Basic Business Statistics
11th Edition

Chapter 16

Time-Series Forecasting and Index Numbers

Learning Objectives

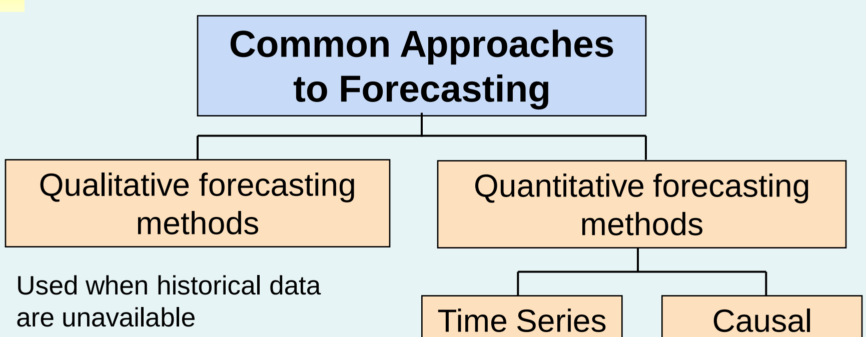
In this chapter, you learn:

- About different time-series forecasting models: moving averages, exponential smoothing, linear trend, quadratic trend, exponential trend, autoregressive models, and least squares models for seasonal data
- To choose the most appropriate time-series forecasting model
- About price indexes and differences between aggregated and unaggregated indexes

The Importance of Forecasting

- Governments forecast unemployment rates, interest rates, and expected revenues from income taxes for policy purposes
- Marketing executives forecast demand, sales, and consumer preferences for strategic planning
- College administrators forecast enrollments to plan for facilities and for faculty recruitment
- Retail stores forecast demand to control inventory levels, hire employees and provide training

Common Approaches to Forecasting



- Used when historical data are unavailable
- Considered highly subjective and judgmental

- Use past data to predict future values

Time-Series Data

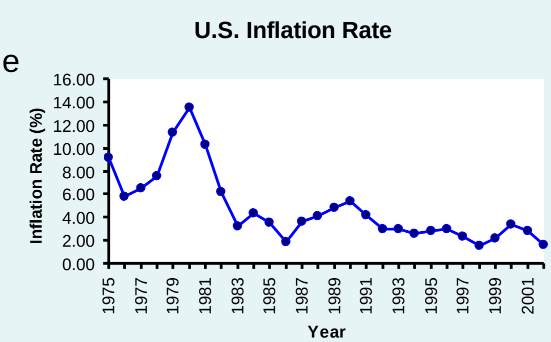
- Numerical data obtained at regular time intervals
- The time intervals can be annually, quarterly, monthly, weekly, daily, hourly, etc.
- Example:

Year:	2000	2001	2002	2003	2004
Sales:	75.3	74.2	78.5	79.7	80.2

Time-Series Plot

A time-series plot is a two-dimensional plot of time series data

- the vertical axis measures the variable of interest
- the horizontal axis corresponds to the time periods



Time-Series Components

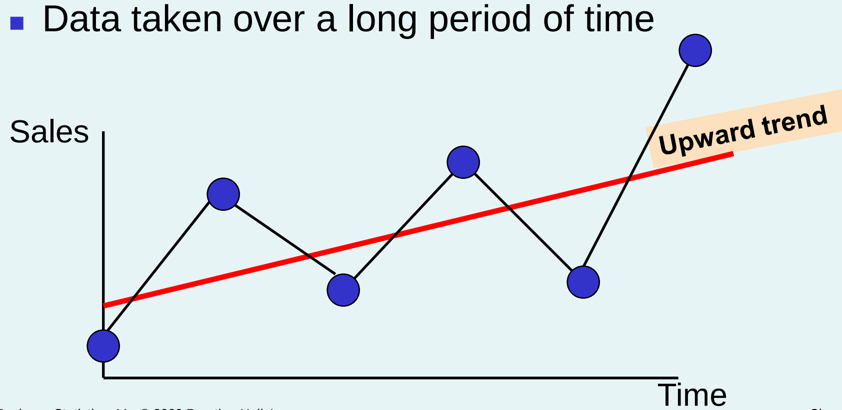
Time Series

Trend Component	Seasonal Component	Cyclical Component	Irregular Component
Overall, persistent, long-term movement	Regular periodic fluctuations, usually within a 12-month period	Repeating swings or movements over more than one year	Erratic or residual fluctuations

