ECON 343
THE ECONOMETRICS OF FINANCIAL MARKETS

Jad Chaaban

Spring 2005-2006

Outline – Lecture 1

- PART I
  - Course description, requirements and syllabus
  - Expected achievements
  - Why do we need econometrics?
  - Forecasting in ancient and modern times
- PART II
  - Types of data
  - Basic properties of time series
  - Properties of random variables
  - Estimation
  - Common probability distributions

PART I

Course description

- Cover most important methods used in the empirical analysis of economic and financial problems
- Emphasis on time series econometrics
- Applications mostly in finance, but other examples explored
- Prerequisite: basic training in statistics
- Requirements:
  - Term paper (40%)
  - Class discussions and tutorials (10%)
  - Final exam (50%)
- Copies of readings and datasets will be provided in class
- Hands-on experience is mandatory (software: Eviews).

Expected achievements

- Learn to apply econometric techniques to:
  - Understand the determinants of economic fundamentals
  - Forecast stock returns, asset prices and interest rates
  - Critically evaluate empirical research on financial markets
- Learn to handle data and use statistical packages
- Be able to conduct empirical research on economic and financial problems
- Some future jobs where you can use these tools:
  - Economist (public sector, central bank)
  - Financial consultant
  - Stock market analyst
  - Investment banker
  - Business Lawyer

Expected achievements

- Some things you will be able to do at the end of this course:
  - Estimate the volatility of Solidere A stock returns
  - Forecast the likely evolution of major stock market indices (Dow Jones, S&P 500) and commodity prices (crude oil)
  - Analyze whether the Lebanese public debt is sustainable
  - Forecast on a monthly basis the interest rate on US 3-month Treasury Bills
  - Estimate the determinants of the flow of savings deposit to a private bank
  - Forecast major macroeconomic indicators (GDP growth, inflation, exchange rate fluctuation…)

Course syllabus

1. Introduction: Properties of random variables and Time Series
3. Review of the basic multiple regression model
4. Application I: The Capital Asset Pricing Model (CAPM)
5. Smoothing and extrapolation of Time Series
6. Application II: Simple forecasting of major asset prices
7. Properties of Stochastic Time Series
8. Application III: Computer Lab session
8. Unit roots and Co-integration
9. Application IV: Sustainability of budget deficits and public debt
10. VAR and VECM models
11. Linear Time Series models (MA, AR, ARMA, ARIMA)
12. Application V: Interest rate forecasting

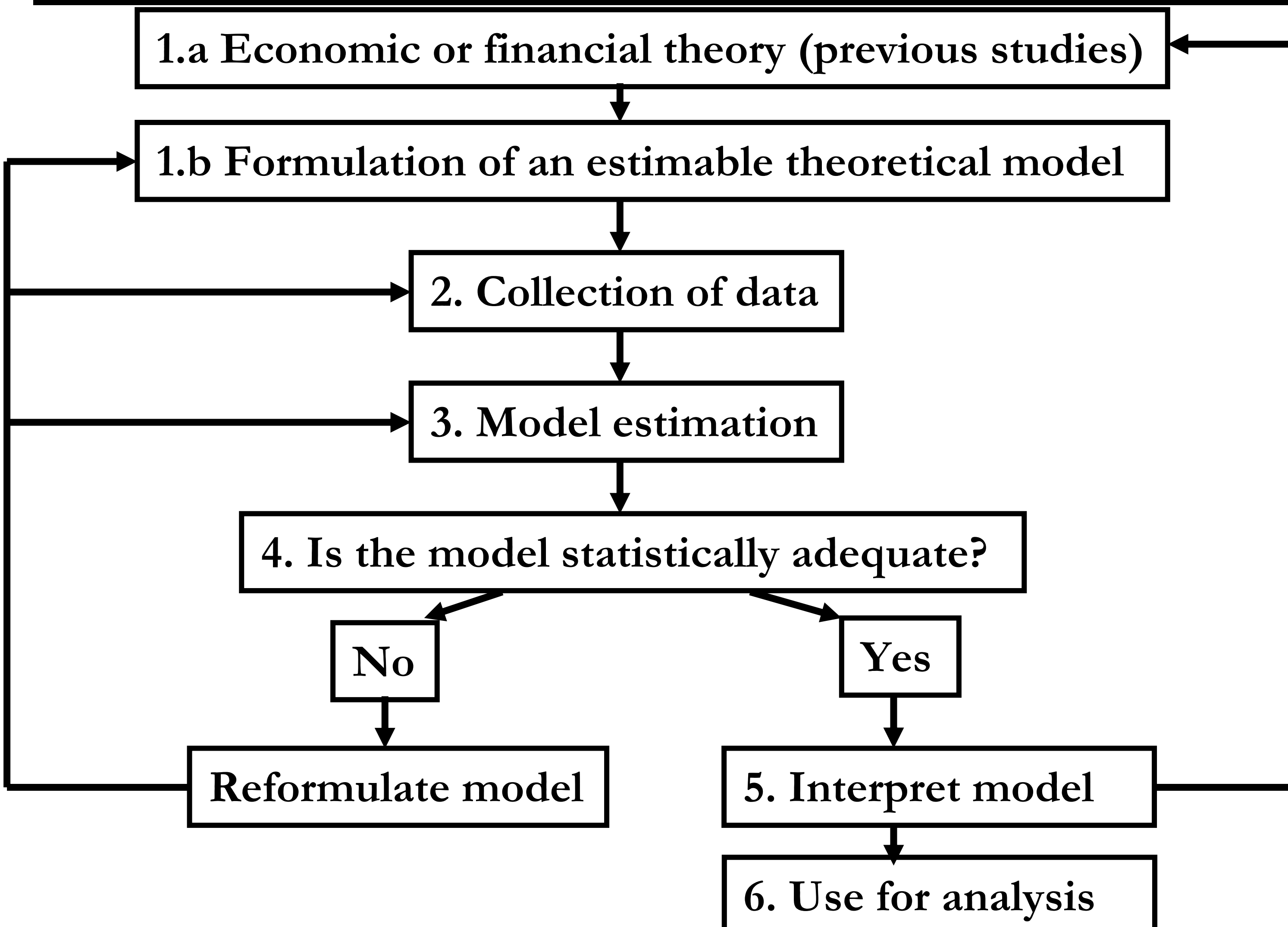
Why do we need econometrics?

