

ECON 343

**Lecture 4 : Smoothing and
Extrapolation of Time Series**



Jad Chaaban

Spring 2005-2006



Outline – Lecture 4

1. Simple extrapolation models
2. Moving-average models
3. Single Exponential smoothing
4. Double Exponential smoothing
5. Seasonal adjustment
6. Structural breaks



1. Simple extrapolation models

- Assumptions:
 - Forecast the time series on the basis of its past behavior
 - *Deterministic* nature: no randomness used in the models here
- Regress the series against a function of time and/or itself lagged
- Types of models:
 - Linear trend model
 - Quadratic trend model
 - Exponential model
 - Autoregressive trend model



Linear trend model

- Assumptions:
 - The series will increase in constant absolute amounts each period
 - Trend line depends only on time t

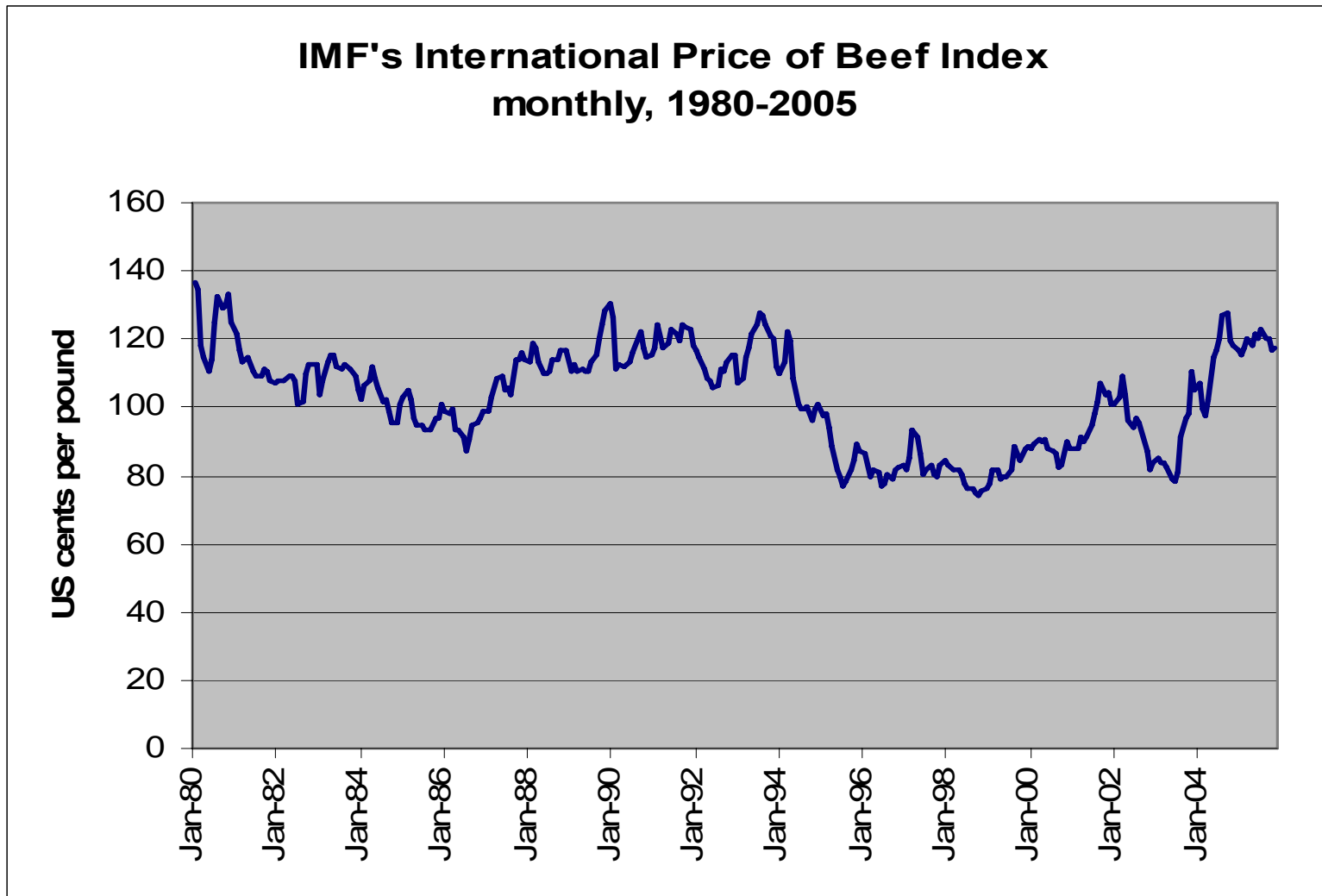
- Formula:

$$Y_t = \alpha + \beta t$$

- t is usually chosen to equal 0 in the base period and to increase by 1 during each successive period
- Simple OLS estimation

