas sales that are a bit weaker.

If the seasonality is deterministic, then we should use dummy variables. If the seasonality varies stochastically, then a seasonal unit root process captures the evolving dynamics quite nicely, and seasonal differencing should be used. Seasonal differencing means that an observation is subtracted from the previous one for the same season. If the data are quarterly and there is quarterly seasonality, then the seasonal first difference is: X_t-X_{t-4}. If the data are monthly, then the seasonal first difference is: X_t-X_{t-12}.

For non-stationary trending data with seasonality, two levels of differencing are required: first differences to remove the unit root in long-run growth, and seasonal differences to remove seasonal unit roots. ACFs and PACFs will prove themselves especially useful in detecting seasonality, as fourth-quarter GDP in one year should tend to be correlated with fourth-quarter GDP in another year.

Seasonality can be additive or multiplicative:

Additive models:

It is easy to add seasonal terms to an ARIMA model. Additive seasonality simply requires adding an AR term or an MA term at a seasonal frequency. If you have quarterly seasonality, then add a lag-4 AR term, or a lag-4 MA term. For a simple example, consider the following model for quarterly data:

$$X_t = \beta_4 X_{t-4} + e_t \qquad \overrightarrow{\text{which is equivallent to}} \quad (1 - \beta_4 B^4)X_t = e_t$$

A simple additive seasonal MA model might be:

$$X_t = a_t - \gamma_4 a_{t-4} \qquad \overrightarrow{\text{which is equivallent to}} \quad (1 - \gamma_4 B^4)a_t = X_t$$

Multiplicative models:

Consider the housing-starts series. Let y_t be the monthly data. Denoting 1959 as year 0, we can write the time index as $t =$ year + month, e.g, $y_1 = y_{0,1}$, $y_2 = y_{0,2}$, and $y_{14} = y_{1,2}$, etc. The multiplicative model is based on the following consideration:

The column dependence is the usual lag-1, lag-2, ... dependence. That is, monthly dependence. We call them the regular dependence. The row dependence is the year-to-year dependence. We call them the seasonal dependence. So, December sales

	Month						
Year	Jan	Feb	Mar	$\cdots$	Oct	Nov	Dec
1959	$y_{0,1}$	$y_{0,2}$	$y_{0,3}$	$\cdots$	$y_{0,10}$	$y_{0,11}$	$y_{0,12}$
1960	$y_{1,1}$	$y_{1,2}$	$y_{1,3}$	$\cdots$	$y_{1,10}$	$y_{1,11}$	$y_{1,12}$
1961	$y_{2,1}$	$y_{2,2}$	$y_{2,3}$	$\cdots$	$y_{2,10}$	$y_{2,11}$	$y_{2,12}$
1962	$y_{3,1}$	$y_{3,2}$	$y_{3,3}$	$\cdots$	$y_{3,10}$	$y_{3,11}$	$y_{3,12}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$

are not always independent of November sales. There is some inertia to human behavior. If November sales are unusually brisk, then we might expect this to carry over into December. For this reason, a purely additive seasonal model would be inadequate. Box and Jenkins (1976) propose a multiplicative model of seasonality.

Multiplicative seasonality allows us to capture a lot of complexity with few parameters: parsimony. A simple example might be:

$$(1-\beta_1 B)(1-\beta_4 B^4)y_t = a_t$$

or

$$y_t = \beta_1 y_{t-1} + \beta_4 y_{t-4} + \beta_1\beta_4 y_{t-5} + a_t$$

Modeled in this way, two parameters (β_1 and β_4) allow for lags at three different lengths: 1, 4, and 5 (parsimony).

Airline model for quarterly series

- Form

  $$(1-B)(1-B^4)y_t = (1-\theta_1 B)(1-\theta_4 B^4)a_t$$

  or

  $$y_t - y_{t-1} - y_{t-4} + y_{t-5} = a_t - \theta_1 a_{t-1} - \theta_4 a_{t-4} + \theta_1\theta_4 a_{t-5}$$

- Define the differenced series w_t as

  $$w_t = y_t - y_{t-1} - y_{t-4} + y_{t-5} = y_t - y_{t-1} - (y_{t-4} - y_{t-5}).$$

  It is called regular and seasonal differenced series.
- ACF of w_t has a nice symmetric structure, i.e. $\rho_3 = \rho_5 = \rho_1\rho_4$. Also, $\rho_\ell = 0$ for $\ell > 5$
- This model is widely applicable to many seasonal time series.
- Multiplicative model means that the regular and seasonal dependences are roughly orthogonal to each other.

