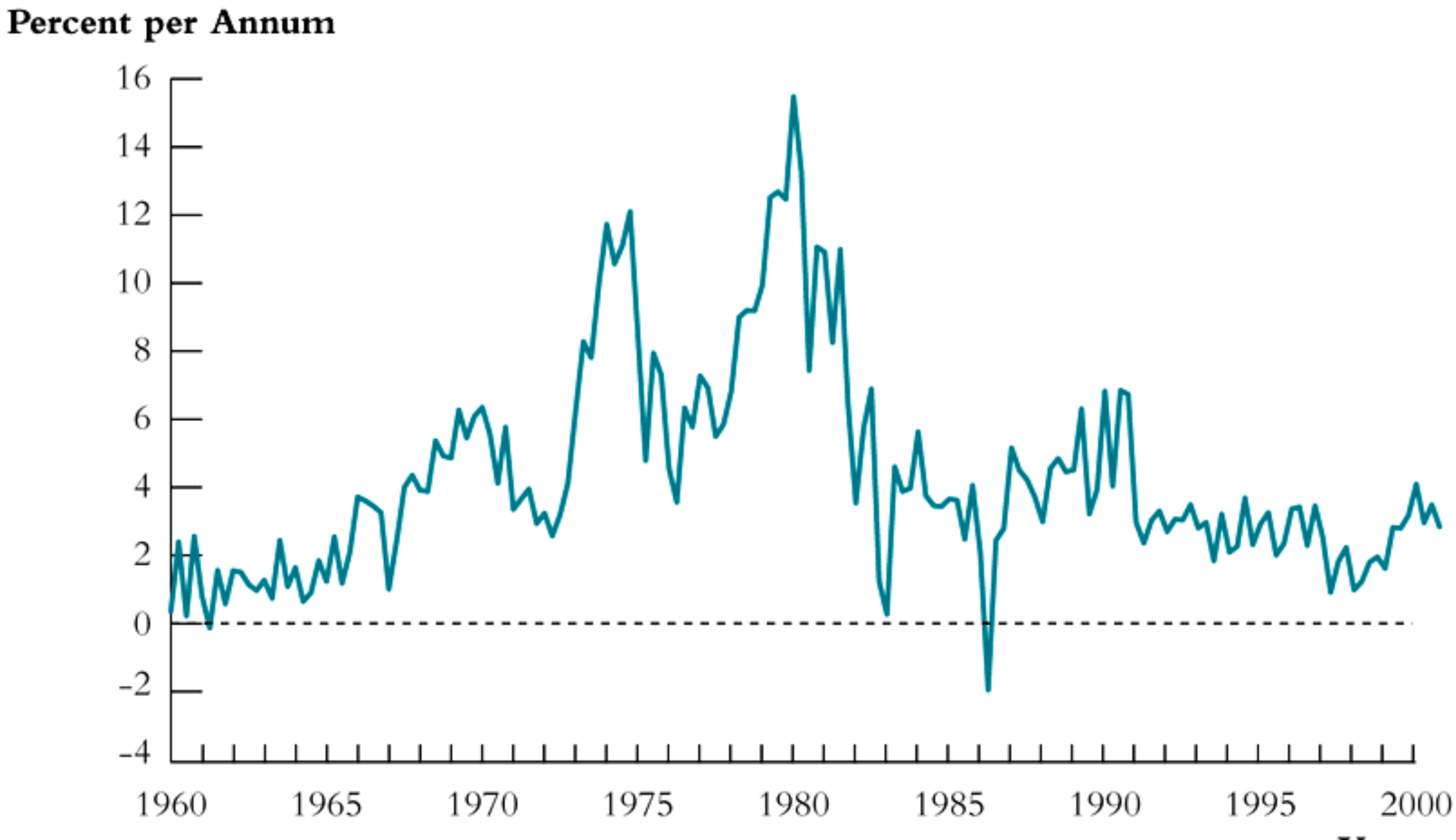
Introduction to Time Series Regression and Forecasting
(SW Chapter 14)

Time series data are data collected on the same observational unit at multiple time periods

- Aggregate consumption and GDP for a country (for example, 20 years of quarterly observations = 80 observations)
- Yen/$, pound/$ and Euro/$ exchange rates (daily data for 1 year = 365 observations)
- Cigarette consumption per capita for a state

Example #1 of time series data: US rate of inflation

FIGURE 12.1 **Inflation and Unemployment in the United States, 1960–1999**

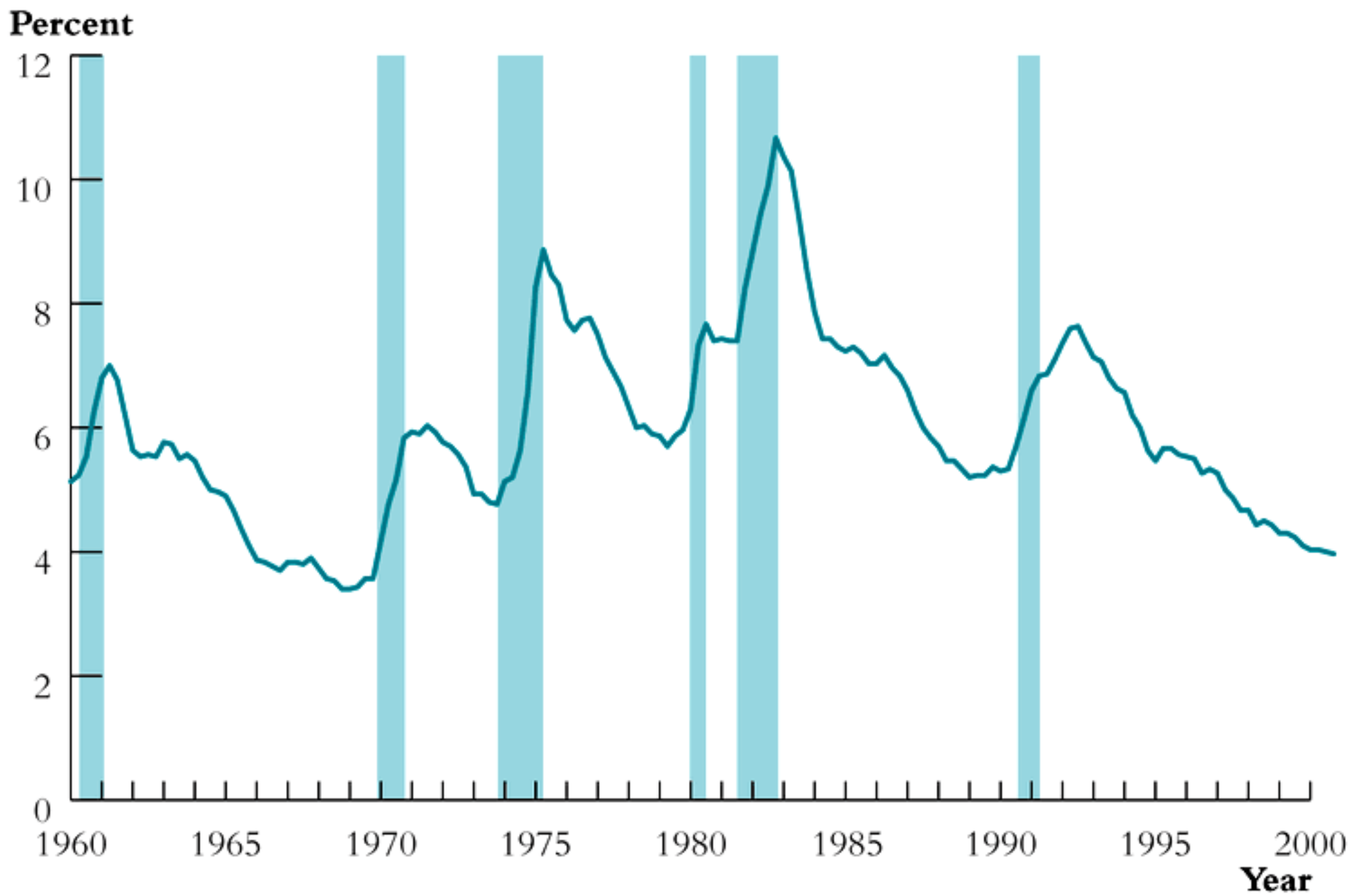


(a) U.S. CPI Inflation Rate

Price inflation in the United States (Figure 12.1a) drifted upwards from 1960 until 1980, and then fell sharply during the early 1980s.The unemployment rate in the United States (Figure 12.1b) rises during recessions (the shaded episodes) and falls during expansions.

Example #2: US rate of unemployment

FIGURE 12.1 **Inflation and Unemployment in the United States, 1960–1999**



(b) U.S. Unemployment Rate

Price inflation in the United States (Figure 12.1a) drifted upwards from 1960 until 1980, and then fell sharply during the early 1980s. The unemployment rate in the United States (Figure 12.1b) rises during recessions (the shaded episodes) and falls during expansions.

Why use time series data?

- To develop forecasting models
  - What will the rate of inflation be next year?
- To estimate *dynamic* causal effects
  - If the Central Bank increases its policy rate now, what will be the effect on the rates of inflation and unemployment in 3 months? in 12 months?
  - What is the effect *over time* on cigarette consumption of a hike in the cigarette tax?
- Plus, sometimes you don't have any choice…
  - Rates of inflation and unemployment in a country can be observed only over time

Time series data raises new technical issues

- Data is ***not i.i.d.***
- Correlation over time (*serial correlation* or *autocorrelation*)
- Trends, cycles, shifts (breaks), and how to model them
- Forecasting models that have no causal interpretation (specialized tools for forecasting):
    - *autoregressive* (AR) models
    - *autoregressive distributed lag* (ADL) models
- Conditions under which dynamic effects can be estimated, and how to estimate them

Using Regression Models for Forecasting

- Forecasting and estimation of causal effects are quite different objectives.
- For forecasting,
  - $\bar{R}^2$ matters (a lot!)
  - Omitted variable bias isn't a problem!
  - We will not worry about interpreting coefficients in forecasting models
  - External validity is paramount: the model estimated using historical data must hold into the (near) future

Introduction to Time Series Data and Serial Correlation

First we must introduce some notation and terminology.

Notation for time series data

- Y_t = value of Y in period t.
- Data set: $Y_1,\ldots,Y_T$ = T observations on the time series random variable Y
- We consider only consecutive, evenly-spaced observations (for example, monthly, 1960 to 2017, no missing months) (else yet more complications...)

We will transform time series variables using lags, first differences, logarithms, & growth rates

Lags, First Differences, Logarithms, and Growth Rates

- The first lag of a time series Y_t is Y_{t-1}; its j^{th} lag is Y_{t-j}.
- The first difference of a series, ΔY_t, is its change between periods $t-1$ and t, that is, $\Delta Y_t = Y_t - Y_{t-1}$.
- The first difference of the logarithm of Y_t is $\Delta \ln(Y_t) = \ln(Y_t) - \ln(Y_{t-1})$.
- The percentage change of a time series Y_t between periods $t-1$ and t is approximately $100\Delta\ln(Y_t)$, where the approximation is most accurate when the percentage change is small.

Example: Quarterly rate of inflation at an annual rate

- CPI in the first quarter of 1999 (1999:I) = 164.87
- CPI in the second quarter of 1999 (1999:II) = 166.03
- Percentage change in CPI, 1999:I to 1999:II

$$= 100 \times \left(\frac{166.03 - 164.87}{164.87} \right) = 100 \times \left(\frac{1.16}{164.87} \right) = 0.703\%$$

- Percentage change in CPI, 1999:I to 1999:II, *at an annual rate* = 4×0.703 = 2.81% (percent per year)
- Like interest rates, inflation rates are (as a matter of convention) reported at an annual rate.
- Using the logarithmic approximation to percent changes yields 4×100× [log(166.03) – log(164.87)] = 2.80%

Example: US CPI inflation – its first lag and its change

CPI = Consumer price index (Bureau of Labor Statistics)

TABLE 12.1 Inflation in the United States in 1999 and the First Quarter of 2000

Quarter	U.S. CPI	Rate of Inflation at an Annual Rate (Inf_t)	First Lag (Inf_{t-1})	Change in Inflation (ΔInf_t)
1999:I	164.87	1.6	2.0	–0.4
1999:II	166.03	2.8	1.6	1.2
1999:III	167.20	2.8	2.8	0.0
1999:IV	168.53	3.2	2.8	0.4
2000:I	170.27	4.1	3.2	0.9

The annualized rate of inflation is the percentage change in the CPI from the previous quarter to the current quarter, times four. The first lag of inflation is its value in the previous quarter, and the change in inflation is the current inflation rate minus its first lag. All entries are rounded to the nearest decimal.