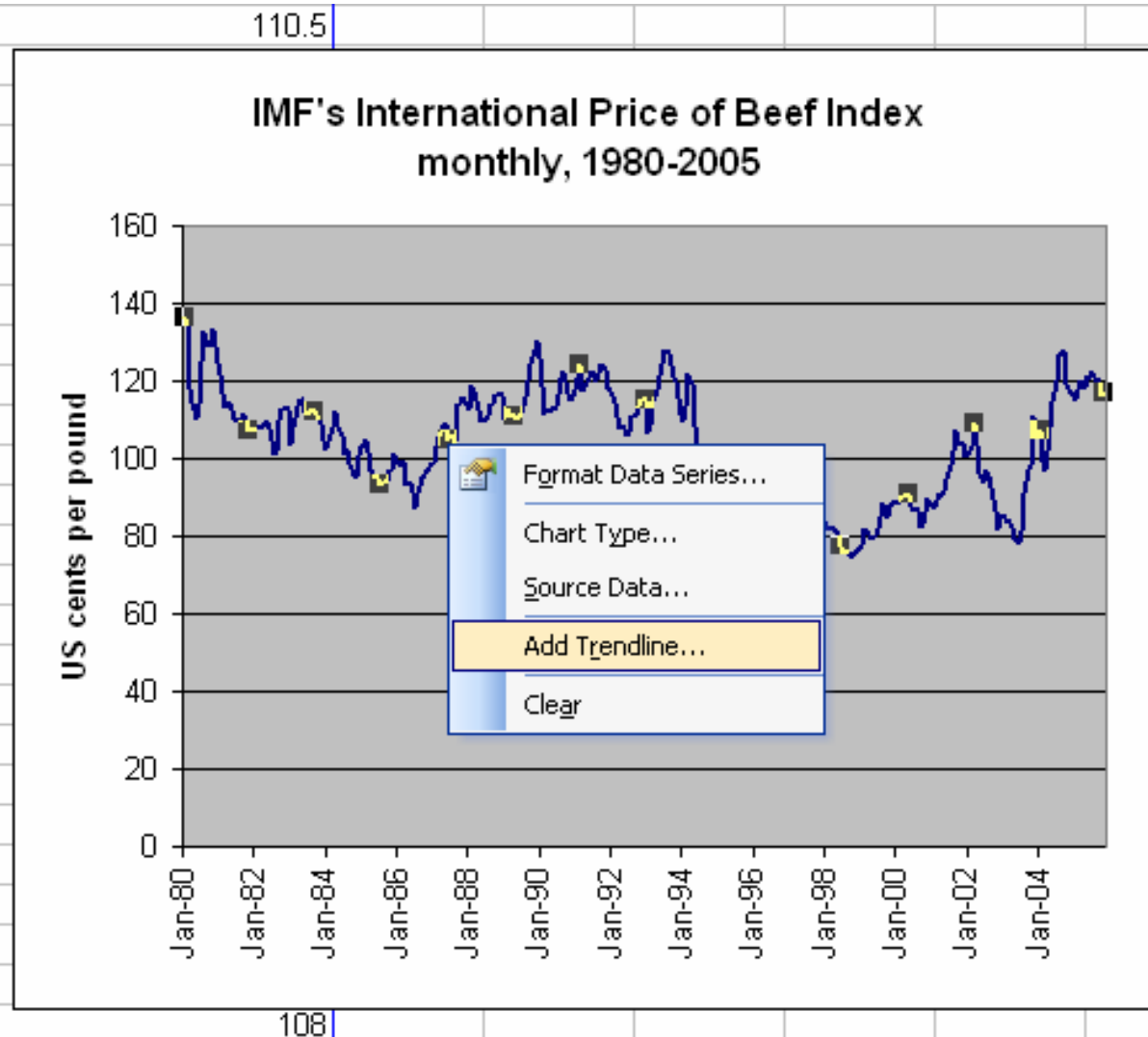


Excel implementation



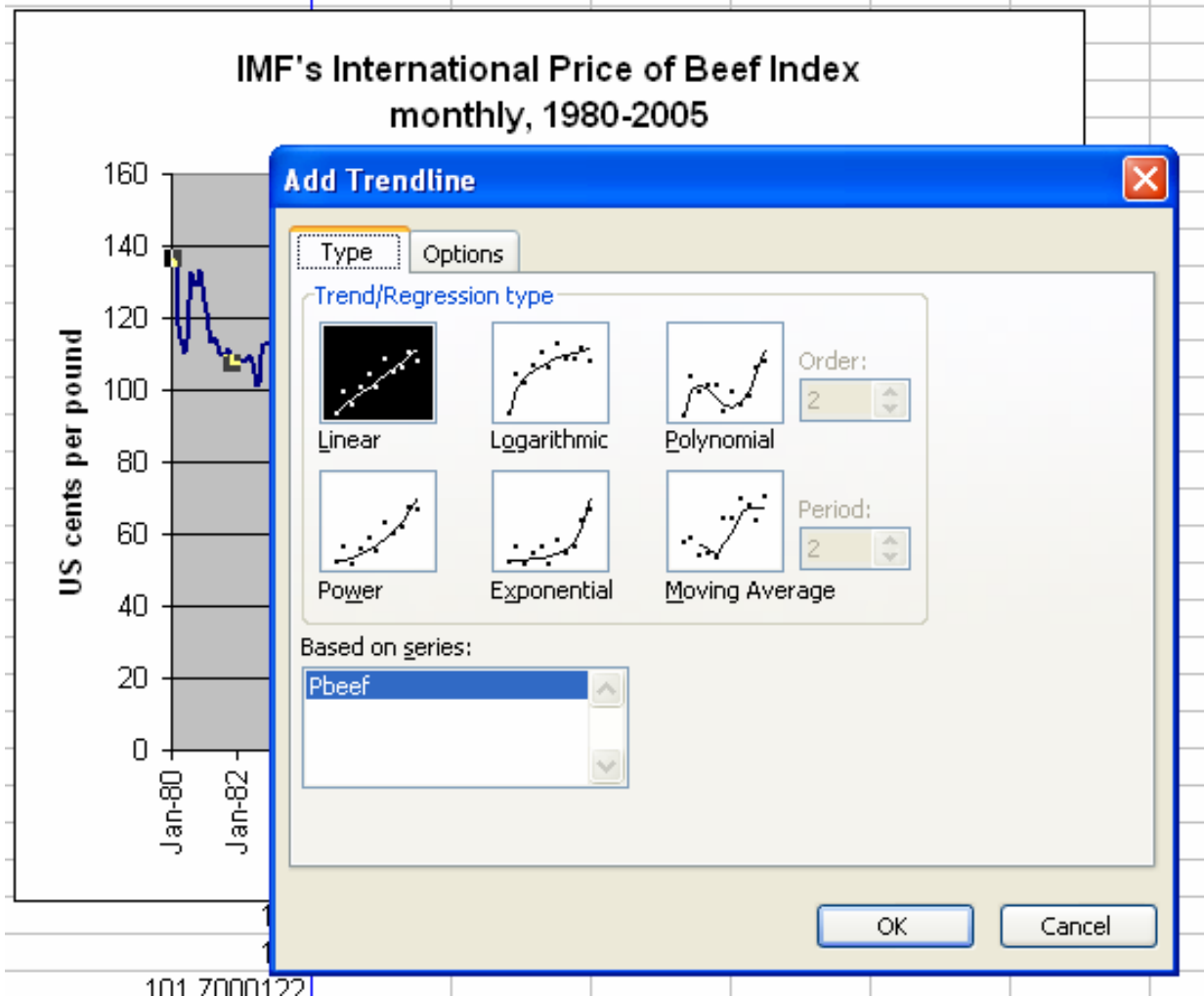


Excel implementation



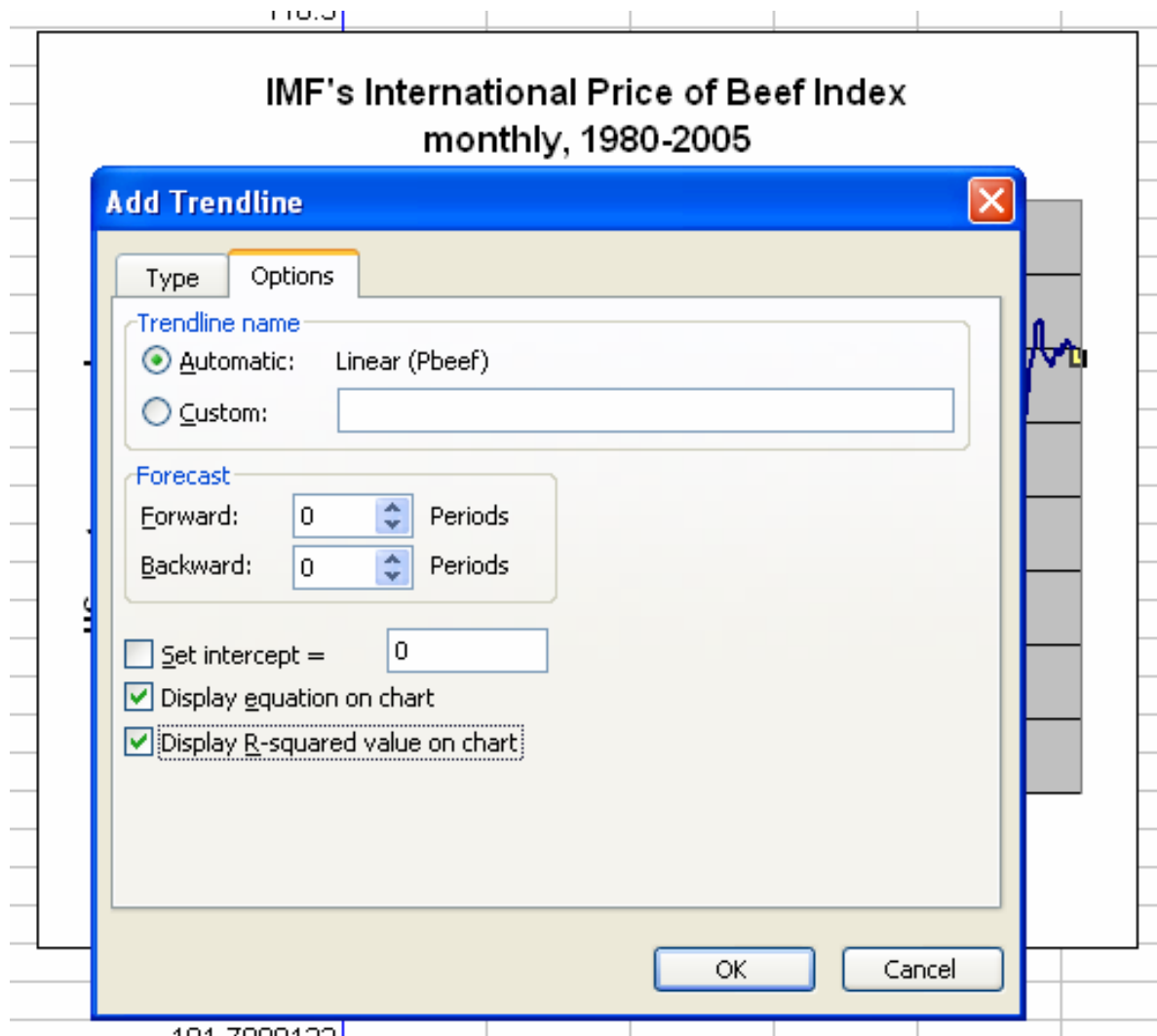


Excel implementation



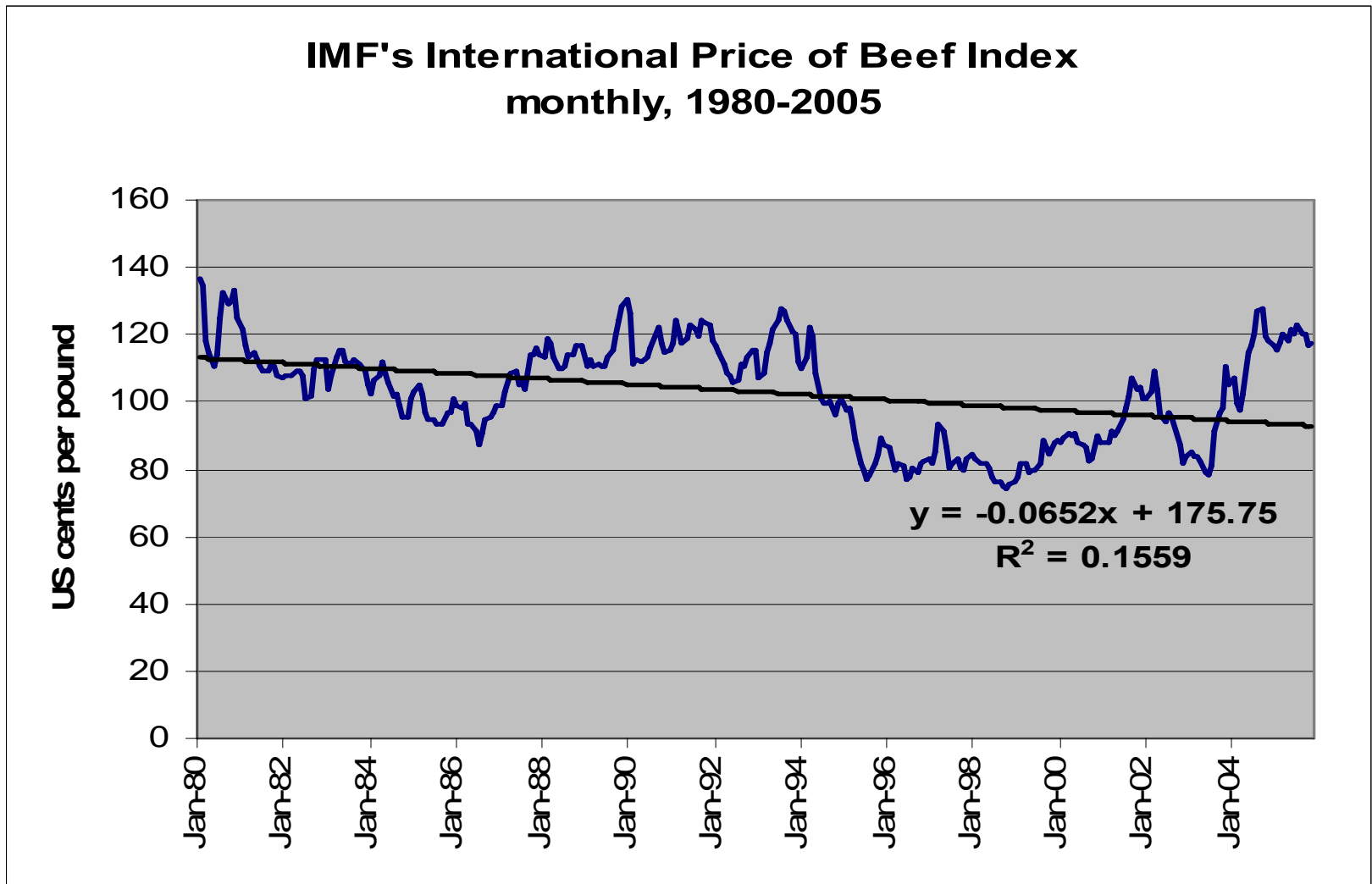


Excel implementation





Excel implementation





Quadratic trend model

- Assumptions:
