

Forecasting Using Eviews 2.0: An Overview

Some Preliminaries

In what follows it will be useful to distinguish between *ex post* and *ex ante* forecasting. In terms of time series modeling, both predict values of a dependent variable beyond the time period in which the model is estimated. However, in an *ex post* forecast observations on both endogenous variables and the exogenous explanatory variables are known with certainty during the forecast period. Thus, *ex post* forecasts can be checked against existing data and provide a means of evaluating a forecasting model. An *ex ante* forecast predicts values of the dependent variable beyond the estimation period, using explanatory variables that may or may not be known with certainty.

Forecasting References

Bowerman and O'Connell, 1993: *Forecasting and Time Series: An Applied Approach*, Third Edition, The Duxbury Advanced Series in Statistics and Decision Sciences, Duxbury Press, Belmont, CA. ISBN: 0-534-93251-7

Gaynor and Kirkpatrick, 1994: *Introduction to Time-Series Modeling and Forecasting in Business and Economics*, McGraw-Hill, Inc. ISBN: 0-07-034913-4

Pankratz, 1994: *Forecasting with Dynamic Regression Models*, Wiley-Interscience. ISBN: 0-471-61528-5.

Pindyck & Rubinfeld, 1991. *Econometric Models and Economic Forecasts*, Third Edition, McGraw-Hill, Inc. ISBN: 0-07-050098-3 (note: Fourth Edition now available)

Univariate Forecasting Methods

The following will illustrate various forecasting techniques using the GA caseload data in the workfile caseload.wf1. I will explain various Eviews commands and statistical output.

Example: *GA caseload data*

Load the data in the Eviews workfile caseload.wf1:

File/Open

Filename: caseload.wf1

Drives: a:

Directories: a:\

Check the **Update Default Directory** box

OK

Notice that the workfile range is 1973:07 2002:12. I have added several variables and equations to the workfile that will be used in the examples to follow.

Modeling trend behavior correctly is most important for good long-run forecasts whereas modeling the deviations around the trend correctly is most important for good short-run forecasts. I will illustrate both of these points in the examples to follow

Trend extrapolation

Trend extrapolation is a very simple forecasting method that is useful if it is believed that the historic trend in the data is smooth and will continue on its present course into the near future. Trend extrapolation is best computed in Eviews using ordinary least squares regression techniques.

Commands to generate a deterministic time trend variable and the natural log of uxcase:

```
genr trend = @trend(1973.07)+1
genr luxcase = log(uxcase)
```

Ordinary least squares (OLS) estimation (see Help/How do I?/Estimate Equations)

Example: *Estimate the trend model for uxcase over the period 1973:07 1996:07 by least squares regression*

Select **Quick/Estimate Equation**, which brings up the equation estimation dialogue box.

Upper box: uxcase c trend
Estimation method: Least squares
Sample: 1973:07 1996:06
OK

You should see the following estimation results in an **Equation Window**:

```
LS // Dependent Variable is UXCASE
Date: 09/16/97   Time: 15:58
Sample: 1973:07 1996:06
Included observations: 276
```

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4931.804	190.4492	25.89564	0.0000
TREND	47.38451	1.191934	39.75429	0.0000

R-squared	0.852244	Mean dependent var	11494.56
Adjusted R-squared	0.851704	S.D. dependent var	4096.928
S.E. of regression	1577.693	Akaike info criterion	14.73466
Sum squared resid	6.82E+08	Schwarz criterion	14.76089

Log likelihood	-2423.010	F-statistic	1580.404
Durbin-Watson stat	0.045668	Prob(F-statistic)	0.000000

Explanation of Standard Regression Output

Regression Coefficients

Each coefficient multiplies the corresponding variable in forming the best prediction of the dependent variable. The coefficient measures the contribution of its independent variable to the prediction. The coefficient of the series called **C** is the constant or intercept in the regression--it is the base level of the prediction when all of the other independent variables are zero. The other coefficients are interpreted as the slope of the relation between the corresponding independent variable and the dependent variable.

Standard Errors

These measure the statistical reliability of the regression coefficients--the larger the standard error, the more statistical noise infects the coefficient. According to regression theory, there are about 2 chances in 3 that the true regression coefficient lies within one standard error of the reported coefficient, and 95 chances out of 100 that it lies within two standard errors.

t-Statistic

This is a test statistic for the hypothesis that a coefficient has a particular value. The t-statistic to test if a coefficient is zero (that is, if the variable does not belong in the regression) is the ratio of the coefficient to its standard error and this is the t-statistic reported by Eviews. If the t-statistic exceeds one in magnitude it is at least two-thirds likely that the true value of the coefficient is not zero, and if the t-statistic exceeds two in magnitude it is at least 95 percent likely that the coefficient is not zero.

Probability

The last column shows the probability of drawing a t-statistic of the magnitude of the one just to the left from a t distribution. With this information, you can tell at a glance if you reject or accept the hypothesis that the true coefficient is zero. Normally, a probability lower than .05 is taken as strong evidence of rejection of that hypothesis.

R-squared

This measures the success of the regression in predicting the values of the dependent variable within the sample. R² is one if the regression fits perfectly, and zero if it fits no better than the simple mean of the dependent variable. R² is the fraction of the variance of the dependent variable explained by the independent variables. It can be negative if the regression does not have an

intercept or constant, or if two-stage least squares is used.

R2 adjusted for degrees of freedom

This is a close relative of R2 in which slightly different measures of the variances are used. It is less than R2 (provided there is more than one independent variable) and can be negative.

S.E. of regression

This is a summary measure of the size of the prediction errors. It has the same units as the dependent variable. About two-thirds of all the errors have magnitudes of less than one standard error. The standard error of the regression is a measure of the magnitude of the residuals. About two-thirds of the residuals will lie in a range from minus one standard error to plus one standard error, and 95 percent of the residuals will lie in a range from minus two to plus two standard errors.

Sum of Squared Residuals

This is just what it says. You may want to use this number as an input to certain types of tests.

Log Likelihood

This is the value of the log likelihood function evaluated at the estimated values of the coefficients. Likelihood ratio tests may be conducted by looking at the difference between the log likelihoods of restricted and unrestricted versions of an equation.

Durbin-Watson Statistic

This is a test statistic for serial correlation. If it is less than 2, there is evidence of positive serial correlation.

Akaike Information Criterion

The Akaike Information Criterion, or AIC, is a guide to the selection of the number of terms in an equation. It is based on the sum of squared residuals but places a penalty on extra coefficients. Under certain conditions, you can choose the length of a lag distribution, for example, by choosing the specification with the lowest value of the AIC.

