Uniwersytet Ekonomiczny

George Matysiak

Modelling and Forecasting
11th January, 2016

Agenda

- Modelling and forecasting
  - The role of models and forecasts
  - Types of forecasting models
    - Behavioural
    - Time series
  - Advantages/disadvantages of different methods
  - Econometric models

The role of models & forecasts

- Understanding of market dynamics/relationships
- Identify growth markets
- Make Buy/Hold/Sell decisions
- Assess supply impacts
- Aid to investment, tactical and strategic decisions
- Portfolio diversification
- 'What if' scenarios

Models and forecasts 1

- Forecasts are necessary because of implementation lead times e.g. it takes time to execute buying/selling decisions or to develop!
- Forecasts are used in order to reduce uncertainty and risk in decision making
- Increasingly being demanded, and used, in commercial real estate applications

Models and forecasts 2

- An abstraction or representation of a real system or problem in hand
- Models can be used to *explain* or *predict*
  e.g. yields, capital values, rental growth, construction costs
- Models can be *deterministic* or *probabilistic*
  - ➢ values certain
  - ➢ values uncertain
    - ⇨ probability distributions

Models and forecasts 3

- Models and forecasts permit an *explicit* analysis
  - ➢ making the underlying model/relationship explicit
  - ➢ Making the underlying assumptions clear
  - ➢ enabling an analysis of potential impact of different variables on property market variables (scenario analysis)
  - ➢ enabling a sensitivity analysis – impact of changes in input values
  - ➢ quantifying the uncertainty surrounding the forecast(s)
  - ➢ providing a basis for *informed decision making*

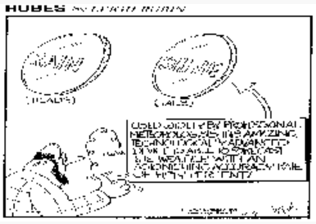
Deterministic and statistical models

- Real world relationships:
  - ➢deterministic or 'exact' relationships
  - ➢statistical relationships have *uncertainty* associated with the values of an outcome ('dependent' variable)
- Many phenomena, particularly 'investments', involve uncertainty
- Statistical relationships, by their nature, are formulated in terms of uncertainty i.e. involve *random* components i.e. outcomes
- Regression analysis commonly used in estimating statistical models for the purposes of forecasting

Probability Distributions

- The set of all possible values that a *random* variable can take and its associated probabilities is called a *probability distribution*
  - ➢*examples*
- Used to determine the degree of confidence (uncertainty) one has in a forecast

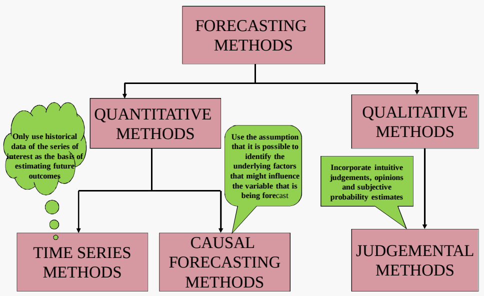
A simple forecasting model



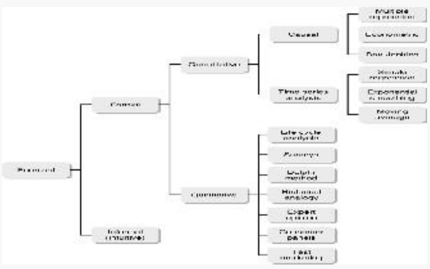
Question: Is there really a 50% chance of getting it right?

Approaches to modelling and forecasting

Forecasting approaches



Forecasting Methods



Approaches to forecasting

- Qualitative Forecasting - gives the *expected direction* of variables of interest
  - Expert Opinion
  - Surveys
  - Indicators
- Quantitative Forecasting – quantifies the *expected direction and magnitude*
  - Deterministic models
  - Stochastic models (random component)

Choosing a model

- Elaborate statistical models are not always required to develop accurate forecasts
- The principle of parsimony suggests that the simpler the model the better (less can go wrong!)
- The main advantage of simple models is that they serve as a *benchmark* with which to assess more sophisticated models
- There are many simple *time series* models

Time series models 1

- Time series models are good tools in forecasting short-term events
- The underlying principle in all time series models is that the historical pattern of the dependent variable can be used as the basis for developing forecasts
- In these models, historical data for the forecast variable are analyzed in an attempt to discern any underlying pattern(s)

Time series models 2

- Time series are best when applied to short-term forecasts
- Time series models prove most satisfactory when historical data contain either no systematic data pattern or when the changes are occurring very slowly or consistently
- Data requirements and ease of implementation are a function of the specific time series technique selected

Time series forecasting 1

- Weights historical observations to forecast future values of the series
- Extrapolates by projecting past patterns into the future
- Time series or autoregressive forecasting models will be most useful when economic conditions can be expected to remain relatively stable.
- Assumes past patterns will continue i.e. the underlying structure generating the values remains unchanged