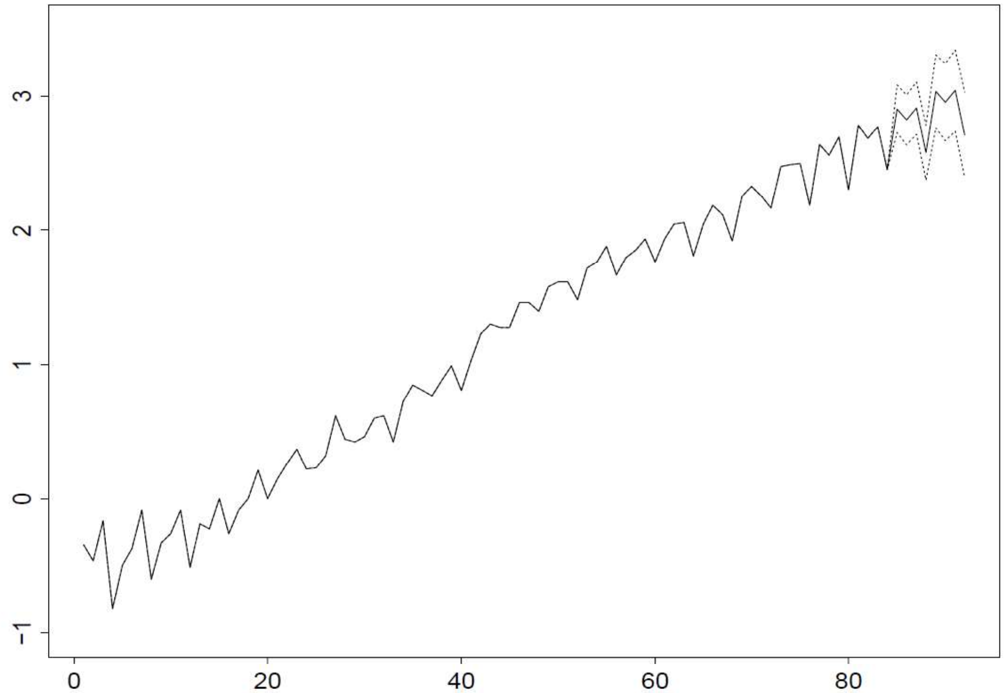

Figure 34: Forecast plot for the quarterly earnings of Johnson and Johnson. Data: 1960-1980, Forecasts: 1981-82.

- Forecasts: exhibit the same pattern as the observed series. See Figure 34
- For data with periodicity s, the Airline model becomes

$$(1-B)(1-B^s)y_t = (1-\theta_1 B)(1-\theta_s B^s)a_t,\ a_t \overset{\text{i.i.d}}{\sim} (0,\sigma_a^2),\ |\theta_1|<1,\ |\theta_s|<1$$

ACF of $w_t = (1-B)(1-B^s)y_t$ has a nice symmetric structure

$$\rho_{s-1} = \rho_{s+1} = \rho_1\rho_s,\ \rho_\ell = 0 \text{ for } \ell > s$$

- The general form of a multiplicative seasonal model is

$$(1-B)(1-B^s)(1-\phi_1 B - \ldots - \phi_p B^p)y_t = (1-\theta_s B^s)(1-\theta_1 B - \ldots - \theta_q B^q)a_t,\ a_t \overset{\text{i.i.d}}{\sim} (0,\sigma_a^2)$$

Example: Detailed analysis of J&J earnings.

R Demonstration:

```
> x=ts(scan("q-earn-jnj.txt"),frequency=4,start=c(1960,1)) % create a time series object.
> plot(x) % Plot data with calendar time
> y=log(x) % Natural log transformation
> plot(y) % plot data
> c1=paste(c(1:4))
> points(y,pch=c1) % put circles on data points.
> par(mfcol=c(2,1)) % two plots per page
> acf(y,lag.max=16)
> y1=as.vector(y) % Creates a sequence of data in R
> acf(y1,lag.max=16)
> dy1=diff(y1) % regular difference
> acf(dy1,lag.max=16)
> sdy1=diff(dy1,4) % seasonal difference
> acf(sdy1,lag.max=12)
> m1=arima(y1,order=c(0,1,1),seasonal=list(order=c(0,1,1),period=4)) % Airline model in R.
> m1
Call:
arima(x = y1, order = c(0, 1, 1), seasonal = list(order = c(0, 1, 1), period = 4))
Coefficients:
ma1 sma1
-0.6809 -0.3146 % The fitted model is (1-B^4)(1-B)R(t) =(1-0.68B)(1-0.31B^4)a(t), var[a(t)] = 0.00793.
s.e. 0.0982 0.1070
sigma^2 estimated as 0.00793: log likelihood = 78.38, aic = -150.75
> par(mfcol=c(1,1)) % One plot per page
> tsdiag(m1) % Model checking
> f1=predict(m1,8) % prediction
> names(f1)
[1] "pred" "se"
> f1
$pred % Point forecasts
Time Series:
Start = 85
End = 92
Frequency = 1
[1] 2.905343 2.823891 2.912148 2.581085 3.036450 2.954999 3.043255 2.712193
$se % standard errors of point forecasts
Time Series:
Start = 85
End = 92
Frequency = 1
[1] 0.08905414 0.09347895 0.09770358 0.10175295 0.13548765 0.14370550
[7] 0.15147817 0.15887102
# You can use ‘‘foreplot’’ to obtain plot of forecasts.
```

