Time Series Introduction

Natalia Khorunzhina

Copenhagen Business School

Contact Information and Office Hours

Natalia Khorunzhina

- Office: Porcelænshaven 16A, 2.67
- Email: nk.eco@cbs.dk
- For office hours see the announcement on Canvas

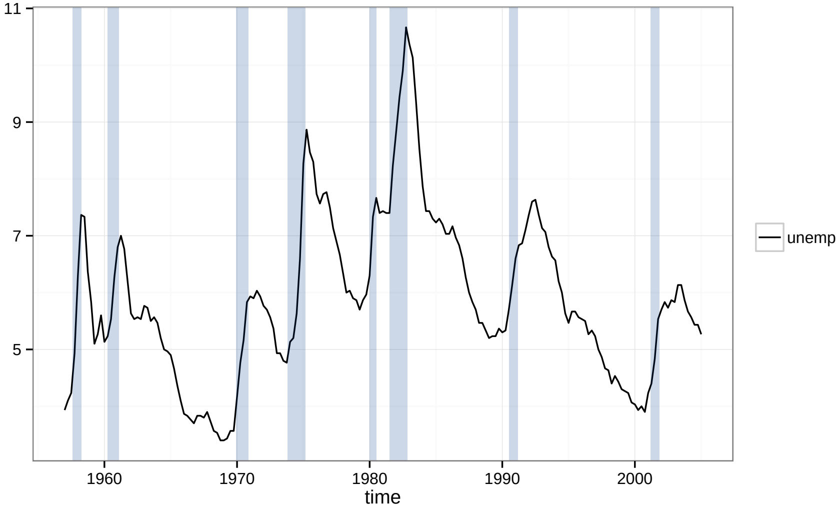
Timetable

Week 43		
Lecture 1	Introduction to Time-Series	Ch.15
Lecture 2	Stationarity / Nonstationarity	Ch.15
Week 44		
Lecture 3	ARIMA	Ch.15
Lecture 4	Estimation of Dynamic Causal Effects	Ch.16
Week 45		
Lecture 5	Vector Autoregressions	Ch.17
Lecture 6	ARCH and GARCH	Ch.17

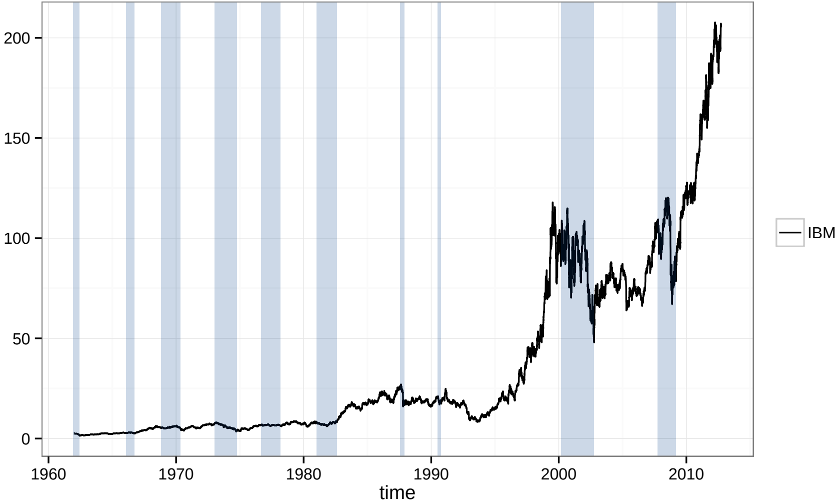
Introduction to Time Series

- Previously you have considered cross-sectional data
- **Time-series data** are data collected on the same observational unit at multiple time periods
- Although there are many cases where time series is similar to cross-sectional, there are some important differences.
- We'll learn how to handle time series data
- The first thing to do with any time series is plot it
- This will help us understand what kind of data we're dealing with
    - Unemployment rate, 50 years of quarterly observations
    - IBM Stock price, 50 years of daily data

Unemployment



IBM Stock Performance



Some uses of time series data

Y_t is the time series at time t, $t = 1, \ldots, T$. We will often use one series, sometimes two series (Y_t and X_t), rarely more than two series.

