Chapter 2: Statistical Tests of Prices and Price Trends

Quantitative Methods for Competition and Regulation

This chapter

- Cross-sectional price test
- Hedonic price analysis
- Price correlation analysis
- Stationarity analysis
- Granger causality tests

- For details on:
  - Co-integration tests

  see OFT report

Cross-sectional price tests

- Data on prices for two different products, two different regions, two points in time…
- Are the two sets of prices uniform? i.e. are they different only due to randomness?
- Applications:
  - Define product market
  - Define geographical market
  - Evidence of competition problems (collusion,…)

Description

- Are the two sets of prices uniform?
- In statistical terms…
  - Are they samples from the same distribution?
  - In particular, are their means equal?
  - Testing equality of means
- Two cases:
  - Samples are independent (e.g. two different areas)
  - Samples are not independent (e.g. before and after alleged price fixing)

Description: Independent Samples

- We have that that:

$$\frac{\overline{X}_1 - \overline{X}_2 - \mu_1 - \mu_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim \text{N}(0,1)$$

- And therefore, if we know the population variances we can compute the statistic:

$$\frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

and check if it falls within the bounds determined by the z's

Description: Independent Samples

- If we don't know the population variances, but the samples are large (n_1,n_2>30), then we can use:

$$\frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

and still compare it with the z's

- If the samples are small (n_1,n_2<30):

a) If the variances are assumed equal then, we can use

$$\frac{\overline{X}_1 - \overline{X}_2}{S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \text{ where } S = \sqrt{\frac{S_1^2(n_1 - 1) + S_2^2(n_2 - 1)}{n_1 + n_2 - 1}}$$

and compare it with the values of a t of student with n_1+n_2-2 degrees

b) If they are not we can use the first statistic in this slide and compare it with a t with a very complicated number of degrees of freedom

Description: Paired Samples

- Prices "before" and "after" are certainly not independent
- In this case, we can use the statistic

$$\frac{\overline{d}}{\sqrt{\frac{S_d^2}{n}}} \text{ where } \overline{d} = \frac{\sum d_i}{n} \text{ and } S_d = \sqrt{\frac{\sum (d_i - \overline{d})^2}{n-1}}$$

and compare it with the values of the t of student with n-1 degrees of freedom (and if the sample is large with the z's)

- This test is more powerful than the previous one

Data and Interpretation

- Need at least 20 observations
- For few observations, population need to be normal (a log transform may help)
- For international comparisons, need to be careful with taxes, exchange rates,…
- Problems:
  - Even if the prices are significantly different in statistical terms, are they in economic terms?
  - Couldn’t it be due to quality or cost differences?

Application: CD Price Comparison

- Consumer association:
  - Aren't CD prices in the UK higher than in US?
- Monopolies and Mergers Commission:
  - Survey comparing retail prices (no tax, in sterling) for UK, US, France, Germany and Denmark
  - Result: average prices 8% higher in the UK wrt US
  - Sony: larger sample for US and UK
  - Result: weighted average prices 5% higher (and statistically significant)
- Conclusion: prices are higher but prices of other goods are also higher!

Table 1: Cost in Pounds Sterling of Pre-selected CD Titles in Europe and the US

Pre-selected Titles	*UK*	*US*	*F**	*G**	*Denmark*
Diva – Annie Lennox	11.78	10.21	13.03	11.23	11.38
Soul Dancing – Taylor Dayne	11.25	9.83	12.87	11.19	11.58
Zooropa – U2	10.22	9.85	11.88	10.72	11.33
Keep the Faith – Bon Jovi	10.56	10.53	12.81	10.89	11.50
River of Dreams – Billy Joel	10.33	9.45	11.74	10.75	11.28
Timeless – Michael Bolton	11.26	10.45	12.41	11.00	11.28
Tubular Bells II – Mike Oldfield	11.71	10.21	12.71	10.92	11.25
What's Love Got to Do With It?					
– Tina Turner	10.06	9.67	13.12	11.15	11.44
Column Average	*10.90*	*10.03*	*12.57*	*10.98*	*11.38*

Further examples

- In many cases two products have been put in two different markets because of price differences:
  - Aerospatiale-Alenia/de Havilland (EC)
    - Jet aircraft and turbo propeller were put in different markets
  - Du Pont/ICI
    - Nylon fibres and polypropylene fibres
  - Nestle/Perrier
    - Mineral waters and soft drinks
- Is it sound to do that?