- Definition: 'Measurement in economics'
- Model building (formalization): construct and test mathematical representations of the real world
- Advantages of models:
  - Forces to think about all the important relationships
  - Easier to discuss economic ideas and criticize them
  - Process of validating the individual relationships in the model
  - Statistical measurement of the accuracy of forecasts
  - Sensitivity analysis is possible (individual effects)
- Formalization is only a complement (and not a substitute) for intuition and informed experience
- Econometrics is very probabilistic, but better than anecdotes

The Data Generating Process

- Central objective in econometrics: determine the Data Generating Process DGP
- Puts aside randomness in seeking models of the true DGP
- The actual economy is very complex to be explained
- Cross-sectional aggregation
  - Aggregation of a very large number of non-identical, non-independent decision making units (households, corporations)
- Temporal aggregation
  - The observation period of the data does not necessarily match the decision making periods

Steps in an econometric model



Steps in an econometric model

- Step 1: Theoretical model to approximate the DGP
- Step 2: Collect data, either from sources or via surveys
- Step 3: Choice of estimation method relevant to the model
- Step 4: Does the model adequately describe the data (goodness of fit)?
- Step 5: Are the parameter estimates in line with the theory?
- Step 6: Propose suggested course of action
  - 'If inflation and GDP rise, buy stocks in sector X'
  - 'The change in expected returns of stock Z is related to the future movements of the Dow-Jones Industrial Average'
  - 'The long-run trend in oil prices is upwards'

Forecasting in ancient and modern times

- The most famous forecaster in known history is not an economist!
- Delphi Oracle: ancient Greece, more than 1000 years of profitable business (700 BC to 300 AD)

- Famous quote: “The best qualification of a prophet is to have a good memory”, Marquis of Halifax
- Modern Oracles: Alan Greenspan, among others

Forecasting in ancient and modern times

- “An economist is an expert who will know tomorrow why the things he predicted yesterday didn't happen today”, Evan Esar

- Federal Reserve Committee sample statement:
  - “The Committee perceives that the upside and downside risks to the attainment of sustainable growth for the next few quarters are roughly equal.
  - In contrast, the probability, though minor, of an unwelcome fall in inflation exceeds that of a rise in inflation from its already low level. The Committee judges that, on balance, the risk of inflation becoming undesirably low remains the predominant concern for the foreseeable future.
  - In these circumstances, the Committee believes that policy accommodation can be maintained for a considerable period”

Forecasting in ancient and modern times

- Each year, the prediction industry showers us with $200 billion in (mostly erroneous) information
- Recent events that caught the forecasters by surprise:
  - 1987 stock market crash and its subsequent rapid recovery to record heights
  - Entry of women into the workforce in massive numbers
  - Fall of communist Eastern Europe
  - Gulf War
  - All recessions, including the crash of 1929 and recent, smaller blips in the financial markets
  - Use of lasers to transmit telephone messages (even though the phone company's researchers at Bell Labs invented it)

PART II

Types of data

- Notations
  - T : amount of data in the time dimension
  - M : number of variables
  - N : Number of 'regions' or 'agents'
- Quantitative vs. Qualitative data
- Continuous vs. Discrete data
- <u>Time series data</u>
  - Data collected over a period of time on one or more variables
  - T large ; M, N small
  - Frequency: measure of the regularity with which the data is collected or recorded (daily, monthly, quarterly, annually…)

Types of data

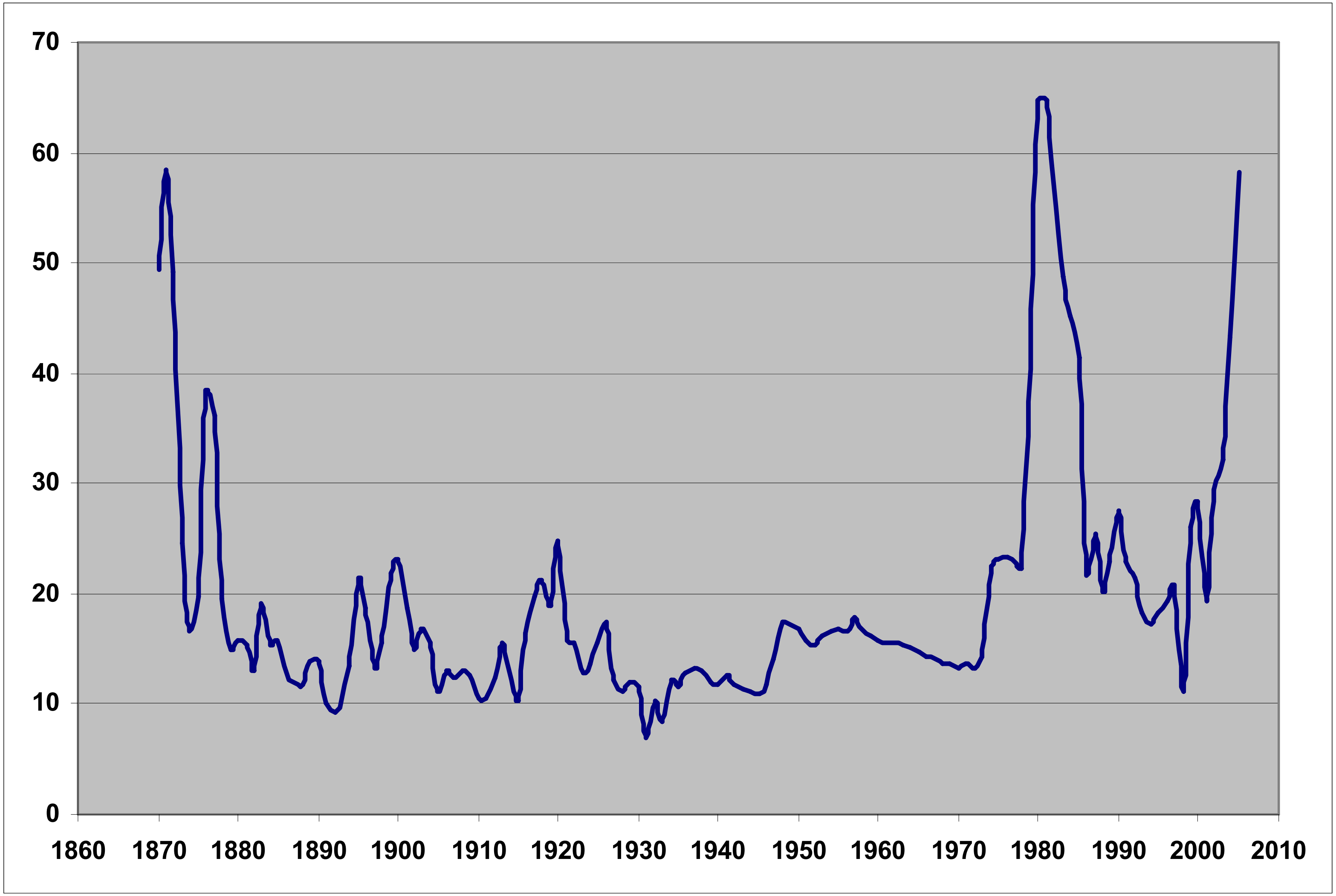
- Problems that could be tackled using time series data:
  - How the value of a country's stock index has varied with that country's macroeconomic fundamentals
  - How the value of a company's stock price has varied when it announced the value of its dividend payment
  - The effect on a country's exchange rate of an increase in its trade deficit
- <u>Cross-sectional data</u>
  - Data collected on one or more variables at a single point in time
  - $T = 1$; M, N large
- Problems that could be tackled using cross-sectional data:
  - The relationship between company size and the return to investing in its shares
  - The relationship between GDP and sovereign debt default