 - Extends the simple linear trend model to potential non-linearity
 - Add a term t^2

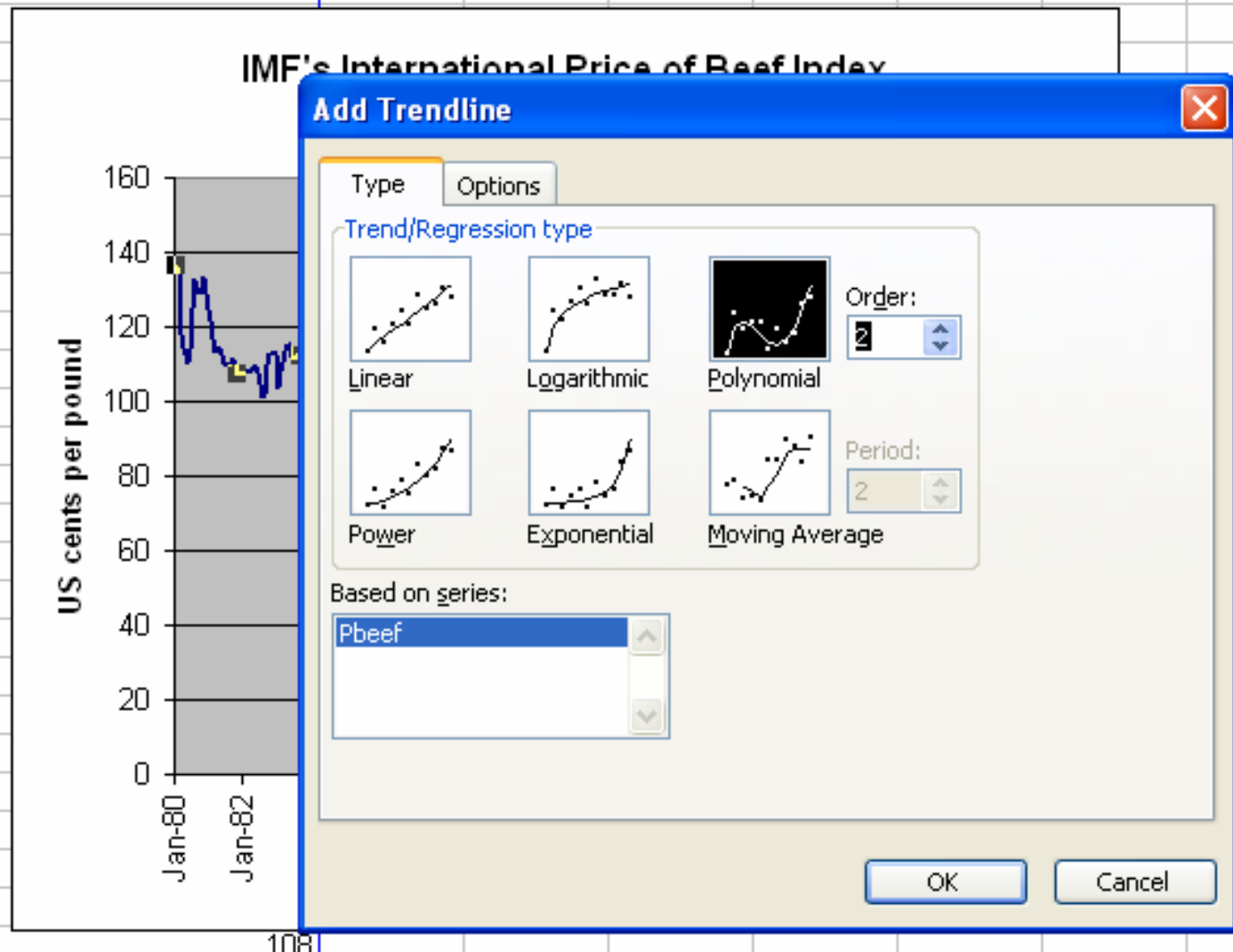
- Formula:

$$Y_t = \alpha + \beta t + \gamma t^2$$

- If β and γ are >0 , then y_t will always increase
- If $\beta <0$ and $\gamma >0$, y_t will first decrease then increase
- If β and γ are <0 , then y_t will always decrease