Schwarz Criterion

The Schwarz criterion is an alternative to the AIC with basically the same interpretation but a larger penalty for extra coefficients.

F-Statistic

This is a test of the hypothesis that all of the coefficients in a regression are zero (except the intercept or constant). If the F-statistic exceeds a critical level, at least one of the coefficients is probably non-zero. For example, if there are three independent variables and 100 observations, an F-statistic above 2.7 indicates that the probability is at least 95 percent that one or more of the three coefficients is non-zero. The probability given just below the F-statistic enables you to carry out this test at a glance.

Equation Object

After you estimate an equation, EViews creates an **Equation Object** and displays the estimation results in an Equation Window. Initially the equation window is untitled (hence it is a temporary object that will disappear when you close the window). You can make the equation a permanent object in the workfile by clicking [**Name**] on the equation toolbar and supplying a name.

Equation Window buttons

[View]	gives you a wide variety of other views of the equation - explained below
[Procs]	gives you a submenu whose first two items are the same as the [Estimate] button and the [Forecast] button. The third choice is Make Regressor Group , which creates a group comprising all of the right-hand variables in the equation.
[Objects]	gives you the standard menu of operations on objects, including Store , which saves the equation on disk under its name, with the extension .DBE, if you have given the equation a name. If you have not given it a name (it is still called Untitled), EViews opens a SaveAs box. You can supply a file name and your estimated equation will be saved on disk as a .DBE file.
[Print]	prints what is currently in the window.
[Name]	gives the estimated equation a name and keeps it in the workfile. The icon in the workfile directory is a little = sign.
[Freeze]	copies the view into a table or graph suitable for further editing.
[Estimate]	opens the estimation box so you can change the equation or sample of observations and re-estimate.
[Forecast]	calculates a forecast from the equation.
[Stats]	brings back the table of standard regression results.
[Resids]	shows the graph of residuals and actual and fitted values.

*Views of the equation object (menu of choices when [**View**] is clicked)*

- **View/Representations** shows the equation in three forms: as a list of series, as an algebraic equation with symbolic coefficients, and as an equation with the estimated values of the coefficients.
- **View/Estimation Output** is the view shown above, with standard estimation results.
- **View/Actual, Fitted, Residual/Table** shows the actual values of the dependent variable, fitted values, and residuals in a table with a plot of the residuals down the right side.
- **View/Actual, Fitted, Residual/Graph** shows a standard EViews graph of the actual values, fitted values, and residuals.

- **View/Covariance Matrix** shows the covariance matrix of the coefficient estimates as a spreadsheet view.
- **View/Coefficient Tests, Residual Tests, and Stability Tests** lead to additional menus for specification and diagnostic tests .

Analysis of Residuals (see **Help/How do I?/Use Specification and Diagnostic Tests**)

The simple deterministic trend model is appropriate provided the residuals are “well behaved” - i.e., provided our assumptions regarding the random error term in the trend model are satisfied. We can check these assumptions by running diagnostic tests on the residuals. Eviews has many such diagnostics built-in and these are available from the [**View**] menu on the equation toolbar. The basic diagnostics available for an estimated equation are:

[View]/Actual,Fitted,Residual/Graph

The **Actual Values** are the values of the dependent variable used in a regression, from the original data. The **Fitted Values** are the predicted values from a regression computed by applying the regression coefficients to the independent variables. The **Residuals** are the differences between the actual and fitted values of the dependent variable. They give an indication of the likely errors that the regression would make in a forecasting application.

Example: residuals from simple trend model

[View]/Residual Tests/Correlogram

The correlogram view of the residuals (forecast errors) shows the autocorrelations of the residuals. These are the correlation coefficients of values of the residuals k periods apart. Nonzero values of these autocorrelations indicate omitted predictability in the dependent variable. If they die off more or less geometrically with increasing lag, k , it is a sign that the series obeys a low-order autoregressive process. For example,

$$e_t = \rho e_{t-1} + u_t$$

where u_t is white noise. If, on the other hand, they drop to close to zero after a small number of lags, it is a sign that the series obeys a low-order moving-average process. For example,

$$e_t = u_t + \theta u_{t-1}$$

where u_t is white noise. Often these simple models for the correlation in the forecast errors can be used to greatly improve the forecasting model. This is the intuition behind the Box-Jenkins ARIMA modeling techniques. See **Help/Serial Correlation** for further explanation of these processes.

The *partial autocorrelation* at lag k is the regression coefficient on e_{t-1} when e_t is regressed on

e_{t-1} . It shows if the pattern of autocorrelation is one that can be captured by an autoregression of order less than k , in which case the partial autocorrelation will be small, or if not, in which case the partial autocorrelation will be large in absolute value.

In the correlogram view, the Ljung-Box Q-statistic can be used to test the hypothesis that all of the autocorrelations are zero; that is, that the series is white noise (unpredictable). Under the null hypothesis, Q is distributed as Chi-square, with degrees of freedom equal to the number of autocorrelations, p , if the series has not previously been subject to ARIMA analysis. If the series is the residuals from ARIMA estimation, the number of degrees of freedom is the number of autocorrelations less the number of autoregressive and moving average terms previously estimated. ARIMA models will be explained later.

Example: residuals from simple trend model

[View]/Residual Tests/Histogram Normality

The histogram displays the frequency distribution of the residuals. It divides the series range (distance between its maximum and minimum values) into a number of equal length intervals or bins and displays a count of the number of observations that fall into each bin.

Skewness is a measure of symmetry of the histogram. The skewness of a symmetrical distribution, such as the normal distribution, is zero. If the upper tail of the distribution is thicker than the lower tail, skewness will be positive.

The **Kurtosis** is a measure of the tail shape of a histogram. The kurtosis of a normal distribution is 3. If the distribution has thicker tails than does the normal distribution, its kurtosis will exceed three.

The **Jarque-Bera** (See **Help/Search/Jarque-Bera**) statistic tests whether a series is normally distributed. Under the null hypothesis of normality, the Jarque-Bera statistic is distributed Chi-square with 2 degrees of freedom.

Nonnormal residuals suggest outliers or general lack of fit of the model.

Example: residuals from simple trend model

[View]/Residual Tests/White Heteroskedasticity

This is a very general test for nonconstancy of the variance of the residuals. If the residuals have nonconstant variance then ordinary least squares is not the best estimation technique - weighted least squares should be used instead.