 Chap 16-7

Trend Component

- Long-run increase or decrease over time (overall upward or downward movement)
- Data taken over a long period of time

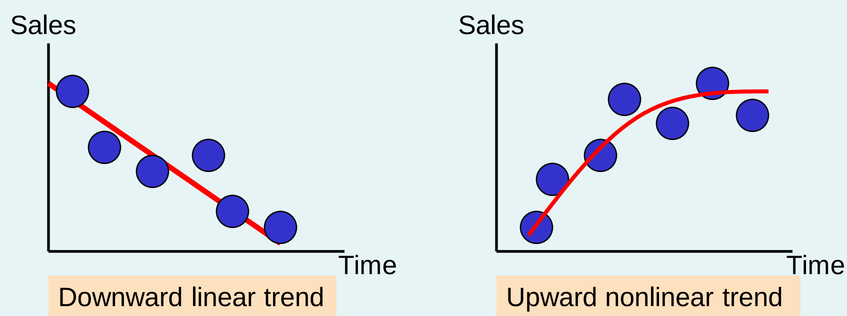


 Chap 16-8

Trend Component

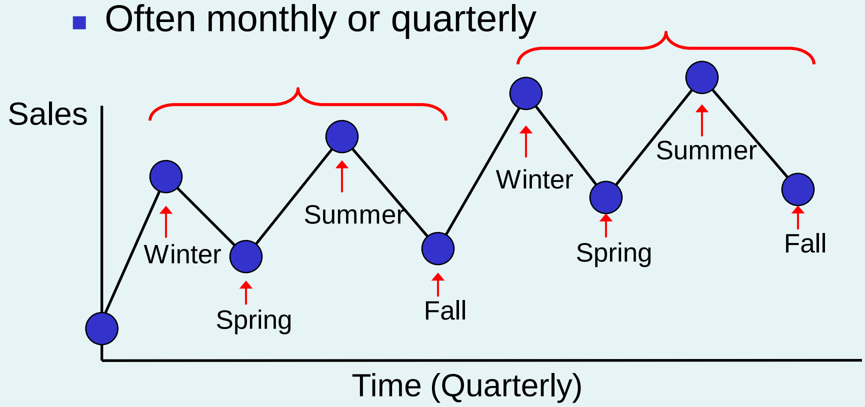
(continued)

- Trend can be upward or downward
- Trend can be linear or non-linear



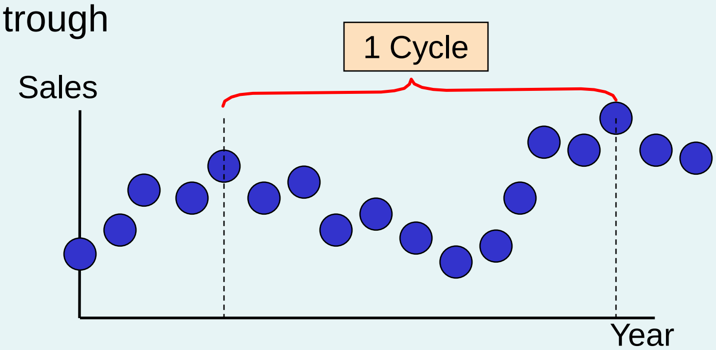
Seasonal Component

- Short-term regular wave-like patterns
- Observed within 1 year
- Often monthly or quarterly



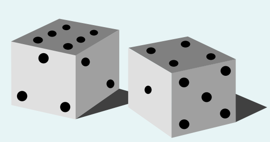
Cyclical Component

- Long-term wave-like patterns
- Regularly occur but may vary in length
- Often measured peak to peak or trough to trough



Irregular Component

- Unpredictable, random, “residual” fluctuations
- Due to random variations of
  - Nature
  - Accidents or unusual events
- “Noise” in the time series



Does Your Time Series Have A Trend Component?

- A time series plot should help you to answer this question.
- Often it helps if you "smooth" the time series data to help answer this question.
- Two popular smoothing methods are moving averages and exponential smoothing.