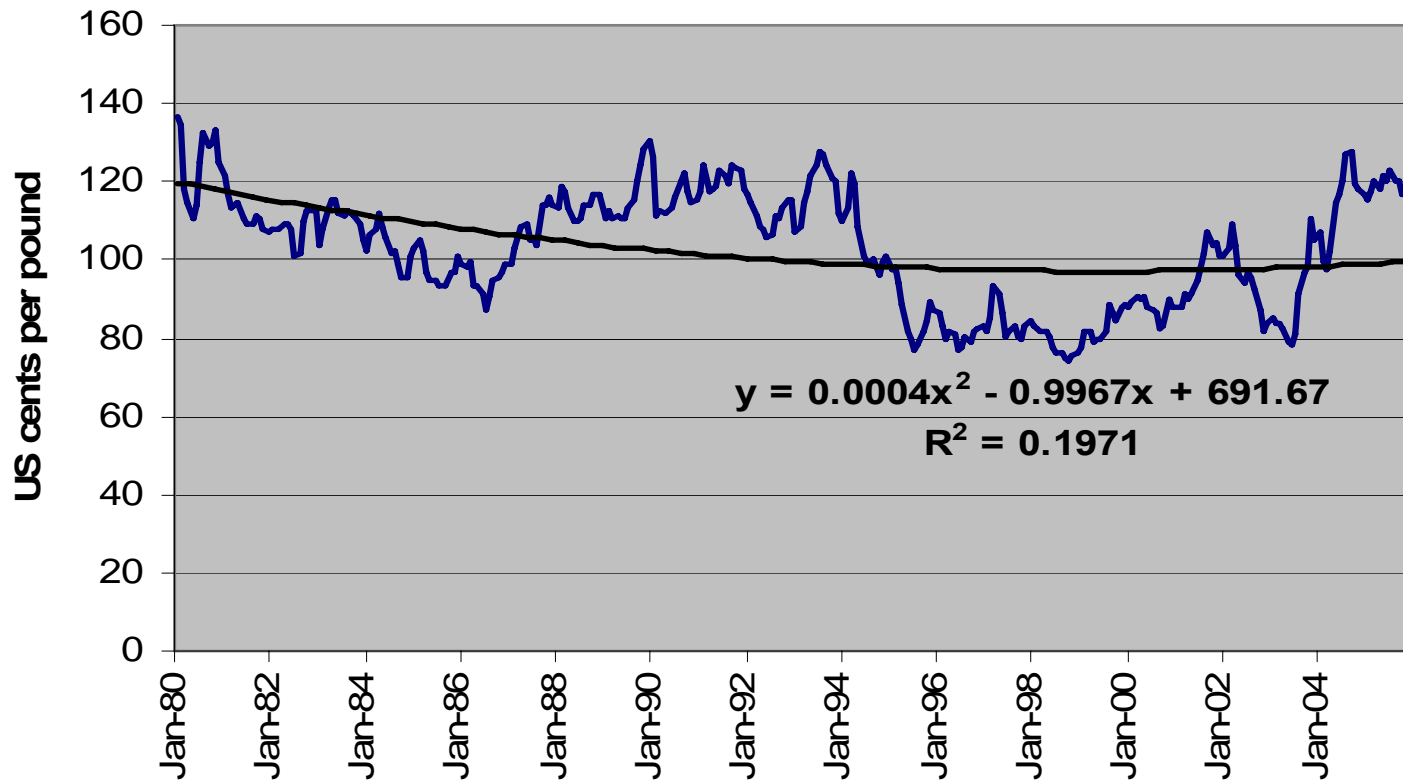
Excel implementation





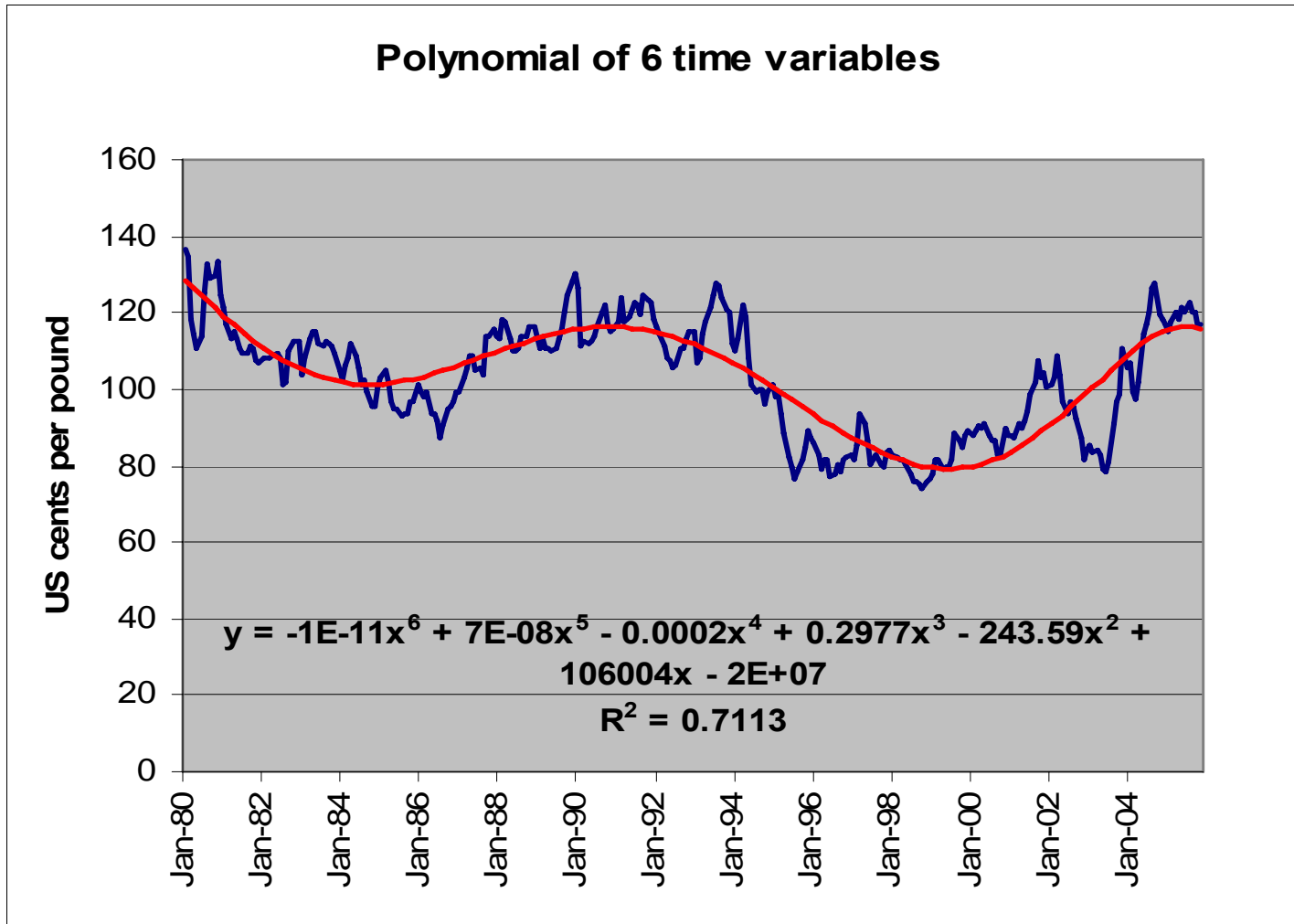
Excel implementation

**IMF's International Price of Beef Index
monthly, 1980-2005**





Excel implementation





Exponential model

- Assumptions:
 - The series will grow with constant percentage increases
 - Use of the exponential function

- Formula:

$$Y_t = Ae^{rt}$$

- The parameters A and r can be estimated by taking the logarithms of both sides of the formula
- Then we fit through OLS the log-linear function:

$$\log(Y_t) = c_1 + c_2 t$$

where $c_1 = \log(A)$ and $c_2 = r$



Excel implementation

110.5

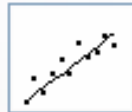
IMF's International Price of Beef Index
monthly 1980-2005

Add Trendline

Type

Options

Trend/Regression type



Linear



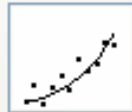
Logarithmic



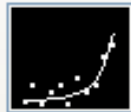
Polynomial

Order:

2



Power



Exponential



Moving Average

Period:

2

Based on series:

Pbeef

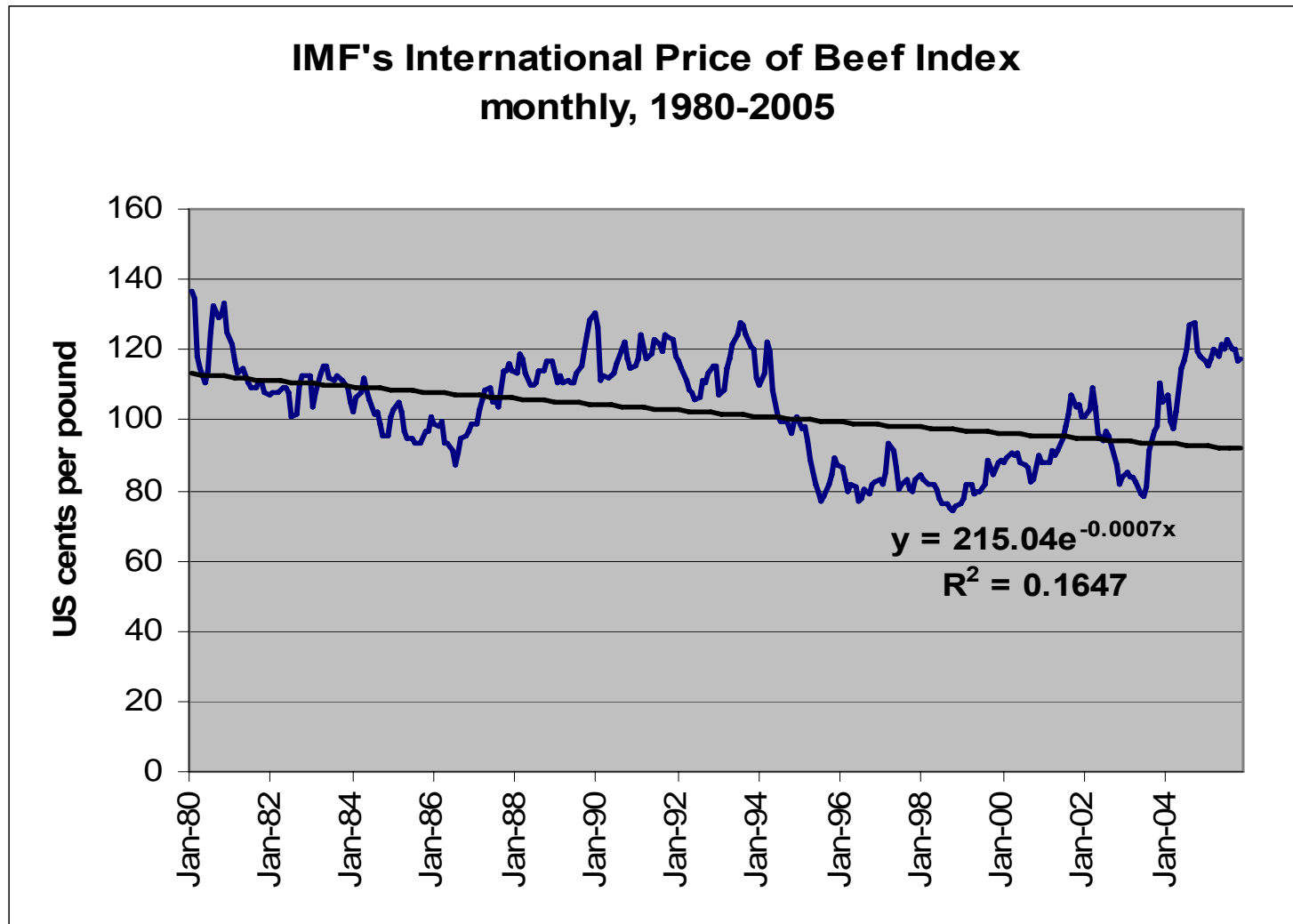
OK

Cancel

101



Excel implementation





Autoregressive trend model

- Assumptions:
 - The series will grow according to its lagged values only

- Formula:

$$Y_t = \alpha + \beta Y_{t-1}$$

- Variation of the model: logarithmic autoregressive trend model:

$$\log(Y_t) = \alpha + \beta \log(Y_{t-1})$$

- If $\alpha=0$ then the value of β is the compounded rate of growth of the series Y



EViews implementation

Workfile Create

Workfile structure type
Dated - regular frequency

Date specification
Frequency: Monthly
Start date: 1/1/1980
End date: 11/1/2005

Names (optional)
WF:
Page:

Irregular Dated and Panel workfiles may be made from Unstructured workfiles by later specifying date and/or other identifier series.