Hedonic Price Analysis

- Price comparison of products with different qualities (changing over time or product space)
- Examples:
  - Are cars more expensive in the UK than in France?
  - Did prices of computers decrease from 2000-3?
- Two steps:
  - Adjust prices accounting for quality differences
  - Compare “standardised” prices

Description

- Example: cars are characterised by their horsepower (X_1), weight (X_2),…
- Data available for different brands (*i*) and for different points in time (*t=1,2,3*)
- Use regression analysis to estimate:

$$\ln(P_{i,t}) = \alpha_1 + \alpha_2 D_{2,i} + \alpha_3 D_{3,i} + \beta_1 X_{1,i} + \beta_2 X_{2,i} + \ldots + \beta_k X_{k,i} + \varepsilon_i$$

where $D_{j,i}$ are dummy variables (i.e. variables that take value 1 or 0), taking value 1 if the observation *i* is in year *j* and 0 otherwise

- Notice that we should never include dummy variables for all years in order to avoid multicolinearity (violation of OLS model)

Description

- Using the estimates on the α's, one can obtain adjusted prices
- What are the relative price changes from year 1 to years 2 and 3?
- "Hedonic price index":
  - If we normalise the base year (year 1) to 1…
  - Then , will be 1 in period 1,
  - $\exp(\alpha_2)$ in year 2 and
  - $\exp(\alpha_3)$ in year 3

  - Example: if $\exp(\alpha_2) = 0.87$, quality-adjusted average prices went down by 13%

Data and Interpretation

- Need combination of cross-section and time series data on prices and characteristics
- Need 20 or 30 observations plus as many observations as the number of regressors
- Need to know demand and supply characteristics of the product
- Regression performance can be measured by the R^2

Application: Car Price Differentials

- Different studies with the following results:
  - EU Commission 89: unadjusted 12% to 50% differential for UK compared to Be, G, Gr, Sp, F, I, L, N
  - Flam and Nordstrom, adjusted differentials at around 12%
  - EU Commission 95: adjusted differentials in excess of 20%