Autocorrelation

The correlation of a series with its own lagged values is called ***autocorrelation*** or ***serial correlation***.

- The first autocorrelation of Y_t is $\text{corr}(Y_t, Y_{t-1})$
- The first ***autocovariance*** of Y_t is $\text{cov}(Y_t, Y_{t-1})$
- Thus

$$\text{corr}(Y_t, Y_{t-1}) = \frac{\text{cov}(Y_t, Y_{t-1})}{\sqrt{\text{var}(Y_t)\,\text{var}(Y_{t-1})}} = \rho_1$$

- These are population correlations – they describe the population joint distribution of (Y_t, Y_{t-1})

Autocorrelation (Serial Correlation) and Autocovariance

The j^{th} autocovariance of a series Y_t is the covariance between Y_t and its j^{th} lag, Y_{t-j}, and the j^{th} autocorrelation coefficient is the correlation between Y_t and Y_{t-j}. That is,

$$j^{\text{th}} \text{ autocovariance} = \operatorname{cov}(Y_t, Y_{t-j}) \tag{12.3}$$

$$j^{\text{th}} \text{ autocorrelation} = \rho_j = \operatorname{corr}(Y_t, Y_{t-j}) = \frac{\operatorname{cov}(Y_t, Y_{t-j})}{\sqrt{\operatorname{var}(Y_t)\operatorname{var}(Y_{t-j})}}. \tag{12.4}$$

The j^{th} autocorrelation coefficient is sometimes called the j^{th} serial correlation coefficient.

Sample autocorrelations

The ***j*^th^ *sample autocorrelation*** is an estimate of the j^{th} population autocorrelation:

$$\hat{\rho}_j = \frac{\widehat{\text{cov}(Y_t, Y_{t-j})}}{\widehat{\text{var}(Y_t)}}$$

where

$$\widehat{\text{cov}(Y_t, Y_{t-j})} = \frac{1}{T-j-1}\sum_{t=j+1}^{T}(Y_t - \bar{Y}_{j+1,T})(Y_{t-j} - \bar{Y}_{1,T-j})$$

where $\bar{Y}_{j+1,T}$ is the sample average of Y_t computed over observations $t = j+1,\ldots,T$

- *Note*: the summation is over $t=j+1$ to T (*why*)?

Example: Autocorrelations of:

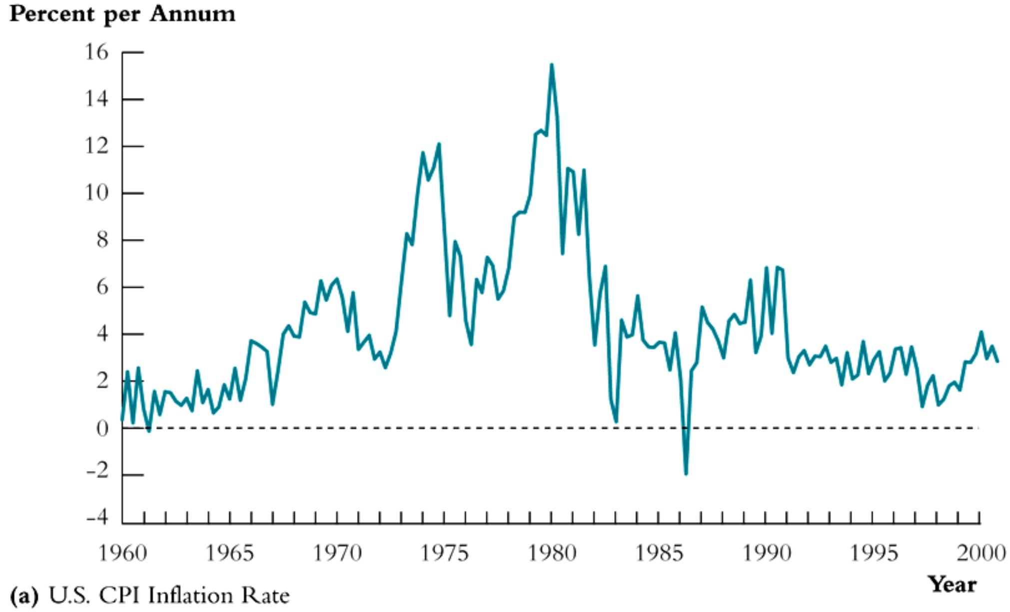
(1) the quarterly rate of U.S. inflation

(2) the quarter-to-quarter change in the quarterly rate of inflation

TABLE 12.2 First Four Sample Autocorrelations of the U.S. Inflation Rate and Its Change, 1960:I–1999:IV

	Autocorrelation of:	
Lag	**Inflation Rate (Inf_t)**	**Change of Inflation Rate (ΔInf_t)**
1	0.85	–0.24
2	0.77	–0.27
3	0.77	0.32
4	0.68	–0.06

FIGURE 12.1 Inflation and Unemployment in the United States, 1960–1999



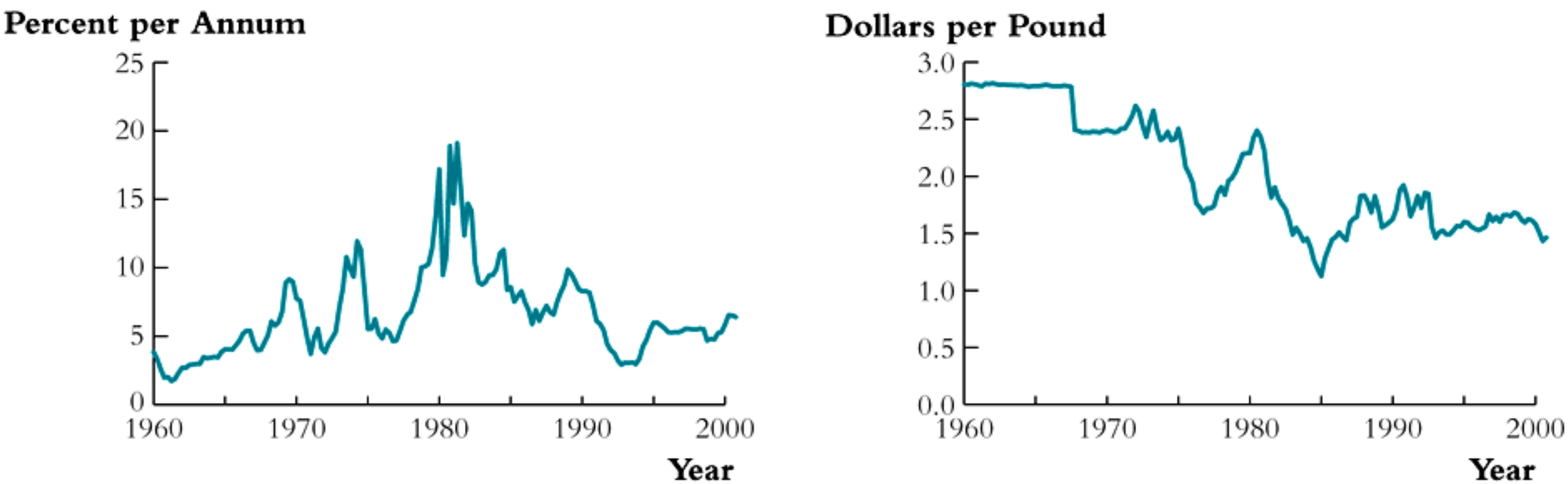
(a) U.S. CPI Inflation Rate

Price inflation in the United States (Figure 12.1a) drifted upwards from 1960 until 1980, and then fell sharply during the early 1980s.The unemployment rate in the United States (Figure 12.1b) rises during recessions (the shaded episodes) and falls during expansions.

- The inflation rate is highly serially correlated ($\rho_1 = .85$)
- Last quarter's inflation rate contains much information about this quarter's inflation rate
- The plot is dominated by multiyear swings
- But there are still surprise movements!

More examples of time series & transformations

FIGURE 12.2 **Four Economic Time Series**



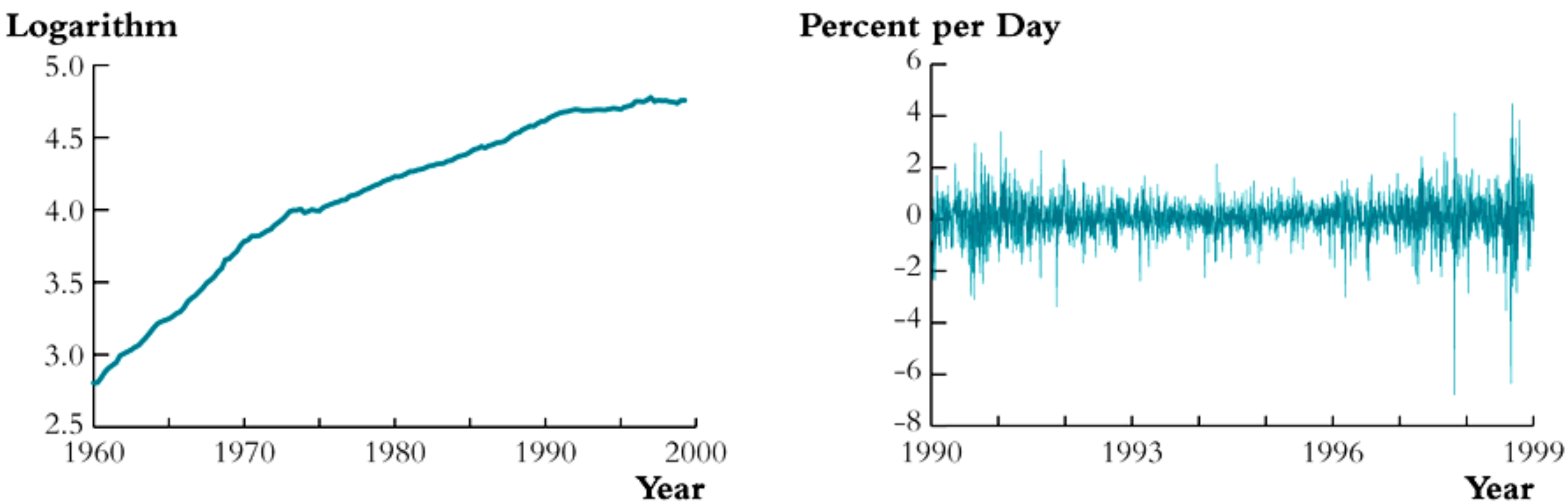
(a) Federal Funds Interest Rate

(b) U.S. Dollar/British Pound Exchange Rate

The four time series have markedly different patterns. The Federal Funds Rate (Figure 12.2a) has a pattern similar to price inflation. The exchange rate between the U.S. dollar and the British pound (Figure 12.2b) shows a discrete change after the 1972 collapse of the Bretton Woods system of fixed exchange rates. The logarithm of real GDP in Japan (Figure 12.2c) shows relatively smooth growth, although the growth rate decreases in the 1970s and again in the 1990s. The daily returns on the NYSE stock price index (Figure 12.2d) are essentially unpredictable, but its variance changes: this series shows "volatility clustering."

More examples of time series & transformations, ctd.

FIGURE 12.2 **Four Economic Time Series**



(c) Logarithm of Real GDP in Japan

(d) Percentage Changes in Daily Values of the NYSE Composite Stock Index

The four time series have markedly different patterns. The Federal Funds Rate (Figure 12.2a) has a pattern similar to price inflation. The exchange rate between the U.S. dollar and the British pound (Figure 12.2b) shows a discrete change after the 1972 collapse of the Bretton Woods system of fixed exchange rates. The logarithm of real GDP in Japan (Figure 12.2c) shows relatively smooth growth, although the growth rate decreases in the 1970s and again in the 1990s. The daily returns on the NYSE stock price index (Figure 12.2d) are essentially unpredictable, but its variance changes: this series shows "volatility clustering."