Example: residuals from simple trend model

Forecasts (See **Help/How do I?/Make Forecasts**)

Once you have estimated an equation, forecasting is a snap. Simply click the [**Forecast**] button on the equation toolbar.

The [**Forecast**] button on the equation toolbar opens a dialog. You should fill in the blank for the name to be given to the forecast. The name should usually be different from the name of the dependent variable in the equation, so that you will not confuse actual and forecasted values. For your convenience in reports and other subsequent uses of the forecast, EViews moves all the available data from the actual dependent variable into the forecasted dependent variable for all the observations before the first observation in the current sample.

Optionally, you may give a name to the standard errors of the forecast; if you do, they will be placed in the workfile as a series with the name you specify.

You must also specify the sample for the forecast. Normally this will be a period in the future, if you are doing true forecasting. But you can also make a forecast for a historical period - which is useful for evaluating a model.

You have a choice of two forecasting methods.

- **Dynamic** calculates forecasts for periods after the first period in the sample by using the previously forecasted values of the lagged left-hand variable. These are also called n-step ahead forecasts.
- **Static** uses actual rather than forecasted values (it can only be used when actual data are available). These are also called 1-step ahead or rolling forecasts.

Both of these methods forecast the value of the disturbance, if your equation has an autoregressive or moving-average error specification. **The two methods will always give identical results in the first period of a multiperiod forecast. They will give identical results in the second and subsequent periods when there are no lagged dependent variables or ARMA terms.**

You can instruct EViews to ignore any ARMA terms in the equation by choosing the Structural option. Structural omits the error specification; it forecasts future errors to be zero even if there is an ARMA specification. More on this later.

You can choose to see the forecast output as a graph or a numerical forecast evaluation, or both.

You are responsible for supplying the values for the independent variables used in forecasting, as well as any lagged dependent variables if you are using static forecasting. You may want to make forecasts based on projections or guesses about the independent variables. In that case, you should use a group window to insert those projections or guesses into the appropriate observations in the independent variables before forecasting.

Example: *Generate forecasts over the horizon 1996:07 - 1997:07:*

[Forecast]

Forecast name: uxcasf

S.E. (Optional): se

Sample range for forecast: 1996:07 1997:07

Method: Dynamic

Output: Forecast evaluation

OK

Note: the dynamic forecasts will be the same as the static forecasts in this case because there are no lagged dependent variables or ARMA terms.

You will see the following forecast evaluation statistics

Actual: UXCASE Forecast: UXCASEF

Sample: 1996:07 1997:07

Include observations: 13

Root Mean Squared Error	1596.253
Mean Absolute Error	1559.850
Mean Absolute Percentage Error	9.317731
Theil Inequality Coefficient	0.045444
Bias Proportion	0.954909
Variance Proportion	0.000085
Covariance Proportion	0.045006

Eviews creates two new series: **uxcasf** (the forecast values of uxcase) and **se** (the standard errors of the forecast). By default, Eviews sets uxcasf equal to uxcase prior to the forecast horizon which can be seen by doubling clicking on the two series and viewing the group spreadsheet.

Forecast Error Variances

Forecasts are made from regressions or other statistical equations. In the case of a regression, given the vector of data on the x-variables, the corresponding forecast of the left-hand variable, y, is computed by applying the regression coefficients b to the x-variables.

Forecasts are made with error. With a properly specified equation there are two sources of forecast error. The first arises because the residuals in the equation are unknown for the forecast period. The best you can do is to set these residuals equal to their expected value of zero. In reality, residuals only average out to zero and residual uncertainty is usually the largest source of forecast error. The equation standard error (called "S.E. of regression" in the output) is a measure of the random variation of the residuals.

In dynamic forecasts, innovation uncertainty is compounded by the fact that lagged dependent variables and ARMA terms depend on lagged innovations. EViews also sets these equal to their expected values, which differ randomly from realized values. This additional source of forecast uncertainty tends to rise over the forecast horizon, leading to a pattern of increasing forecast errors.

The second source of forecast error is coefficient uncertainty. The estimated coefficients of the equation deviate from the true coefficients in a random fashion. The standard error of the coefficient, given in the regression output, is a measure of the precision with which the estimated coefficients measure the true coefficients. Since the estimated coefficients are multiplied by the exogenous variables in the computation of forecasts, the more the exogenous variables deviate from their mean values, the greater forecast uncertainty.

In a properly specified model, the realized values of the endogenous variable will differ from the forecasts by less than plus or minus two standard errors 95 percent of the time. A plot of this 95 percent confidence interval is produced when you make forecasts in EViews.

EViews computes a series of forecast standard errors when you supply a series name in the standard error box in the forecast dialog box.

Normally the forecast standard errors computed by EViews account for both innovation and coefficient uncertainty. One exception is in equations estimated by nonlinear least squares and equations that include PDL (polynomial distributed lag) terms. For forecasts made from these equations the standard errors measure only innovation uncertainty.

Evaluation of forecasts

Once forecasts are made they can be evaluated if the actual values of the series to be forecast are observed. Since we computed ex post forecasts we can compute forecast errors and these errors can tell us a lot about the quality of our forecasting model.

Example: *Generate forecast errors:*

```
smpl 1996.07 1997.07  
genr error = uxcase - uxcasf  
genr abserror = @ABS(error)  
genr pcterror = error/uxcase  
genr abspcterror = abserror/uxcase
```

Highlight the series uxcase, uxcasf error abserror pcterror and abspcterror, double click and open a group.

Let Y_t = actual values, f_t = forecast values, $e_t = Y_t - f_t$ = forecast errors and n = number of forecasts.. Eviews reports the following evaluation statistics if forecasts are computed ex post.

Root Mean Square Error

$$RMSE = \sqrt{\sum \frac{e_t^2}{n}}$$

Mean Absolute Error

$$MAE = \frac{\sum |e_t|}{n}$$

Mean Absolute Percentage Error

$$MAPE = \frac{1}{n} \cdot \sum \frac{|e_t|}{Y_t}$$

Theil Inequality Coefficient

$$U = \frac{\sqrt{\frac{1}{n} \cdot \sum (Y_t - f_t)^2}}{\sqrt{\frac{1}{n} \cdot \sum f_t^2} + \sqrt{\frac{1}{n} \cdot \sum Y_t^2}}$$

The scaling of U is such that it will always lie between 0 and 1. If U = 0, $Y_t = f_t$ for all forecasts and there is a perfect fit; if U = 1 the predictive performance is as bad as it possibly could be.