Types of data

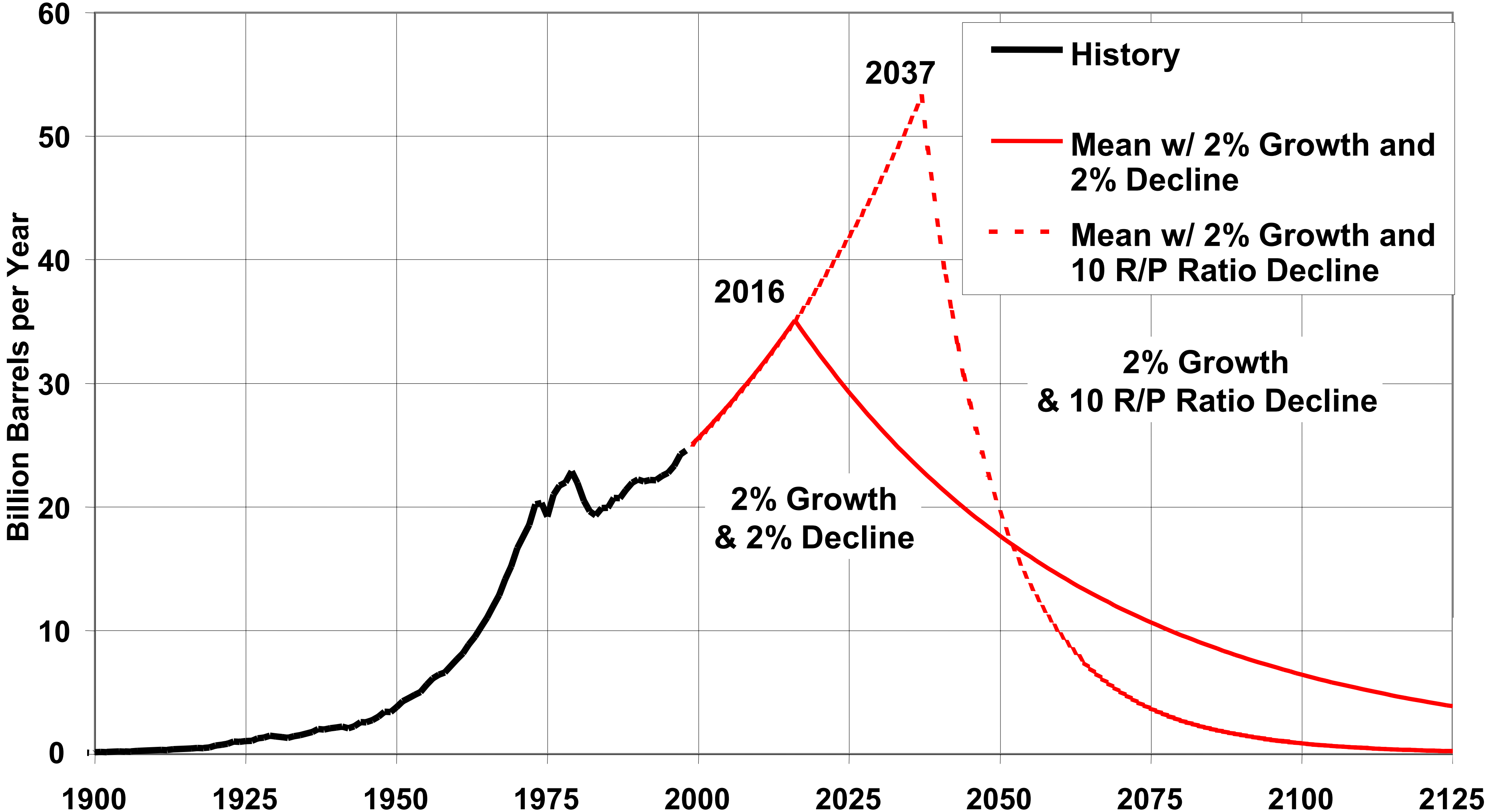
- <u>Panel data</u>
  - Data having both time and cross-sectional dimensions
  - T small , M small , N large
  - Not within the scope of this course

- Differences between financial and economic data
  - More data available, less measurement error in finance
  - Frequency: asset prices can be obtained at a high frequency
  - Accuracy: less revision problems involved
  - Seasonality: many patterns involved in financial data

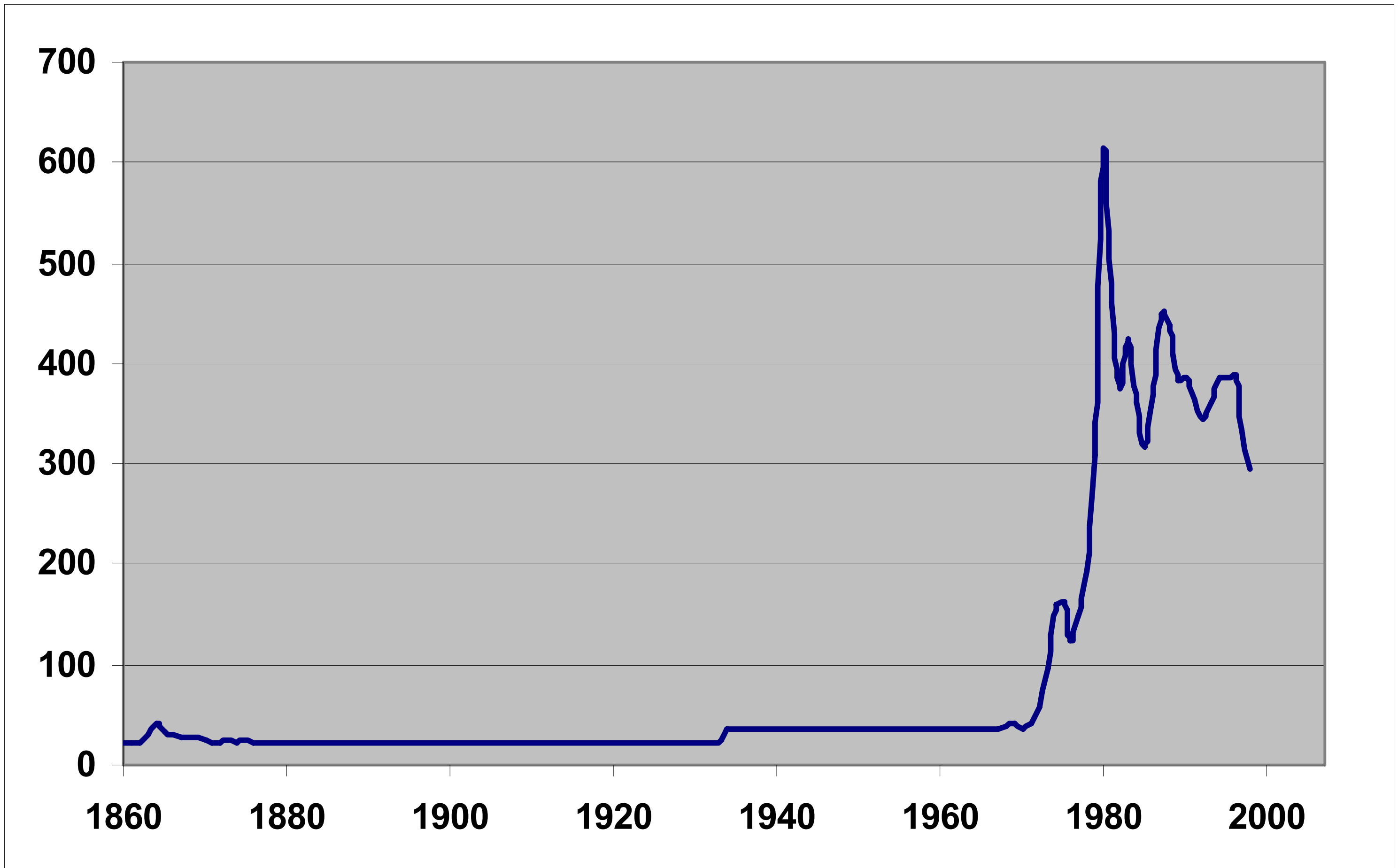
Examples:
Real Oil price, Texas Brent, 1870-2005