Stationarity: a key idea for external validity of time series regression

Stationarity says that the past is like the present and the future, at least in a probabilistic sense.

Stationarity

A time series Y_t is **stationary** if its probability distribution does not change over time, that is, if the joint distribution of $(Y_{s+1}, Y_{s+2}, \ldots, Y_{s+T})$ does not depend on s; otherwise, Y_t is said to be **nonstationary**. A pair of time series, X_t and Y_t, are said to be **jointly stationary** if the joint distribution of $(X_{s+1}, Y_{s+1}, X_{s+2}, Y_{s+2}, \ldots, X_{s+T}, Y_{s+T})$ does not depend on s. Stationarity requires the future to be like the past, at least in a probabilistic sense.

The above definition is called **strict stationarity**

A generally weaker notion of stationarity is **weak stationarity.**

- A time series Y_t is weakly stationary if its first and second moments exist and are time invariant
- More specifically, $E(Y_t)$ and $E(Y_t Y_{t-j})$ are finite for all t and j and do not depend on t (they may depend on j)
- Does strict stationarity imply weak stationarity?
- Most economic time series are nonstationary
- But we may be able to transform them such that they become stationary
- To do so, we need to know more about the nature of nonstationarity (*more on this later*)

Software for analyzing time series data

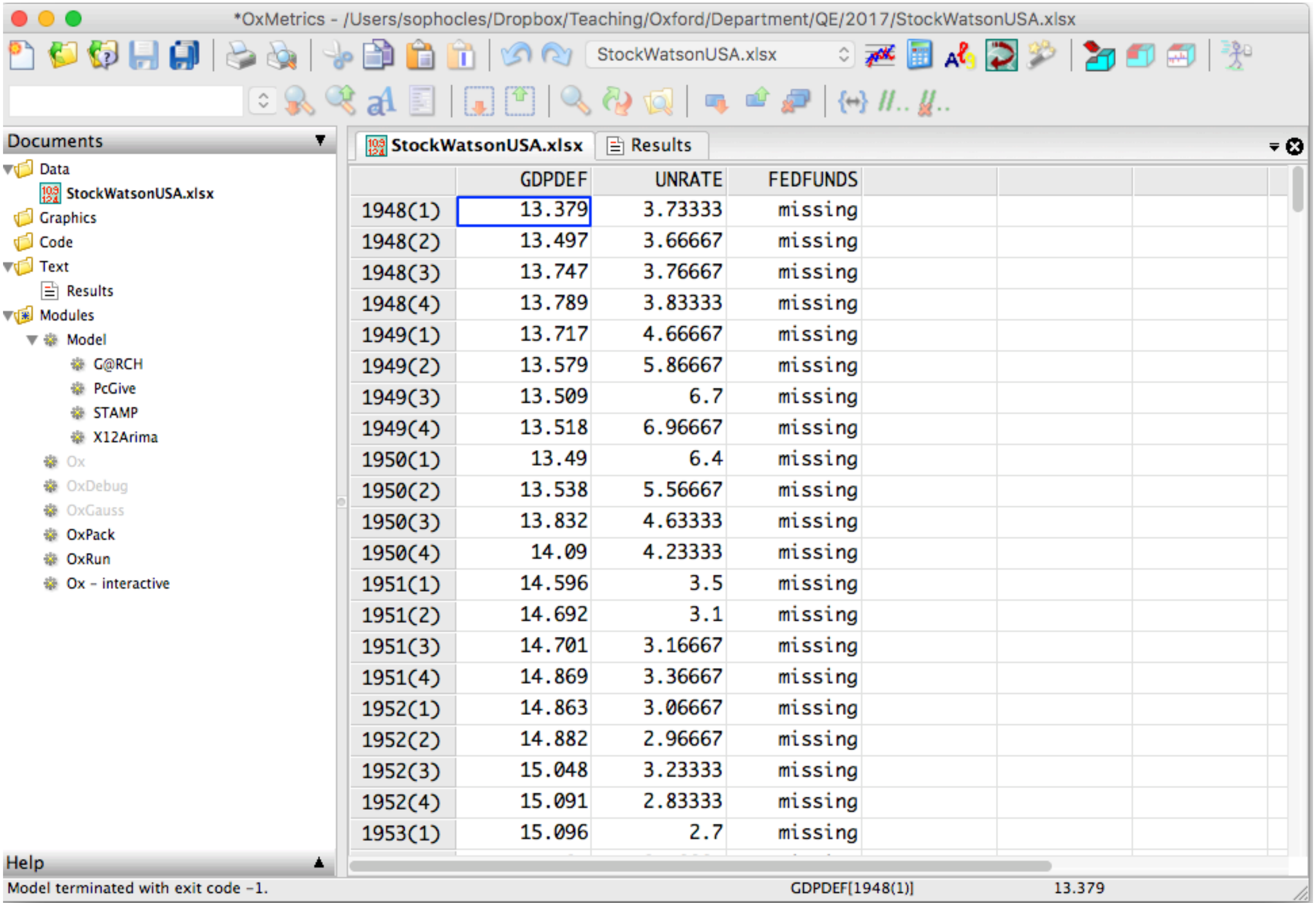
- Most statistical software will be able to analyze TS data
- But some are easier to use than others
- **OxMetrics** is especially good at time series
- And it's also free to all Oxford university members!

  Go to: http://www.doornik.com/download/Oxford/

  (you must access it from OU domain – use VPN if necessary)
- Some online tutorials by the creators:
  - https://www.youtube.com/watch?v=Rc7I9adfISU
  - https://www.youtube.com/watch?v=WCmDqI9eCCg
  - https://www.youtube.com/watch?v=AT0uaFYchZI
  - https://www.youtube.com/watch?v=rsMJLdjBqrc

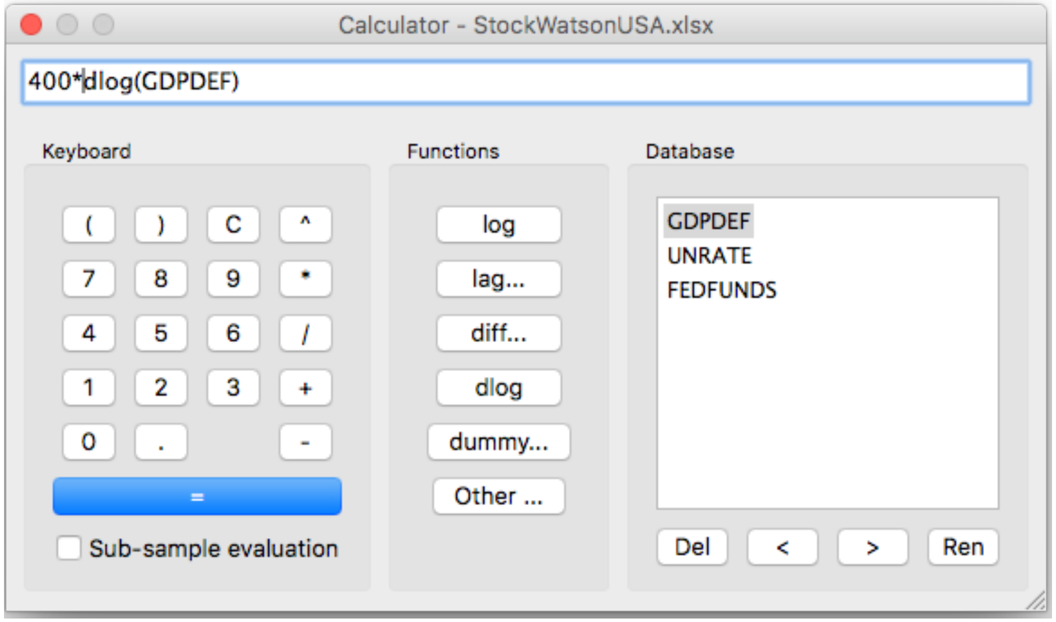
Using OxMetrics

This is what an OxMetrics window looks like…

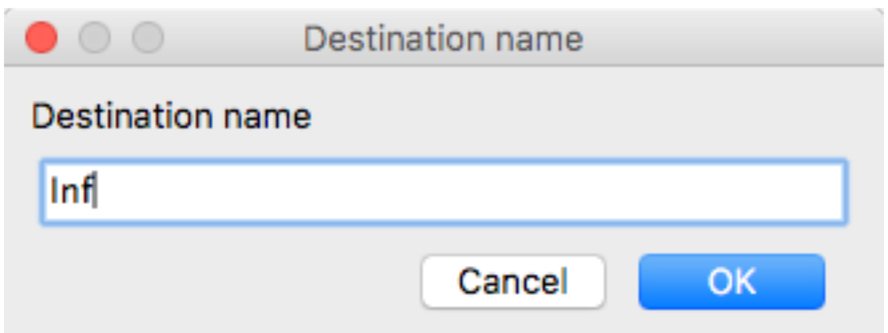


- We have uploaded the data file StockWatsonUSA.xlsx
- Note: data must be in the first worksheet and the first row must contain variable names
- If the date format is not recognized properly, use "Edit/Change Sample…" to adjust the frequency
- The present data file contains quarterly data on GDP deflator, Unemployment rate and Fed Funds rate from 1948q1 to 2016q4, obtained from the St. Louis Fed database "FRED"
- We will transform the data to get quarterly inflation at annualized rate:
- The "Algebra"  syntax for this is:
- `Inf = 400*dlog(GDPDEF);`