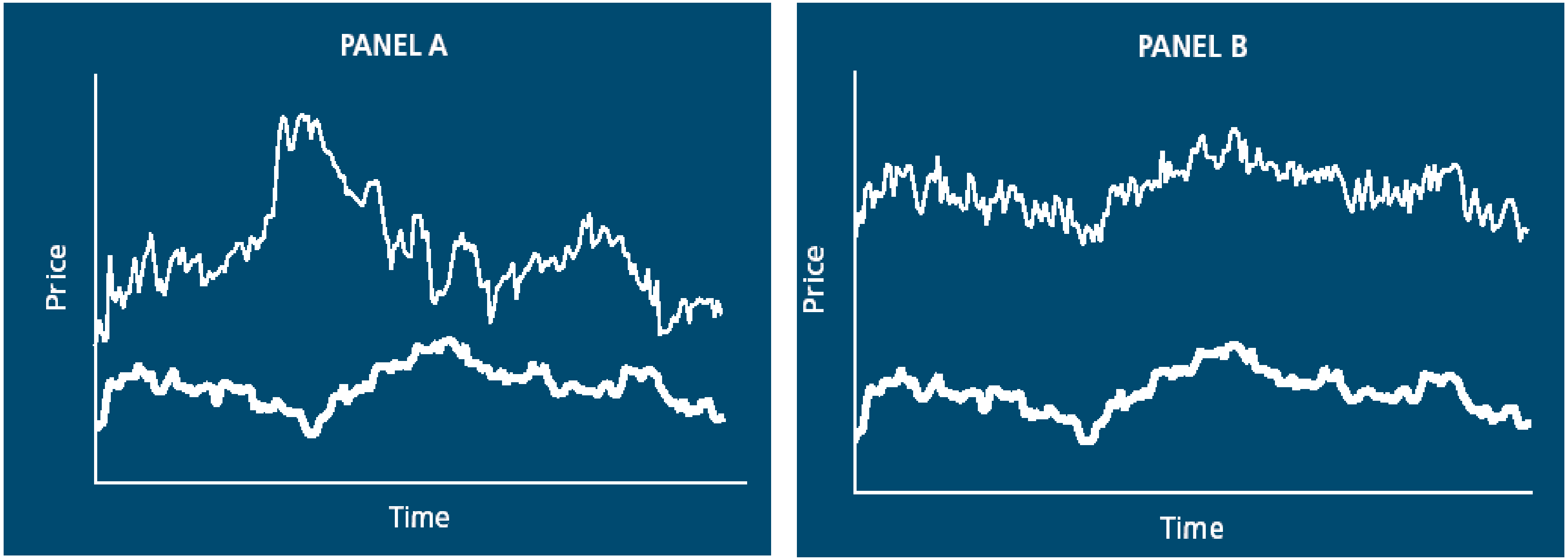
Price Correlation

- Are two products in the same *economic* market?
  - If so, there is a limit as to how far prices can diverge before supply and demand effects kick in
  - Then, prices should be roughly moving together
- Measure of the co-movement: correlation coefficient

$$\rho_{i,j} = \frac{\sigma_{i,j}}{\sigma_i \sigma_j} \text{ where } \sigma_{i,j} \equiv \text{cov}(p_i, p_j) = E[(p_i - \overline{p_i})(p_j - \overline{p_j})]$$

  - $-1 \leq \rho_{i,j} \leq 1$
  - When $\rho_{i,j} = 1$ (or -1), prices are perfectly positively (negatively) correlated
  - When $\rho_{i,j} = 0$, prices are uncorrelated

Example



Source: Lexecon 2005

Estimating standard deviation, covariance and correlation coefficient

Suppose that you have T observations on any two prices, $p_{i,1}\ldots p_{i,T}$ and $p_{j,1}\ldots p_{j,T}$.

Remember that an estimate for the variance is $\hat{\sigma}_i^2 = \frac{1}{T-1}\sum_{t=1}^{T}(p_{i,t}-\overline{p_i})^2$

An estimate for the covariance is $\hat{\sigma}_{i,j} = \frac{1}{T-1}\sum_{t=1}^{T}(p_{i,t}-\overline{p_i})(p_{j,t}-\overline{p_j})$

Therefore an estimate for the correlation coefficient is

$$\hat{\rho}_{i,j} = \frac{\hat{\sigma}_{i,j}}{\hat{\sigma}_i\hat{\sigma}_j}$$

Applications

- BP/E.ON (Aral):
  - Are wholesale petrol and diesel markets wider than Germany? Are they in the same market as ARA (Antwerp, Rotterdam, Amsterdam) region?
  - Price correlation analysis on quoted wholesale prices at different locations: high correlation within Germany and also with ARA region
  - Bundeskartellamt: markets are wider than national

- Nestle/Perrier:
  - EU Commission (initially): high and low minelarised still water are two different markets
  - Nestle: all soft drinks are in the same market (water, colas,…)
  - Price correlation results: high correlation between water prices (sparkling and still) but low or negative between water and soft drinks
  - Commission: market is all bottled water

Data and Interpretation

- Requires at least 20 observations

- Problem 1: How correlated do they need to be to be considered in the same market?
  - Possible solution: benchmarking
  - Example: compare the average correlation of still-sparkling brands with the average of within-still or within-sparkling (accepted to be in the same market)
  - Results: within-still(0.89), within-sparkling(0.94), still-sparkling(0.90)