Theil's U statistic can be rescaled and decomposed into 3 proportions of inequality - bias, variance and covariance - such that bias + variance + covariance = 1. The interpretation of these three proportions is as follows:

Bias Indication of systematic error. Whatever the value of U, we would hope that bias is close to 0. A large bias suggests a systematic over or under prediction.

Variance Indication of the ability of the forecasts to replication degree of variability in the variable to be forecast. If the variance proportion is large then the actual series has fluctuated considerably whereas the forecast has not.

Covariance This proportion measures unsystematic error. Ideally, this should have the highest proportion of inequality.

such that these proportions sum to 1.

Plotting the forecasts and confidence intervals

Example: *plotting the trend forecasts*

To plot the forecast vs. the actual values with standard error bands do the following. First compute the standard error bands:

[Genr]

Upper box: lower = uxcasf - 2*se

Lower box: 1996:07 1997:07

[Genr]

Upper box: upper = uxcase + 2*se

Lower box: 1996:07 1997:07

Next, highlight the four series **uxcase**, **uxcasf**, **lower** and **upper** and then double click to open a group. Then click [**View**]/**Graph** to produce a graph of the data. Click [**Sample**] and change the sample to **1996:07 1997:07**.

Revising the estimated model

Based on the residual diagnostics and the evaluation of the forecasts the simple trend model does not look very good. What follows is an illustration of revising the simple trend model so that the statistical assumptions appear satisfied and the forecasts look reasonable.

Example: *Estimate the model excluding the observations prior to 1988*

[Estimate]

Upper box: uxcase c trend

Estimation method: Least squares

Sample: 1988:01 1996:06

OK

You should see the following estimation results:

LS // Dependent Variable is UXCASE

Date: 09/16/97 Time: 16:32

Sample: 1988:01 1996:06

Included observations: 102

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3663.266	316.4566	11.57589	0.0000
TREND	50.61940	1.391544	36.37644	0.0000
R-squared	0.929738	Mean dependent var	15077.94	

Adjusted R-squared	0.929035	S.D. dependent var	1553.334
S.E. of regression	413.7953	Akaike info criterion	12.07016
Sum squared resid	17122658	Schwarz criterion	12.12163
Log likelihood	-758.3097	F-statistic	1323.245
Durbin-Watson stat	0.281048	Prob(F-statistic)	0.000000

[Name] the equation EQ2.

Do residual diagnostics

Notice the serial correlation in the residuals.

Generate forecasts over the horizon 1996:07 - 1997:07:

[Forecast]

Forecast name: uxcasef

S.E. (Optional): se

Sample range for forecast: 1996:07 1997:07

Method: Dynamic

Output: Forecast evaluation

OK

You will see the following forecast evaluation statistics

Actual: UXCASE Forecast: UXCASEF
Sample: 1996:07 1997:07
Include observations: 13

Root Mean Squared Error	1256.522
Mean Absolute Error	1206.788
Mean Absolute Percentage Error	7.214177
Theil Inequality Coefficient	0.036136
Bias Proportion	0.922404
Variance Proportion	0.000004
Covariance Proportion	0.077591

Plot the forecast vs. the actual values with standard error bands

[Genr]

Upper box: lower = uxcasef - 2*se

Lower box: 1996:07 1997:07

[Genr]

Upper box: upper = uxcase + 2*se

Lower box: 1996:07 1997:07

Highlight the four series **uxcase**, **uxcasef**, **lower** and **upper** and then double click to open a group. Then click [View]/Graph to produce a graph of the data. Click [Sample] and change the sample to **1996:07 1997:07**.

Example: *Correcting for error autocorrelation using Cochrane-Orcutt*

Since the residuals are highly autocorrelated and the partial autocorrelation function cuts at lag 1, the forecast error can be modeled using a first order autoregression or AR(1). Modeling the residual using an AR(1) is also called the Cochrane-Orcutt correction for serial correlation.

[Estimate]

Upper box: uxcase c trend AR(1)

Estimation method: Least squares

Sample: 1988:01 1996:06

OK

You should see the following estimation results:

LS // Dependent Variable is UXCASE
Date: 09/16/97 Time: 16:42
Sample: 1988:01 1996:06
Included observations: 102
Convergence achieved after 3 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3233.859	1056.078	3.062140	0.0028
TREND	52.19044	4.524451	11.53520	0.0000
AR(1)	0.838200	0.051141	16.38992	0.0000
R-squared	0.981079	Mean dependent var		15077.94
Adjusted R-squared	0.980697	S.D. dependent var		1553.334
S.E. of regression	215.8144	Akaike info criterion		10.77781
Sum squared resid	4611011.	Schwarz criterion		10.85501
Log likelihood	-691.4000	F-statistic		2566.634
Durbin-Watson stat	1.948854	Prob(F-statistic)		0.000000
Inverted AR Roots	.84			

Static vs. Dynamic Forecasts

Given that there is an autoregressive (AR) term in the forecast, there will now be differences between static (one-step ahead or rolling) and dynamic (n-step ahead) forecasts. Compute static forecasts if interest is in the performance of rolling one-step ahead forecasts. Otherwise compute dynamic forecasts.

Also, you have the choice of Ignoring ARMA terms in the computation of the forecasts. This option is useful if you want to compare forecasts that use ARMA corrections versus forecasts that do not use such corrections.

Example: *Comparison of static and dynamic forecasts*

Example: *Allowing for a break in the trend*

Since it looks like there is a break in the trend around the beginning of 1995 (using the eyeball metric), let's modify the model to allow the trend to drop after the beginning of 1995. First, generate a new trend variable to pick up the fall in trend after the beginning of 1995:

[Genr]

Upper window: trend2 = 0

Lower window: 1988:01 1997:07

[Genr]

Upper window: trend2 = trend-259

Lower window: 1995:01 1997:07

Now, re-estimate the model assuming a broken trend using Cochrane-Orcutt:

[Estimate]

Upper box: uxcase c trend trend2 AR(1)

Estimation method: Least squares

Sample: 1988:01 1996:06

OK

You should see the following estimation results:

LS // Dependent Variable is UXCASE

Date: 09/16/97 Time: 17:04

Sample(adjusted): 1988:02 1996:06

Included observations: 101 after adjusting endpoints

Convergence achieved after 4 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2093.445	973.5779	2.150259	0.0340
TREND	57.90439	4.348501	13.31594	0.0000
TREND2	-63.70321	24.64620	-2.584708	0.0112
AR(1)	0.789057	0.061619	12.80547	0.0000
R-squared	0.982218	Mean dependent var		15099.39
Adjusted R-squared	0.981668	S.D. dependent var		1545.832
S.E. of regression	209.2987	Akaike info criterion		10.72632
Sum squared resid	4249179.	Schwarz criterion		10.82989
Log likelihood	-680.9921	F-statistic		1785.986
Durbin-Watson stat	1.974058	Prob(F-statistic)		0.000000
Inverted AR Roots	.79			

Note: The difference between the trend and trend2 coefficients gives the slope of the trend after 1995.01.