- Or using the "calculator"  from the menu, type

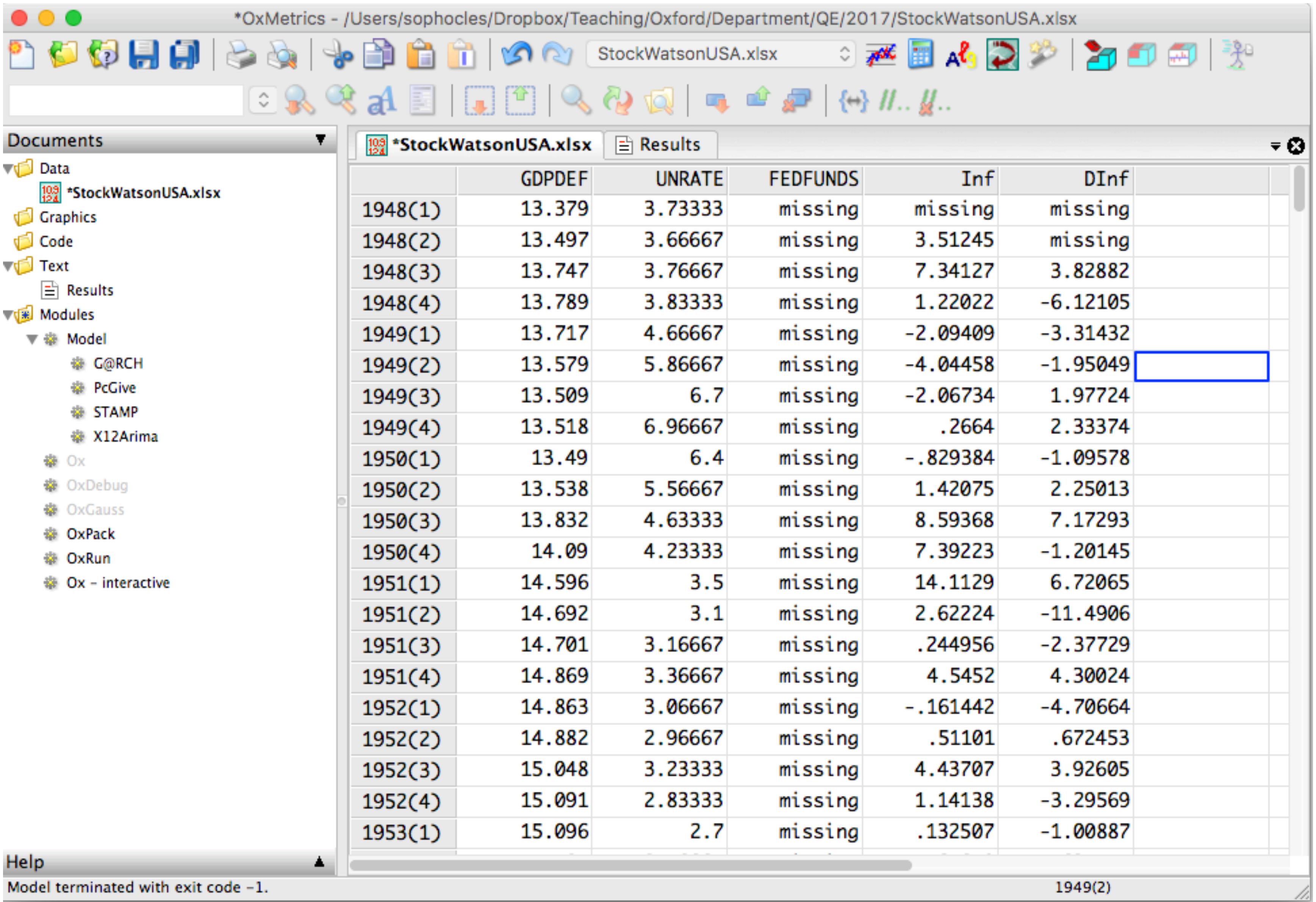


hit "=" and give a variable name in the next window

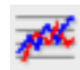


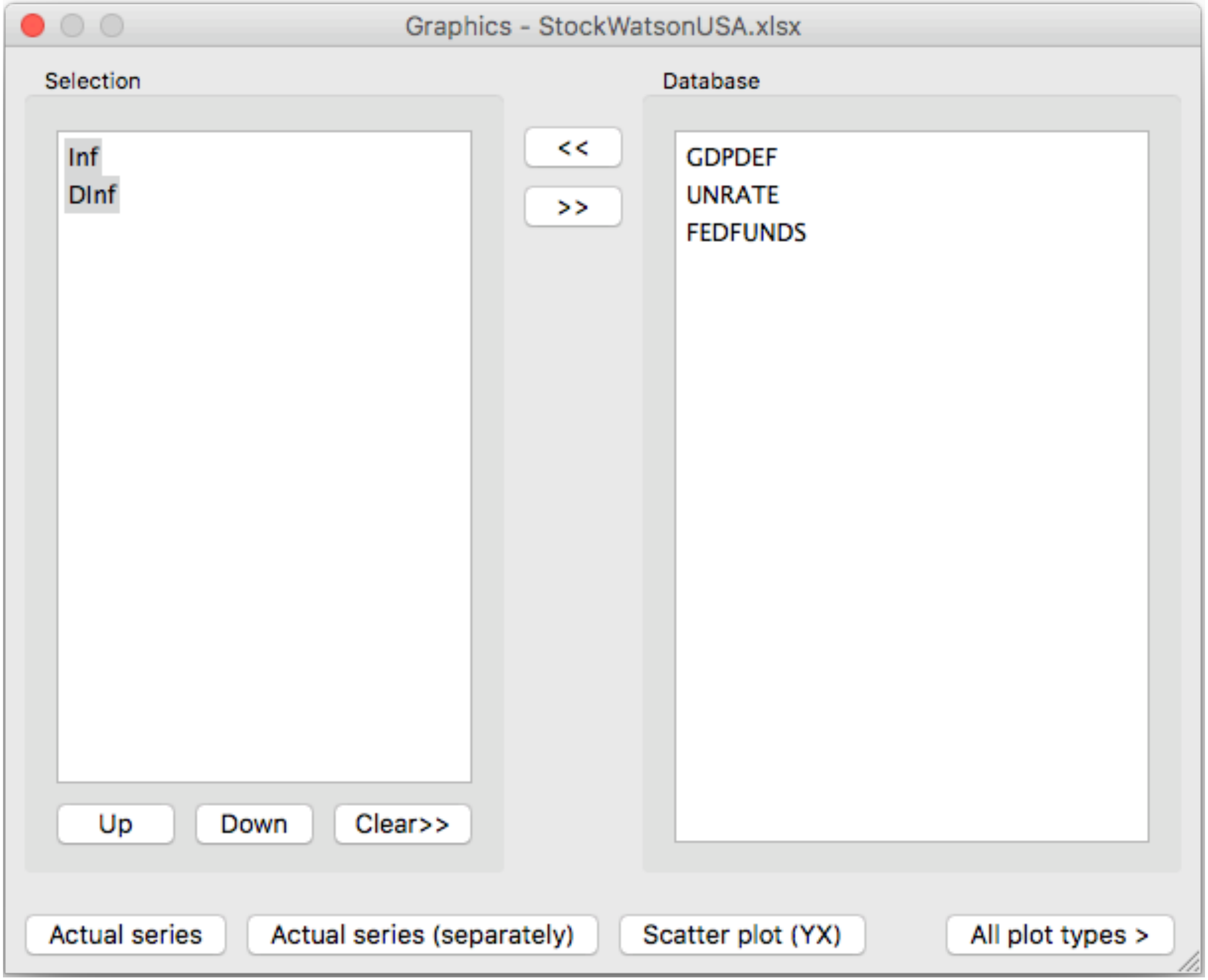
- Also get change in inflation: `DInf = diff(Inf,1);`

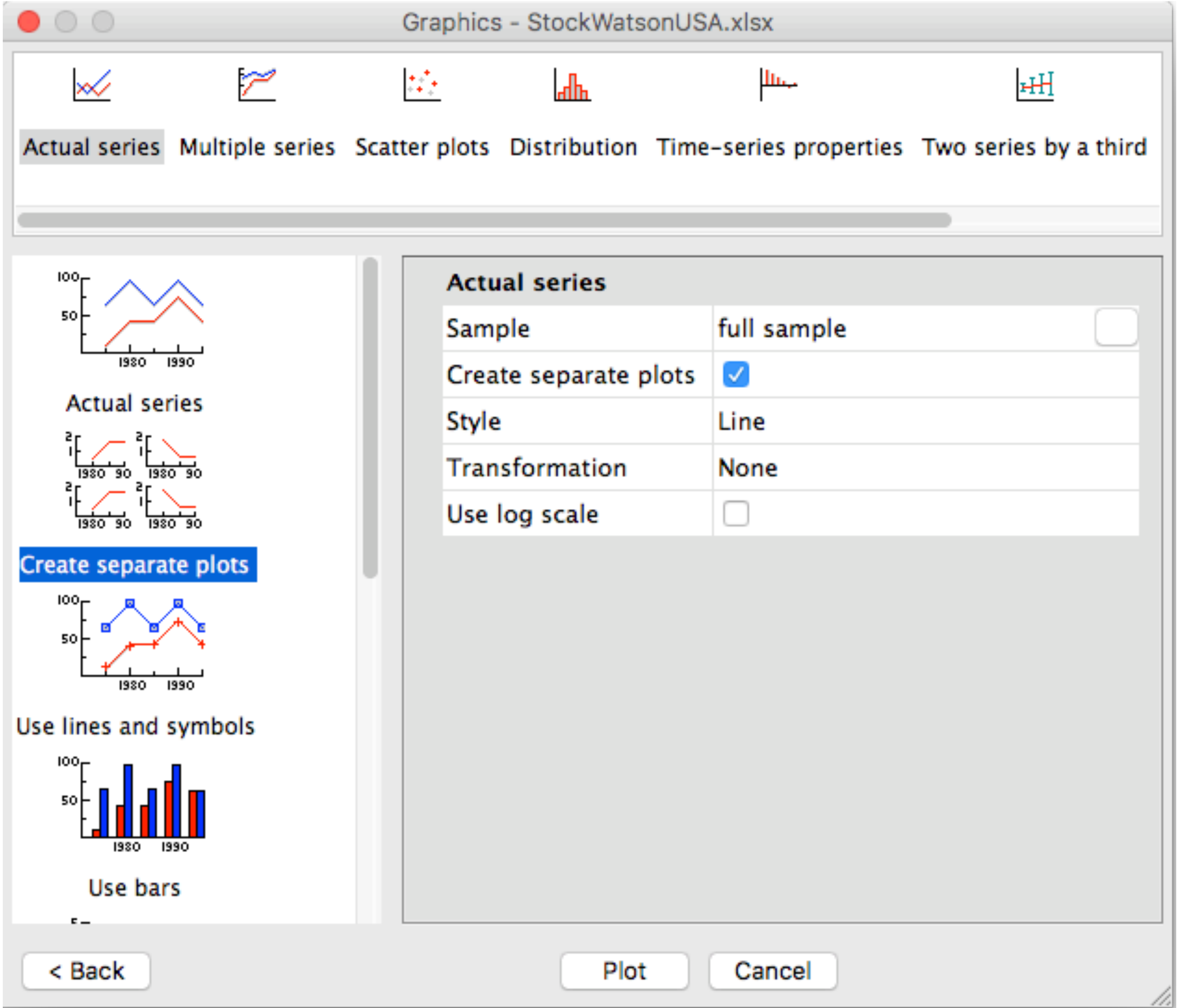
- We lose one obs for `Inf` and two for `DInf`