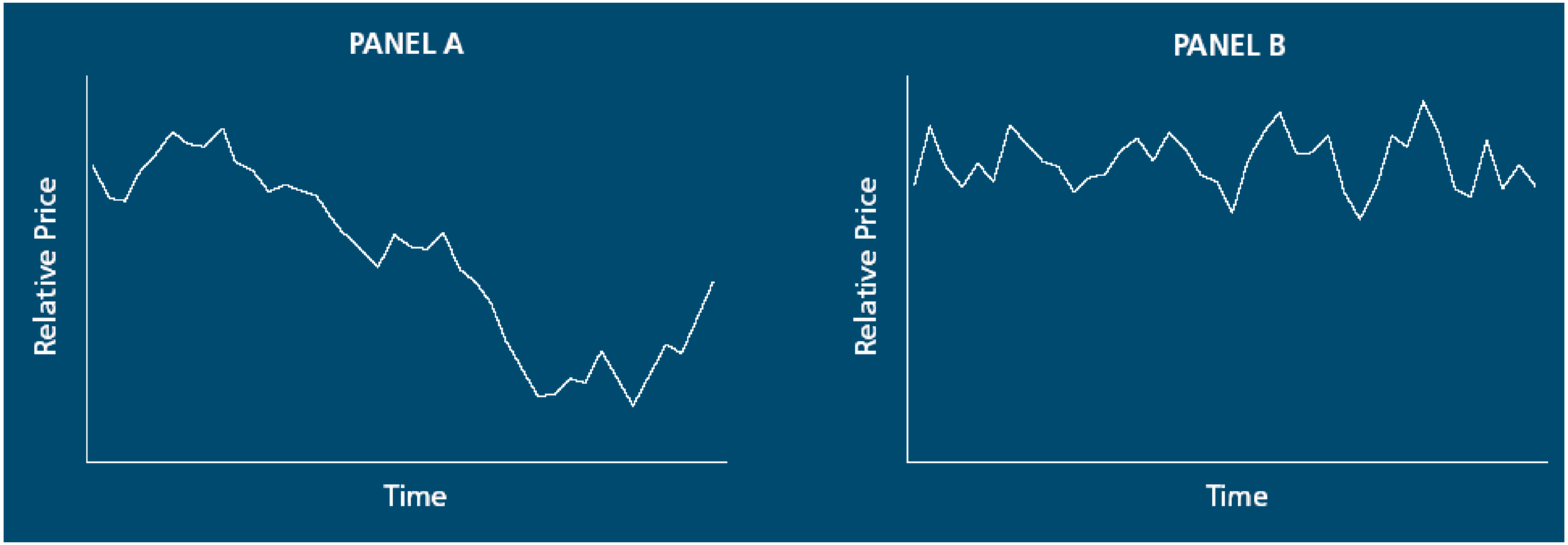
- Problem 2: couldn't it be due to common factors (e.g. input costs, inflation,…)? (spurious correlation)
  - Example: previous wholesale petrol markets may only be correlated because of common use of crude oil
  - Possible solution: purge the data by regressing prices on input prices and use the residuals

Stationarity Analysis

- Same idea as in price correlation analysis
- Again, measure the extent of price levels moving together over time
- Here, use relative prices directly
- Do they tend to return to a stable value over time? If they do, how quickly?

- Informally, one can define a series as stationary if shocks in particular points in time die over time

Example



Source: Lexecon 2005

Data and Interpretation

- Advantages:
  - Not affected by common influences (common costs, inflation,…)
  - Not affected by the choice of base currency
  - Captures delayed responses
- Problems:
  - Need significant variation in price data, and prices themselves should be non-stationary
  - Changes in market structure may cause changes in the relative prices (and even make them non-stationary)

Description: Testing for Stationarity

- Is Y_t stationary?
  - Model: $Y_t = ß_o + ß_1 Y_{t-1} + \varepsilon_t$,
  - If $ß_1<1$ the series is stationary (shocks die out over time)
  - If $ß_1=1$ the series is non-stationary (random walk)
  - If $ß_1>1$ the series is non-stationary
- Dicky Fuller Test: is $ß_1<1$? (available in statistical packages):
  - Null hypothesis: $ß_1=1$ and alternative hypothesis $ß_1<1$
  - Statistic:

$$\left|\frac{\hat{\beta}_1 - 1}{s^2_{\hat{\beta}_1}}\right|$$

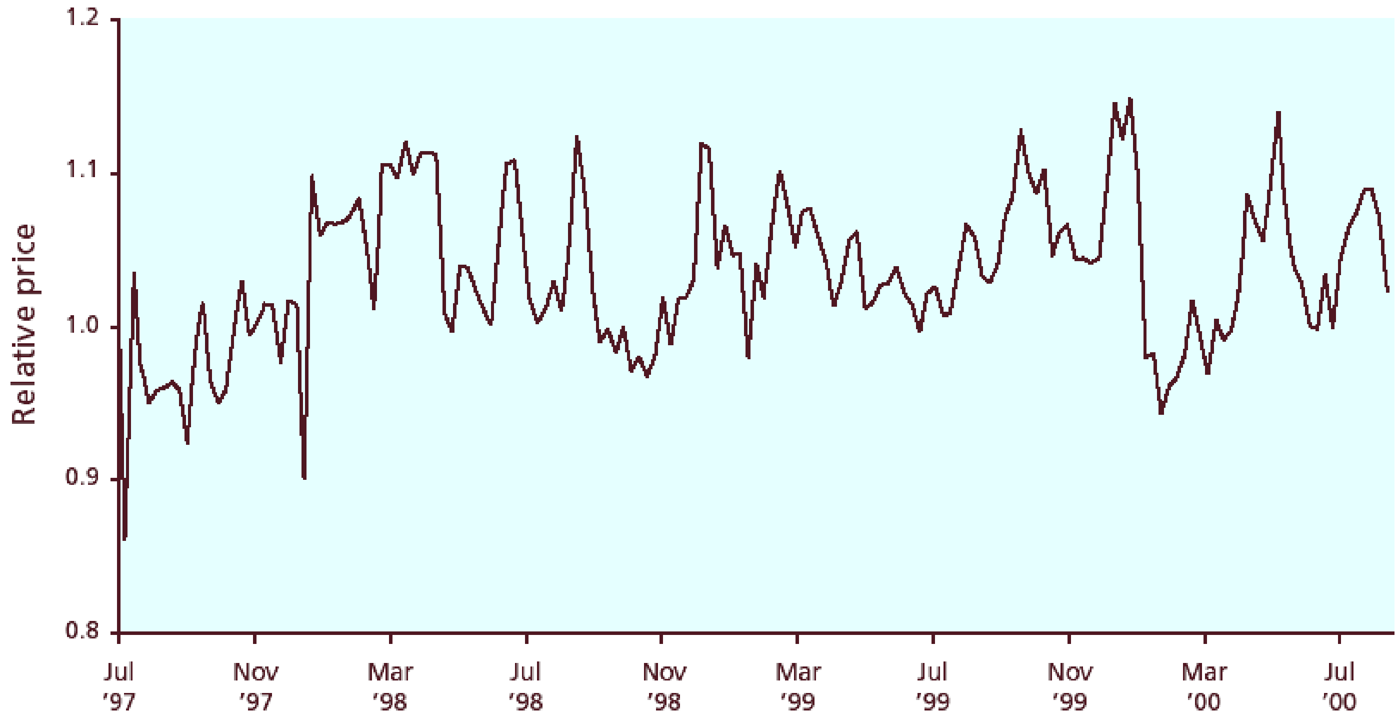
  - But under the null, it does not follow a t, but another distribution (due to non-stationarity of the variables), computed by Dicky and Fuller
  - Proceed as usual
- Augmented Dicky-Fuller (also available)
  - Same idea, controlling for potential correlation of the errors

Applications

- Nutreco Holding NV/Hydro Seafood GSP:
  - Two producers of glutted salmon
  - Product market: are Scottish and Norwegian salmon in the same market?
  - Geographic market: if so, is the market national or EEA-wide?
- Schumann Sasol/Price's Daelite
  - Vertical merger (wax and candle producers)
  - Do imported wax compete with domestic?
  - Stationarity analysis showed that relative prices were stationary