- Forecasting
- Estimation of dynamic causal effects
    - If the Fed increases the Federal Funds rate now, what will be the effect on the rates of inflation and unemployment in 3 months?
- Modeling risks, which is used in financial markets (one aspect of this, modeling changing variances and "volatility clustering"

Sources for time series data

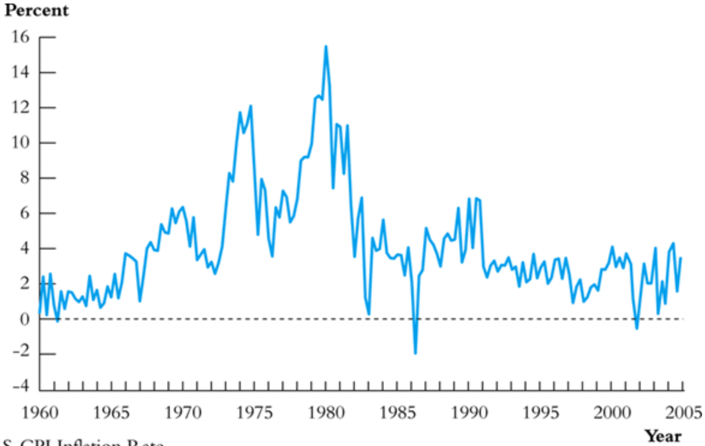
- Federal Reserve Bank of St. Louis Economic Database (FRED)
- Yahoo Finance
- Google Finance
- Oanda (FX and metals prices)
- Quandl (Commodities and Futures, even Bitcoins)
- TrueFX (FX tick data – intraday data)
- Bloomberg (Available from the Library)
- Datastream (Available from the Library)

Consumer Price Index (Bureau of Labor Statistics)

Quarterly rate of inflation at an annual rate (U.S.)

- CPI in the first quarter of 2004 (2004:I) = 186.57
- CPI in the second quarter of 2004 (2004:II) = 188.60
- Percentage change in CPI, 2004:I to 2004:II is $100 \times (188.60\text{-}186.57)/186.57 = 1.09\%$
- Like interest rates, inflation rates are (as a matter of convention) reported at an annual rate.
- Percentage change in CPI, 2004:I to 2004:II, at an annual rate $= 4 \times 1.09 = 4.4\%$ (percent per year)
- Using the logarithmic approximation to percent changes yields $4 \times 100 \times [\ln(188.60) - \ln(186.57)] = 4.3\%$

US Inflation Rate



(a) U.S. CPI Inflation Rate

Inflation and Its Lags

TABLE 14.1 Inflation in the United States in 2004 and the First Quarter of 2005

Quarter	U.S. CPI	Rate of Inflation at an Annual Rate (Inf_t)	First Lag (Inf_{t-1})	Change in Inflation (ΔInf_t)
2004:I	186.57	3.8	0.9	2.9
2004:II	188.60	4.4	3.8	0.6
2004:III	189.37	1.6	4.4	−2.8
2004:IV	191.03	3.5	1.6	1.9
2005:I	192.17	2.4	3.5	−1.1

The annualized rate of inflation is the percentage change in the CPI from the previous quarter to the current quarter, multiplied by 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.

Lags, First Differences, Logarithms, and Growth Rates

To summarize

- The first lag of Y_t is Y_{t-1}; its jth lag is Y_{t-j}
- The first difference of a series, ΔY, 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.