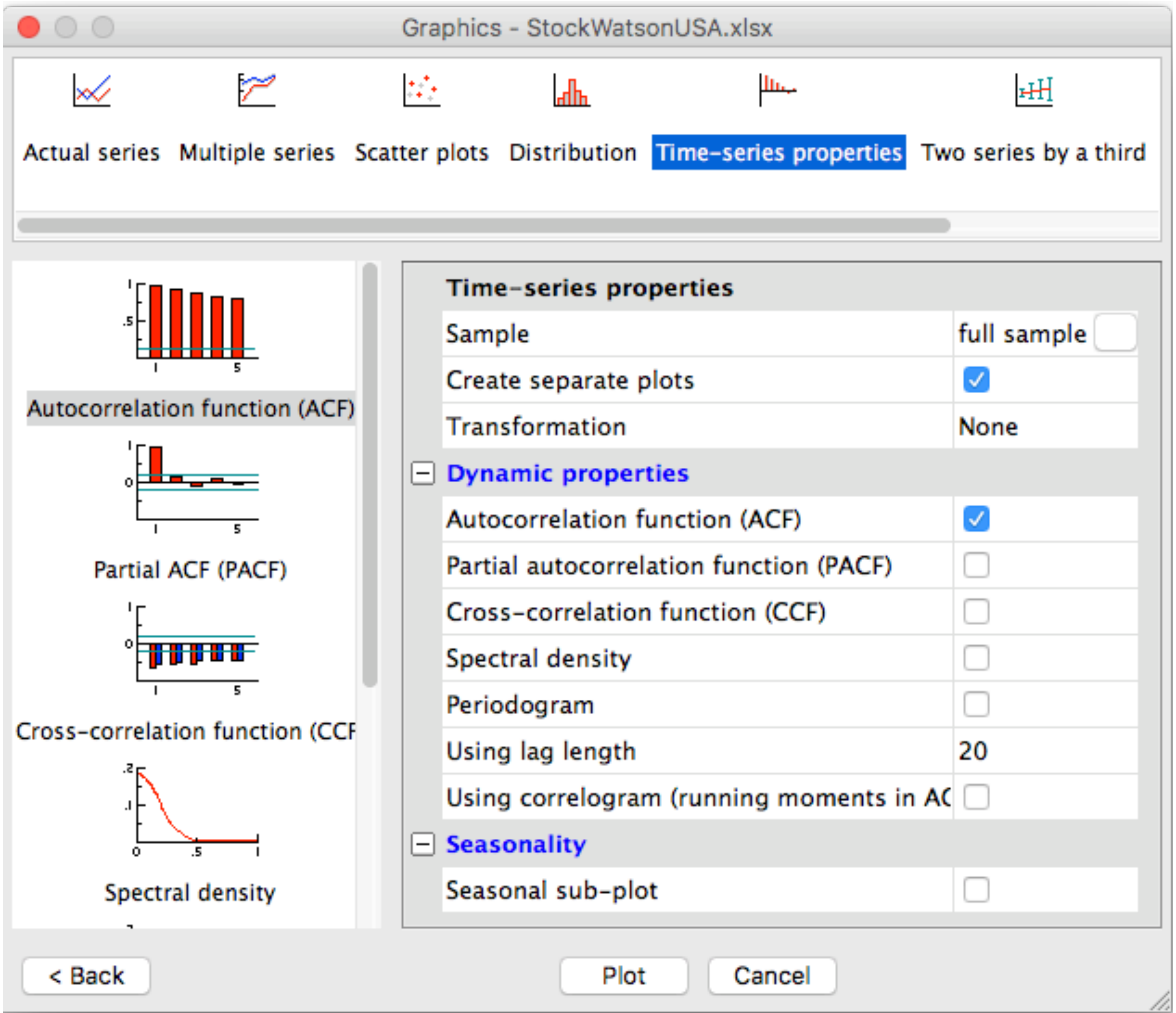


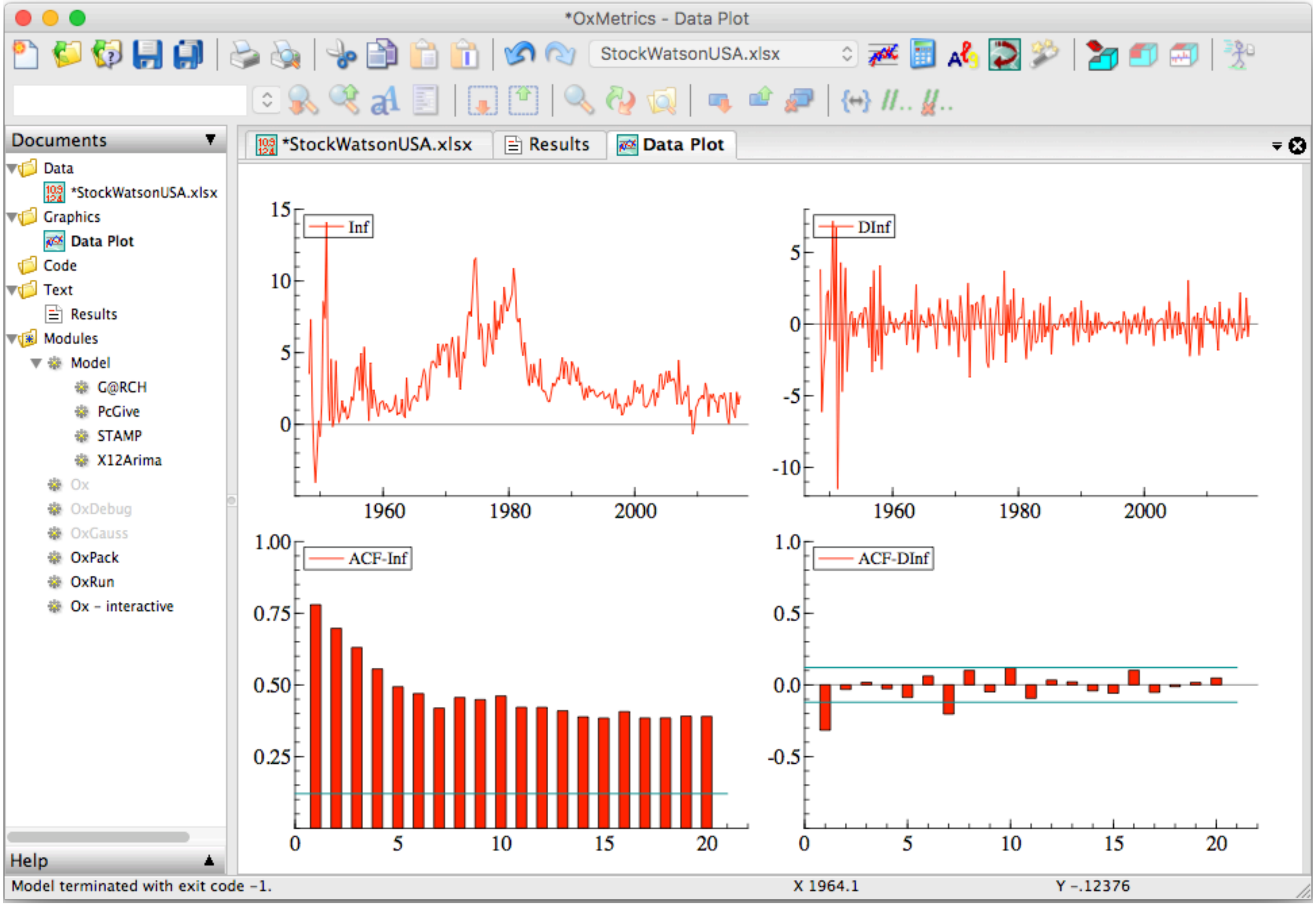
	GDPDEF	UNRATE	FEDFUNDS	Inf	DInf
1948(1)	13.379	3.73333	missing	missing	missing
1948(2)	13.497	3.66667	missing	3.51245	missing
1948(3)	13.747	3.76667	missing	7.34127	3.82882
1948(4)	13.789	3.83333	missing	1.22022	-6.12105
1949(1)	13.717	4.66667	missing	-2.09409	-3.31432
1949(2)	13.579	5.86667	missing	-4.04458	-1.95049
1949(3)	13.509	6.7	missing	-2.06734	1.97724
1949(4)	13.518	6.96667	missing	.2664	2.33374
1950(1)	13.49	6.4	missing	-.829384	-1.09578
1950(2)	13.538	5.56667	missing	1.42075	2.25013
1950(3)	13.832	4.63333	missing	8.59368	7.17293
1950(4)	14.09	4.23333	missing	7.39223	-1.20145
1951(1)	14.596	3.5	missing	14.1129	6.72065
1951(2)	14.692	3.1	missing	2.62224	-11.4906
1951(3)	14.701	3.16667	missing	.244956	-2.37729
1951(4)	14.869	3.36667	missing	4.5452	4.30024
1952(1)	14.863	3.06667	missing	-.161442	-4.70664
1952(2)	14.882	2.96667	missing	.51101	.672453
1952(3)	15.048	3.23333	missing	4.43707	3.92605
1952(4)	15.091	2.83333	missing	1.14138	-3.29569
1953(1)	15.096	2.7	missing	.132507	-1.00887

- Always a good idea to start your analysis by **plotting the data**, and their *correlogram* (aka *autocorrelation function -- ACF*)
- Hit the graphics toolbar  and follow the instructions in the popup window:
- Select series from Database and hit “<<”
- Hit “All plot types”
- Select “Actual series” from top menu, “Create Separate Plots” from left menu, and hit “Plot”
- Next, select “Time Series Properties” and then “Plot”
- Finally, hit “Cancel” to view the results…


Graphics - StockWatsonUSA.xlsx
Selection
Inf
DInf
<<
>>
Database
GDPDEF
UNRATE
FEDFUNDS
Up
Down
Clear>>
Actual series
Actual series (separately)
Scatter plot (YX)
All plot types >


Graphics - StockWatsonUSA.xlsx
Actual series
Multiple series
Scatter plots
Distribution
Time-series properties
Two series by a third
Actual series
Create separate plots
Use lines and symbols
Use bars
Actual series
Sample
full sample
Create separate plots
Style
Line
Transformation
None
Use log scale
< Back
Plot
Cancel


Graphics - StockWatsonUSA.xlsx
Actual series
Multiple series
Scatter plots
Distribution
Time-series properties
Two series by a third
Autocorrelation function (ACF)
Partial ACF (PACF)
Cross-correlation function (CCF
Spectral density
Time-series properties
Sample
full sample
Create separate plots
Transformation
None
Dynamic properties
Autocorrelation function (ACF)
Partial autocorrelation function (PACF)
Cross-correlation function (CCF)
Spectral density
Periodogram
Using lag length
20
Using correlogram (running moments in AC
Seasonality
Seasonal sub-plot
< Back
Plot
Cancel



*OxMetrics - Data Plot
StockWatsonUSA.xlsx
Documents
Data
*StockWatsonUSA.xlsx
Graphics
Data Plot
Code
Text
Results
Modules
Model
G@RCH
PcGive
STAMP
X12Arima
Ox
OxDebug
OxGauss
OxPack
OxRun
Ox - interactive
Help
*StockWatsonUSA.xlsx
Results
Data Plot
Inf
DInf
ACF-Inf
ACF-DInf
Model terminated with exit code -1.
X 1964.1
Y -.12376

Autoregressions

A natural starting point for a forecasting model is to use past values of Y (that is, Y_{t-1}, Y_{t-2},…) to forecast Y_t.

- An ***autoregression*** is a regression model in which Y_t is regressed against its own lagged values.
- The number of lags used as regressors is called the ***order*** of the autoregression.
  - In a ***first-order autoregression***, Y_t is regressed against Y_{t-1}
  - In a ***p^th^-order autoregression***, Y_t is regressed against $Y_{t-1}, Y_{t-2}, \ldots, Y_{t-p}$.

The First-Order Autoregressive (AR(1)) Model

The population AR(1) model is

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + u_t$$

- β_0 and β_1 *do not* have causal interpretations
- if $\beta_1 = 0$, Y_{t-1} is not useful for forecasting Y_t
- The AR(1) model can be estimated by OLS regression of Y_t against Y_{t-1}
- Testing $\beta_1 = 0$ v. $\beta_1 \neq 0$ provides a test of the hypothesis that Y_{t-1} is not useful for forecasting Y_t

Example: AR(1) model of the change in inflation

Estimated using data from 1960:I – 2016:IV:

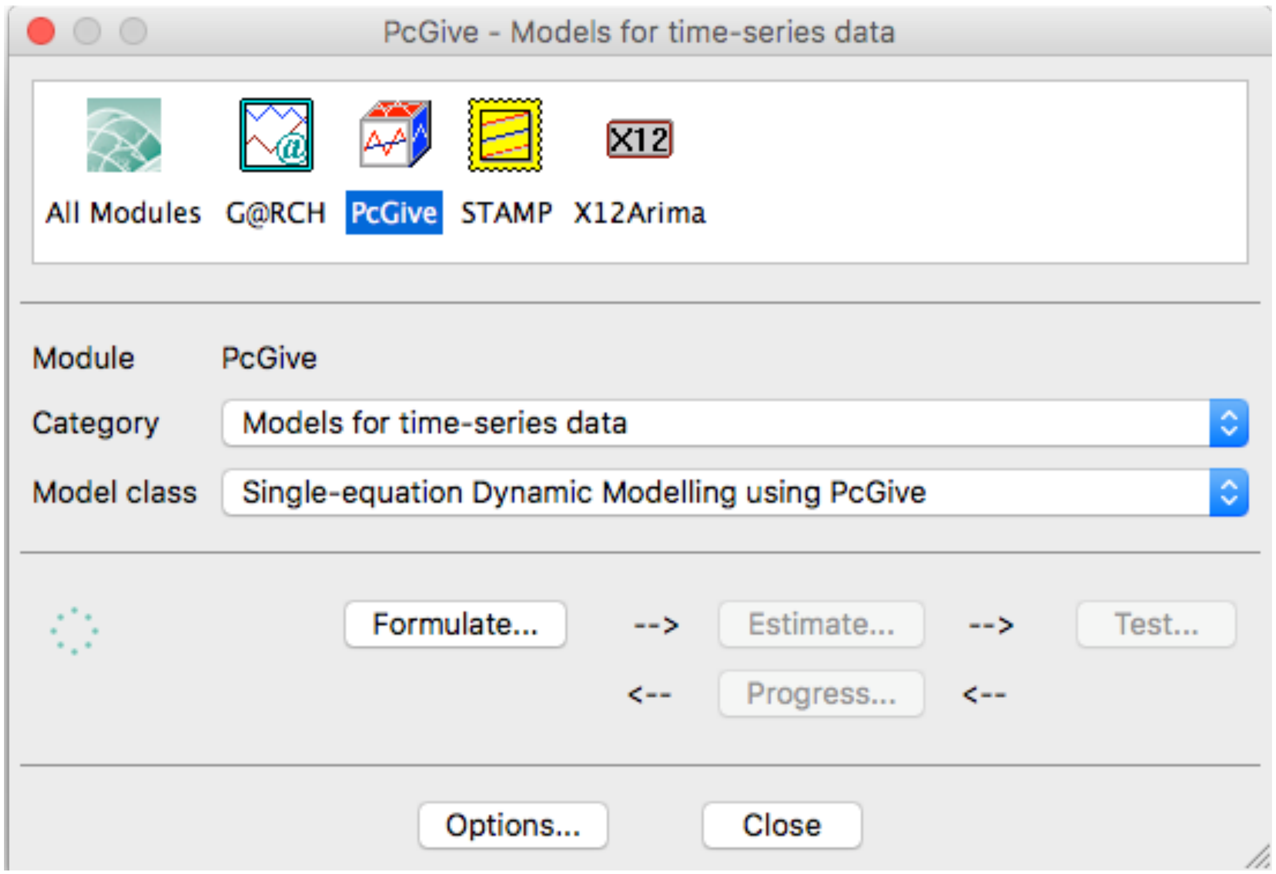
$$\widehat{\Delta Inf_t} = 0.002 - 0.305\Delta Inf_{t-1} \quad \bar{R}^2 = 0.09$$
$$(0.069) \quad (0.063)$$

Is the lagged change in inflation a useful predictor of the current change in inflation?