Example

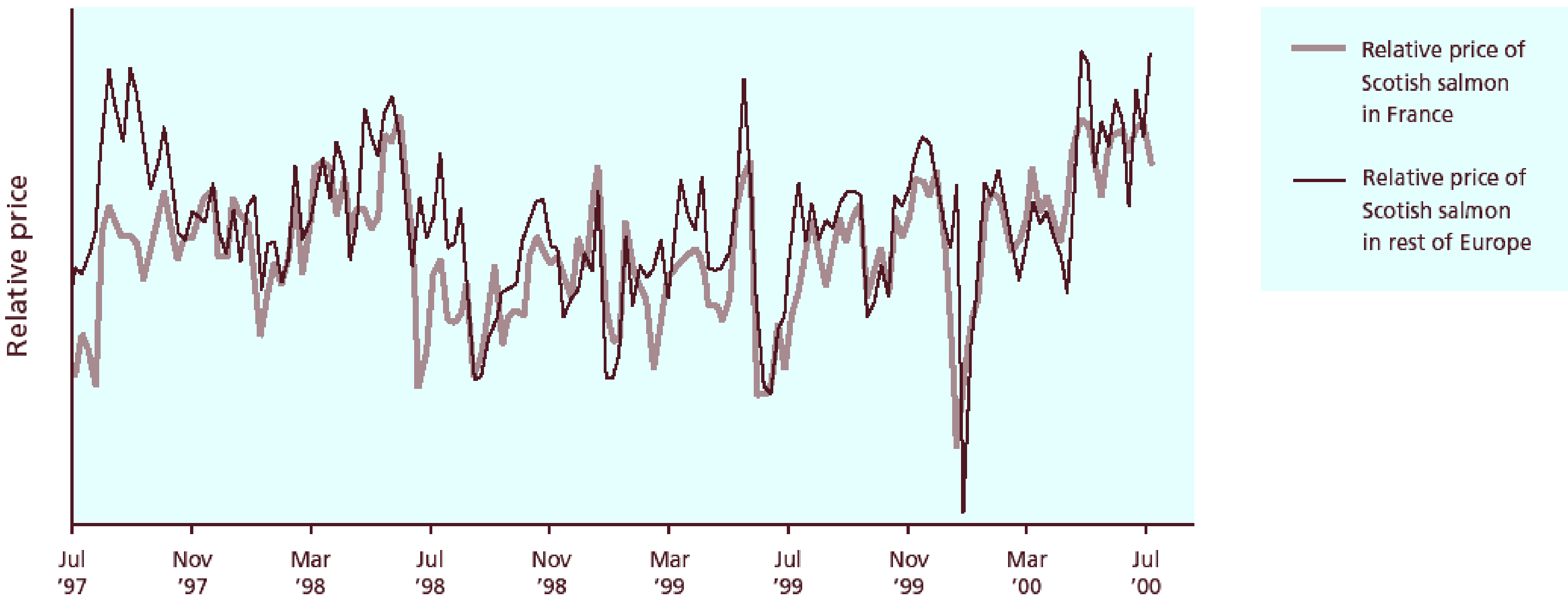
Chart 3.2: Price of Scottish Salmon relative to the price of Norwegian salmon in the UK



Source: Lexecon 2005

Example

Chart 3.3: Prices of Scottish salmon in France and the rest of Europe relative to the UK



Source: Lexecon 2005

Granger Causality Test

- Do a price series "causes" the other? Do they mutually determine each other?
- Two price series, P_1 and P_2, estimate:

$$P_{1,t} = \sum_{s=1}^{T} \beta_s P_{1,t-s} + \sum_{s=1}^{T} \gamma_s P_{2,t-s} + u_t$$

- If past levels of P_2 have no influence on P_1 then γ coefficients are 0

Description

- Estimate the previous equation
- Notice that the choice of lags (T) is arbitrary
- Perform an F-test: are all the γ coefficients statistically equal to 0?
- If they are not, then we say that P_2 Granger-causes P_1

- Same in the other way around: does P_1 Granger-causes P_2?

Data, interpretation and application

- Substantially long time series data needed
- Interpretation:
  - Again a negative result is easier to interpret
  - Need to purge from common factors (e.g. adding them in the previous equation)
  - Need to be careful with autocorrelation (see next chapter): F-test is then not valid
- Application:
  - Determining petrol markets in the US (Margaret Slade,1986)