Smoothing Methods

- Moving Averages
  - Calculate moving averages to get an overall impression of the pattern of movement over time
  - Averages of consecutive time series values for a chosen period of length L
- Exponential Smoothing
  - A weighted moving average

Moving Averages

- Used for smoothing
- A series of arithmetic means over time
- Result dependent upon choice of L (length of period for computing means)
- Last moving average of length L can be extrapolated one period into future for a short term forecast
- Examples:
  - For a 5 year moving average, L = 5
  - For a 7 year moving average, L = 7
  - Etc.

Moving Averages

(continued)

- Example: Five-year moving average
  - First average:

$$MA(5) = \frac{Y_1 + Y_2 + Y_3 + Y_4 + Y_5}{5}$$

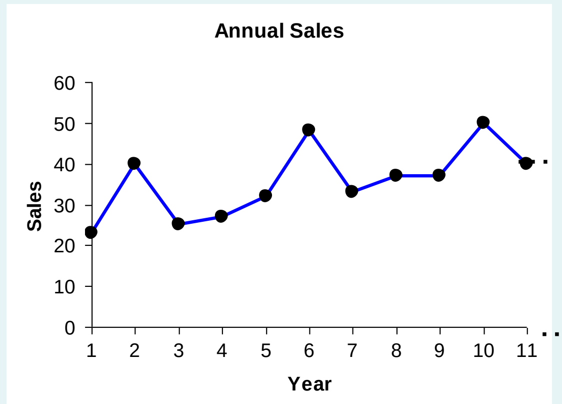
  - Second average:

$$MA(5) = \frac{Y_2 + Y_3 + Y_4 + Y_5 + Y_6}{5}$$

  - etc.

Example: Annual Data

Year	Sales
1	23
2	40
3	25
4	27
5	32
6	48
7	33
8	37
9	37
10	50
11	40
etc…	etc…



Calculating Moving Averages

Year	Sales
1	23
2	40
3	25
4	27
5	32
6	48
7	33
8	37
9	37
10	50
11	40

etc…

Average Year	5-Year Moving Average
3	29.4
4	34.4
5	33.0
6	35.4
7	37.4
8	41.0
9	39.4
…	…

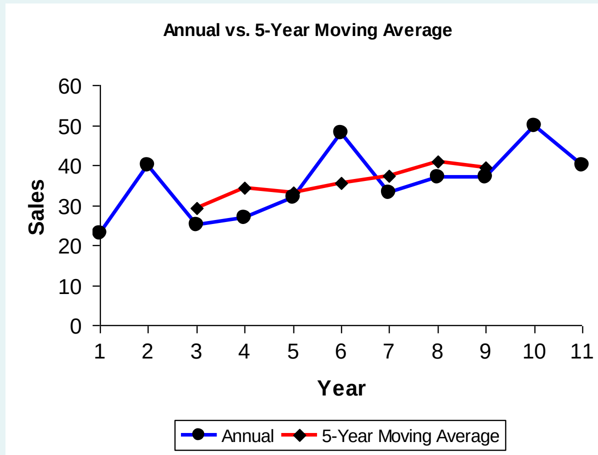
$$3 = \frac{1+2+3+4+5}{5}$$

$$29.4 = \frac{23+40+25+27+32}{5}$$

- Each moving average is for a consecutive block of 5 years

Annual vs. Moving Average

The 5-year moving average smoothes the data and makes it easier to see the underlying trend



Exponential Smoothing

- Used for smoothing and short term forecasting (one period into the future)
- A weighted moving average
  - Weights decline exponentially
  - Most recent observation weighted most

Exponential Smoothing

(continued)

- The weight (smoothing coefficient) is W
  - Subjectively chosen
  - Ranges from 0 to 1
  - Smaller W gives more smoothing, larger W gives less smoothing
- The weight is:
  - Close to 0 for smoothing out unwanted cyclical and irregular components
  - Close to 1 for forecasting

Exponential Smoothing Model

- Exponential smoothing model

$$E_1 = Y_1$$

$$E_i = WY_i + (1 - W)E_{i-1}$$

For i = 2, 3, 4, …

where:

E_i = exponentially smoothed value for period i

E_{i-1} = exponentially smoothed value already computed for period i - 1

Y_i = observed value in period i

W = weight (smoothing coefficient), 0 < W < 1

Exponential Smoothing Example

- Suppose we use weight W = 0.2

Time Period (i)	Sales (Y_i)	Forecast from prior period (E_{i-1})	Exponentially Smoothed Value for this period (E_i)
1	23	--	23
2	40	23	(.2)(40)+(.8)(23)=26.4
3	25	26.4	(.2)(25)+(.8)(26.4)=26.12
4	27	26.12	(.2)(27)+(.8)(26.12)=26.296
5	32	26.296	(.2)(32)+(.8)(26.296)=27.437
6	48	27.437	(.2)(48)+(.8)(27.437)=31.549
7	33	31.549	(.2)(48)+(.8)(31.549)=31.840
8	37	31.840	(.2)(33)+(.8)(31.840)=32.872
9	37	32.872	(.2)(37)+(.8)(32.872)=33.697
10	50	33.697	(.2)(50)+(.8)(33.697)=36.958
etc.	etc.	etc.	etc.

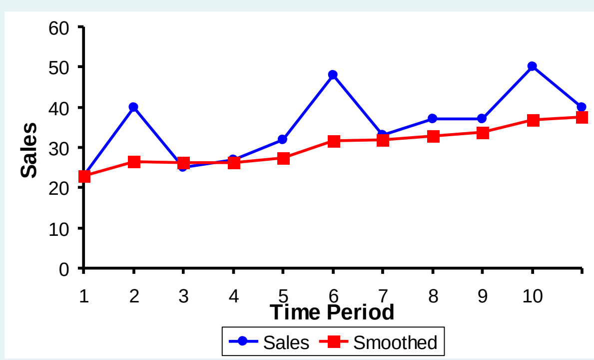
$E_1 = Y_1$ since no prior information exists

$E_i = WY_i + (1 - W)E_{i-1}$

Basic Business Statistics, 11e © 2009 Prentice-Hall, Inc.. Chap 16-23

Sales vs. Smoothed Sales

- Fluctuations have been smoothed
- NOTE: the smoothed value in this case is a little low, generally a little low, since the trend is upward sloping and the weighting factor is only .2



Basic Business Statistics, 11e © 2009 Prentice-Hall, Inc.. Chap 16-24

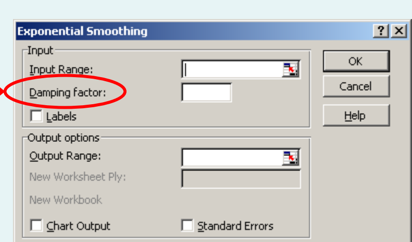
Forecasting Time Period i + 1

- The smoothed value in the current period (i) is used as the forecast value for next period (i + 1) :

$$\hat{Y}_{i+1} = E_i$$

Exponential Smoothing in Excel

- Use data analysis / exponential smoothing
  - The “damping factor” is (1 - W)



There Are Three Popular Methods For Trend-Based Forecasting

- Linear Trend Forecasting
- Nonlinear Trend Forecasting
- Exponential Trend Forecasting

Linear Trend Forecasting

Estimate a trend line using regression analysis

Year	Time Period (X)	Sales (Y)
1999	0	20
2000	1	40
2001	2	30
2002	3	50
2003	4	70
2004	5	65

- Use time (X) as the independent variable:

$$\hat{Y} = b_0 + b_1 X$$

In least squares linear, non-linear, and exponential modeling, time periods are numbered starting with 0 and increasing by 1 for each time period.

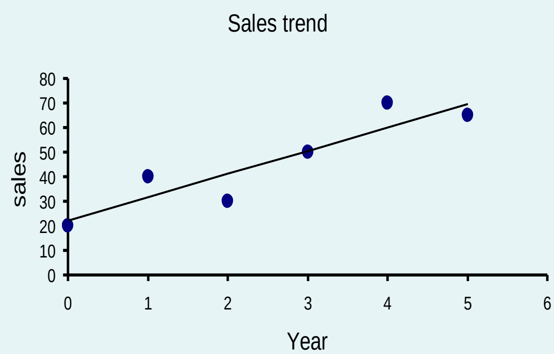
Linear Trend Forecasting

(continued)

The linear trend forecasting equation is:

Year	Time Period (X)	Sales (Y)
1999	0	20
2000	1	40
2001	2	30
2002	3	50
2003	4	70
2004	5	65

$$\hat{Y}_i = 21.905 + 9.5714\, X_i$$



 Chap 16-29

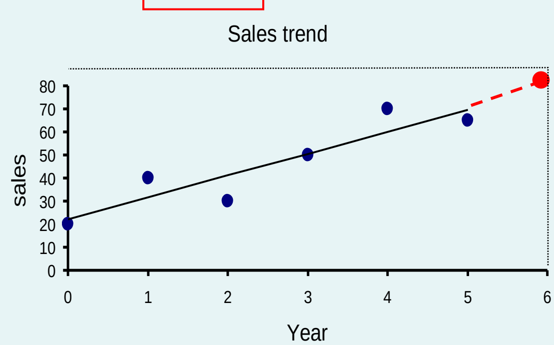
Linear Trend Forecasting

(continued)

- Forecast for time period 6:

Year	Time Period (X)	Sales (y)
1999	0	20
2000	1	40
2001	2	30
2002	3	50
2003	4	70
2004	5	65
2005	6	??

$$\hat{Y} = 21.905 + 9.5714(6)$$
$$= 79.33$$



 Chap 16-30

Nonlinear Trend Forecasting

- A nonlinear regression model can be used when the time series exhibits a nonlinear trend
- Quadratic form is one type of a nonlinear model:

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 X_i^2 + \varepsilon_i$$

- Compare adj. r^2 and standard error to that of linear model to see if this is an improvement
- Can try other functional forms to get best fit

Exponential Trend Model

- Another nonlinear trend model:

$$Y_i = \beta_0 \beta_1^{X_i} \varepsilon_i$$

- Transform to linear form:

$$\log(Y_i) = \log(\beta_0) + X_i \log(\beta_1) + \log(\varepsilon_i)$$

Exponential Trend Model

(continued)

- Exponential trend forecasting equation:

$$\log(\hat{Y}_i) = b_0 + b_1 X_i$$

where b_0 = estimate of $\log(\beta_0)$
b_1 = estimate of $\log(\beta_1)$

Interpretation:

$(\hat{\beta}_1 - 1) \times 100\%$ is the estimated annual compound growth rate (in %)

Trend Model Selection Using Differences

- Use a linear trend model if the first differences are approximately constant

$$(Y_2 - Y_1) = (Y_3 - Y_2) = \cdots = (Y_n - Y_{n-1})$$

- Use a quadratic trend model if the second differences are approximately constant

$$[(Y_3 - Y_2) - (Y_2 - Y_1)] = [(Y_4 - Y_3) - (Y_3 - Y_2)] = \cdots = [(Y_n - Y_{n-1}) - (Y_{n-1} - Y_{n-2})]$$

Trend Model Selection Using Differences

(continued)

- Use an exponential trend model if the percentage differences are approximately constant

$$\frac{(Y_2 - Y_1)}{Y_1} \times 100\% = \frac{(Y_3 - Y_2)}{Y_2} \times 100\% = \cdots = \frac{(Y_n - Y_{n-1})}{Y_{n-1}} \times 100\%$$

Autoregressive Modeling

- Used for forecasting
- Takes advantage of autocorrelation
  - 1st order - correlation between consecutive values
  - 2nd order - correlation between values 2 periods apart
- p^{th} order Autoregressive model:

$$Y_i = A_0 + A_1 Y_{i-1} + A_2 Y_{i-2} + \cdots + A_p Y_{i-p} + \delta_i$$

Random Error

Autoregressive Model: Example

The Office Concept Corp. has acquired a number of office units (in thousands of square feet) over the last eight years. Develop the second order Autoregressive model.