Time series forecasting 2

- We consider two time series forecasting methods, namely:
  - Averaging methods
    - equally weighted observations
  - Exponential smoothing methods
    - unequal set of weights to past data, where the weights decay exponentially from the most recent to the most distant data points
    - all methods in this group require that certain parameters to be defined
    - these parameters (with values between 0 and 1) will determine the weights to be applied to past data

Examples: simple time series forecasts

- Naïve forecasts - the forecast for any period equals the previous period's actual value
- Moving Averages - technique that averages a number of recent/past actual values, updated as new values become available
- Exponential Smoothing - weighted averaging method based on previous forecast plus a percentage of the forecast error

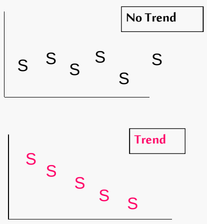
Naïve forecast

Naive Forecast

$F_{t+1} = A_t$

Method best when there is no trend, only random error

Graphs of actual sales, S, over time with and without trends



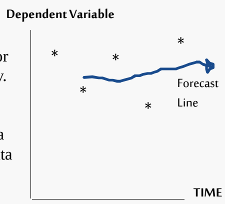
Moving average forecasting

A smoothing forecast method for data that jumps around

Best when there is no trend. For example a 3-Period Moving Av.

$F_{t+1} = [A_t + A_{t-1} + A_{t-2}]/3$

Averaging the observations is a way of "smoothing" out the data by eliminating much of the "noise" (random effects)



Motivation for exponential smoothing

- Simple moving average method assigns equal weights (1/*n*) to all *n* data points
- Arguably, recent observations provide more relevant information than do observations in the past
- So we want a weighting scheme that assigns decreasing weights to the more distant observations

Simple exponential smoothing

$$F_{t+1} = F_t + \alpha(A_t - F_t)$$

Next Forecast = Previous Forecast + (smoothing constant) x(Actual - Previous Forecast)

Premise: The most recent observations might have the highest predictive value.

- therefore, we should give more weight to the more recent time periods when forecasting
- smoothing constant is a percentage of the forecast error. Each new forecast is equal to the previous forecast plus a percentage of the previous error

Simple exponential smoothing

- Using exponential smoothing, the forecast for the next period is equal to the forecast for the current period plus a proportion (α) of the forecast error in the current period.
- Using exponential smoothing, the forecast is calculated by:

$$F_{t+1} = \alpha A_t + (1-\alpha)F_t$$ ←

This is the same as $F_{t+1} = F_t + \alpha(A_t - F_t)$

where:

α is the smoothing constant (a number between 0 and 1)

F_t is the forecast for period t

F_{t+1} is the forecast for period t+1

A_t is the actual data value for period t

Simple exponential smoothing

- The implication of exponential smoothing can be better seen if the previous equation is expanded by replacing F_t with its components as follows:

$$
\begin{aligned}
F_{t+1} &= \sim y_t + (1-\sim)F_t \\
&= \sim y_t + (1-\sim)[\sim y_{t-1} + (1-\sim)F_{t-1}] \\
&= \sim y_t + \sim(1-\sim)y_{t-1} + (1-\sim)^2 F_{t-1}
\end{aligned}
$$

Simple exponential smoothing

- If this substitution process is repeated by replacing F_{t-1} by its components, F_{t-2} by its components, and so on the result is:

$$F_{t+1} = \sim y_t + \sim(1-\sim)y_{t-1} + \sim(1-\sim)^2 y_{t-2} + \sim(1-\sim)^3 y_{t-3} + \cdots + \sim(1-\sim)^{t-1} y_1$$

- Therefore, F_{t+1} is the weighted moving average of all past observations.

Simple exponential smoothing

- The value of smoothing constant α must be between 0 and 1
- α can not be *equal* to 0 or 1
- If stable predictions with smoothed random variation is desired then a small value of α is used
- If a rapid response to a real change in the pattern of observations is desired, a large value of α is appropriate

Simple exponential smoothing

- Calculate the smallest RMSE (later) e.g. using the *Solver* facility in *Excel*
- The value of α with the smallest RMSE is chosen for producing the future forecasts

Simple exponential smoothing

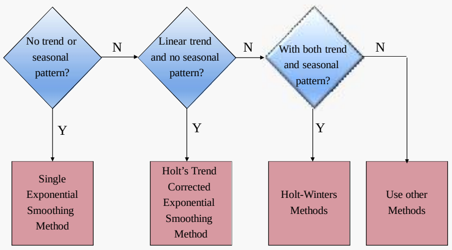
- To start the algorithm, we need F_1 because

$$F_2 = \sim y_1 + (1 - \sim)F_1$$

- Since F_1 is not known, we can
  - ➢set the first estimate equal to the first observation or
  - ➢use the average of a number of the first observations (the first half of the observations is commonly used) for the initial smoothed value
  - combine with the smallest RMSE alpha

Alternatives to simple exponential smoothing

- Regular exponential smoothing will always lag behind the trend, and so the trend needs to be taken into account

