Chapter 1

Introduction

Outline

- Econometrics. Financial Econometrics. Examples of problems.
- Financial Data
- Continuous <> Discrete Data.
- Cardinal <> Ordinal <> Nominal Data.
- Time-Series <> Cross-Sectional <> Panel Data.
- Simple Return <> Log Return
- Steps in formulating an econometrics model.
- Points to consider when reading academic papers

Introduction:
The Nature and Purpose of Econometrics

- What is Econometrics?

- Literal meaning is “measurement in economics”.

- Definition of Financial econometrics:

  The application of Statistical and Mathematical techniques to problems in Finance using Computational tools.

Examples of the kind of problems that may be solved by an Econometrician

1. Testing whether the CAPM or APT represent superior models for the determination of returns on risky assets.
2. Measuring and forecasting the volatility of bond returns.
3. Explaining the determinants of bond credit ratings used by the ratings agencies.
4. Modelling long-term relationships between prices and exchange rates.
5. Testing the hypothesis that earnings or dividend announcements have no effect on stock prices.

What are the Special Characteristics of Financial Data?

- ***Frequency & Quantity of data:***

  High Frequency **&** ***Large amount of data***

  Example: Stock market prices are measured every time there is a trade or somebody posts a new quote: high frequency and large data

- ***Quality***

  Recorded asset prices are usually those at which the transaction took place. ***No measurement error*** **but financial data are "*noisy*".**

  Non-normality problem

Types of data: Continuous and Discrete

- **Continuous data** can take on any value in a range

  Note: their values are limited only by precision.

  For example, the rental yield on a property could be 6.2%, 6.24%, or 6.238%.

- On the other hand, **discrete data** can only take on certain values, usually integers

  Example: the number of people in a particular underground carriage, the number of shares traded during a day…

Types of data: Cardinal, Ordinal, Nominal Numbers

- ***Cardinal numbers***: actual numerical values that a variable takes; with equal distance between values. (12 is twice as good as 6)
  - Example: the price of a share or of a building, the number of houses in a street.

- ***Ordinal numbers*** can only be interpreted as providing a position or an ordering.
  - Example: Ranking of a runner in a race (1^{st}, 2^{nd}…), Ranks associated to credit rating levels (1^{st}, 2^{nd},…)

- ***Nominal numbers*** occur where there is no natural ordering of the values at all.
  - Examples: Telephone numbers, Coding assigned to qualitative data (e.g. when describing the exchange that a US stock is traded on: 1 means NYSE, 2 means NASDAQ, 3 means AMEX).

Types of Data and Notation

- There are 3 types of data which econometricians might use for analysis:
  1. Time series data
  2. Cross-sectional data
  3. Panel data

- **Time series**: data collected over a period of time on one or more variables

- Examples:

Series	*Frequency*
GNP or unemployment rate	monthly, or quarterly
money supply	weekly
exchange rate	daily
value of a stock market index	as transactions occur

Time Series versus Cross-sectional Data

- Examples of *Time Series Regression*

  - How the value of a country's stock index has varied with that country's macroeconomic fundamentals.
  - How the value of a company's stock price has varied when it announced the value of its dividend payment.
  - The effect on a country's exchange rate of an increase in its trade deficit.

- **Cross-sectional data**: data on one or more variables collected at a single point in time, e.g.
  - Cross-section of stock returns on the New York Stock Exchange
  - A sample of bond credit ratings for UK banks

Cross-sectional and Panel Data

- Examples of *Cross-Sectional Regression*

  - The relationship between company size and the return to investing in its shares
  - The relationship between a country's GDP level and the probability that the government will default on its sovereign debt.

- **Panel Data**: has the dimensions of both time series and cross-sections, e.g. the daily prices of a number of blue chip stocks over two years.

- It is common to denote each observation by the letter t and the total number of observations by T for time series data, and to to denote each observation by the letter i and the total number of observations by N for cross-sectional data.

Returns in Financial Modelling

- It is preferable not to work directly with asset prices (one reason: non-stationarity), so we usually convert the raw prices into a series of returns. There are two ways to do this:

Simple returns or **Log returns**

$$R_t = \frac{p_t - p_{t-1}}{p_{t-1}} \times 100\%$$

$$r_t = \ln\left(\frac{p_t}{p_{t-1}}\right) \times 100\%$$

where, R_t denotes the return at time t

p_t denotes the asset price at time t

ln denotes the natural logarithm

- **Note**: We assume that the price series have been already adjusted to account for dividend payments and stock splits (**adjusted price**).