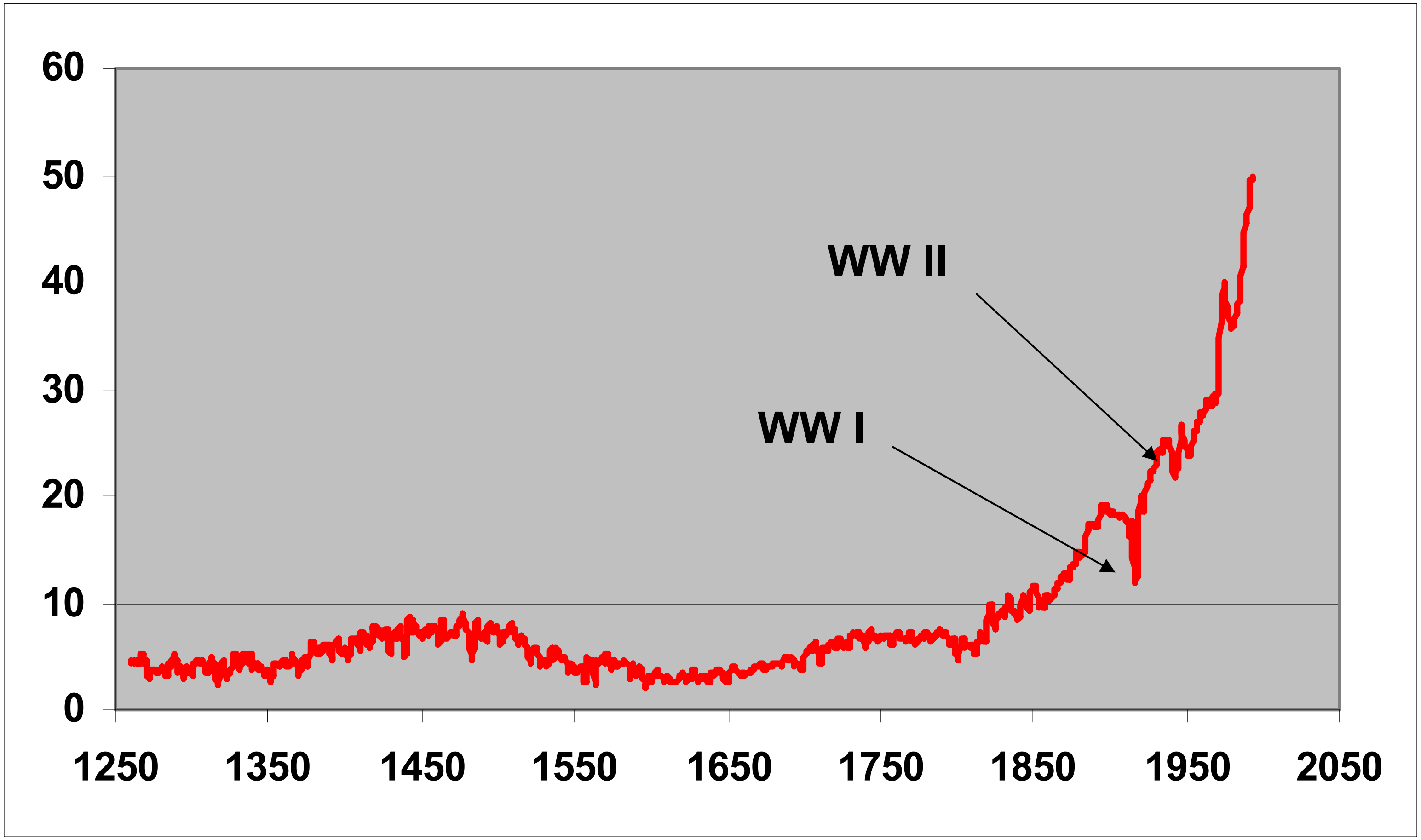
An example: Forecasting crude oil production with declining reserves