Year	Units
97	4
98	3
99	2
00	3
01	2
02	2
03	4
04	6

Autoregressive Model: Example Solution

- **Develop the 2nd order table**
- **Use Excel or Minitab to estimate a regression model**

Year	Y_i	Y_{i-1}	Y_{i-2}
97	4	--	--
98	3	4	--
99	2	3	4
00	3	2	3
01	2	3	2
02	2	2	3
03	4	2	2
04	6	4	2

Excel Output

	Coefficients
Intercept	3.5
X Variable 1	0.8125
X Variable 2	-0.9375

$$\hat{Y}_i = 3.5 + 0.8125 Y_{i-1} - 0.9375 Y_{i-2}$$

Autoregressive Model Example: Forecasting

Use the second-order equation to forecast number of units for 2005:

$$\hat{Y}_i = 3.5 + 0.8125 Y_{i-1} - 0.9375 Y_{i-2}$$

$$\hat{Y}_{2005} = 3.5 + 0.8125(Y_{2004}) - 0.9375(Y_{2003})$$

$$= 3.5 + 0.8125(6) - 0.9375(4)$$

$$= 4.625$$

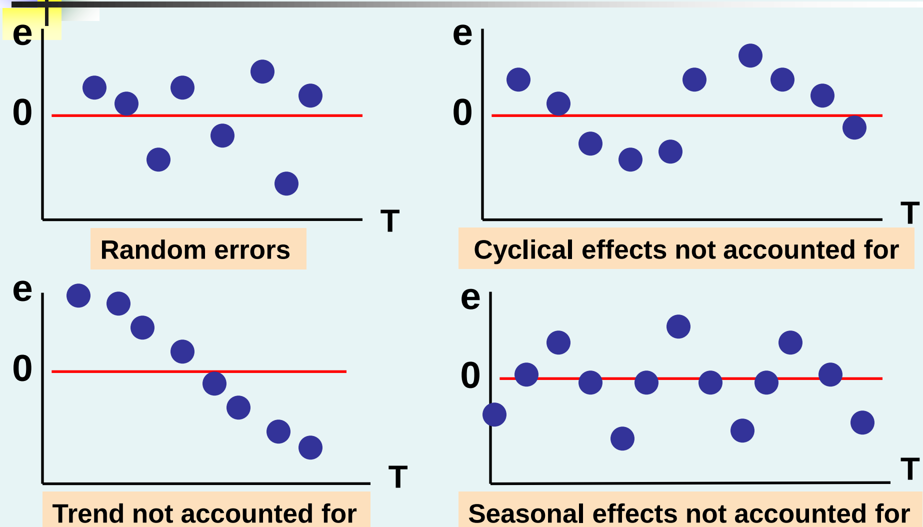
Autoregressive Modeling Steps

1. Choose p (note that df = n – 2p – 1)
2. Form a series of "lagged predictor" variables Y_{i-1} , Y_{i-2} , … , Y_{i-p}
3. Use Excel or Minitab to run regression model using all p variables
4. Test significance of A_p
   - If null hypothesis rejected, this model is selected
   - If null hypothesis not rejected, decrease p by 1 and repeat

Choosing A Forecasting Model

- Perform a residual analysis
  - Look for pattern or trend
- Measure magnitude of residual error using squared differences
- Measure magnitude of residual error using absolute differences
- Use simplest model
  - Principle of parsimony

Residual Analysis



Measuring Errors

- Choose the model that gives the smallest measuring errors

- Sum of squared errors (SSE)

$$SSE = \sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$$

  - Sensitive to outliers

- Mean Absolute Deviation (MAD)

$$MAD = \frac{\sum_{i=1}^{n}\left|Y_i - \hat{Y}_i\right|}{n}$$

  - Less sensitive to extreme observations

Principal of Parsimony

- Suppose two or more models provide a good fit for the data
- Select the simplest model
  - Simplest model types:
    - Least-squares linear
    - Least-squares quadratic
    - 1st order autoregressive
  - More complex types:
    - 2nd and 3rd order autoregressive
    - Least-squares exponential

Forecasting With Seasonal Data

- Time series are often collected monthly or quarterly
- These time series often contain a trend component, a seasonal component, and the irregular component
- Suppose the seasonality is quarterly
  - Define three new dummy variables for quarters:

Q_1 = 1 if first quarter, 0 otherwise

Q_2 = 1 if second quarter, 0 otherwise

Q_3 = 1 if third quarter, 0 otherwise

(Quarter 4 is the default if $Q_1 = Q_2 = Q_3 = 0$)

Exponential Model with Quarterly Data

$$Y_i = \beta_0 \beta_1^{X_i} \beta_2^{Q_1} \beta_3^{Q_2} \beta_4^{Q_3} \varepsilon_i$$

