EC 413
Economic Forecast and Analysis
(Professor Lee)

Lecture 1

Introduction to Forecasting

Read:

(WK Ch 1)

Major Questions in this lecture

1. Why and what to forecast?
2. Based on What?
3. Is your forecast good or bad?

1. Why and what to forecast?

Examples of Forecasting

(Ex 1) Sales of the Gap Store for each month for the next two years.

(Ex 2) Sales of SUVs to be sold for the next 5 years in the US.

(Ex 3) Revenue forecasts for TWA
(industry passenger traffic
$\rightarrow$ industry share of TWA)

(Ex 4) # of expected passengers for tomorrow for each flight of the Southwest Airlines

(Ex 5) Who will pay only the minimum payment for credit cards?

(Ex 6) Gas / electricity peak-load demand for each month in a certain region

(Ex 7) How many donuts to prepare each morning?

(Ex 8) Demand for police patrol services

(Ex 9) How many spare parts to carry for a submarine?

(Ex 10) Economic and financial forecasts for the state in the coming year??!
(national economic variables → regional forecast)

(Ex 11) Tomorrow's return (or price) of the Disney stock

2. Based on What?

- Past values
- Past errors
- Time series components
- Relationships with other variables
-

Statistical Model

$E(y_{t+h} \mid \Omega_t)$ = conditional expectation

where

y_{t+h} = future values at time t+h

Ω_t= all information available at time t

Write:

$y_{t+h} = E(y_{t+h} \mid \Omega_t) + u_{t+h}$

where

u_{t+h} = forecase error

Examples of Forecasting Models

(1) Naïve Model

$$y_{t+1} = y_t + u_{t+1}$$

or equivalently,

$$y_t = y_{t-1} + u_t \quad \text{(random walk)}$$

Implying

(i) $E(y_t \mid \Omega_{t-1}) = \tilde{y}_t = y_{t-1}$

(ii) $E(u_t \mid \Omega_{t-1}) = 0$

(martingale hypothesis)

(2) Modified Naïve Model

$$y_{t+1} = y_t + \alpha (y_t - y_{t-1}) + u_{t+1}$$

That is,

$$\tilde{y}_{t+1} = y_t + \alpha (y_t - y_{t-1})$$

where $(y_t - y_{t-1}) = \Delta y_t$ refers to as the first difference and α is the proportion of its contribution

(3) Moving Average Models

3 days, 5 days, 30 days, 90 days..

(eg) $\tilde{y}_{t+1} = \frac{1}{3}(y_t + y_{t-1} + y_{t-2})$

.. Lecture 2

(4) Purely based on Trends and Seasonal components

.. Lecture 2

(5) Autoregressive (AR) Model

$y_t = b_0 + b_1 y_{t-1} + b_2 y_{t-2} + .. + b_p y_{t-p} + u_t$

# of AR terms = p

.. Lecture 4

(6) Moving Average (MA) Error model

$y_t = b_0 + b_1 u_{t-1} + b_2 u_{t-2} + .. + b_q u_{t-q} + u_t$

# of MA terms = q

.. Lecture 4

(7) Time Series Components

$Y = T \times S \times C \times I$

T = Trend
S = Seasonality
C = Cycle
I = Irregular term

.. Lecture 3

(8) Regression

Demand for police patrol services =
$a_0 + a_1$(population) + a_2(# of arrests) +
a_3(income level) + a_4(population density) + u

.. Lecture 3

(9) Multivariate regression

$y_t = b_0 + b_1 y_{t-1} + b_2 y_{t-2} + b_3 X_{t-1} + b_4 X_{t-2} + u_{1t}$
$X_t = c_0 + c_1 y_{t-1} + c_2 y_{t-2} + c_3 X_{t-1} + c_4 X_{t-2} + u_{2t}$
y_t = Fed Ex sales
X_t = UPS sales
.. Lecture 6

(10) Subjective forecasting

.. Surveys, Jury of executive opinion, Delphi method,..)

.. Read WK, Ch 1.

(11) Composite Index

Leading
Coincident
Lagging Indicatiors

http://www.globalindicators.org (US)
http://biz.yahoo.com/rf/030723/full_text_canada_s_june_leading_indicator_1.html (Canada)

U.S. COMPOSITE INDEXES FOR JUNE 2003

The Conference Board announced today that the U.S. leading index increased 0.1 percent, the coincident index increased 0.1 percent, and the lagging index decreased 0.5 percent in June.

- The leading index increased in June for the third consecutive month. This suggests the flat trend in the leading index over the past year may have ended, but additional months of growth are needed to determine if an upward trend has indeed developed.
- After declining for eleven months, the coincident index began increasing in December 2001, consistent with the officially declared November 2001 trough of the last recession. Following a moderate increase through most of 2002, this measure of current economic activity has since been essentially flat, as the leading index signaled.
- The recent improvement in the growth rate of the leading index is consistent with near-term improvement in the growth rate of the coincident index and real GDP. However, three months of increases in the leading index is not enough to signal the beginning of a sustained period of above-trend economic growth.