■ **Regression Models with Time Series Errors**

In many applications, the relationship between two time series is of major interest. An obvious example is the market model in finance that relates the excess return of an individual stock to that of a market index.

$$y_t = \alpha + \beta x_t + e_t$$

The least squares (LS) method is often used to estimate model. If $\{e_t\}$ is a white noise series, then the LS method produces consistent estimates. In practice, however, it is common to see that the error term $\{e_t\}$ is serially correlated. In this case, we have a regression model with time series errors, and the LS estimates of α and β may not be consistent. A regression model with time series errors is widely applicable in economics and finance, but it is one of the most commonly misused econometric models because the serial dependence in $\{e_t\}$ is often overlooked. It pays to study the model carefully.

Example: U.S. weekly interest rate data: 1-year and 3-year constant maturity rates. Data are shown in Figure 35.

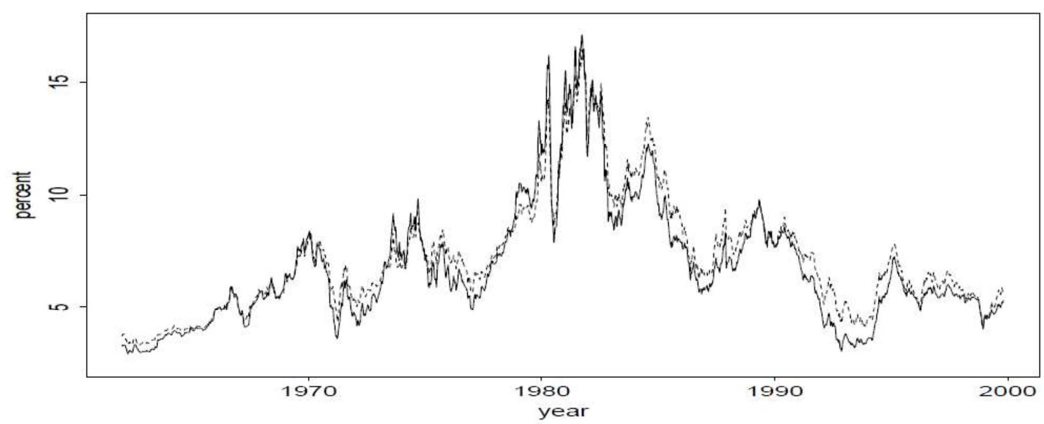

Figure 35: Time plots of U.S. weekly interest rates: 1-year constant maturity rate (solid line) and 3-year rate (dashed line).

We outline a general procedure for analyzing linear regression models with time series errors:

1. Fit the linear regression model and check serial correlations of the residuals.

2. If the residual series is unit-root non-stationary, take the first difference of both the dependent and explanatory variables. Go to step 1. If the residual series appears to be stationary, identify an ARMA model for the residuals and modify the linear regression model accordingly.