Advantages of Log Returns

1. They have the nice property that they can be interpreted as **continuously compounded returns**.

2. Can add them up, e.g. if we want a weekly return and we have calculated **daily log returns**:

$$r_1 = \ln p_1/p_0 = \ln p_1 - \ln p_0$$
$$r_2 = \ln p_2/p_1 = \ln p_2 - \ln p_1$$
$$r_3 = \ln p_3/p_2 = \ln p_3 - \ln p_2$$
$$r_4 = \ln p_4/p_3 = \ln p_4 - \ln p_3$$
$$r_5 = \ln p_5/p_4 = \ln p_5 - \ln p_4$$

$$\ln p_5 - \ln p_0 \quad = \quad \ln p_5/p_0$$

Portfolio: Simple Returns

- Portfolio Return: is based on the total value of the portfolio

- We can check that the simple return on a portfolio of assets is a weighted average of the simple returns on the individual assets:

$$R_{pt} = \sum_{i=1}^{N} w_{ip} R_{it}$$

- But this does not work for the continuously compounded returns, since log is not a linear transformation.

Aggregated Simple Return (page 128)

Over K periods, if the asset is purchased at time (**t-K**) at price $\mathbf{P_{t-K}}$, and sold at time t at price $\mathbf{P_t}$, then aggregate return over K periods is:

$$R_{Kt} = \frac{p_t - p_{t-K}}{p_{t-K}} = \frac{p_t}{p_{t-K}} - 1 = \left[\frac{p_t}{p_{t-1}} \times \frac{p_{t-1}}{p_{t-2}} \times \ldots \times \frac{p_{t-K+1}}{p_{t-K}}\right] - 1$$

$$= \left[(1+R_t)(1+R_{t-1})\ldots(1+R_{t-K+1})\right] - 1$$

$$= \left[\prod_{i=0}^{K-1}(1+R_{t-i})\right] - 1$$

where R_t is the *simple return* for each period. Then the *annualized holding period return* R_H is given by:

$$(1+R_H) = (R_{Kt}+1)^{1/K}$$

Real Versus Nominal Series (pages 130-131)

- The general level of prices has a tendency to rise most of the time because of **inflation**
- We may wish to **transform nominal series into real ones** to adjust them for inflation
- This is called **deflating a series** or displaying a series *at constant prices*
- We only deflate series that are in nominal price terms, not quantity terms.

Deflating a Series

- If we wanted to convert a series into a particular year's figures (e.g. house prices in 2010 figures), we would use:

$$\textit{Real series}_t = \textit{Nominal series}_t \times \textit{Deflator}_{\textit{reference year}} / \textit{Deflator}_t$$

- This is the same equation as the previous slide except with the deflator for the reference year replacing the assumed deflator base figure of 100