[Name] the equation EQ3.

Forecast evaluation

[Forecast]

Forecast name: uxcasf

S.E. (Optional): se

Sample range for forecast: 1996:07 1997:07

Method: Dynamic

Output: Forecast evaluation

OK

Actual: UXCASE Forecast: UXCASEF

Sample: 1996:07 1997:07

Include observations: 13

Root Mean Squared Error	283.6561
Mean Absolute Error	238.8897
Mean Absolute Percentage Error	1.433460
Theil Inequality Coefficient	0.008391
Bias Proportion	0.709269
Variance Proportion	0.162318
Covariance Proportion	0.128413

Example: *Adding Deterministic Seasonal Dummy Variables to the broken trend model*

The caseload data appear to exhibit some seasonality. A simple way to model this is to use deterministic seasonal dummy variables.

Generate the seasonal dummies using the **@Seas** function:

```
smpl 1973.07 2002.12  
genr feb = @seas(2)  
genr march = @seas(3)  
genr April = @seas(4)  
genr may = @seas(5)  
genr June = @seas(6)  
genr July = @seas(7)  
genr Aug = @seas(8)  
genr sept = @seas(9)  
genr oct = @seas(10)  
genr nov = @seas(11)  
genr dec = @seas(12)
```

Note: The omitted season, January, is the control season. The coefficient on the **C** variable picks up the effect on the omitted seasonal.

Create the group variable called seasons by highlighting all of the series, double clicking, choosing group variable, clicking [**Name**] and typing the name **seasons**. The group object seasons containing the eleven seasonal dummy variables is now part of the workfile.

Example: *Re-estimating the model with seasonal dummy variables added*

[Estimate]

Upper box: uxcase c seasons trend trend2 AR(1)

Estimation method: Least squares

Sample: 1988:01 1996:06

OK

You should see the following estimation results

LS // Dependent Variable is UXCASE

Date: 09/17/97 Time: 13:16

Sample(adjusted): 1988:02 1996:06

Included observations: 101 after adjusting endpoints

Convergence achieved after 4 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1710.137	784.1761	2.180807	0.0319
FEB	175.2242	71.20292	2.460913	0.0159
MARCH	356.1129	94.34259	3.774679	0.0003
APRIL	387.6474	108.0614	3.587287	0.0006
MAY	448.2834	116.5315	3.846887	0.0002
JUNE	409.6114	121.4306	3.373214	0.0011
JULY	264.9249	123.8598	2.138909	0.0353
AUG	193.7078	123.0769	1.573877	0.1192
SEPT	96.21790	119.0508	0.808209	0.4212
OCT	23.33588	111.1483	0.209953	0.8342
NOV	3.816040	97.73644	0.039044	0.9689
DEC	52.16161	74.45884	0.700543	0.4855
TREND	59.00916	3.482863	16.94272	0.0000
TREND2	-77.60918	20.63200	-3.761593	0.0003
AR(1)	0.748731	0.067005	11.17433	0.0000

R-squared	0.985894	Mean dependent var	15099.39
Adjusted R-squared	0.983598	S.D. dependent var	1545.832
S.E. of regression	197.9757	Akaike info criterion	10.71254
Sum squared resid	3370716.	Schwarz criterion	11.10093
Log likelihood	-669.2963	F-statistic	429.3418
Durbin-Watson stat	2.177351	Prob(F-statistic)	0.000000

Inverted AR Roots .75

Forecast evaluation

[Forecast]

Forecast name: uxcasf

S.E. (Optional): se
Sample range for forecast: 1996:07 1997:07
Method: Dynamic
Output: Forecast evaluation
OK

Actual: UXCASE Forecast: UXCASEF
Sample: 1996:07 1997:07
Include observations: 13

Root Mean Squared Error	184.8134
Mean Absolute Error	147.0950
Mean Absolute Percentage Error	0.879516
Theil Inequality Coefficient	0.005513
Bias Proportion	0.045557
Variance Proportion	0.070801
Covariance Proportion	0.883642

Ex Ante Forecasting

Having chosen the forecasting model, it should be re-estimated over the entire available sample and the full sample model should be used for ex ante forecasts.

Example: *Ex ante forecasts of uxcase*

[Estimate]
Upper box: uxcase c seasons trend trend2 AR(1)
Estimation method: Least squares
Sample: 1988:01 1997:07
OK

You should see the following estimation results

LS // Dependent Variable is UXCASE
Date: 09/17/97 Time: 13:24
Sample(adjusted): 1988:02 1997:07
Included observations: 114 after adjusting endpoints
Convergence achieved after 4 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1643.227	757.6141	2.168950	0.0325
FEB	166.3855	65.21835	2.551207	0.0123
MARCH	349.7923	86.24490	4.055803	0.0001
APRIL	353.7506	98.64415	3.586128	0.0005
MAY	425.0117	106.2347	4.000688	0.0001
JUNE	383.3416	110.5326	3.468133	0.0008
JULY	214.7922	112.2271	1.913906	0.0585
AUG	154.2950	112.1279	1.376062	0.1719

SEPT	69.54968	108.7271	0.639672	0.5239
OCT	33.97228	101.6013	0.334368	0.7388
NOV	-10.31782	89.34644	-0.115481	0.9083
DEC	32.88135	68.03769	0.483281	0.6300
TREND	59.43654	3.351398	17.73485	0.0000
TREND2	-81.24704	11.82829	-6.868873	0.0000
AR(1)	0.752334	0.062499	12.03761	0.0000

R-squared	0.986563	Mean dependent var	15291.24
Adjusted R-squared	0.984663	S.D. dependent var	1551.586
S.E. of regression	192.1501	Akaike info criterion	10.63863
Sum squared resid	3655246.	Schwarz criterion	10.99866
Log likelihood	-753.1611	F-statistic	519.2131
Durbin-Watson stat	2.180060	Prob(F-statistic)	0.000000

Inverted AR Roots .75

Check out residuals

[View]/Actual,Fitted,Residuals/Graph
[View]/Residual Tests/Histogram-Normality Test
[View]/Residual Tests/Correlogarm-Q-stat

Generate out-of-sample forecasts from 1997:08 to 2002:12

[Forecast]
Forecast name: uxcasef
S.E. (Optional): se
Sample range for forecast: 1997:08 2002:12
Method: Dynamic
Output: Graph
OK

Save the graph if you like by clicking **[Name]**.