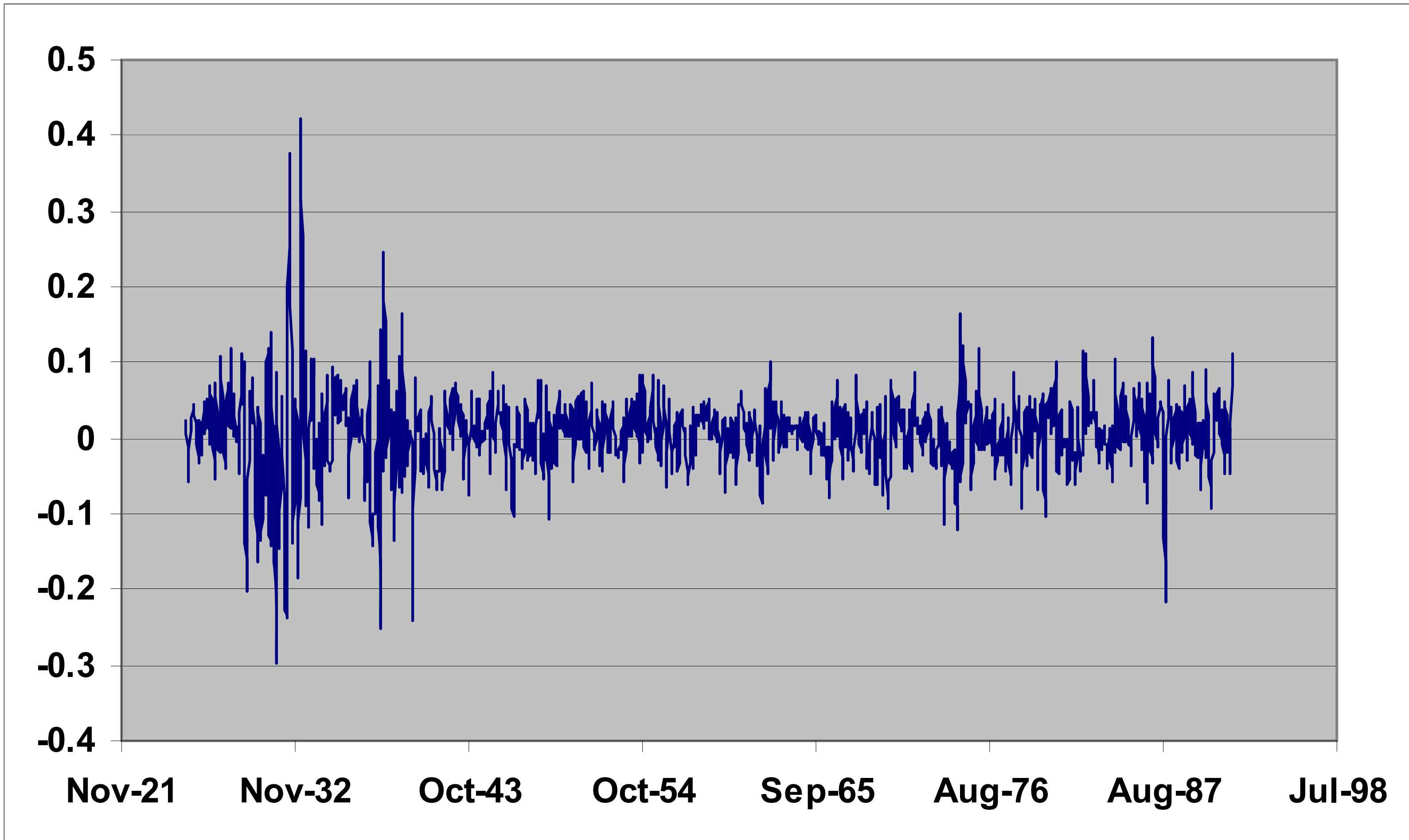
Gold price, NYSE, 1860-1999



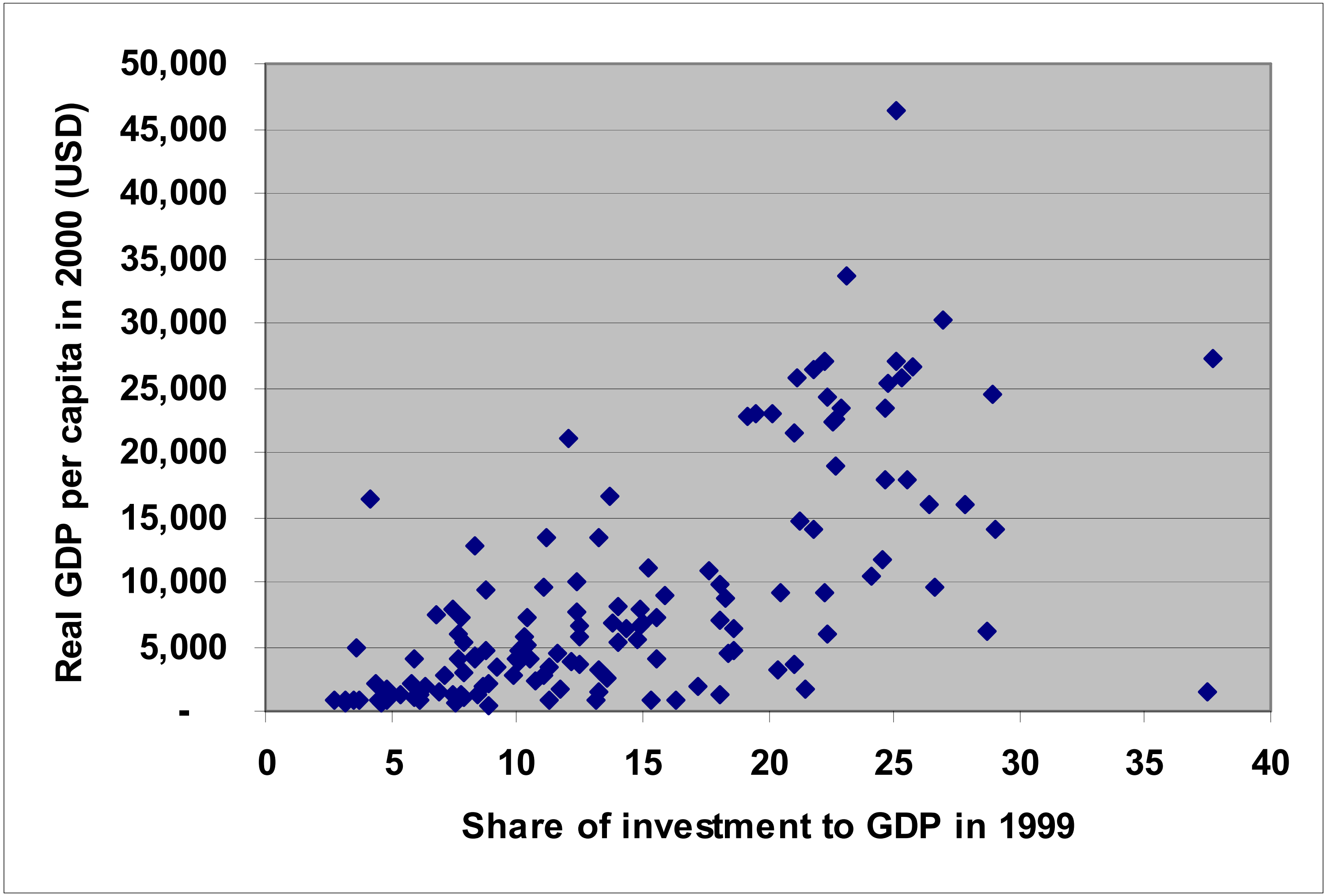
Real daily wage in £, UK, 1260-1994



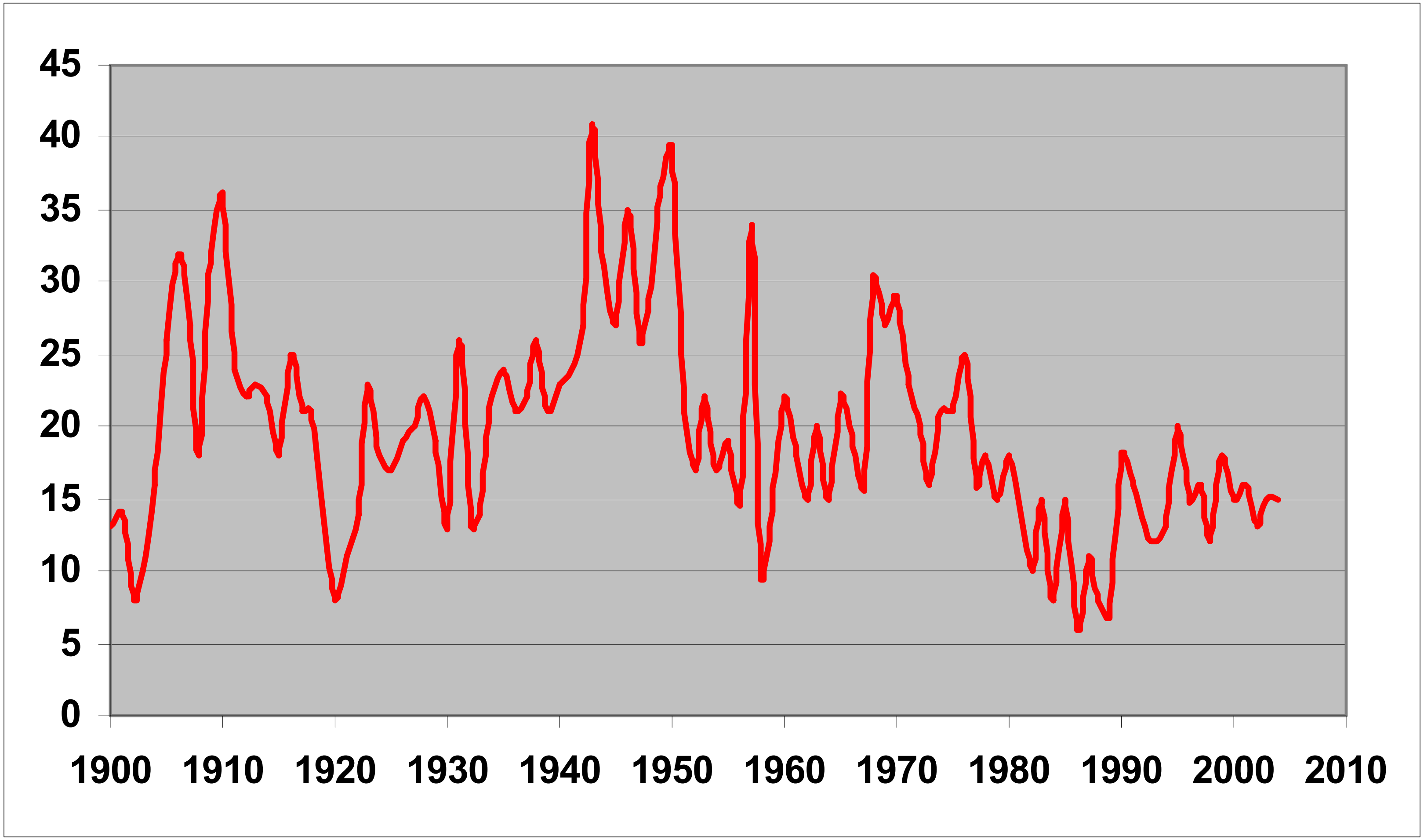
Monthly returns of value-weighted Standard and Poor 500 stock from 1926-1991