3. Perform a joint estimation via the maximum-likelihood method and check the fitted model for further improvement.

R Demonstration:

```
> da=read.table("w-gs1n36299.txt") % load the data
> r1=da[,1] % 1-year rate
> r3=da[,2] % 3-year rate
> plot(r1,type='l') % Plot the data
> lines(1:1967,r3,lty=2)
> plot(r1,r3) % scatter plot of the two series
> m1=lm(r3~r1) % Fit a regression model with likelihood method.
> summary(m1)
Call: lm(formula = r3 ~ r1)
Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept) 0.910687 0.032250 28.24 <2e-16 ***
r1 0.923854 0.004389 210.51 <2e-16 ***
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 0.538 on 1965 degrees of freedom
Multiple R-Squared: 0.9575, Adjusted R-squared: 0.9575
F-statistic: 4.431e+04 on 1 and 1965 DF, p-value: < 2.2e-16
> acf(m1$residuals)
> c3=diff(r3)
> c1=diff(r1)
> plot(c1,c3)
> m2=lm(c3~c1) % Fit a regression with likelihood method.
> summary(m2)
Call:
lm(formula = c3 ~ c1)
Residuals:
Min 1Q Median 3Q Max
-0.3806040 -0.0333840 -0.0005428 0.0343681 0.4741822
Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept) 0.0002475 0.0015380 0.161 0.872
c1 0.7810590 0.0074651 104.628 <2e-16 ***
---
Residual standard error: 0.06819 on 1964 degrees of freedom
Multiple R-Squared: 0.8479, Adjusted R-squared: 0.8478
F-statistic: 1.095e+04 on 1 and 1964 DF, p-value: < 2.2e-16
> acf(m2$residuals)
> plot(m2$residuals,type='l')
> m3=arima(c3,xreg=c1,order=c(0,0,1)) % Residuals follow an MA(1) model
> m3
Call: arima(x = c3, order = c(0, 0, 1), xreg = c1)
Coefficients:
ma1 intercept c1 % Fitted model is c3 = 0.0002+0.782c1 + a(t)+0.212a(t-1) with var[a(t)] = 0.00446.
0.2115 0.0002 0.7824 a(t) = 0.192a(t-1)+e(t).
s.e. 0.0224 0.0018 0.0077 a(t) = 0.192a(t-1)+e(t).
sigma^2 estimated as 0.004456: log likelihood = 2531.84, aic = -5055.69
> acf(m3$residuals)
> tsdiag(m3)
> m4=arima(c3,xreg=c1,order=c(1,0,0)) %Residuals follow an AR(1) model.
> m4
Call:
arima(x = c3, order = c(1, 0, 0), xreg = c1)
Coefficients:
ar1 intercept c1 % Fitted model is c3 = 0.0003 + 0.783c1 + a(t), a(t) = 0.192a(t-1)+e(t).
0.1922 0.0003 0.7829
s.e. 0.0221 0.0019 0.0077
sigma^2 estimated as 0.004474: log likelihood = 2527.86, aic = -5047.72
```

- Long-memory models (for reference only)
  - Meaning? ACF decays to zero very slowly!
  - Example: ACF of squared or absolute log returns
    ACFs are small, but decay very slowly.
  - How to model long memory? Use "fractional" difference: namely, $(1-B^d)r_t$, where $-0.5 < d < 0.5$.
  - Importance? In theory, Yes. In practice, yet to be determined.
  - In R, the package *rugarch* may be used to estimate the fractionally integrated ARMA models. The package can also be used for GARCH modeling.

Summary of the chapter Little research is conducted these days using only univariate ARIMA models, but they are quite important. The concepts surrounding ARIMA modeling are foundational to time series; AR and MA processes are the component pieces of many more complicated models, and the problems of integration and non-stationarity must be dealt with in any time-series setting. Mastering these ingredients ensures that the more complicated material will be more digestible.

1. Sample ACF $\longrightarrow$ MA order
2. Sample PACF $\longrightarrow$ AR order
3. Some packages have "automatic" procedure to select a simple model for "conditional mean" of a FTS, e.g., R uses "*ar*" for AR models.
4. Check a fitted model before forecasting, e.g. residual ACF and hetroscedasticity (chapter 3)
5. Interpretation of a model, e.g. constant term
   For an AR(1) with coefficient ϕ_1 the speed of mean reverting as measured by half-life is
   $$k = \frac{ln(0.5)}{ln(|\phi_1|)}$$
   For an MA(q) model, forecasts revert to the mean in $q+1$ steps.
6. Make proper use of regression models with time series errors, e.g. regression with AR(1) residuals
   Perform a joint estimation instead of using any two-step procedure, e.g. Cochrane-Orcutt (1949).
7. Basic properties of a random-walk model
8. Multiplicative seasonal models, especially the so-called airline model.

Question: Why don't we use R-square in this course?

- Impact of serial correlations in regression is often overlooked.
- It may introduce biases in estimates and in standard errors, resulting in unreliable t-ratios.
- Detecting residual serial correlation: Use Q-stat instead of DW-statistic, which is not sufficient!
- Avoid the problem of spurious regression. R-square can be misleading!!!