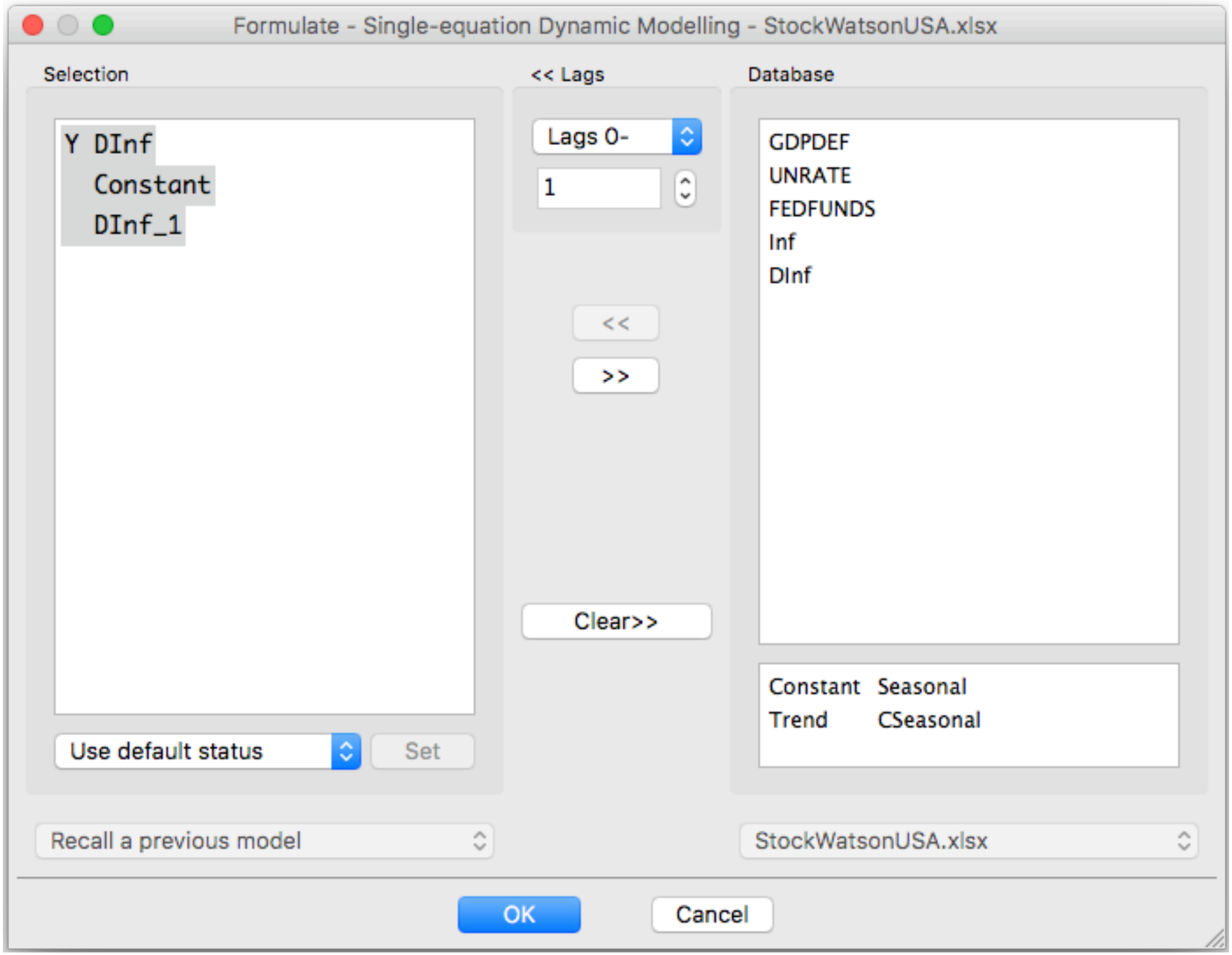
- t = -.305/.063 = -4.81 < -1.96
- Reject H_0: $\beta_1 = 0$ at the 5% significance level
- Yes, the lagged change in inflation is a useful predictor of current change in infl. (but low $\bar{R}^2$!)

Example: AR(1) model of Δinflation – OxMetrics

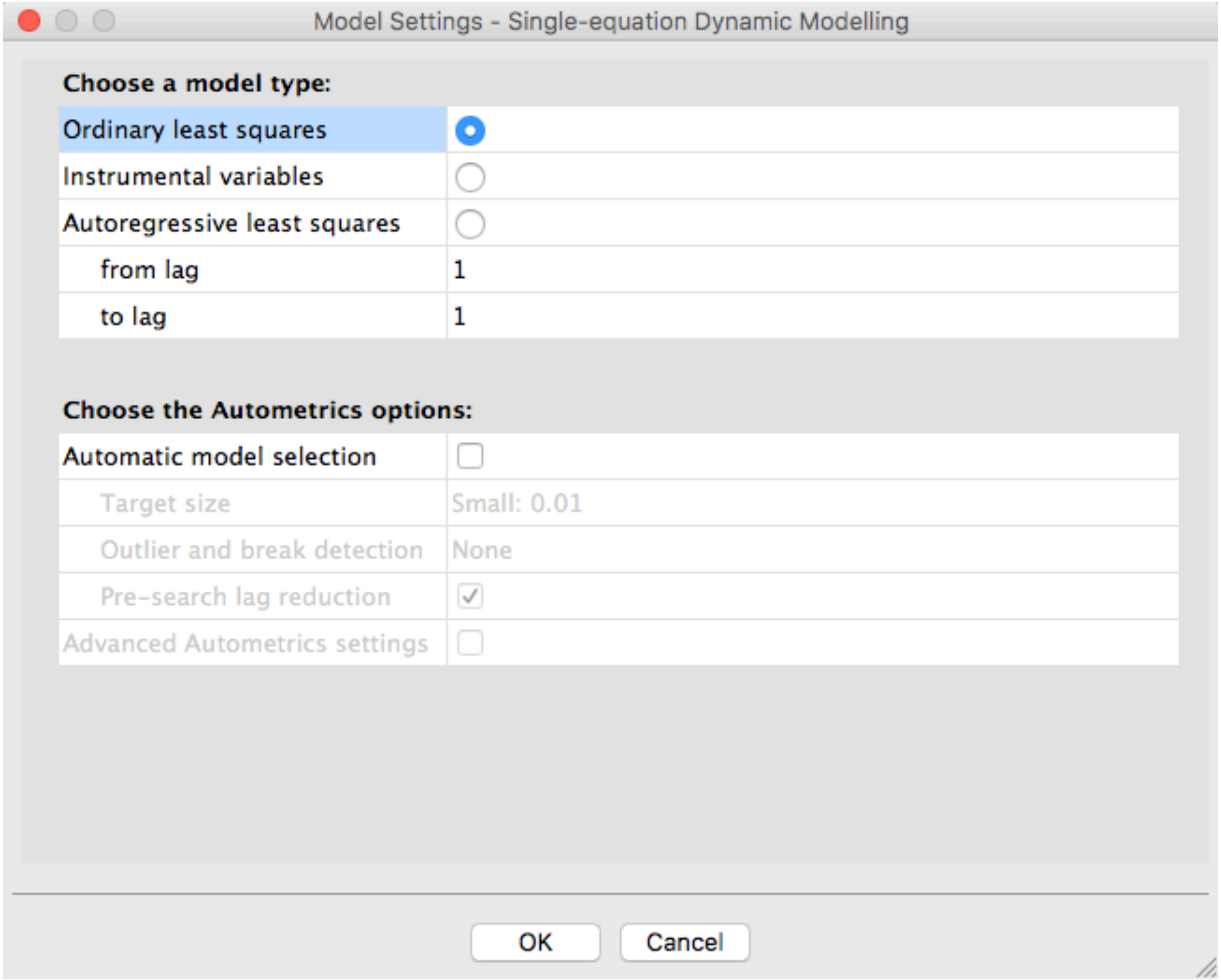
- To estimate a new model, hit toolbar  (or "Model\Model" from the menu)
- Select the choices shown below and hit "Formulate"



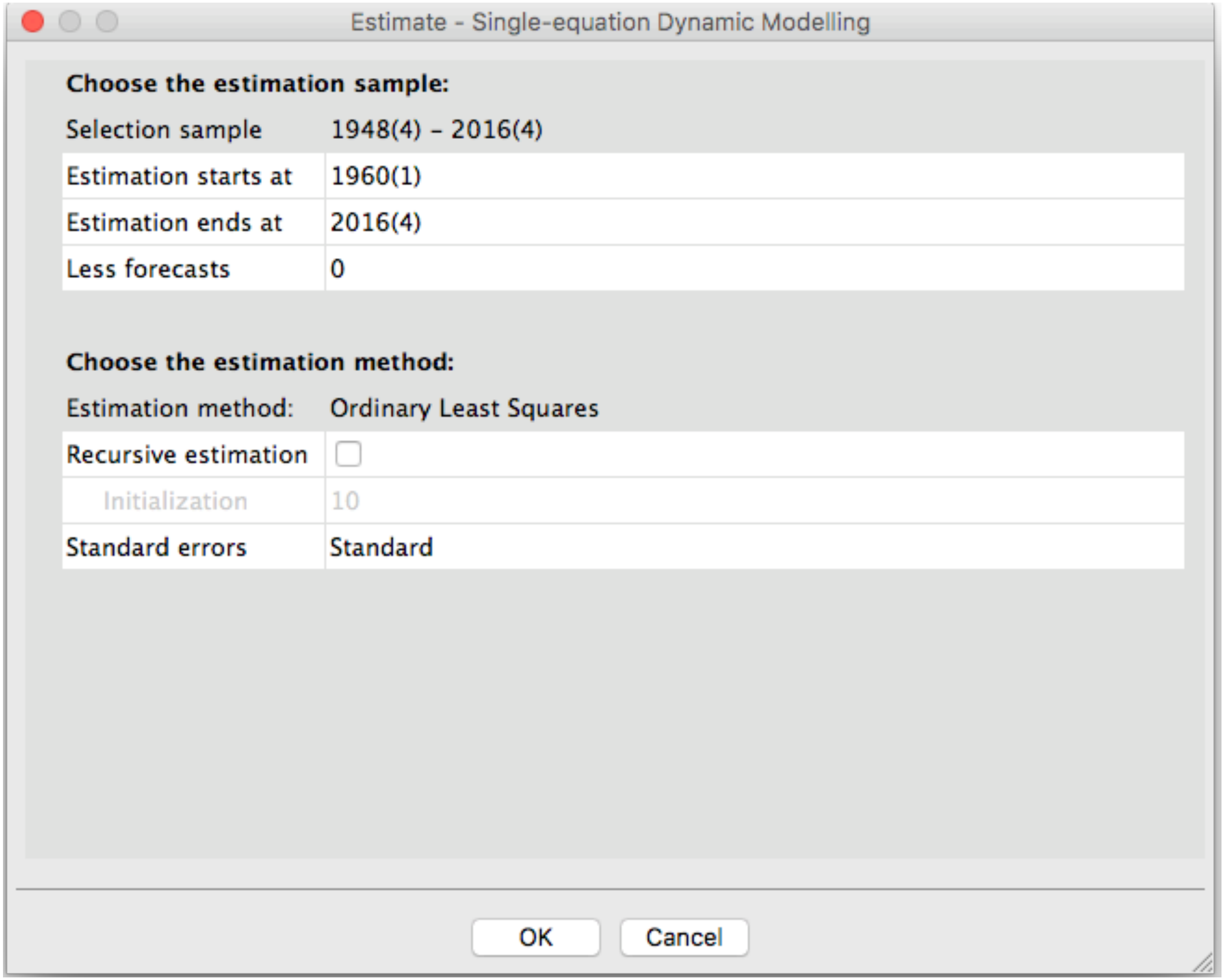
- Choose “DInf”, hit “<<” and then “OK”