OK Cancel

EViews

File Edit Object View Proc Quick Options Window Help

- Sample...
- Generate Series...
- Show ...
- Graph
- Empty Group (Edit Series)
- Series Statistics
- Group Statistics
- Estimate Equation...
- Estimate VAR...

Workfile: UNTITLED

Group: UNTITLED Monthly

View Proc Object Print Name Freeze Default Sort Transpose Edit+/- Smpl+/- InsDe

obs					
1980M01					
1980M02					
1980M03					
1980M04					
1980M05					
1980M06					
1980M07					
1980M08					
1980M09					
1980M10					



EViews implementation

The screenshot displays the EViews software interface. The main window is titled "Workfile: UNTITLED" and contains a sub-window titled "Group: UNTITLED Workfile: UNTITLEDMonthly". The sub-window has a menu bar with options: View, Proc, Object, Print, Name, Freeze, Default, Sort, Transpose, Edit+/-, Smpl+/-, and InsD. Below the menu bar, a numerical value "136.3599854" is displayed. The main area shows a data table with the following content:

obs	SER01				
1980M01	136.3600				
1980M02	134.5500				
1980M03	118.0000				
1980M04	114.5100				
1980M05	110.5000				
1980M06	113.8900				
1980M07	124.9600				
1980M08	132.5100				
1980M09	129.1500				
1980M10					

A context menu is open over the first row of the table, showing the following options: Copy, Paste (highlighted), Paste Special..., Display format..., Select all (Ctrl+A), Edit +/-, Sample +/-, Transpose+/-, InsDel, Title..., Sort..., and Save table to disk... The bottom status bar shows "Monthly" and "New Page".



EViews implementation

Workfile: BEEF - (f:\econ\beef.wf1)

View Proc Object Print Save Details+/- Show

Range: 1980M01 2005M11 -- 311 obs
Sample: 1980M01 2005M11 -- 311 obs

c
 pbeef
 resid

renamed

**normal
autoregressive**

Equation Estimation

Specification Options

Equation specification
Dependent variable followed by list of regressors including ARMA and PDL terms, OR an explicit equation like $Y=c(1)+c(2)X$.

pbeef c pbeef(-1)

Estimation settings
Method: LS - Least Squares (NLS and ARMA)
Sample: 1980M01 2005M11

OK Cancel

Monthly New Page



EViews implementation

Workfile: BEEF - (f:\econ\beef.wf1)

Equation: UNTITLED Workfile: BEEFMonthly

View Proc Object Print Name Freeze Estimate Forecast Stats Resids

Dependent Variable: PBEEF
Method: Least Squares
Date: 02/19/06 Time: 18:31
Sample (adjusted): 1980M02 2005M11
Included observations: 310 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3.921160	1.481744	2.646314	0.0086
PBEEF(-1)	0.961316	0.014245	67.48573	0.0000

R-squared 0.936656 Mean dependent var 102.8954
Adjusted R-squared 0.936450 S.D. dependent var 14.76059
S.E. of regression 3.721012 Akaike info criterion 5.472299
Sum squared resid 4264.548 Schwarz criterion 5.496406
Log likelihood -846.2064 F-statistic 4554.324
Durbin-Watson stat 1.537813 Prob(F-statistic) 0.000000



EViews implementation

**Logarithmic
autoregressive**

Equation Estimation [X]

Specification | Options

Equation specification
Dependent variable followed by list of regressors including ARMA and PDL terms, OR an explicit equation like $Y=c(1)+c(2)*X$.

log(pbeef) c log(pbeef(-1))

Estimation settings
Method: LS - Least Squares (NLS and ARMA)
Sample: 1980M01 2005M11

OK Cancel



EViews implementation

**Logarithmic
autoregressive**

Equation: UNTITLED Workfile: BEEFMonthly

View Proc Object Print Name Freeze Estimate Forecast Stats Resids

Dependent Variable: LOG(PBEEF)
Method: Least Squares
Date: 02/19/06 Time: 18:40
Sample (adjusted): 1980M02 2005M11
Included observations: 310 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.156398	0.062990	2.482891	0.0136
LOG(PBEEF(-1))	0.966068	0.013617	70.94683	0.0000

R-squared	0.942338	Mean dependent var	4.623059
Adjusted R-squared	0.942151	S.D. dependent var	0.147716
S.E. of regression	0.035528	Akaike info criterion	-3.830534
Sum squared resid	0.388780	Schwarz criterion	-3.806427
Log likelihood	595.7327	F-statistic	5033.453
Durbin-Watson stat	1.527922	Prob(F-statistic)	0.000000



2. Moving Average models

- **Assumptions:**
 - The likely value for our series next month is a simple average of its values over the past 12 months
 - No regression used in the model
- **Moving Average**
 - An n -period moving average is the average value over the previous n time periods. As you move forward in time, the oldest time period is dropped from the analysis.
- **Weighted Moving Average**
 - An n -period weighted moving average allows you to place more weight on more recent time periods by weighting those time periods more heavily.



Moving Average models

- Formulas:

- Moving Average
$$MA = \frac{\sum (y_{t-1}, \dots, y_{t-n})}{n}$$

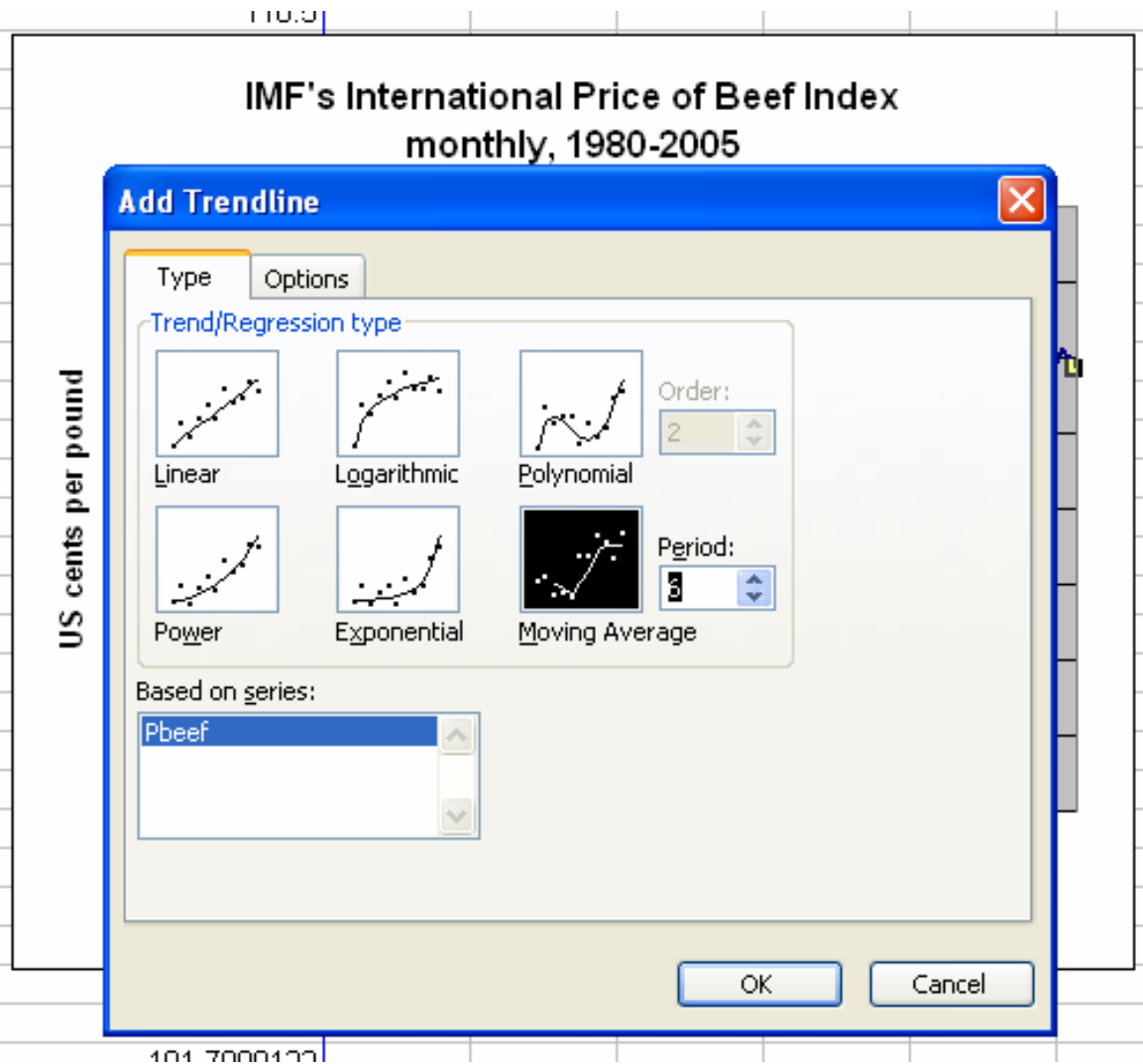
with n time periods

- Weighted Moving Average

$$WMA = \frac{\sum (\alpha_i y_{t-1}, \dots, \alpha_M y_{t-n})}{\sum_m \alpha_i}$$