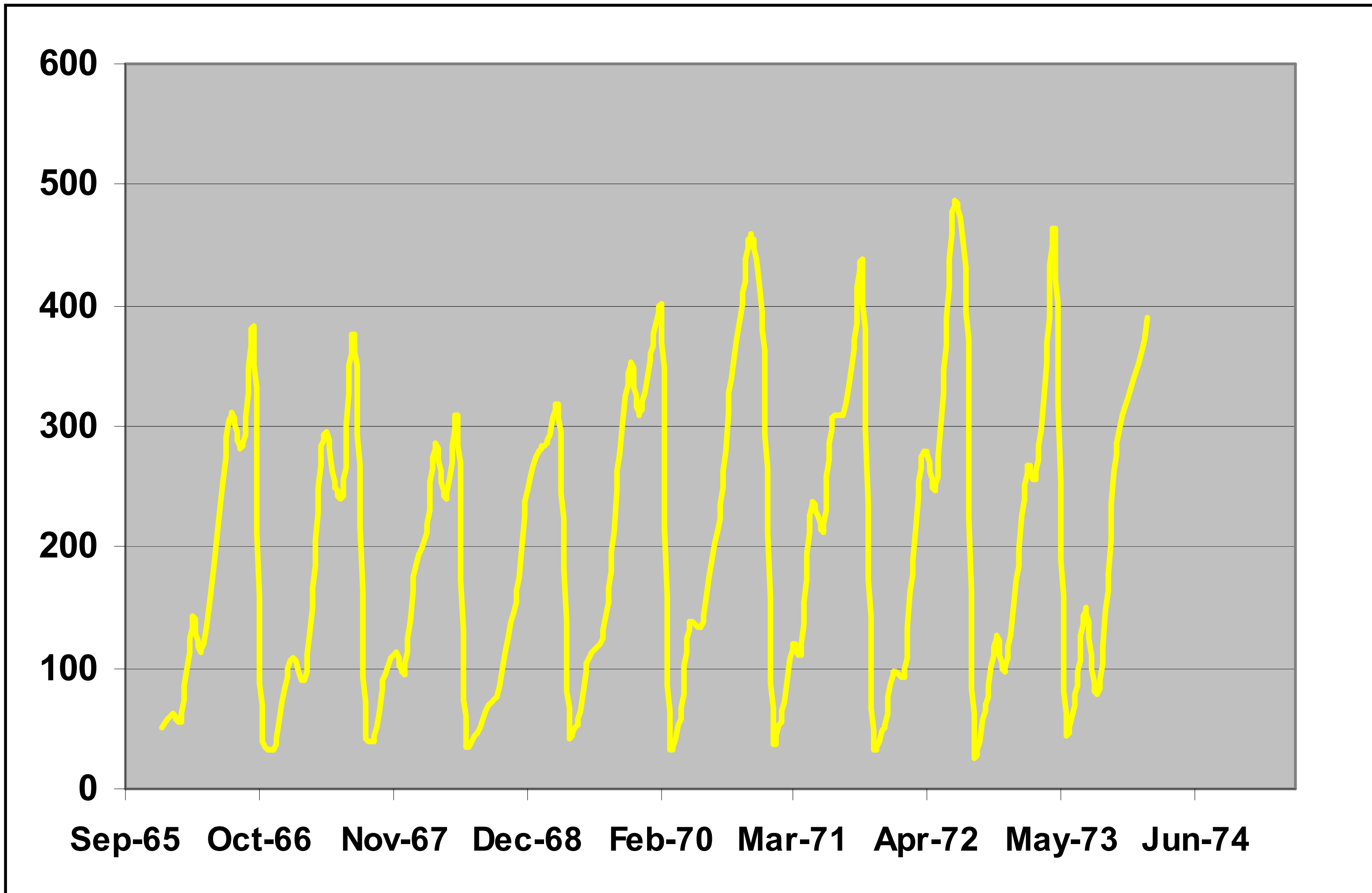
Investment (1999) vs. GDP (2000), 134 countries



Number of magnitude 7.0 and greater earthquakes per year, since 1900



Monthly Boston armed robberies Jan.1966-Oct.1975



Basic properties of time series

- It is traditional to decompose time series into a variety of components, some or all of which may be present in a particular instance
- If Y_t is the sequence of values of an economic index, then its generic element can be expressed as