Autocorrelation

- A big difference between time series and cross-sectional data is autocorrelation
- A series exhibiting autocorrelation is related to its own past values
- First, recall the formula for the covariance

$$cov(x, y) = E(y - Ey)(x - Ex)$$

For time series, we use the autocovariance

$$\gamma_0 = cov(y_t, y_t) = var(y_t) = E(y_t - Ey_t)^2$$

$$\gamma_1 = cov(y_t, y_{t-1}) = E(y_t - Ey_t)(y_{t-1} - Ey_{t-1})$$

$$\gamma_j = cov(y_t, y_{t-j}) = E(y_t - Ey_t)(y_{t-j} - Ey_{t-j})$$

Autocorrelation

Properties of the Autocovariance

- $var(y_t) = \gamma_0$

Recall the correlation coefficient (between x and y)

$$\rho_{x,y} = \frac{cov(x,y)}{\sqrt{var(x)var(y)}}$$

$$-1 \leq \rho_{x,y} \leq 1$$

Autocorrelation

For time series

$$\rho_j = \frac{cov(y_t, y_{t-j})}{\sqrt{var(y_t)var(y_{t-j})}}$$

Assume for now that $var(y_t) = var(y_{t-1}) = \cdots = var(y_{t-j})$ (we'll see why later)

$$\rho_j = \frac{cov(y_t, y_{t-j})}{var(y_t)} \tag{1}$$

$$= \frac{\gamma_j}{\gamma_0} \tag{2}$$

Autocorrelation

Properties of the autocorrelation function (ACF)

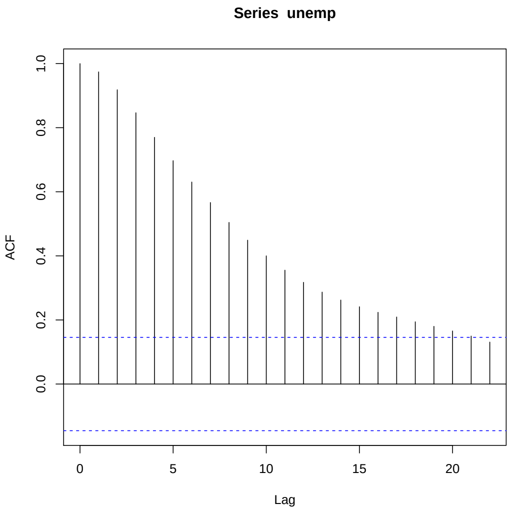
- $\rho_0 = \frac{\gamma_0}{\gamma_0} = 1$
- $\rho_{-j} = \rho_j$
- $-1 \leq \rho_j \leq 1$
- $\rho_j = 0$ if y_t is not serially correlated

Autocorrelations of U.S. inflation and change in the inflation rate

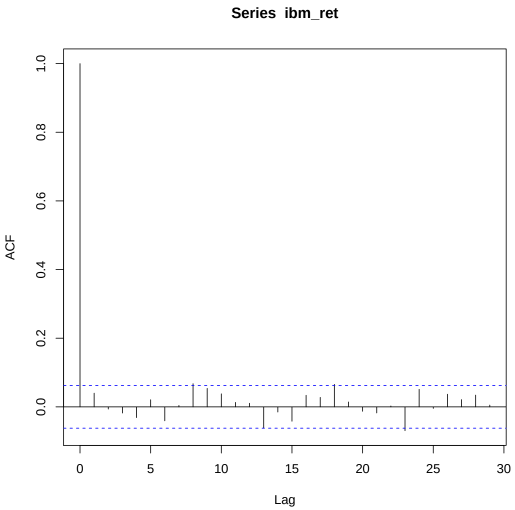
TABLE 14.2 First Four Sample Autocorrelations of the U.S. Inflation Rate and Its Change, 1960:I–2004:IV

	Autocorrelation of:	
Lag	**Inflation Rate (Inf_t)**	**Change of Inflation Rate (ΔInf_t)**
1	0.84	−0.26
2	0.76	−0.25
3	0.76	0.29
4	0.67	−0.06

ACF of unemployment



ACF of IBM returns



Introduction to Autoregressive AR processes

The AR(1) – autoregressive model of order 1

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

- In the AR(1), Y depends on 1 lag of its own past values
- The "1" in AR(1) stands for one lag
- You estimate autoregressive models just as you would a normal regression but you need an extra condition that we will discuss later

Example:

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

In the above model, the growth in inflation depends on 1 past lag

Example: AR(1) for Inflation Growth

$$\widehat{\Delta Inf}_t = 0.015 - 0.273 \Delta Inf_{t-1}$$

```
Time series regression with "ts" data:
Start = 1955(4), End = 2017(4)

Call:
dynlm(formula = dinfl ~ L(dinfl, 1))

Residuals:
     Min       1Q   Median       3Q      Max
-15.1433  -0.8295   0.0747   1.1312   6.6426

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.01478    0.12658   0.117    0.907
L(dinfl, 1) -0.27280    0.06123  -4.455 1.27e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.997 on 247 degrees of freedom
Multiple R-squared:  0.07438,	Adjusted R-squared:  0.07063
F-statistic: 19.85 on 1 and 247 DF,  p-value: 1.272e-05
```

Forecasts: terminology and notation

- Predicted values are "in-sample" (the usual definition)
- Forecasts are "out-of-sample" – predict future data
- Notation:
  - $Y_{T+1|T}$ = forecast of Y_{T+1} based on $Y_T, Y_{T-1}, \ldots$ using the population (true unknown) coefficients
  - $\hat{Y}_{T+1|T}$ = forecast of Y_{T+1} based on $Y_T, Y_{T-1}, \ldots$ using the estimated coefficients, which are estimated using data through period T.
- For AR(1):
  - $Y_{T+1|T} = \beta_0 + \beta_1 Y_T$
  - $\hat{Y}_{T+1|T} = \hat{\beta}_0 + \hat{\beta}_1 Y_T$, where $\hat{\beta}_0$ and $\hat{\beta}_1$ are estimated using data through period T

Forecasts of the AR(1) model

The forecast of the AR(1) (one-step ahead **point forecast** into the future)

$$\hat{Y}_{T+1|T} = \hat{\beta}_0 + \hat{\beta}_1 Y_T$$

The forecast error

$$\hat{e}_{T+1} = Y_{T+1} - \hat{Y}_{T+1|T}$$

Root mean squared forecast error (RMSFE):

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

If the error in estimation the coefficients β_0 and β_1 is small (will be so in large samples), then RMSFE is estimated by the standard error of the regression. We can construct 95% **interval forecast** as follows:

$$\hat{Y}_{T+1|T} \pm 1.96 SE(Y_{T+1} - \hat{Y}_{T+1|T})$$

Example: forecasting inflation using an AR(1)

$$\widehat{\Delta Inf}_t = 0.015 - 0.273 \Delta Inf_{t-1}$$

Last period in my data is 2017:IV. What is the forecast of inflation in the 2018:I?

$$\widehat{\Delta Inf}_{2018:I} = 0.015 - 0.273 \Delta Inf_{2017:IV} = -0.3$$

This forecast is still poor, because $\bar{R}^2 = 0.07$ and RMSFE=2.0
The predicted rate of inflation

$$\widehat{Inf}_{T+1|T} = Inf_T + \widehat{\Delta Inf}_{T+1|T}$$

in 2018:I is:

$$\widehat{Inf}_{2018:I} = 3.25 - 0.3 = 2.95$$

AR(p) - Autoregressive model of order p

Just like the AR(1), but has p lags:

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

Examples

AR(2): $Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + u_t$

AR(3): $Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \beta_3 Y_{t-3} + u_t$

The point forecast is constructed in a similar way:

$$\hat{Y}_{T+1|T} = \hat{\beta}_0 + \hat{\beta}_1 Y_T + \hat{\beta}_2 Y_{T-1} + \ldots + \hat{\beta}_p Y_{T-p+1}$$

If we want to forecast, we should be careful with large p models, because the error in estimation of the coefficients will be larger, strongly affecting RMSFE and forecast uncertainty.