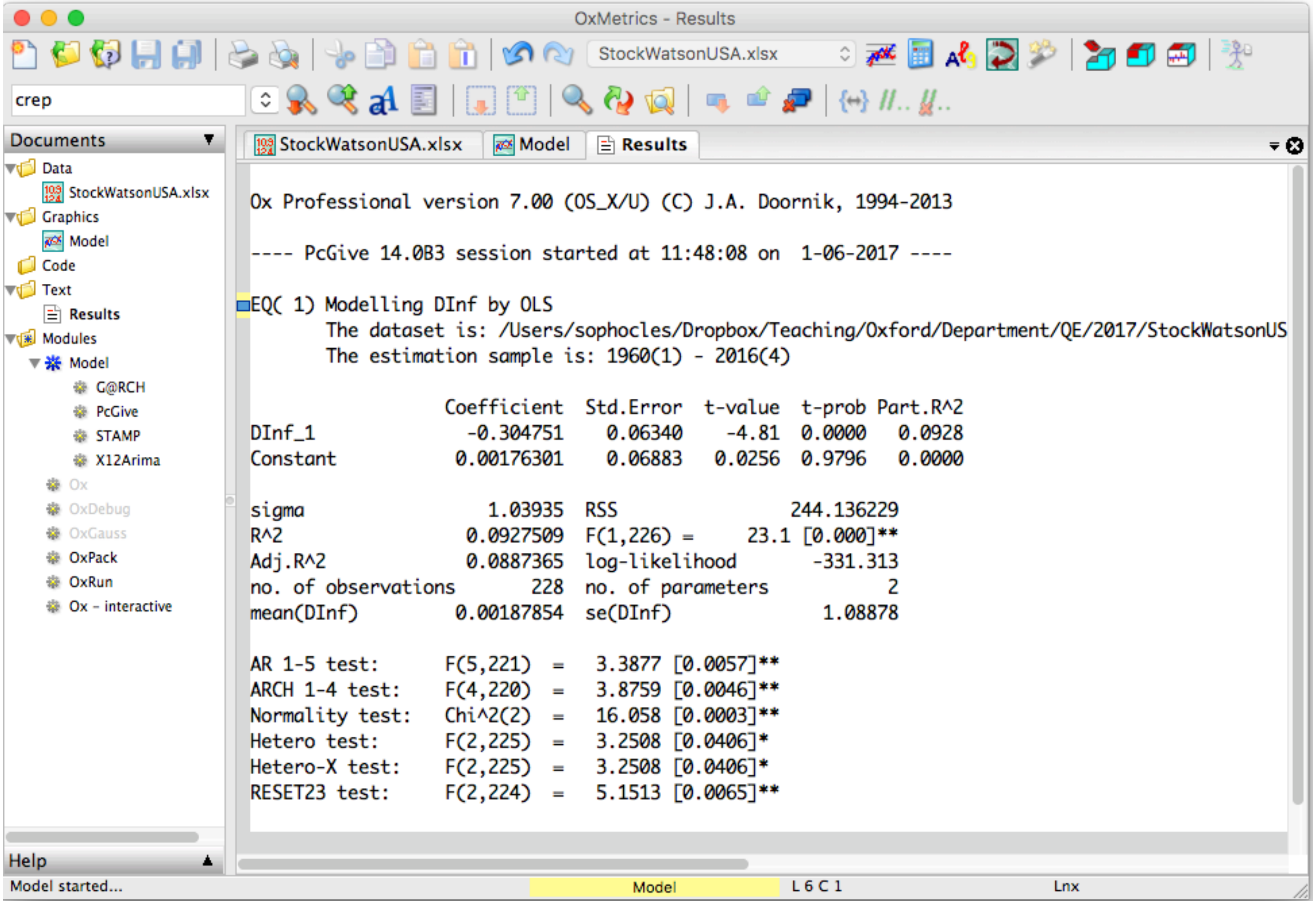
- Choose OLS and hit “OK”



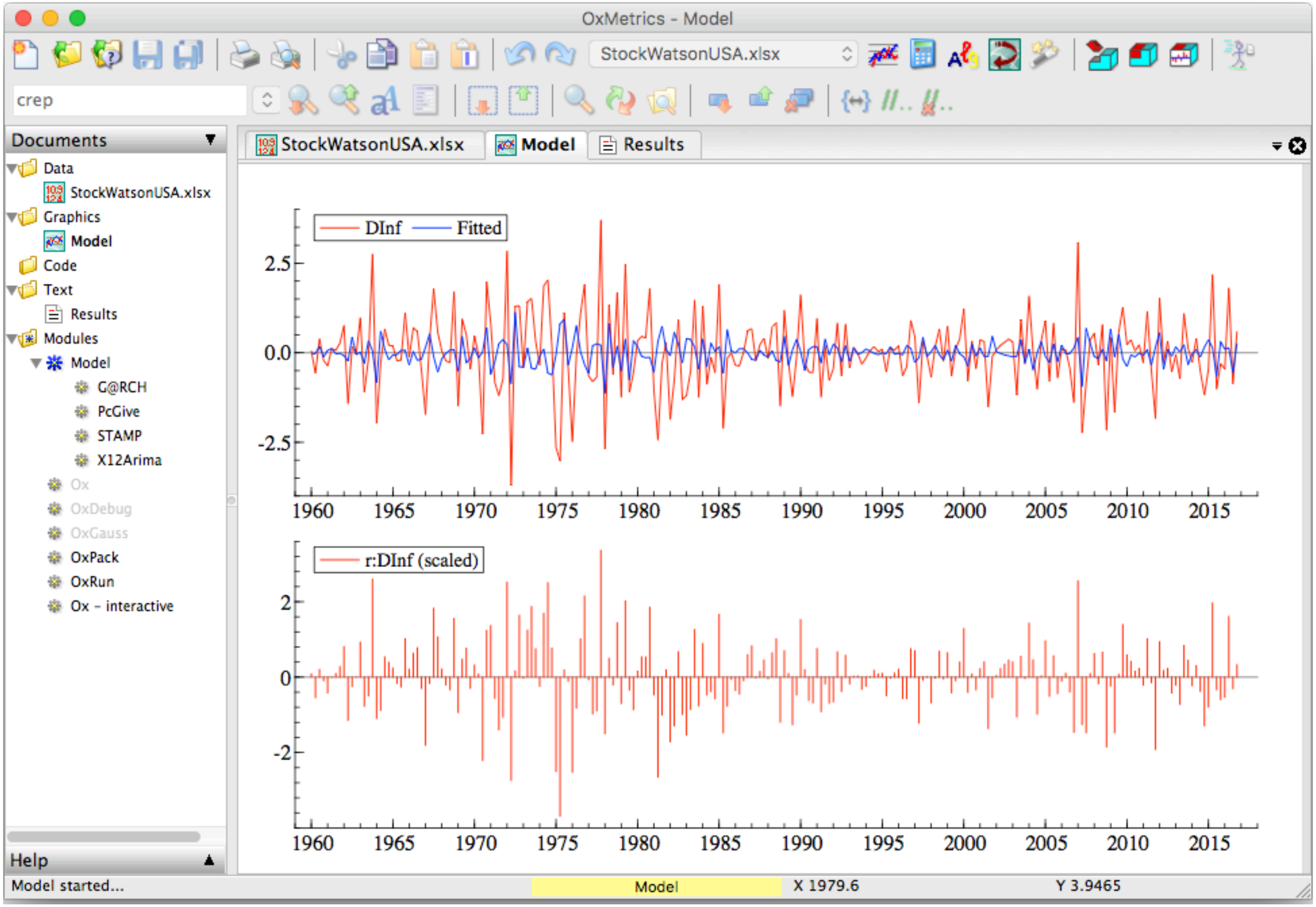
- Select estimation sample and hit “OK”



- The estimation results are printed in the “Results” window



- Residuals and model fit are shown in the "Model" window



Forecasts and forecast errors

A note on terminology:

- A *predicted value* refers to the value of *Y* predicted (using a regression) for an observation in the sample used to estimate the regression – this is the usual definition
- A *forecast* refers to the value of *Y* forecasted for an observation *not* in the sample used to estimate the regression.
- Predicted values are “in-sample”
- Forecasts are forecasts of the future – which cannot have been used to estimate the regression.

Forecasts: notation

- $Y_{t|t-1}$ = forecast of Y_t based on $Y_{t-1}, Y_{t-2}, \ldots$, using the population (true unknown) coefficients
- $\hat{Y}_{t|t-1}$ = forecast of Y_t based on $Y_{t-1}, Y_{t-2}, \ldots$, using the estimated coefficients, which were estimated using data through period $t-1$.

For an AR(1),

- $Y_{t|t-1} = \beta_0 + \beta_1 Y_{t-1}$
- $\hat{Y}_{t|t-1} = \hat{\beta}_0 + \hat{\beta}_1 Y_{t-1}$, where $\hat{\beta}_0$ and $\hat{\beta}_1$ were estimated using data through period $t-1$.

Forecast errors

The one-period-ahead forecast error is,

$$\text{forecast error} = Y_t - \hat{Y}_{t|t-1}$$

The distinction between a forecast error and a residual is the same as between a forecast and a predicted value:

- a residual is "in-sample"
- a forecast error is "out-of-sample" – the value of Y_t isn't used in the estimation of the regression coefficients

The root mean squared forecast error (RMSFE)

$$\text{RMSFE} = \sqrt{E[(Y_t - \hat{Y}_{t|t-1})^2]}$$

- The RMSFE is a measure of the spread of the forecast error distribution.
- The RMSFE is like the standard deviation of u_t, except that it explicitly focuses on the forecast error using estimated coefficients, not using the population regression line.
- The RMSFE is a measure of the magnitude of a typical forecasting "mistake"

***Example*:** forecasting inflation using and AR(1)

AR(1) estimated using data from 1960:I – 2016:IV:

$$\widehat{\Delta Inf_t} = 0.002 - 0.305\Delta Inf_{t-1}$$

$Inf_{2016:\mathrm{III}} = 1.4$ (units are percent, at an annual rate)

$Inf_{2016:\mathrm{IV}} = 2.0$

$\Delta Inf_{2016:\mathrm{IV}} = 0.6$

So the forecast of $\Delta Inf_{2017:\mathrm{I}}$ is,

$$\widehat{\Delta Inf}_{2017:I|2016:IV} = 0.002 - 0.305\times 0.6 = -\ 0.185$$

so

$$\widehat{Inf}_{2017:I|2016:IV} = Inf_{2016:\mathrm{IV}} + \widehat{\Delta Inf}_{2017:I|2016:IV} = 2 - 0.185 = 1.815$$

The p^{th} order autoregressive model (AR(p))

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \ldots + \beta_p Y_{t-p} + u_t$$

- The AR(p) model uses p lags of Y as regressors
- The AR(1) model is a special case
- The coefficients do not have a causal interpretation
- To test the hypothesis that $Y_{t-2},\ldots,Y_{t-p}$ do not further help forecast Y_t, beyond Y_{t-1}, use an F-test
- Use t- or F-tests to determine the lag order p
- Or determine p using an "information criterion" (*see SW Section 14.5*)

Example: AR(4) model of Δinflation

$$\widehat{\Delta Inf_t} = \underset{(.07)}{.004} - \underset{(.07)}{.37}\Delta Inf_{t-1} - \underset{(.07)}{.21}\Delta Inf_{t-2} - \underset{(.07)}{.07}\Delta Inf_{t-3}$$

$$+ \underset{(.07)}{.14}\Delta Inf_{t-4},\ \bar{R}^2 = 0.21$$

- F-statistic testing lags 2, 3, 4 is 5.82 (p-value < .001)
- $\bar{R}^2$ increased from .09 to .16 by adding lags 2, 3, 4
- Lags 2, 3, 4 (jointly) help to predict the change in inflation, above and beyond the first lag

Example: AR(4) model of inflation – OxMetrics

- Follow same steps as for AR(1) above, except for choosing Lags 0-4, and you get:

```
EQ( 3) Modelling DInf by OLS
       The dataset is: /Users/sophocles/Dropbox/Teaching/Oxford/Depa
       The estimation sample is: 1960(1) - 2016(4)

                 Coefficient  Std.Error  t-value  t-prob Part.R^2
DInf_1             -0.373685    0.06637    -5.63  0.0000   0.1245
DInf_2             -0.210212    0.07074    -2.97  0.0033   0.0381
DInf_3            -0.0658262    0.07098   -0.927  0.3547   0.0038
DInf_4              0.139512    0.06659     2.09  0.0373   0.0193
Constant          0.00369924    0.06673   0.0554  0.9558   0.0000

sigma                1.00759  RSS                 226.397191
R^2                 0.158672  F(4,223) =     10.51 [0.000]**
Adj.R^2             0.143581  log-likelihood        -322.714
no. of observations      228  no. of parameters            5
mean(DInf)        0.00187854  se(DInf)               1.08878
```

Example: AR(4) model of inflation – OxMetrics, ctd.