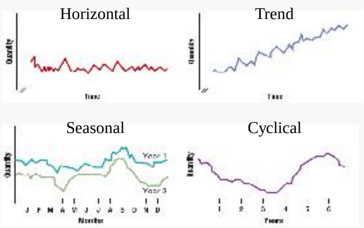


Time series components

Time Series is a time-ordered sequence of observations taken at regular interval over time (compare against *cross-sectional* data)

- *Trend* - long-term movement in data
- *Cycles* - wavelike variations lasting more than one year
- *Seasonality* - short-term regular variations in data
- *Irregular variations* - caused by unusual (one off) circumstances
- *Random variations* - caused by chance

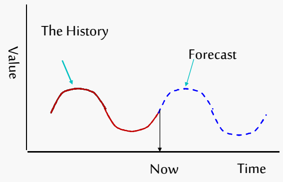
Time-Series patterns



Trend analysis and projection

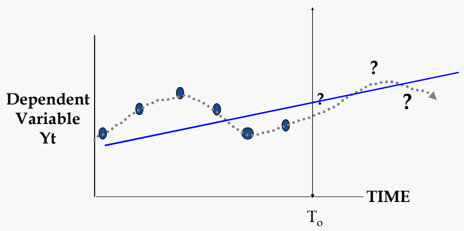
- Linear Trend Analysis: Assumes a constant unit change in an important variable over time
- Growth Trend Analysis: Growth trend analysis assumes constant percentage change in an important economic variable over time

Extrapolative methods



Recognising a pattern in the past data history - assuming the same pattern applies to the future

Extrapolative methods

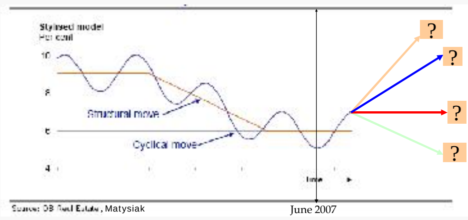


Y_t = Trend +Cyclical + Seasonal + Random Noise

Pure time series approach

- Advantages
  - Best in immediate or short-term
  - Simple techniques easy to understand and communicate
  - Need data on only one variable for univariate techniques
- Disadvantages
  - Vulnerable to changes in structure
  - “Black box”--no theory!
  - Poor in predicting cyclical turning points
  - Cannot be used for “what if” (different assumptions) simulations

Can capital values (yields) be explained by a pattern?



Econometric models

- Application of statistics to economic/property data in order to establish relationships/models
- Models should be based on *causal* relationships

The research objective

- Enhancing our understanding of the forces driving Property Markets
  - ➢ Underlying causal relationships
  - ➢ Trends
  - ➢ Cycles
  - ➢ Secular Change
- Leveraging market knowledge and expertise with data and rigorous analysis
- In a commercial setting, providing clients with the best possible advice

Establishing "relationships"

- We are often interested in examining relationships between two or more numerical variables
- Two approaches, for example, are:
  - correlation analysis
    - correlation does not imply causality – it's a measure of association!
    - does not quantify the degree of response
  - regression analysis
    - needs to be conditioned on theory, knowledge, experience for a regression analysis formulation

Econometric models

- Single equation models
  Focus on forecasting a single variable (can ignore 'simultaneity' issues)

- Structural models
  Recognize simultaneity issues (e.g. RICS City office market model)

Econometric/explicit approach:advantages

- **Logical and explicit framework enabling relationships to be investigated**
- **May have a well-developed body of theory and existing empirical evidence to draw upon**
- **Possible to investigate hypothesised relationships**
- **Permit an understanding of relationships and subsequent modification when underlying dynamic structure changes**
- **Alternative modelling methodologies/strategies which can be referred to**
- **Permits a simplified representation of the real world to be postulated to any degree of dis-aggregation**

Econometric/explicit approach:advantages (cont)

- **Influence of each variable in a relationship may be assessed (simulation)**
- **Whole battery of statistical tests available**
- **Possible to trace and reproduce the causes of both successful and inaccurate forecasts when outcome known**
- **Identification of turning points**
- **Policy simulations possible**
- **Forecast best in the long run**

Econometric approach:issues

- **Prone to "data mining"**
- **Data problems e.g. measurement/revision**
- **Unobservable variables e.g. expectations**
- **No consensus about relationships**
- **Simultaneity problems**
- **Wide margins of uncertainty. Implies that true relationship may be little understood**
- **A "good fit" on past data is no guarantee of continuing relationship**
- **Non-constant parameters i.e. changing relationships**
- **Inadequate sample sizes**
- **Omitted variable bias (missing variables)**

Regression models, cause and effect

- When we formulate a regression model relating the inputs X to the output Y, we often think of the model as describing a cause and effect relationship between the inputs and outputs
- Because of the ease of use of modern statistical software, regression analysis has become grossly abused by naïve users, who formulate regression models that are virtually meaningless, and sometimes, downright hilarious (to those understanding the limits of regression modeling)
- Many examples of regression models being suspect
  - ⇒ correlation/significant regression results does not imply causation or meaningful relationship
  - ⇒ apply theory/knowledge/experience and common sense tests, ***not just statistical tests!***

Uniwersytet Ekonomiczny

George Matysiak

Modelling and Forecasting

11th January, 2016