Generate upper and lower confidence limits and export data to excel file:

[Genr]
Upper box: lower = uxcasef - 2*se
Lower box: 1997:08 2002:12
[Genr]
Upper box: upper = uxcase + 2*se
Lower box: 1997:08 2002:12

[Sample] 1997.08 2002.12
[Procs]/Export Data
File type: Excel
Filename: forecast.xls

OK

Series to write: by observation (in columns)

List series by name: lower uxcasef upper

OK

Exponential Smoothing (see [Help/What is?/Exponential Smoothing](#))

Exponential smoothing is a method of forecasting based on a simple statistical model of a time series. Unlike regression models, it does not make use of information from series other than the one being forecast. In addition, the exponential smoothing models are special cases of the more general Box-Jenkins ARIMA models described below.

The simplest technique of this type, single exponential smoothing, is appropriate for a series that moves randomly above and below a constant mean. It has no trend and no seasonal pattern. The single exponential smoothing forecast of a series is calculated recursively as

$$f_t = \alpha y_{t-1} + (1 - \alpha)f_{t-1}$$

where f_t is the forecast at time t and y_{t-1} is the actual series at time $t-1$. The damping coefficient, α , is usually a fairly small number, such as 0.05. The forecast adapts slowly to the actual value of the series.

In typical use, you will include all of the historical data available for the series you are interested in forecasting. After calculating the smoothed forecast for the entire period, your forecast will be in the next observation in the output series.

You can ask EViews to estimate the damping parameter α for you by finding the value that minimizes the sum of squared forecast errors. Don't be surprised, however, if the result is a very large value of the damping parameter. That is a sign that your series is close to a random walk, in which case the best forecast gives high weight to the most recent value and little weight to a longer-run average.

If your series has a trend, then you may want to use double exponential smoothing, which includes a trend in the forecast. That is, single smoothing makes a forecast which is the same for every future observation, whereas double smoothing makes a forecast that grows along a trend. You can also carry out the Holt-Winters forecasting procedure, which is related to exponential smoothing. The basic idea of the Holt-Winters method is to forecast with an explicit linear trend model with seasonal effects. The method computes recursive estimates of the intercept or permanent component, the trend coefficient, and the seasonal effects.

You are not required to estimate seasonal effects. If your series does not have seasonal variation, you should use the simplest form of the Holt-Winters command. In this case, there are two damping parameters, one for the permanent component and one for the trend coefficient. You may specify the values of either or both of these parameters. If you do not specify values for

them, EViews will estimate them, again by finding the values that minimize the sum of squared forecasting errors.

If your series has seasonal effects, and the magnitudes of the effects do not grow along with the series, then you should use the Holt-Winters method with additive seasonals. In this case, there are three damping parameters, which you may specify or estimate, in any combination.

If your series has seasonal effects that grow along with the series, then you should use the Holt-Winters method with multiplicative seasonals. Again, there are three damping parameters which can be specified or estimated.

A good reference for more information about exponential smoothing is: Bowerman, B. and O'Connell, R. *Forecasting and Time Series*, Third Edition, Duxbury Press, North Scituate, Massachusetts, 1993.

How to Use Exponential Smoothing

To calculate a forecast using exponential smoothing, click **Procs/Exponential Smoothing** on the toolbar of the window for the series you want to forecast. You will see a dialog box. Your first choice is the smoothing method:

- | | |
|--|---|
| Single | calculates the smoothed series as a damping coefficient times the actual series plus 1 minus the damping coefficient times the lagged value of the smoothed series. The extrapolated smoothed series is a constant, equal to the last value of the smoothed series during the period when actual data on the underlying series are available. |
| Double | applies the process described above twice. The extrapolated series has a constant growth rate, equal to the growth of the smoothed series at the end of the data period. |
| Holt-Winters
No seasonal | uses a method with an explicit linear trend. The method computes recursive estimates of permanent and trend components, each with its own damping parameter. |
| Holt-Winters
Additive | uses the same method with additive seasonal effects. |
| Holt-Winters
Multiplicative | uses the same method with multiplicative seasonal effects. |

The next step is to specify whether you want to estimate or prescribe the values of the smoothing parameters. If you request estimation, EViews will calculate and use the least squares estimates. Don't be surprised if the estimated damping parameters are close to one--that is a sign that your series is close to a random walk, where the most recent value is the best estimate of future values. If you don't want to estimate any of the parameters, type in the value you wish to prescribe in the fields in the dialog box.

Next you should enter the name you want to give the smoothed series. You can leave it unchanged if you like the default, which is the series name with SM appended.

You can change the sample of observations for estimation if you do not want the default, which is the current workfile sample.

Finally, you can change the number of seasons per year from its default of 12 for monthly or 4 for quarterly. Thus you can work with unusual data with an undated workfile and enter the appropriate number in this field.

After you click OK, you will see a window with the results of the smoothing procedure. You will see the estimated parameter values, the sum of squared residuals, the root mean squared error of the forecast, and the mean and trend at the end of the sample that describes the post-sample smoothed forecast. In addition, the smoothed series will be in the workfile. It will run from the beginning of the data period to the last date in the workfile range. All of its values after the data period are forecasts.

Example: *Double Exponential Smoothing forecasts of uxcase*

! Highlight uxcase and double click to open the series window
! **[Procs]/Exponential Smoothing**
! **Smoothing Method:** Double
! **Smoothing Parameters:** Enter E for all
! **Smoothed Series:** UXCASESM
! **Estimation Sample:** 1988.01 1996.06
! **Cycle for seasonal:** 12
! **OK**

You should see the following output displayed in the series window:

Date: 09/18/97 Time: 20:35
Sample: 1988:01 1996:06
Included observations: 102
Method: Double Exponential
Original Series: UXCASE
Forecast Series: UXCASESM

Parameters:	Alpha	0.5320
Sum of Squared Residuals		6017742.
Root Mean Squared Error		242.8940
End of Period Levels:	Mean	17264.19
	Trend	88.50120

and the series UXCASESM has been added to the workfile. Highlight uxcase and uxcasesm and double click to open a group. Plot the two series and change the sample to 1988.01 1997.07.

Note: The double exponential smoothing model is equivalent to the Box-Jenkins ARIMA(0,1,1)

model. That is, it is equivalent to a MA(1) model for the first difference of a series.