- To test that lags 2 to 4 are statistically significant, go to "Model\Test", choose "Exclusion Restrictions", select DInf_2 to DInf_4, and hit OK to get

```
Test for excluding:
[0] = DInf_2
[1] = DInf_3
[2] = DInf_4
Subset F(3,223)  =   5.8243 [0.0008]**
```

- Lags 2, 3, 4 (jointly) help to predict the change in inflation, above and beyond the first lag

Digression: we used ΔInf, not Inf, in the AR's. ***Why?***

The AR(1) model of ΔInf_t is an AR(2) model of Inf_t:

$$\Delta Inf_t = \beta_0 + \beta_1 \Delta Inf_{t-1} + u_t$$

or

$$Inf_t - Inf_{t-1} = \beta_0 + \beta_1(Inf_{t-1} - Inf_{t-2}) + u_t$$

or

$$Inf_t = Inf_{t-1} + \beta_0 + \beta_1 Inf_{t-1} - \beta_1 Inf_{t-2} + u_t$$

so

$$Inf_t = \beta_0 + (1+\beta_1) Inf_{t-1} - \beta_1 Inf_{t-2} + u_t$$

So why use ΔInf_t, not Inf_t?

AR(1) model of ΔInf: $\Delta Inf_t = \beta_0 + \beta_1 \Delta Inf_{t-1} + u_t$

AR(2) model of Inf: $Inf_t = \gamma_0 + \gamma_1 Inf_{t-1} + \gamma_2 Inf_{t-2} + v_t$

- When Y_t is strongly serially correlated, the OLS estimator of the AR coefficient is biased towards zero.
- In the extreme case that the AR coefficient = 1, Y_t isn't stationary: the u_t's accumulate and Y_t blows up.
- If Y_t isn't stationary, the regression theory we are working with here breaks down
- Here, Inf_t is strongly serially correlated – so to keep ourselves in a framework we understand, the regressions are specified using ΔInf

Time Series Regression with Additional Predictors and the Autoregressive Distributed Lag (ADL) Model

- So far we have considered forecasting models that use only past values of *Y*
- It makes sense to add other variables (*X*) that might be useful predictors of *Y*, above and beyond the predictive value of lagged values of *Y*:

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \ldots + \beta_p Y_{t-p} + \delta_1 X_{t-1} + \ldots + \delta_r X_{t-r} + u_t$$

- This is an ***autoregressive distributed lag*** (***ADL***) ***model***

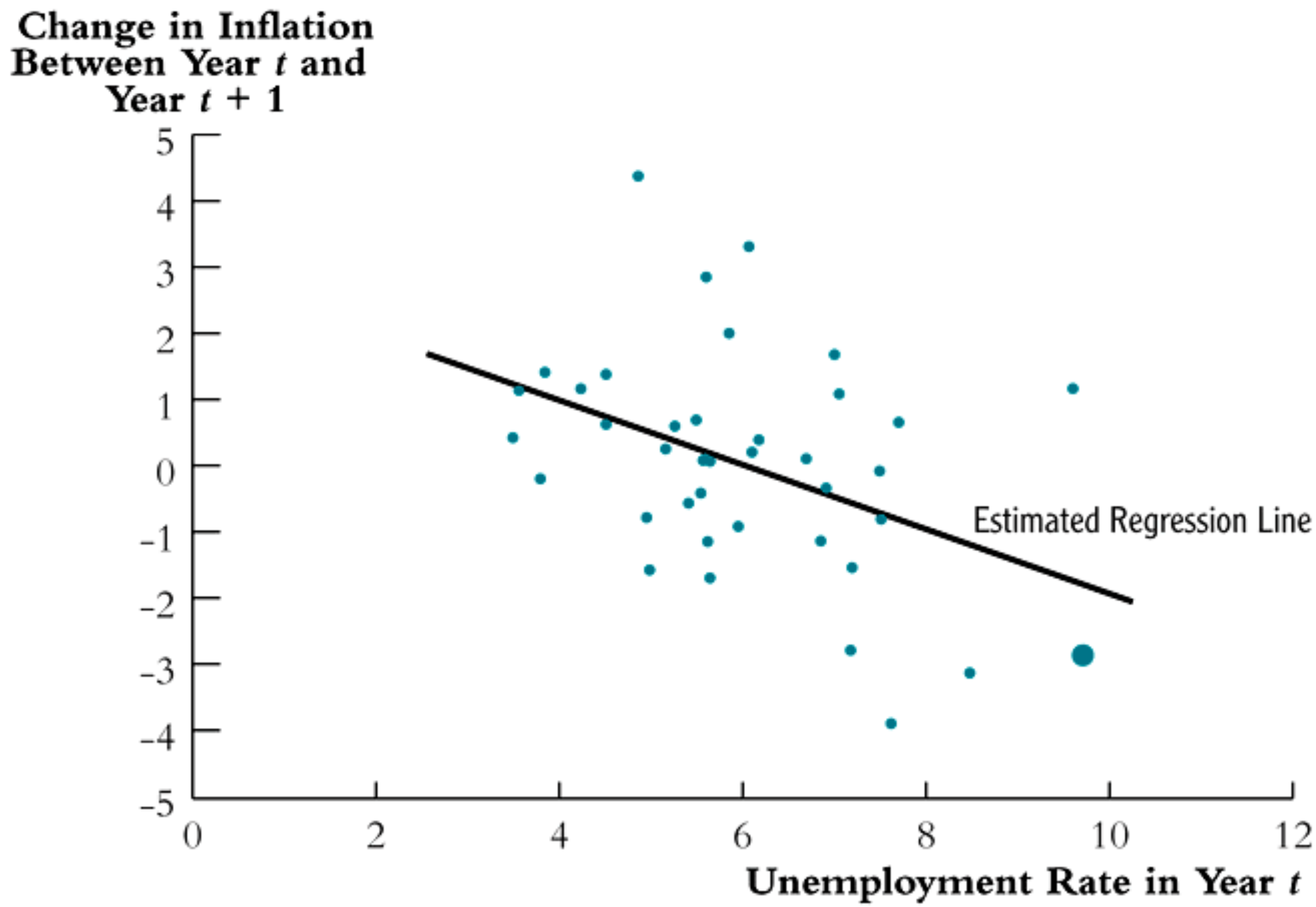
Example: lagged unemployment and inflation

- According to the "Phillips curve", if unemployment is below its equilibrium, or "natural," rate, then the rate of inflation will increase.
- That is, ΔInf_t should be related to lagged values of the unemployment rate, with a negative coefficient
- The rate of unemployment at which inflation neither increases nor decreases is often called the "non-accelerating rate of inflation" unemployment rate: the NAIRU
- Is this relation found in US economic data?
- Can this relation be exploited for forecasting inflation?

The empirical “Phillips Curve”

FIGURE 12.3 **Scatterplot of Change in Inflation Between Year *t* and Year *t* + 1 vs. the Unemployment Rate in Year *t***

In 1982, the U.S. unemployment rate was 9.7% and the rate of inflation in 1983 fell by 2.9% (the large dot). In general, high values of the unemployment rate in year *t* tend to be followed by decreases in the rate of price inflation in the next year, year *t* + 1, with a correlation of –0.40.



The NAIRU is the value of u for which $\Delta Inf = 0$

Example: ADL(4,4) model of ΔInf

$$\widehat{\Delta Inf_t} = \underset{(.28)}{.65} - \underset{(.07)}{.43}\Delta Inf_{t-1} - \underset{(.07)}{.26}\Delta Inf_{t-2} - \underset{(.07)}{.11}\Delta Inf_{t-3} + \underset{(.07)}{.12}\Delta Inf_{t-4}$$

$$- \underset{(.26)}{.99}Unrate_{t-1} + \underset{(.51)}{1.12}Unrate_{t-2} - \underset{(.50)}{.15}Unrate_{t-3} - \underset{(.26)}{.09}Unrate_{t-4}$$

- $\bar{R}^2 = 0.23$ – an improvement over the AR(4), for which $\bar{R}^2 = .16$

Example: ADL(4,4) model of ΔInf – OxMetrics

```
EQ( 4) Modelling DInf by OLS
       The dataset is: /Users/sophocles/Dropbox/Teaching/Oxford/Departme
       The estimation sample is: 1960(1) - 2016(4)

                     Coefficient  Std.Error  t-value  t-prob Part.R^2
DInf_1                 -0.428425    0.06709    -6.39  0.0000   0.1570
DInf_2                 -0.257780    0.07252    -3.55  0.0005   0.0545
DInf_3                 -0.114421    0.07202    -1.59  0.1135   0.0114
DInf_4                  0.123130    0.06660     1.85  0.0658   0.0154
Constant                0.655400     0.2825     2.32  0.0213   0.0240
UNRATE_1               -0.987856     0.2639    -3.74  0.0002   0.0602
UNRATE_2                 1.12369     0.5055     2.22  0.0273   0.0221
UNRATE_3               -0.150854     0.4982   -0.303  0.7624   0.0004
UNRATE_4              -0.0923871     0.2603   -0.355  0.7230   0.0006

sigma                   0.973242  RSS                 207.436613
R^2                     0.229132  F(8,219) =     8.137 [0.000]**
Adj.R^2                 0.200973  log-likelihood        -312.743
no. of observations          228  no. of parameters            9
mean(DInf)            0.00187854  se(DInf)               1.08878
```

Example: ADL(4,4) model of ΔInf – OxMetrics, ctd.

```
Test for excluding:
[0] = DInf_2
[1] = DInf_3
[2] = DInf_4
Subset F(3,219)  =   7.5793 [0.0001]**
```

- Lags 2, 3, 4 of ΔInf (jointly) help to predict the change in inflation, above and beyond the first lag of ΔInf:

```
Test for excluding:
[0] = UNRATE_1
[1] = UNRATE_2
[2] = UNRATE_3
[3] = UNRATE_4
Subset F(4,219)  =   5.0044 [0.0007]**
```

- Lags 1, 2, 3, 4 of *Unrate* (jointly) help to predict the change in inflation, above and beyond lags of ΔInf

The test of the joint *hypothesis* that none of the *X*'s is a useful predictor, above and beyond lagged values of *Y*, is called a ***Granger causality test***

Granger Causality Tests (Tests of Predictive Content)

The Granger causality statistic is the *F*-statistic testing the hypothesis that the coefficients on all the values of one of the variables in Equation (12.20) (for example, the coefficients on $X_{1t-1}, X_{1t-2}, \ldots, X_{1t-q_1}$) are zero. This null hypothesis implies that these regressors have no predictive content for Y_t beyond that contained in the other regressors, and the test of this null hypothesis is called the Granger causality test.

"*causality*" *is an unfortunate term here: Granger Causality simply refers to (marginal) predictive content.*

Summary: Time Series Forecasting Models

- For forecasting purposes, it isn't important to have coefficients with a causal interpretation!
- Simple and reliable forecasts can be produced using AR(p) models – these are common "benchmark" forecasts against which more complicated forecasting models can be assessed
- Additional predictors (X's) can be added; the result is an autoregressive distributed lag (ADL) model
- Stationarity means that the models can be used outside the range of data with which they were estimated