$(\beta_1 - 1)$x100% is the quarterly compound growth rate

β_i provides the multiplier for the i^{th} quarter relative to the 4th quarter (i = 2, 3, 4)

- Transform to linear form:

$$\log(Y_i) = \log(\beta_0) + X_i \log(\beta_1) + Q_1 \log(\beta_2) + Q_2 \log(\beta_3) + Q_3 \log(\beta_4) + \log(\varepsilon_i)$$

Estimating the Quarterly Model

- Exponential forecasting equation:

$$\log(\hat{Y}_i) = b_0 + b_1 X_i + b_2 Q_1 + b_3 Q_2 + b_4 Q_3$$

where b_0 = estimate of $\log(\beta_0)$, so $10^{b_0} = \hat{\beta}_0$
b_1 = estimate of $\log(\beta_1)$, so $10^{b_1} = \hat{\beta}_1$
etc…

Interpretation:

$(\hat{\beta}_1 - 1) \times 100\%$ = estimated quarterly compound growth rate (in %)

$\hat{\beta}_2$ = estimated multiplier for first quarter relative to fourth quarter

$\hat{\beta}_3$ = estimated multiplier for second quarter rel. to fourth quarter

$\hat{\beta}_4$ = estimated multiplier for third quarter relative to fourth quarter

Quarterly Model Example

- Suppose the forecasting equation is:

$$\log(\hat{Y}_i) = 3.43 + .017X_i - .082Q_1 - .073Q_2 + .022Q_3$$

b_0 = 3.43, so $10^{b_0} = \hat{\beta}_0 = 2691.53$

b_1 = .017, so $10^{b_1} = \hat{\beta}_1 = 1.040$

b_2 = -.082, so $10^{b_2} = \hat{\beta}_2 = 0.827$

b_3 = -.073, so $10^{b_3} = \hat{\beta}_3 = 0.845$

b_4 = .022, so $10^{b_4} = \hat{\beta}_4 = 1.052$

Quarterly Model Example

(continued)

Value:	Interpretation:
$\hat{\beta}_0 = 2691.53$	Unadjusted trend value for first quarter of first year
$\hat{\beta}_1 = 1.040$	4.0% = estimated quarterly compound growth rate
$\hat{\beta}_2 = 0.827$	Average sales in Q_1 are 82.7% of average 4^{th} quarter sales, after adjusting for the 4% quarterly growth rate
$\hat{\beta}_3 = 0.845$	Average sales in Q_2 are 84.5% of average 4^{th} quarter sales, after adjusting for the 4% quarterly growth rate
$\hat{\beta}_4 = 1.052$	Average sales in Q_3 are 105.2% of average 4^{th} quarter sales, after adjusting for the 4% quarterly growth rate

Index Numbers

- Index numbers allow relative comparisons over time
- Index numbers are reported relative to a Base Period Index
- Base period index = 100 by definition
- Used for an individual item or group of items

Simple Price Index

- Simple Price Index:

$$I_i = \frac{P_i}{P_{base}} \times 100$$

where

I_i = index number for year i

P_i = price for year i

P_{base} = price for the base year

Index Numbers: Example

Airplane ticket prices from 1995 to 2003:

Year	Price	Index (base year = 2000)
1995	272	85.0
1996	288	90.0
1997	295	92.2
1998	311	97.2
1999	322	100.6
2000	320	100.0
2001	348	108.8
2002	366	114.4
2003	384	120.0

$$I_{1996} = \frac{P_{1996}}{P_{2000}} \times 100 = \frac{288}{320}(100) = 90$$

Base Year:

$$I_{2000} = \frac{P_{2000}}{P_{2000}} \times 100 = \frac{320}{320}(100) = 100$$

$$I_{2003} = \frac{P_{2003}}{P_{2000}} \times 100 = \frac{384}{320}(100) = 120$$

Index Numbers: Interpretation

$$I_{1996} = \frac{P_{1996}}{P_{2000}} \times 100 = \frac{288}{320}(100) = 90$$

- Prices in 1996 were 90% of base year prices

$$I_{2000} = \frac{P_{2000}}{P_{2000}} \times 100 = \frac{320}{320}(100) = 100$$

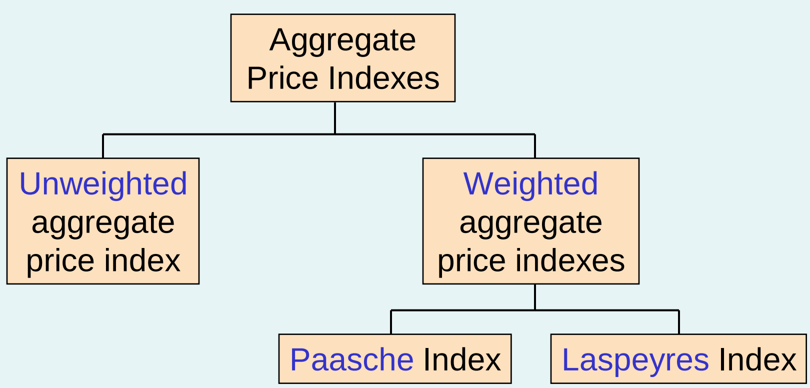
- Prices in 2000 were 100% of base year prices (by definition, since 2000 is the base year)

$$I_{2003} = \frac{P_{2003}}{P_{2000}} \times 100 = \frac{384}{320}(100) = 120$$

- Prices in 2003 were 120% of base year prices

Aggregate Price Indexes

- An aggregate index is used to measure the rate of change from a base period for a group of items



Unweighted Aggregate Price Index

- Unweighted aggregate price index formula:

$$I_U^{(t)} = \frac{\sum_{i=1}^{n} P_i^{(t)}}{\sum_{i=1}^{n} P_i^{(0)}} \times 100$$

i = item
t = time period
n = total number of items

$I_U^{(t)}$ = unweighted price index at time t

$\sum_{i=1}^{n} P_i^{(t)}$ = sum of the prices for the group of items at time t

$\sum_{i=1}^{n} P_i^{(0)}$ = sum of the prices for the group of items in time period 0

 Chap 16-55

Unweighted Aggregate Price Index: Example

	Automobile Expenses: Monthly Amounts ($):				
Year	**Lease payment**	**Fuel**	**Repair**	**Total**	**Index (2001=100)**
2001	260	45	40	345	100.0
2002	280	60	40	380	110.1
2003	305	55	45	405	117.4
2004	310	50	50	410	118.8

$$I_{2004} = \frac{\sum P_{2004}}{\sum P_{2001}} \times 100 = \frac{410}{345}(100) = 118.8$$

- Unweighted total expenses were 18.8% higher in 2004 than in 2001

 Chap 16-56

Weighted Aggregate Price Indexes

- Laspeyres index

$$I_L^{(t)} = \frac{\sum_{i=1}^{n} P_i^{(t)} Q_i^{(0)}}{\sum_{i=1}^{n} P_i^{(0)} Q_i^{(0)}} \times 100$$

$Q_i^{(0)}$: weights based on period 0 quantities

- Paasche index

$$I_P^{(t)} = \frac{\sum_{i=1}^{n} P_i^{(t)} Q_i^{(t)}}{\sum_{i=1}^{n} P_i^{(0)} Q_i^{(t)}} \times 100$$

$Q_i^{(t)}$: weights based on current period quantities

$P_i^{(t)}$ = price in time period t

$P_i^{(0)}$ = price in period 0

Common Price Indexes

- Consumer Price Index (CPI)
- Producer Price Index (PPI)
- Stock Market Indexes
  - Dow Jones Industrial Average
  - S&P 500 Index
  - NASDAQ Index

Pitfalls in Time-Series Analysis

- Assuming the mechanism that governs the time series behavior in the past will still hold in the future
- Using mechanical extrapolation of the trend to forecast the future without considering personal judgments, business experiences, changing technologies, and habits, etc.

Chapter Summary

- Discussed the importance of forecasting
- Addressed component factors of the time-series model
- Performed smoothing of data series
  - Moving averages
  - Exponential smoothing
- Described least square trend fitting and forecasting
  - Linear, quadratic and exponential models

Chapter Summary

(continued)

- Addressed autoregressive models
- Described procedure for choosing appropriate models
- Addressed time series forecasting of monthly or quarterly data (use of dummy variables)
- Discussed pitfalls concerning time-series analysis
- Discussed index numbers and index number development