Example: *Holt-Winters (no seasonal) forecasts of uxcase*

! Highlight uxcase and double click to open the series window
! **[Procs]/Exponential Smoothing**
! **Smoothing Method:**Holt-Winters- No Seasonal
! **Smoothing Parameters:** Enter E for all
! **Smoothed Series:** UXCASESM
! **Estimation Sample:** 1988.01 1996.06
! **Cycle for seasonal:** 12
! **OK**

You should see the following output displayed in the series window:

Date: 09/18/97 Time: 20:32
Sample: 1988:01 1996:06
Included observations: 102
Method: Holt-Winters No Seasonal
Original Series: UXCASE
Forecast Series: UXCASESM

Parameters:	Alpha	0.9400
	Beta	0.0000
Sum of Squared Residuals		4821891.
Root Mean Squared Error		217.4246
End of Period Levels:	Mean	17245.80
	Trend	61.13725

Example: *Holt-Winters(Additive seasonal) forecasts of uxcase*

! Highlight uxcase and double click to open the series window
! **[Procs]/Exponential Smoothing**
! **Smoothing Method:**Holt-Winters-Additive Seasonal
! **Smoothing Parameters:** Enter E for all
! **Smoothed Series:** UXCASESM
! **Estimation Sample:** 1988.01 1996.06
! **Cycle for seasonal:** 12
! **OK**

You should see the following output displayed in the series window:

Date: 09/18/97 Time: 20:46
Sample: 1988:01 1996:06
Included observations: 102
Method: Holt-Winters Additive Seasonal
Original Series: UXCASE

Forecast Series: UXCASESM

Parameters:	Alpha	0.8100
	Beta	0.0000
	Gamma	0.0000
Sum of Squared Residuals		4144597.
Root Mean Squared Error		201.5771
End of Period Levels:	Mean	17057.38
	Trend	47.69841
Seasonals:	1995:07	36.31746
	1995:08	-33.38095
	1995:09	-129.3294
	1995:10	-200.6528
	1995:11	-218.6012
	1995:12	-168.6746
	1996:01	-121.6171
	1996:02	29.68452
	1996:03	203.8611
	1996:04	185.9127
	1996:05	236.9643
	1996:06	179.5159

Box-Jenkins ARIMA models

Stationary Time Series

The Box-Jenkins methodology only applies to what are called *stationary* time series. Stationary time series have constant means and variances that do not vary over time. Stationary time series can be correlated, however, and it is this autocorrelation that the Box-Jenkins methods try to capture. Consequently, any time series that trends up or down is not stationary since the mean of the series depend on time. Two techniques are generally used to transform a trending series to stationary form: linear detrending (regressing on a time trend) and differencing. Box and Jenkins recommend differencing a time series, rather than detrending by regressing on a time trend, to remove the trend and achieve stationarity. This approach views the trend in a series as erratic and not very predictable.

Unit root tests

Unit root tests are used to test the null hypothesis that an observed trend in a time series is erratic and not predictable versus the alternative that the trend is smooth and deterministic. They can be viewed as a statistical test to see if a series should be differenced or linearly detrended to achieve stationarity.

ARIMA Models

An ARIMA model uses three tools for modeling the disturbance. The first is AR--autoregressive terms. An autoregressive model of order p has the form

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + e_t$$

The second tool is the integrated term. When a forecasting model contains an integrated term, it can handle a series that tends to drift over time. A pure first-order integrated process is called a random walk:

$$y_t = y_{t-1} + e_t$$

It is a good model for stock prices and some other financial variables. Notice that the first difference of the random walk is white noise. Each integrated term corresponds to differencing the series being forecast. A first-order integrated component means that the forecasting model is built for the first difference of the original series. A second-order component is for second differences, and so on. First differencing removes a linear trend from the data and second differencing removes a quadratic trend from the data and so on. Seasonal differencing (e.g., annually differencing monthly data) removes seasonal trends

The third tool is the moving average term. A moving average forecasting model uses lagged values of the forecast error to improve the current forecast. A first-order moving average term uses the most recent forecast error, a second-order term uses the forecast error from two periods ago, and so on. The algebraic statement of an MA(q) model is

$$y_t = e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \dots + \theta_q e_{t-q}$$

Some authors and software use the opposite sign convention for the q coefficients. The only time this might become a factor is in comparing EViews' results to results obtained with other software. The other results should be the same except for the signs of the moving-average coefficients.

Although econometricians typically use ARIMA as a model of the residuals from a regression model, the specification can also be applied directly to a series.

Box-Jenkins ARIMA modeling consists of several steps: identification, estimation, diagnostic checking and forecasting.

Identification

In this step, the tentative form of the ARIMA model is determined. The order of differencing is determined and the next step in building an ARIMA model for a series is to look at its autocorrelation properties. The way that current values of the residuals are correlated with past values is a guide to the type of ARIMA model you should estimate. For this purpose, you can use the **correlogram** view of a series which shows the *autocorrelations* and *partial autocorrelations* of the series.

The autocorrelations are easy to explain--each one is the correlation coefficient of the current value of the series with the series lagged a certain number of periods. The partial autocorrelations

are a little more complicated. They measure the correlation of the current and lagged series after taking account of the predictive power of all the values of the series with smaller lags.

The next step is to decide what kind of ARIMA model to use. If the autocorrelation function dies off smoothly at a geometric rate, and the partial autocorrelations were zero after one lag, then a first-order autoregressive model would be suggested. Alternatively, if the autocorrelations were zero after one lag and the partial autocorrelations declined geometrically, a first-order moving average process would come to mind.

Estimation

EViews estimates ARIMA models using conditional maximum likelihood estimation. EViews gives you complete control over the ARMA error process in your regression equation. In your estimation specification, you may include as many MA and AR terms as you want in the equation. For example, to estimate a second-order autoregressive and first-order moving average error process, you would use **AR(1) AR(2)**, and **MA(1)**.

You need not use the terms consecutively. For example, if you want to fit a fourth-order autoregressive model to take account of seasonal movements, you could use **AR(4)** by itself.

For a pure moving average specification, just use the MA terms.

The standard Box-Jenkins or ARMA model does not have any right-hand variables except for the constant. In this case, your specification would just contain a C in addition to the AR and MA terms, for example, **C AR(1) AR(2) MA(1) MA(2)**.

Diagnostic Checking

The goal of ARIMA analysis is a parsimonious representation of the process governing the residual. You should use only enough AR and MA terms to fit the properties of the residuals. After fitting a candidate ARIMA specification, you should check that there are no remaining autocorrelations that your model has not accounted for. Examine the autocorrelations and the partial autocorrelations of the innovations (the residuals from the ARIMA model) to see if any important forecasting power has been overlooked.