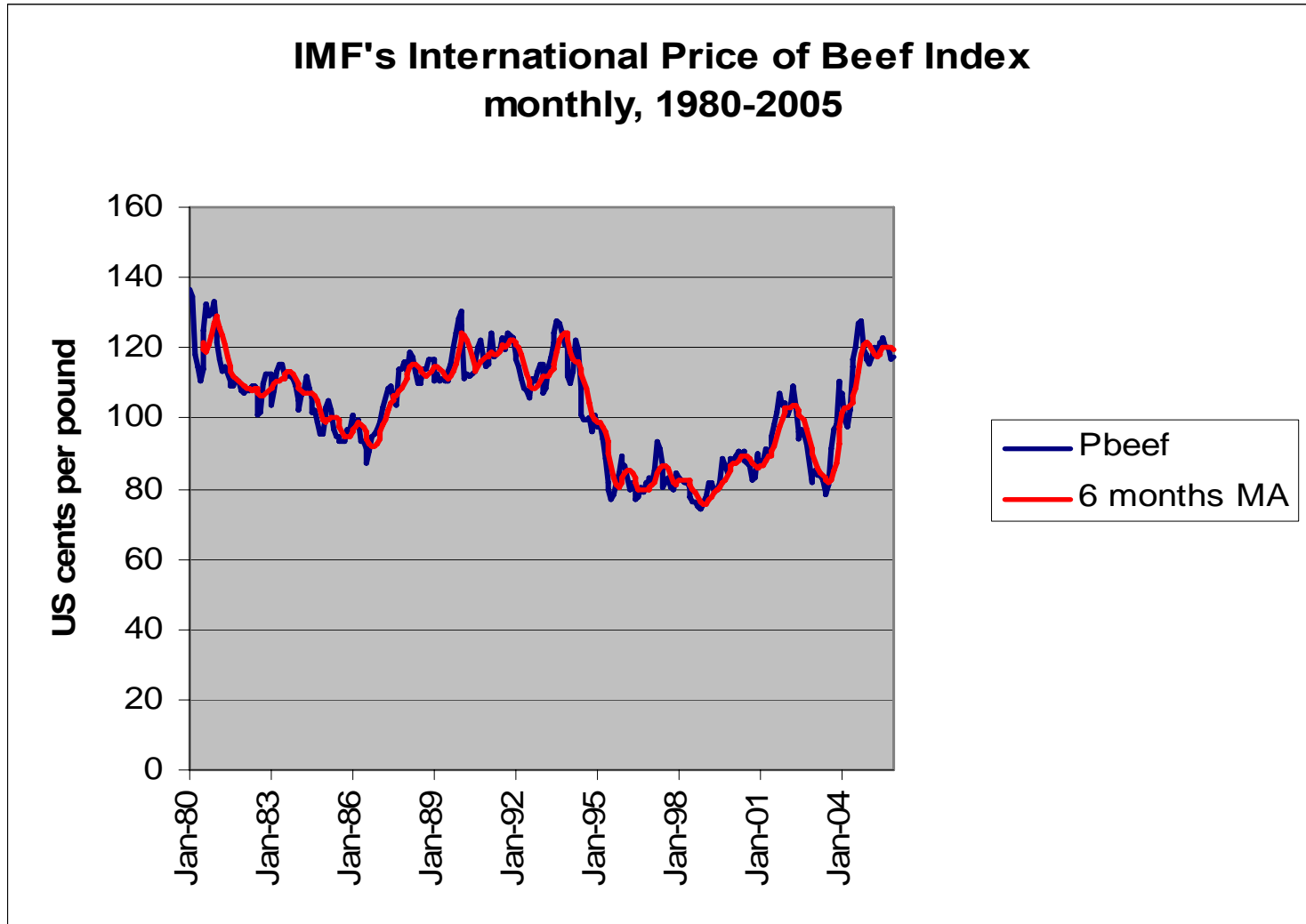


Excel implementation





Excel implementation





3. Single Exponential Smoothing

- This single exponential smoothing method is appropriate for series that move randomly above and below a constant mean with no trend nor seasonal patterns.
- The smoothed series \hat{y}_t of y_t is computed recursively, by evaluating:
$$\hat{y}_t = \alpha y_t + (1 - \alpha)\hat{y}_{t-1}$$

where $0 < \alpha \leq 1$ is the damping (or smoothing) factor. The smaller is the α , the smoother is the \hat{y}_t series. By repeated substitution, we can rewrite the recursion as

$$\hat{y}_t = \alpha \sum_{s=0}^{t-1} (1 - \alpha)^s y_{t-s}$$



Single Exponential Smoothing

- Exponential smoothing-the forecast of y_t is a weighted average of the past values of y_t , where the weights decline exponentially with time.

- The forecasts from single smoothing are constant for all future observations. This constant is given by:

$$\hat{y}_{T+k} = \hat{y}_T \quad \text{for all } k > 0$$

Where T is the end of the estimation sample.

- EViews uses the mean of the initial observations of y_t to start the recursion. Values of α around 0.01 to 0.30 work quite well. You can also let EViews estimate α to minimize the sum of squares of one-step forecast errors



4. Double Exponential Smoothing

- This method applies the single smoothing method twice (using the same parameter)
- It is appropriate for series with a linear trend
- Double smoothing of a series y is defined by the recursions:

$$S_t = \alpha y_t + (1 - \alpha)S_{t-1}$$

$$D_t = \alpha S_t + (1 - \alpha)D_{t-1}$$

where S is the single smoothed series and D is the double smoothed series. Note that double smoothing is a single parameter smoothing method with damping factor $0 < \alpha \leq 1$.



Double Exponential Smoothing

- Forecasts from double smoothing are computed as:

$$\begin{aligned}\hat{y}_{T+k} &= \left(2 + \frac{\alpha k}{1-\alpha}\right)S_T - \left(1 + \frac{\alpha k}{1-\alpha}\right)D_T \\ &= \left(2S_T - D_T + \frac{\alpha}{1-\alpha}(S_T - D_T)k\right)\end{aligned}$$

- This expression shows that forecasts from double smoothing lie on a linear trend with intercept $2S_T - D_T$

and slope $\alpha(S_T - D_T)/(1 - \alpha)$



Eviews Exponential Smoothing

The screenshot displays the EViews software interface. The main menu bar includes 'File', 'Edit', 'Object', 'View', 'Proc', 'Quick', 'Options', 'Window', and 'Help'. The 'Quick' menu is open, showing options: 'Sample...', 'Generate Series...', 'Show ...', 'Graph', 'Empty Group (Edit Series)', 'Series Statistics', 'Group Statistics', 'Estimate Equation...', and 'Estimate VAR...'. The 'Series Statistics' option is highlighted, and its sub-menu is open, showing: 'Histogram and Stats', 'Correlogram...', 'Unit Root Test...', 'Exponential Smoothing...', and 'Hodrick-Prescott Filter...'. The 'Exponential Smoothing...' option is selected. Below the menu, the workfile 'Workfile: BEEF - (f)' is visible, with a sub-menu containing 'View', 'Proc', 'Object', 'Print', and 'S...'. The workfile range is '1980M01 2005M11 -- 311 obs' and the sample is '1980M01 2005M11 -- 311 obs'. The object list shows 'c', 'pbeef', and 'resid' with checkboxes next to them.



Eviews Exponential Smoothing

Exponential Smoothing

Smoothing method — # of params

- Single** 1
- Double 1
- Holt-Winters - No seasonal 2
- Holt-Winters - Additive 3
- Holt-Winters - Multiplicative 3

Smoothing parameters

Alpha: (mean) Enter number between 0 and 1, or E to estimate.

Beta: (trend)

Gamma: (seasonal)

Smoothed series

Series name for smoothed and forecasted values.