Inflation Growth - AR(4)

$$\widehat{\Delta Inf}_t = 0.01 - 0.37\Delta Inf_{t-1} - 0.41\Delta Inf_{t-2} - 0.06\Delta Inf_{t-3} - 0.15\Delta Inf_{t-4}$$

```
Time series regression with "ts" data:
Start = 1956(3), End = 2017(4)

Call:
dynlm(formula = dinfl ~ L(dinfl, 1:4))

Residuals:
     Min       1Q   Median       3Q      Max
-14.3545  -0.7631   0.2202   0.9976   4.9513

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)     0.006596   0.118572   0.056    0.956
L(dinfl, 1:4)1 -0.372070   0.063432  -5.866 1.47e-08 ***
L(dinfl, 1:4)2 -0.413981   0.067745  -6.111 3.95e-09 ***
L(dinfl, 1:4)3 -0.065366   0.067887  -0.963    0.337
L(dinfl, 1:4)4 -0.155987   0.063649  -2.451    0.015 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.86 on 241 degrees of freedom
Multiple R-squared:  0.2107,	Adjusted R-squared:  0.1976
F-statistic: 16.09 on 4 and 241 DF,  p-value: 1.087e-11
```

Example: forecasting inflation using an AR(4)

$$\widehat{\Delta Inf}_t = 0.01 - 0.37\Delta Inf_{t-1} - 0.41\Delta Inf_{t-2} - 0.06\Delta Inf_{t-3} - 0.15\Delta Inf_{t-4}$$

What is the forecast of inflation in the 2018:I?

$$\begin{aligned}\widehat{\Delta Inf}_{2018:I} = 0.01 &- 0.37\Delta Inf_T - 0.41\Delta Inf_{T-1} \\ &- 0.06\Delta Inf_{T-2} - 0.15\Delta Inf_{T-3} = -1.1\end{aligned}$$

This forecast should be better, because $\bar{R}^2 = 0.20$ and RMSFE=1.86

The predicted rate of inflation

$$\widehat{Inf}_{T+1|T} = Inf_T + \widehat{\Delta Inf}_{T+1|T}$$

in 2018:I is:

$$\widehat{Inf}_{2018:I} = 3.25 - 1.1 = 2.15$$

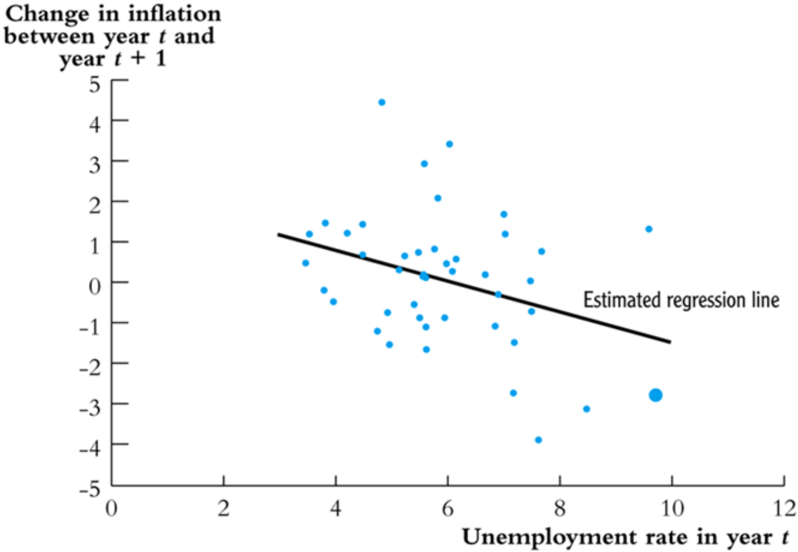
Autoregressive Distributed Lag Model

- Sometimes we may want to consider outside predictors
- For example, the Phillips Curve states that the unemployment is negatively related to the changes in inflation rate. So, we may want to use the unemployment rate
- This leads us to the Autoregressive Distributed Lag Model ADL(p,q)

$$
\begin{aligned}
Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \cdots + \beta_p Y_{t-p} \\
+ \delta_1 X_{t-1} + \delta_2 X_{t-2} + \cdots + \delta_q X_{t-q} + u_t
\end{aligned}
$$

where we assume that $E(u_t | Y_{t-1}, Y_{t-2}, \ldots, X_{t-1}, X_{t-2}, \ldots) = 0$. Thus the residuals must have no autocorrelation (must be white noise)