U.S. Composite Indexes: Components and Standardization Factors

	Leading Index	Factor
1.	Average weekly hours, manufacturing	.1946
2.	Average weekly initial claims for unemployment insurance	.0268
3.	Manufacturers' new orders, consumer goods and materials	.0504
4.	Vendor performance, slower deliveries diffusion index	.0296
5.	Manufacturers' new orders, nondefense capital goods	.0139
6.	Building permits, new private housing units	.0205
7.	Stock prices, 500 common stocks	.0309
8.	Money supply, M2	.2775
9.	**Interest rate spread, 10-year Treasury bonds less federal funds**	**.3364**
10.	Index of consumer expectations	.0193
	Coincident Index	
1.	Employees on nonagricultural payrolls	.5186
2.	Personal income less transfer payments	.2173
3.	Industrial production	.1470
4.	Manufacturing and trade sales	.1170
	Lagging Index	
1.	Average duration of unemployment	.0368
2.	Inventories to sales ratio, manufacturing and trade	.1206
3.	Labor cost per unit of output, manufacturing	.0693
4.	Average prime rate	.2692
5.	Commercial and industrial loans	.1204
6.	Consumer installment credit to personal income ratio	.1951
7.	Consumer price index for services	.1886

(Example)

8. Interest rate spread, 10-year Treasury bonds less federal funds **.3364**

It is called *yield curve*. If it becomes negative (short-term rates are higher than long-run rates), it is an indicative of recession.

(12) Volatility Models
(Combined with risk measures)
.. Lecture 5

(13) Long-run Relationship Models
.. Lecture 7

(14) Dynamic econometric simulation Models

http://fairmodel.econ.yale.edu/ **FAIR Model**
http://www.globalinsight.com/ **Global Insight (DRI-WEFA)**
.. beyond our scope

3. Is your forecast good or bad?

Evaluating Forecasting Models

(a) Mean Square Errors (MSE)

$$\mathbf{MSE} = \frac{\mathbf{1}}{\mathbf{n}} \sum (\mathbf{y_t} - \tilde{\mathbf{y}}_\mathbf{t})^2$$

$\mathbf{y_t}$ = Actual values
$\tilde{\mathbf{y}}_\mathbf{t}$ = Forecasted values

(b) Root MSE (RMSE) = $\sqrt{\mathbf{MSE}}$

(eg) Retail Store Sales (Naïve and Modified Models)

Root-mean-squared error for years 2 through 7 is found using the following data.

Year	Retail Store Sales	Naïve Forecast	Squared Error
1	1,225		na
2	1,285	1,225	3,600
3	1,359	1,285	5,476
4	1,392	1,359	1,089
5	1,443	1,392	2,601
6	1,474	1,443	961
7	1,467	1,474	49
8		1,467	
		RMSE =	47.917

Using the second-naive forecasting model, forecasts were generated with the adjustment parameter set equal to .20: $F_t = A_{t-1} + .20[A_{t-1} - A_{t-2}]$.

Year	Retail Store Sales	First Difference in Annual Sales	0.2 x First Difference in Annual Sales	Modified Naïve Forecast	Squared Error
1	1,225	na	na		na
2	1,285	60	12.0		na
3	1,359	74	14.8	1,297	3,844
4	1,392	33	6.6	1,374	331
5	1,443	51	10.2	1,399	1,971
6	1,474	31	6.2	1,453	433
7	1,467	-7	-1.4	1,480	174
8		na	na	1,466	na
				RMSE =	36.752

Note: Models with <u>smaller</u> values are better.

Note: <u>In-sample</u> (historic) MSE vs.
<u>Out-of-sample</u> (hold-out period) MSE

(c) Mean <u>Absolute</u> Errors (MAE)

$$\mathbf{MAE} = \frac{1}{n} \sum |y_t - \tilde{y}_t|$$

.. Less frequently used, but robust in the presence of <u>outliers</u> (extreme values)

(d) Mean <u>Absolute</u> <u>Percentage</u> Errors (MAPE)

$$\mathbf{MAE} = \frac{1}{n} \sum \left| (y_t - \tilde{y}_t) / y_t \right|$$

(e) Mean Errors (ME)

$$\mathbf{ME} = \frac{1}{n} \sum (y_t - \tilde{y}_t)$$

.. Not a good measure due to cancellation, but it provides the information on the direction of bias of forecasts; if positive, forecasts are lower than actual values, and vice versa.

(f) Theil's U statistic

$$U = \sqrt{\frac{\sum (y_t - \tilde{y}_t)^2}{\sum (y_t - y_{t-1})^2}}$$

.. compared with the naïve model (or any basic model)

If U < 1, the model forecasts <u>better</u> than the naïve model.

If U > 1, the model forecasts <u>worse</u> than the naïve model.

If U = 0, the model forecasts perfectly.

Evaluating Econometric Models Using Information Criteria

Akaike IC

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

where

$$\hat{\sigma}^2 \text{ (error variance)} = \frac{1}{T-k} \sum_{t=1}^{T} \hat{u}_t^2$$

T = # of observations
k = # of parameters (regressors)
$\hat{u}_t$ **= residuals**

Schwarz's IC (SIC or BIC)

$$\mathbf{SIC = \log(\hat{\sigma}^2) + (k/T)\log(T)}$$

Note: Models with <u>smaller</u> values are better.