Estimation sample

Forecasts begin in period following estimation endpoint.

Cycle for seasonal

OK Cancel

Series: PBEEF Workfile: BEEFMonthly

View Proc Object Properties Print Name Freeze Sample Genr Sheet Stats Ident Line Bar

Date: 02/19/06 Time: 23:47
Sample: 1980M01 2005M11
Included observations: 311
Method: Single Exponential
Original Series: PBEEF
Forecast Series: PBEEFSM

Parameters:	Alpha	0.9990
	Sum of Squared Residuals	5026.981
	Root Mean Squared Error	4.020439

End of Period Levels:	Mean	117.2493
-----------------------	------	----------

A red curved arrow points from the 'OK' button in the 'Exponential Smoothing' dialog box to the 'Series: PBEEF' properties window.



5. Seasonal adjustment

- If y_t is a time series, then its generic element can be expressed as

$$y_t = L_t \times C_t \times S_t \times \varepsilon_t$$

where L_t : the global trend,

C_t : a secular cycle (long term cyclical component),

S_t : the seasonal variation,

ε_t : an irregular component.

- The objective is to eliminate S
- First, estimate $L \times C$ by deriving an average. For monthly data, a 12-month average:

$$\tilde{y}_t = \frac{1}{12} (y_{t+6} + \dots + y_t + y_{t-1} + \dots + y_{t-5})$$

which is free of irregular and seasonal fluctuations



Seasonal adjustment

- Then divide the original data by this estimate:

$$\frac{L \times C \times S \times \varepsilon}{L \times C} = S \times \varepsilon = \frac{y_t}{\tilde{y}_t} = z_t$$

- The next step is to eliminate ε_t

Average the values of $S_t \times \varepsilon_t$ corresponding to the same month

For monthly data, suppose there are 48 months:

$$\tilde{z}_1 = \frac{1}{4} (z_1 + z_{13} + z_{25} + z_{37})$$

$$\tilde{z}_2 = \frac{1}{4} (z_2 + z_{14} + z_{26} + z_{38})$$

.....

$$\tilde{z}_1 = \frac{1}{4} (z_{12} + z_{24} + z_{36} + z_{48})$$



5. Seasonal adjustment

- The sum of these seasonal indices should be 12 in this case
- If not, adjust by a factor
- The final seasonal indices are used to deflate the data as follows:

$$y_1^a = y_1 / \bar{z}_1$$

$$y_2^a = y_2 / \bar{z}_2$$

....

$$y_{13}^a = y_{13} / \bar{z}_1$$

$$y_{14}^a = y_{14} / \bar{z}_2$$

etc .



Eviews seasonal adjustment

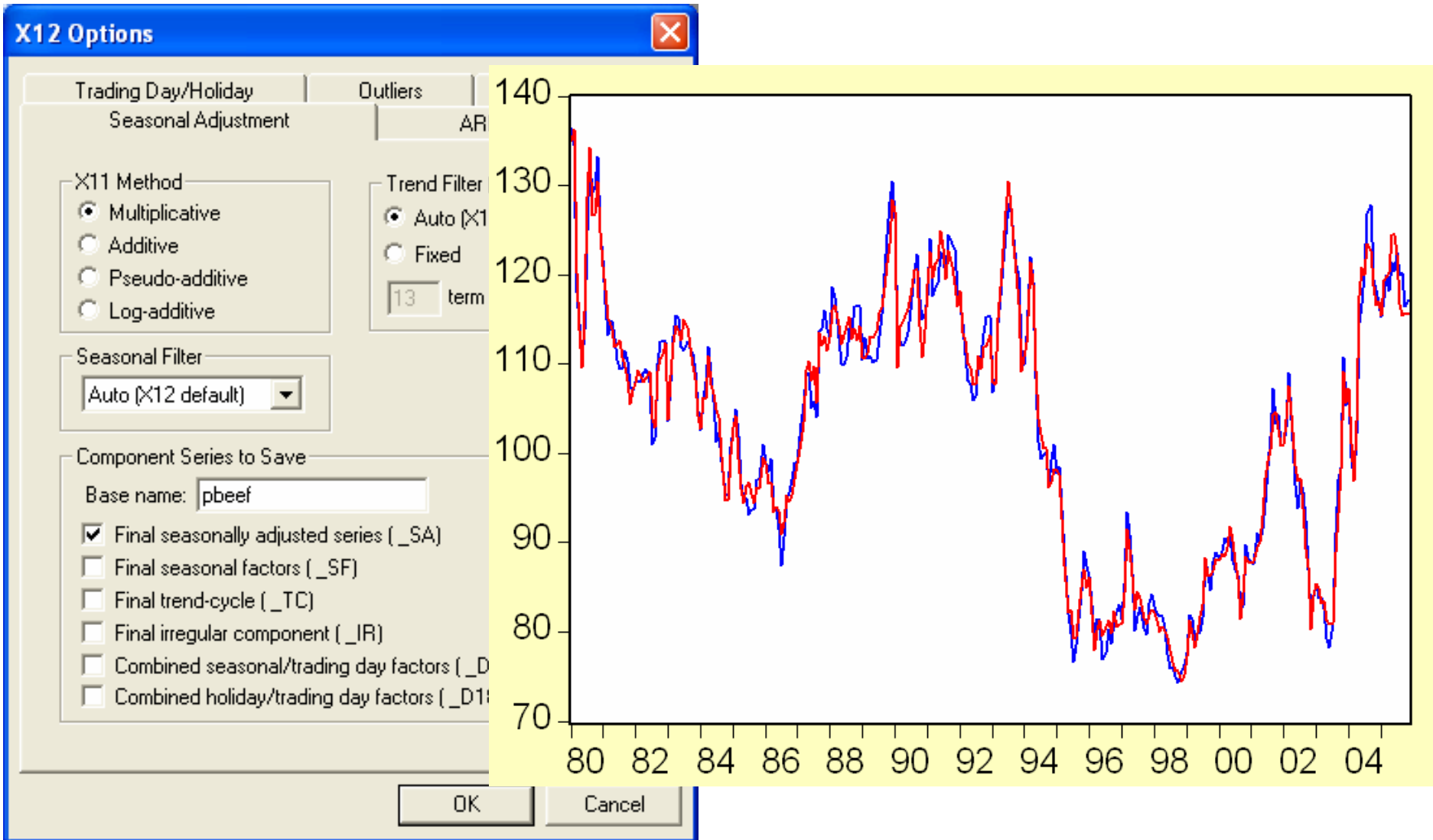
The screenshot shows the EViews software interface. The main window is titled 'EViews' and has a menu bar with 'File', 'Edit', 'Object', 'View', 'Proc', 'Quick', 'Options', 'Window', and 'Help'. The 'Proc' menu is open, showing options: 'Generate by Equation...', 'Resample...', 'Seasonal Adjustment', 'Exponential Smoothing...', 'Hodrick-Prescott Filter...', and 'Frequency Filter...'. The 'Seasonal Adjustment' option is selected, and a sub-menu is open with options: 'Census X12...', 'X11 (Historical)...', 'Tramo/Seats...', and 'Moving Average Methods...'. The 'Census X12...' option is highlighted.

Below the main window, a smaller window titled 'Workfile: BEEF - (f:\econ\beef.wf1)' is open. Inside this window, a sub-window titled 'Series: PBEEF Workfile: BEEFMonthly' is active. The sub-window has a menu bar with 'View', 'Proc', 'Object', 'Properties', 'Print', 'Name', 'Freeze', 'Default', 'Sort', 'Edit+/-', 'Smpl+/-', and 'Label'. The 'View' menu is open, showing a table of data for the series 'PBEEF'. The table has two columns: the first column contains monthly dates from 1980M01 to 1980M08, and the second column contains numerical values. The table is titled 'PBEEF' and has a status bar indicating 'Last updated: 02/19/06 - 23:31'.

