# Forecasting Methods / Métodos de Previsão

## Week 7

ISCTE - IUL, Gestão, Econ, Fin, Contab.

**Diana Aldea Mendes**

*diana.mendes@iscte.pt*

March 24, 2011

# Forecasting Methods

Forecasting with Linear Regression Models

1. Point forecasts
2. Interval forecasts
3. Evaluation the forecast performance: MAD, MSE, MAPE,

- If the regression model was validated, then we can use the model for:
  - estimate the mean value of $y$ for given $x$ (several experiences)
  - forecast a particular value of $y$ for given $x$ (one experience)

# Forecasting Methods

- Estimating the Population Mean Response when $x = x_i$

$$\text{Parameter:} \qquad E[y|x = x_i] = \beta_0 + \beta_1 x_i$$

- Point Estimate:

$$\hat{y}_i = \beta_0 + \beta_1 x_i$$

- $(1-\alpha)100\%$ Confidence Interval for $E[y|x = x_i] = \beta_0 + \beta_1 x_i$:

$$\widehat{y} \pm t_{\alpha/2} s \sqrt{\frac{1}{n} + \frac{(x_p - \overline{x})^2}{SS_{xx}}}$$

# Forecasting Methods

- Predicting a Future Outcome when $x = x_i$
- Future Outcome:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

- Point Prediction:

$$\hat{y}_i = \beta_0 + \beta_1 x_i$$

- $(1-\alpha)100\%$ Prediction Interval for $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$

$$\widehat{y} \pm t_{\alpha/2} s \sqrt{1 + \frac{1}{n} + \frac{(x_p - \overline{x})^2}{SS_{xx}}}$$

# Forecasting Methods

- Forecasting Performance (error): measures how accurate the forecast was
- For time period $t$:

**Forecast error = Actual value – Forecast value**

$$F_e = A_t - F_t$$

- **Mean Forecast Error** (MFE or **Bias**): the arithmetic sum of the errors (average deviation of forecast from actuals)
- **Mean Absolute Deviation** (MAD), Measures average absolute deviation of forecast from actuals ($T$ - the number of time periods)

$$MAD = \frac{\sum_{t=1}^{T} |A_t - F_t|}{T}$$

# Forecasting Methods

- **Mean Square Error** (MSE): measures variance of forecast error

$$MSE = \frac{\sum_{t=1}^{T} (A_t - F_t)^2}{T}$$

- **Mean Absolute Percentage Error** (MAPE): measures absolute error as a percentage of the forecast

$$MAPE = 100 \frac{\sum_{t=1}^{T} \left| \frac{A_t - F_t}{A_T} \right|}{T}$$

# Forecasting Methods - eviews

- Once you have estimated an equation, forecasting is simple: 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
- Optionally, you may give a name to the standard errors of the forecast;
- You must also specify the sample for the forecast. Normally this will be a period in the future, if you are doing true forecasting.
- You can also make a forecast for a historical period - which is useful for evaluating a model.

# Forecasting Methods - eviews

- 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.
- The two methods will always give identical results in the first period of a multiperiod forecast.
- Compute static forecasts if interest is in the performance of rolling one-step ahead forecasts. Otherwise compute dynamic forecasts.

# Forecasting Methods - eviews

- 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.

# Forecasting Methods - eviews

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

# Forecasting Methods - eviews

- Eviews creates two new series: **uxcasef** (the forecast values of uxcase) and **se** (the standard errors of the forecast).
- By default, Eviews sets **uxcasef** 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 $\beta$ to the $x$-variables.

# Forecasting Methods - eviews

- 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.
  - 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

# Forecasting Methods - eviews

- 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% 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.

# Forecasting Methods - eviews

- 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 - uxcasef
genr abserror = @ABS(error)
genr pcterror = error/uxcase
genr abspcterror = abserror/uxcase
```

- Highlight the series **uxcase, uxcasef error abserror pcterror** and **abspcterror**, double click and open a group. Plot the series

# Forecasting Methods - eviews

- 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.

$$\text{Root Mean Square Error} \quad RMSE = \sqrt{\sum \frac{e_t^2}{n}}$$

$$\text{Mean Absolute Value} \quad MAE = \sum \frac{|e_t|}{n}$$

$$\text{Mean Absolute Percentage Error} \quad MAPE = \frac{1}{n} \sum \frac{|e_t|}{y_t}$$

$$\text{Theil Inequality Coefficient} \quad U = \frac{\sqrt{\frac{1}{n}\sum (y_t - f_t)^2}}{\sqrt{\frac{1}{n}\sum f_t^2} + \sqrt{\frac{1}{n}\sum y_t^2}}$$

# Forecasting Methods - eviews

- 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 $\boxed{\textbf{bias + variance + covariance = 1}}$.
- 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.

# Forecasting Methods - eviews

- 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 = uxcasef - 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, 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.