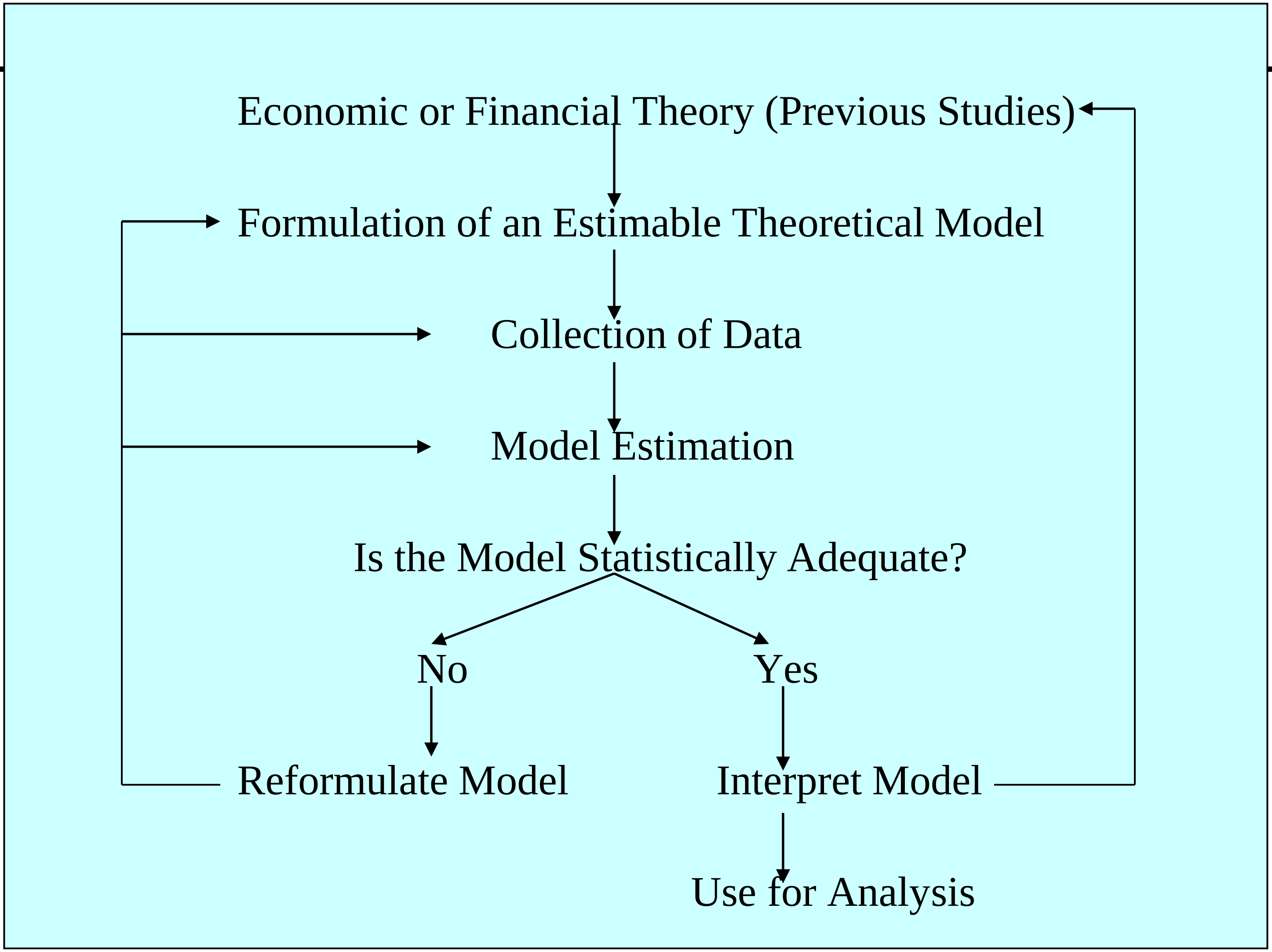
- Often the **consumer price index, CPI (price level change)** is used as the deflator series.

Table 1.1 How to construct a series in real terms from a nominal one

Year	Nominal house prices	CPI (2004 levels)	House prices (2004 levels)	House prices (2013) levels
2001	83,450	97.6	85,502	105,681
2002	93,231	98.0	95,134	117,585
2003	117,905	98.7	119,458	147,650
2004	134,806	100.0	134,806	166,620
2005	151,757	101.3	149,810	185,165
2006	158,478	102.1	155,218	191,850
2007	173,225	106.6	162,500	200,850
2008	180,473	109.4	164,966	165,645
2009	150,501	112.3	134,017	173,147
2010	163,481	116.7	140,086	167,162
2011	161,211	119.2	135,244	155,472
2012	162,228	121.1	133,962	165,577
2013	162,245	123.6	131,266	162,245

Notes: All prices in British pounds; house price figures taken in January of each year from Nationwide (see appendix 1 for the source). CPI figures are for illustration only.

Steps involved in the formulation of econometric models



Some Points to Consider when reading papers in the academic finance literature

1. Does the paper involve the development of a theoretical model or is it merely a technique looking for an application, or an exercise in data mining, so that the motivation is poor?

2. Is the data of "good quality"? Is it from a reliable source? Is the size of the sample sufficiently large for asymptotic theory to be invoked?

3. Have the techniques been validly applied? Have diagnostic tests for violations of been conducted for any assumptions made in the estimation of the model?

Some Points to Consider when reading papers in the academic finance literature (cont'd)

4. Have the results been interpreted sensibly? Is the strength of the results exaggerated? Do the results actually address the questions posed by the authors? Can the results be replicated by other researchers?

5. Are the conclusions drawn appropriate given the results, or has the importance of the results of the paper been overstated?