PBEEF	
Last updated: 02/19/06 - 23:31	
1980M01	136.3600
1980M02	134.5500
1980M03	118.0000
1980M04	114.5100
1980M05	110.5000
1980M06	113.8900
1980M07	124.9600
1980M08	132.5100

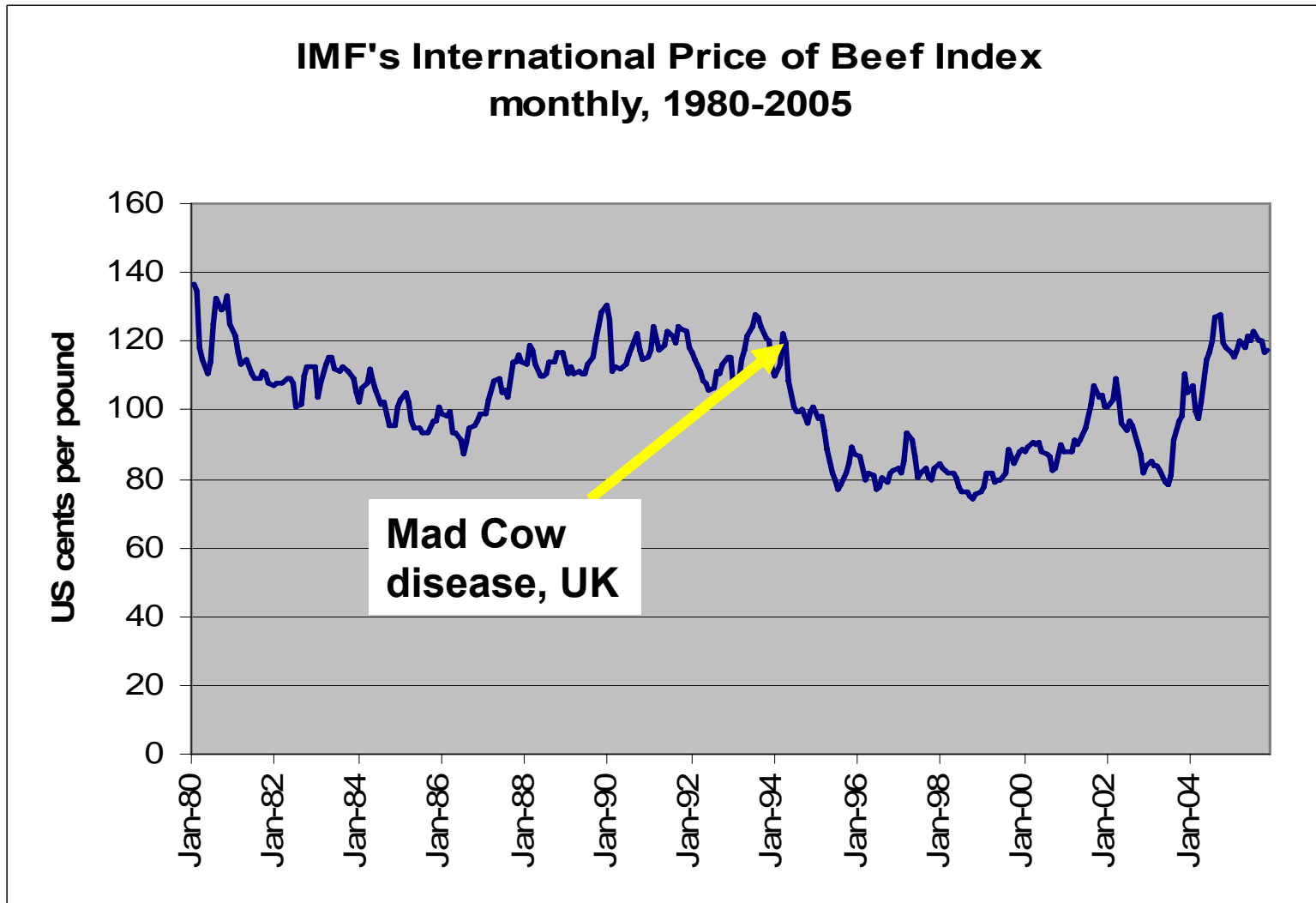


Eviews seasonal adjustment





6. Importance of historical accidents





Incorporating structural breaks

- **Mad Cow timeline:**
 - **Early 1990s:** The British government insists the disease poses no threat to humans
 - **1993:** 120,000 cattle have been diagnosed with Bovine Spongiform Encephalopathy BSE in Britain
 - **May 1995:** Stephen Churchill, 19, becomes the first victim of a new version of Creutzfeldt-Jakob Disease (vCJD). His is one of three vCJD deaths in 1995
 - **April 2002:** First confirmed case of vCJD appears in the US, in a 22-year-old British woman living in Florida
- **Structural break: end of 1994**





Dummy variable

Generate Series by Equation ✕

Enter equation

dum=0

Sample

1980M01 1994M12

OK Cancel

before

Generate Series by Equation ✕

Enter equation

dum=1

Sample

1995M01 2005M11

OK Cancel

after



Eviews estimation

Equation Estimation ✕

Specification | Options

Equation specification
Dependent variable followed by list of regressors including ARMA and PDL terms. OR an explicit equation like $Y=c(1)+c(2)*X$.

pbeef c @trend @trend*@trend dum

cons **t** **t²**

Estimation settings
Method: LS - Least Squares (NLS and ARMA)
Sample: 1980M01 2005M11

OK Cancel



Eviews estimation

Equation: UNTITLED Workfile: BEEFMonthly

View Proc Object Print Name Freeze Estimate Forecast Stats Resids

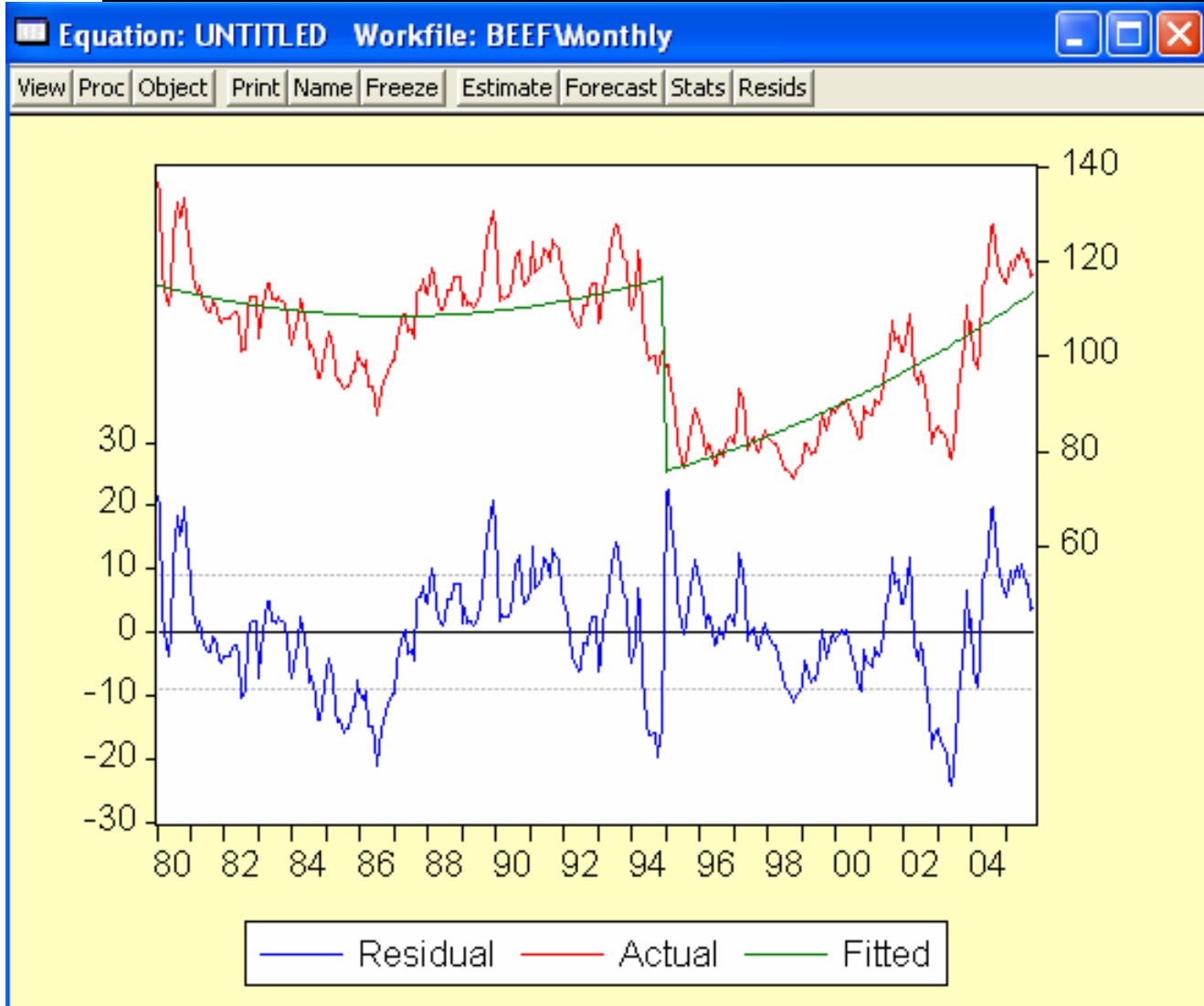
Dependent Variable: PBEEF
Method: Least Squares
Date: 02/19/06 Time: 20:28
Sample: 1980M01 2005M11
Included observations: 311

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	114.9830	1.539961	74.66617	0.0000
@TREND	-0.153494	0.022748	-6.747514	0.0000
@TREND*@TREND	0.000904	7.51E-05	12.04327	0.0000
DUM	-40.83223	2.115033	-19.30571	0.0000

R-squared	0.637372	Mean dependent var	103.0030
Adjusted R-squared	0.633828	S.D. dependent var	14.85844
S.E. of regression	8.991161	Akaike info criterion	7.243139
Sum squared resid	24818.18	Schwarz criterion	7.291240
Log likelihood	-1122.308	F-statistic	179.8655
Durbin-Watson stat	0.231766	Prob(F-statistic)	0.000000



Eviews estimation graph





Comparison