$$Y_t = T_t + C_t + S_t + \varepsilon_t$$

where

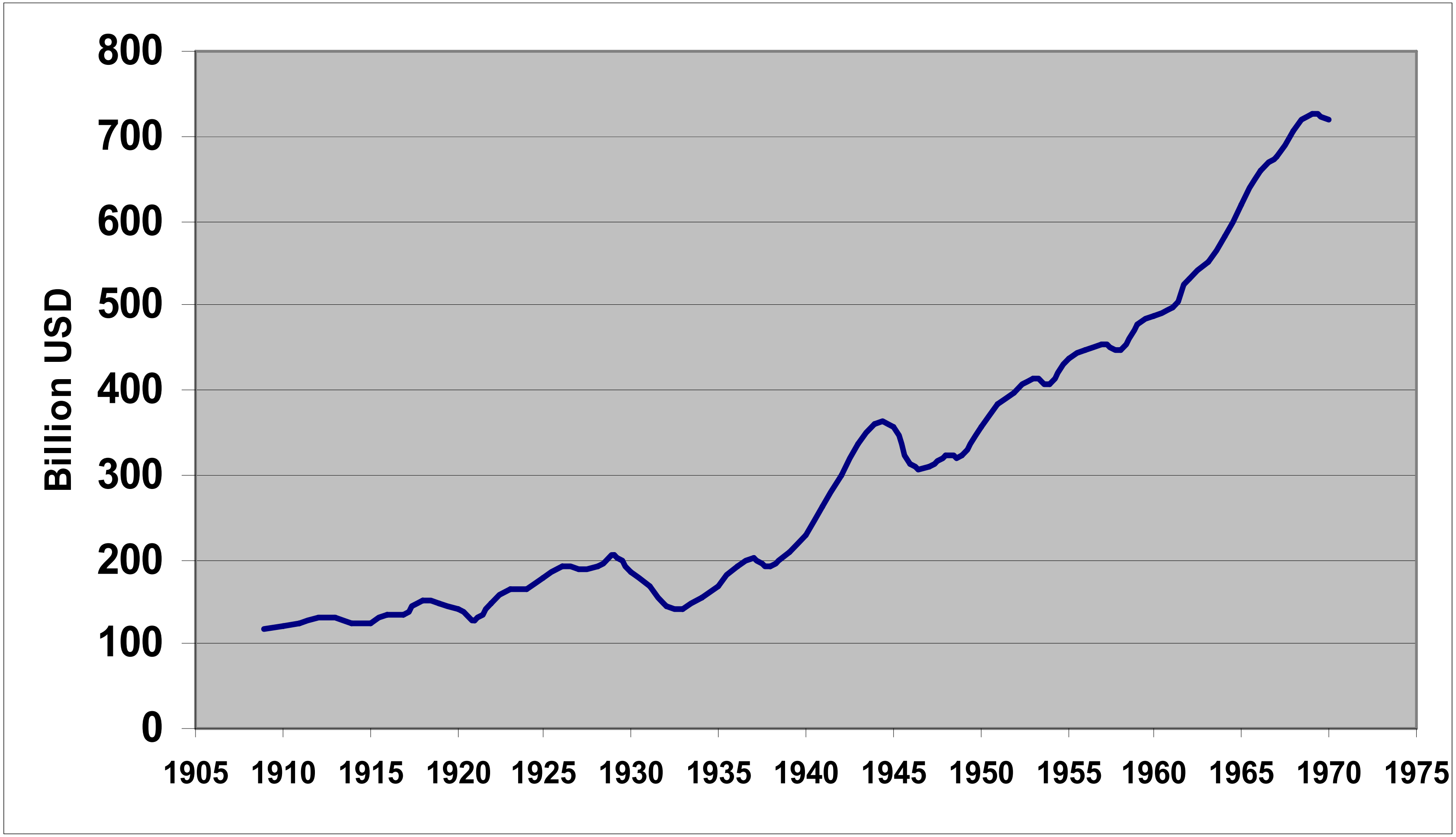
T_t : the global trend,

C_t : a secular cycle,

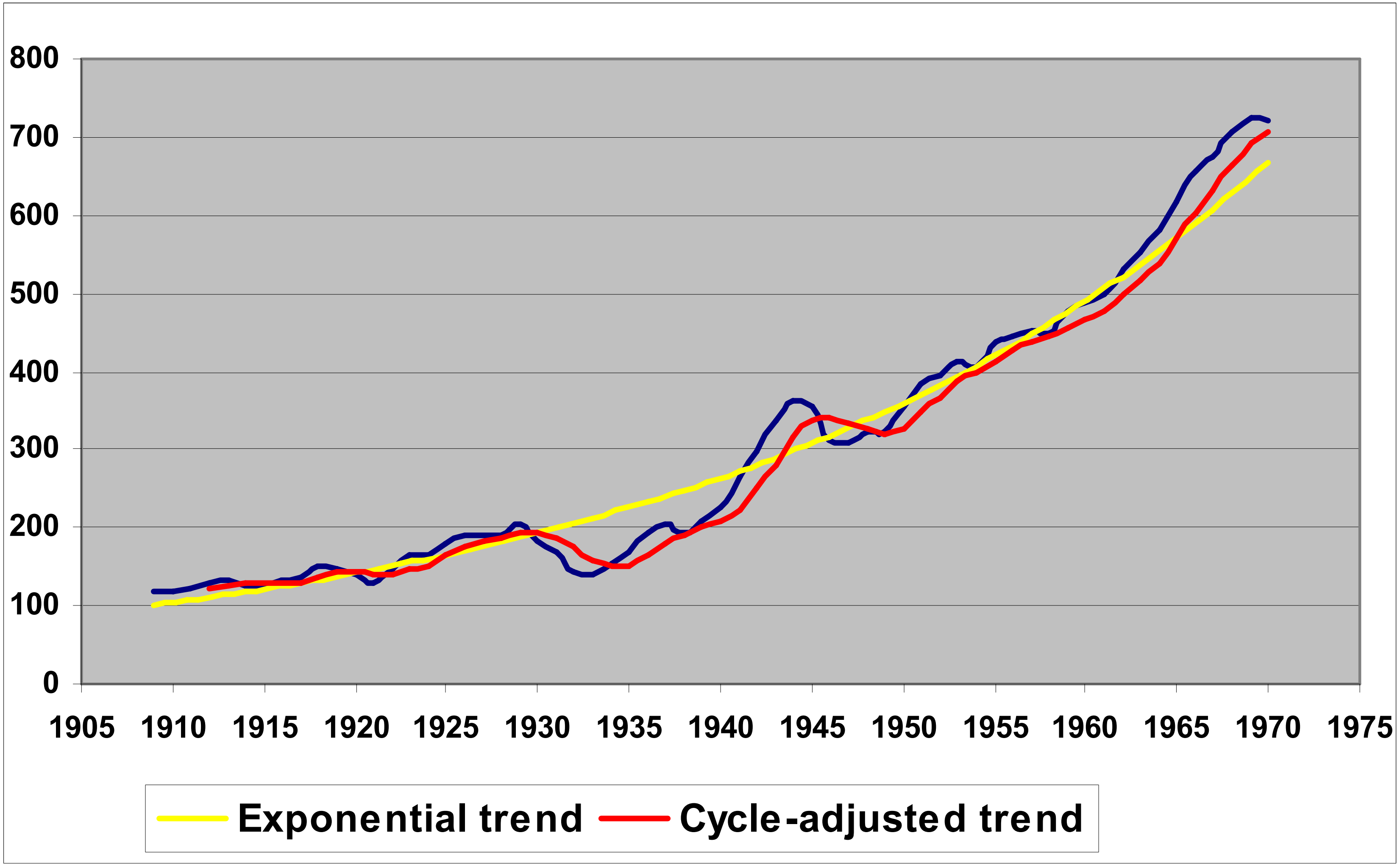
S_t : the seasonal variation

ε_t: an irregular component.