Forecasting

After you have estimated an equation with an ARMA error process, you can make forecasts from the model. EViews automatically makes use of the ARMA error model to improve the forecast.

A Digression on the Use of the Difference Operator in Eviews

The **D** operator is handy for differencing. To specify first differencing, simply include the series name in parentheses after **D**. For example **D(UXCASE)** specifies the first difference of **UXCASE** or **UXCASE-UXCASE(-1)**.

More complicated forms of differencing may be specified with two optional parameters, n and s. **D(X,n)** specifies n-th order difference of the series X. For example, **D(UXCASE,2)** specifies the second order difference of **UXCASE**, or **D(UXCASE)-D(UXCASE(-1))**, or **UXCASE - 2*UXCASE(-1) + UXCASE(-2)**.

D(X,n,s) specifies n-th order ordinary differencing of X with a seasonal difference at lag s. For example, **D(UXCASE,0,4)** specifies zero ordinary differencing with a seasonal difference at lag 4, or **UXCASE - UXCASE(-4)**.

There are two ways to estimate models with differenced data. First, you may generate a new differenced series and then estimate using the new data. For example, to incorporate a first-order integrated component into a model for **UXCASE**, that is to model **UXCASE** as an **ARIMA(1,1,1)** process, first use **Genr** to compute

DUXCASE=D(UXCASE)

and then fit an ARMA model to **DUXCASE**, for example in the equation window type

DUXCASE C AR(1) MA(1)

Alternatively, you may include the difference operator directly in the estimation specification. For example, in the equation window type

D(UXCASE) C AR(1) MA(1)

There is an important reason to prefer this latter method. If you define a new variable, such as **DUXCASE** above, and use it for estimation, then the forecasting procedure will make forecasts of **DUXCASE**. That is, you will get a forecast of the differenced series. If you are really interested in forecasts of the level variable, in this case **UXCASE**, you will have to manually undifference. However, if you estimate by including the difference operator in the estimation command, the forecasting procedure will automatically forecast the level variable, in this case **UXCASE**.

The difference operator may also be used in specifying exogenous variables and can be used in equations without ARMA terms. Simply include them after the endogenous variables as you would any simple series names. For example,

D(UXCASE,2) C D(INCOME,2) D(UXCASE(-1),2) D(UXCASE(-2),2) TIME

Example: ARIMA modeling of uxcase

Determining the order of differencing

First plot **uxcase**. Notice the trending behavior. Clearly, **uxcase** is not stationary. Next, take the first (monthly) difference of **uxcase**:

smpl 1973.07 2002.12
genr duxcase = d(uxcase)

Plot the first difference. The trend has been eliminated so there is no need to difference again. However, the mean of the first differences appears to have fallen after 1995.

Identification of ARMA model for first difference of uxcase

Plot the autocorrelations and partial autocorrelations:

[View]/Correlogram
Correlogram of: Level
Lags to include: 36
OK

Notice that there is not a lot of obvious short-run correlations in the data. There are some non-zero correlations at what look like seasonal frequencies. It is not clear what kind of ARMA model should be fit to this data.

Capturing Seasonality with seasonal dummy variables.

Try regressing the first difference of uxcase on a constant and seasonal dummies:

Quick/Estimate Equation
Upper box: d(uxcase) c seasons
Estimation method: Least squares
Sample: 1988:01 1996:06
OK

Notice that I used d(uxcase) instead of duxcase. This allows me to forecast uxcase from the estimated model instead of forecasting duxcase.

You should see the following estimation results

LS // Dependent Variable is D(UXCASE)
Date: 09/18/97 Time: 21:09
Sample: 1988:01 1996:06
Included observations: 102

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-56.77778	72.11312	-0.787343	0.4331
FEB	247.6667	101.9834	2.428501	0.0172
MARCH	259.8889	101.9834	2.548346	0.0125
APRIL	115.4444	101.9834	1.131993	0.2606
MAY	148.2222	101.9834	1.453396	0.1496
JUNE	51.66667	101.9834	0.506619	0.6137
JULY	-38.72222	105.1220	-0.368355	0.7135

AUG	34.77778	105.1220	0.330832	0.7415
SEPT	8.527778	105.1220	0.081123	0.9355
OCT	33.15278	105.1220	0.315374	0.7532
NOV	86.52778	105.1220	0.823117	0.4126
DEC	154.4028	105.1220	1.468795	0.1454

R-squared	0.171690	Mean dependent var	37.68627
Adjusted R-squared	0.070452	S.D. dependent var	224.3880
S.E. of regression	216.3394	Akaike info criterion	10.86383
Sum squared resid	4212245.	Schwarz criterion	11.17265
Log likelihood	-686.7869	F-statistic	1.695901
Durbin-Watson stat	2.268005	Prob(F-statistic)	0.086885

Note: The fit is not very good. Also, the constant in the regression for d(uxcase) will turn into the slope of the trend in the forecast for uxcase.

The residual correlogram does not show much significant correlation. The tentative ARIMA thus has no AR or MA terms in it!

Generating forecasts from the model:

[Forecast]

Forecast name: uxcasef

S.E. (Optional): se

Sample range for forecast: 1996:07 1997:07

Method: Dynamic

Output: Graph

OK

Actual: UXCASE Forecast: UXCASEF

Sample: 1996:07 1997:07

Include observations: 13

Root Mean Squared Error	645.6601
Mean Absolute Error	534.7671
Mean Absolute Percentage Error	3.207238
Theil Inequality Coefficient	0.018934
Bias Proportion	0.685996
Variance Proportion	0.004394
Covariance Proportion	0.309610

Notice that the forecast variable is UXCASE and not DUXCASE. Plotting the forecast shows that the ARIMA forecast does not pick up the drop in trend after 1995. Further notice that the forecast looks a lot like the Holt-Winters forecast with the additive seasonals.

Example: A Seasonal differenced model (see **Help/Search/Seasonal Terms in ARMA**)

Box and Jenkins suggest modeling seasonality by seasonally differencing a model. Therefore, let's take the annual (12th) difference of the duxcase (note: it does not matter what order you difference

- you can 12th difference first and then 1st difference to get the same result):

```
smpl 1973.07 2002.12
```

```
Genr d112uxcase = d(uxcase,1,12)
```

Now plot d112uxcase and compute the correlogram. Notice that plot of the series is greatly affected by the periods when uxcase increased unusually. It is not clear at all what to do with this series.