An example: US GNP index evolution



An example: GNP index evolution



An example: GNP index evolution

- T_t exponential growth trend
- C_t a superimposed cycle with a period of roughly four and a half years, which happens to correspond, more or less, to the average lifetime of the legislative assembly
- S_t an annual cycle; a reflection of the fact that some economic activities, such as building construction, are significantly affected by the weather and by the duration of sunlight
- ε_t the residue, corresponding to an irregular component for which no unique explanation can be offered
- This component ought to resemble a time series generated by a so-called *stationary stochastic process*
- Important to extract the trends before handling this process

Properties of random variables

- **Definition:**
  - A variable that takes on alternative values, each with a probability less than or equal to 1
- **Probability distribution:**
  - Lists all possible outcomes and the probability that each will occur
- **Population vs. Sample**
- **Population summary statistics**
  - Mean: Variance:

$$\mu_X = \sum_{i=1}^{N} p_i X_i \qquad \sigma_X^2 = \sum_{i=1}^{N} p_i \left[X_i - E(X)\right]^2$$

  - Standard deviation $\sigma = +\sqrt{\sigma_X^2}$

Properties of random variables

- **Joint distribution of Random Variables (Population):**
  - **Covariance of two variables *X* and *Y***

$$Cov(X,Y) = E\left[(X - E(X))(Y - E(Y))\right]$$

  - **Correlation coefficient (lies between -1 and +1)**

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$$

- **If *X* and *Y* are independent, *E(XY)=E(X)E(Y)***
- **If *X* and *Y* are independent, *Cov (X,Y)*=0**
- **If X and Y are correlated, beware of <u>Spurious Correlation</u>**
  - **Two variables are correlated, but one is not the cause of the other**

Estimation

- Estimate the statistics of a population through a sample
- Sample of *N* observations
- Unbiased estimator: expected value of estimator equal to true parameter
- Estimator of the mean → $\bar{X} = \frac{1}{N}\sum_{i=1}^{N} X_i$
- Estimator of the variance → $\frac{1}{N-1}\sum_{i=1}^{N}(X_i - \bar{X})^2$
- Sample correlation coefficient

$$r_{XY} = \frac{\sum_{i=1}^{N}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{N}(X_i - \bar{X})^2 \sum_{i=1}^{N}(Y_i - \bar{Y})^2}}$$

Estimation

- Desirable properties of estimators
  - <u>Lack of bias</u>: zero bias in the estimation

$$Bias = E(\hat{\beta}) - \beta$$

  - <u>Efficiency</u>: small variance of the estimator
  - <u>Minimum Mean Square Error MSE</u>

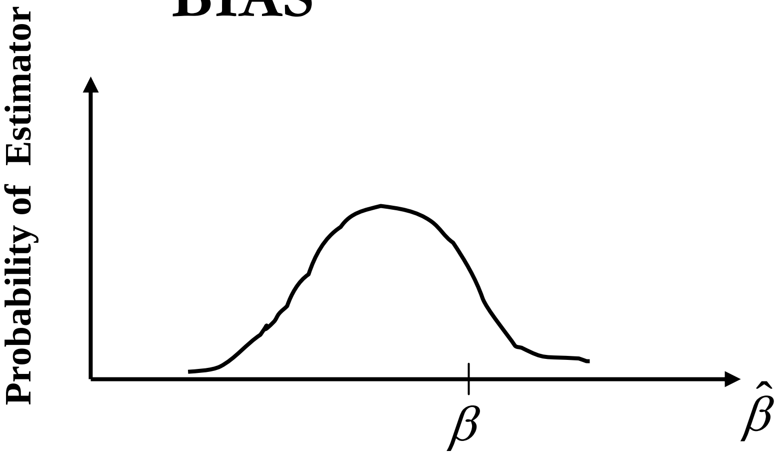
$$MSE = \left[Bias(\hat{\beta})\right]^2 + Var(\hat{\beta})$$

  - <u>Consistency</u>

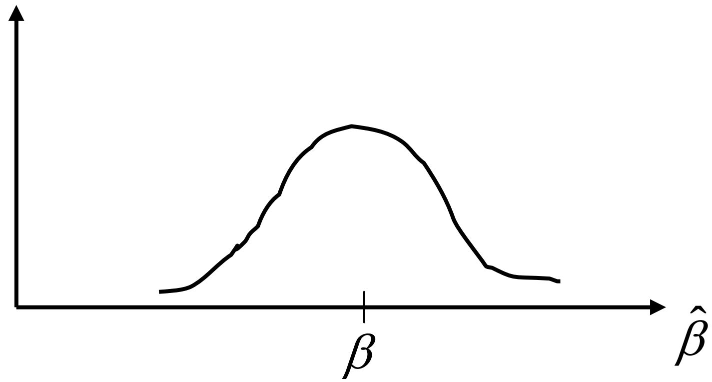
    The estimator gets close to the true parameter as the sample size increases

$$\underset{N \to \infty}{p\lim} \left|\beta - \hat{\beta}\right| < \varepsilon$$

BIAS

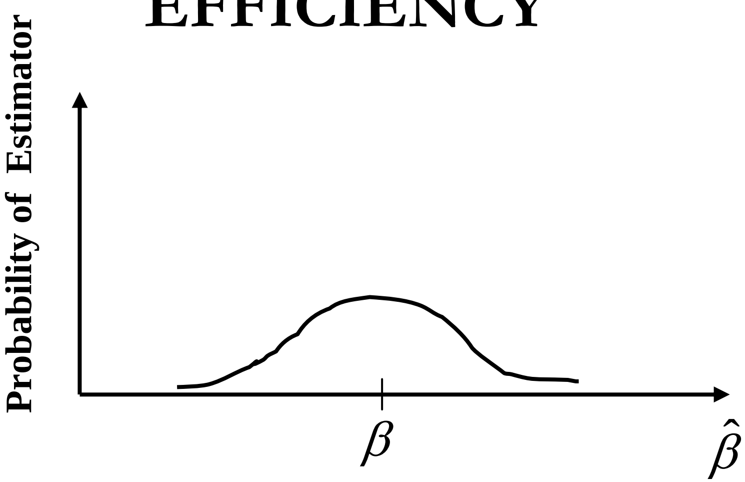


Biased Estimator

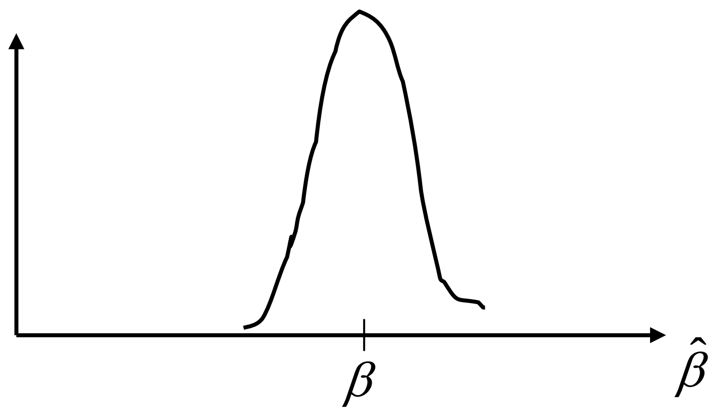


Unbiased Estimator

EFFICIENCY



Inefficient Estimator



Efficient Estimator

Common probability distributions

- Normal distribution $N(0,\sigma^2)$
- Chi-Square distribution
  - The sum of the squares of *N* independently distributed *normal* random variables (with mean 0 and variance 1) is distributed as *chi* square with *N* degrees of freedom
- The Student t distribution
  - If X is normally distributed with mean 0 and variance 1 and *Z* is distributed as chi square, then if *X* and *Z* are independent, *X/squroot(Z/N)* follows a t distribution with *N* degrees of freedom
- The F distribution
  - If *X* and *Z* independent and follow chi square distribution, then $(X/N_1)/(Z/N_2)$ is distributed according to an F distribution with N_1 and N_2 degrees of freedom