Phillips Curve



ADL(4,4) model for the Phillips Curve

$$\Delta Inf_t = 1.09 - 0.42\Delta Inf_{t-1} - 0.47\Delta Inf_{t-2} - 0.13\Delta Inf_{t-3} - 0.19\Delta Inf_{t-4}$$
$$- 1.94 Unemp_{t-1} + 2.53 Unemp_{t-2} - 1.03 Unemp_{t-3} + 0.26 Unemp_{t-4}$$

```
Time series regression with "ts" data:
Start = 1956(3), End = 2017(4)

Call:
dynlm(formula = dinfl ~ L(dinfl, 1:4) + L(uerate, 1:4))

Residuals:
     Min       1Q   Median       3Q      Max
-13.1584  -0.7690   0.0472   0.9283   6.0808

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)      1.09692    0.48532   2.260  0.02472 *
L(dinfl, 1:4)1  -0.42056    0.06364  -6.609 2.54e-10 ***
L(dinfl, 1:4)2  -0.47527    0.06843  -6.945 3.61e-11 ***
L(dinfl, 1:4)3  -0.13377    0.06817  -1.962  0.05091 .
L(dinfl, 1:4)4  -0.19193    0.06262  -3.065  0.00243 **
L(uerate, 1:4)1 -1.94435    0.42154  -4.613 6.51e-06 ***
L(uerate, 1:4)2  2.53464    0.81384   3.114  0.00207 **
L(uerate, 1:4)3 -1.03227    0.82632  -1.249  0.21281
L(uerate, 1:4)4  0.26082    0.43140   0.605  0.54604
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.777 on 237 degrees of freedom
Multiple R-squared:  0.2916,	Adjusted R-squared:  0.2677
F-statistic: 12.19 on 8 and 237 DF,  p-value: 1.393e-14
```

Example: forecasting inflation using an ADL(4,4)

What is the forecast of inflation in the 2018:I?

$$\begin{aligned}
\widehat{\Delta Inf}_{2018:I} &= 1.09 - 0.42\Delta Inf_T - 0.47\Delta Inf_{T-1} - 0.13\Delta Inf_{T-2} - 0.19\Delta Inf_{T-3} \\
&\quad - 1.94 Unemp_T + 2.53 Unemp_{T-1} - 1.03 Unemp_{T-2} + 0.26 Unemp_{T-3} \\
&= -0.29
\end{aligned}$$

This forecast should be even better, because $\bar{R}^2 = 0.27$ and RMSFE=1.77

The predicted rate of inflation

$$\widehat{Inf}_{T+1|T} = Inf_T + \widehat{\Delta Inf}_{T+1|T}$$

in 2018:I is:

$$\widehat{Inf}_{2018:I} = 3.25 - 0.29 = 2.96$$

Granger Causality Test

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**

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \cdots + \beta_k Y_{t-j} \\ + \gamma_1 X_{t-1} + \cdots + \gamma_k X_{t-k} + u_t$$

Null Hypothesis: $\gamma_1 = \cdots = \gamma_k = 0$ F-test.
From our inflation modeling example ADL(4,4): Does unemployment rate help explaining changes in inflation?

```
Hypothesis:
L(uerate,4)1 = 0
L(uerate,4)2 = 0
L(uerate,4)3 = 0
L(uerate,4)4 = 0

Model 1: restricted model
Model 2: dinfl ~ L(dinfl, 1:4) + L(uerate, 1:4)

  Res.Df    RSS Df Sum of Sq      F    Pr(>F)
1    241 833.41
2    237 748.04  4    85.375 6.7623 3.602e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Determining the order of an autoregression

Methods to determine the order of an autoregression

- The F/t statistic approach
- R^2
- Adjusted R^2
- The Bayes Information Criterion (BIC)
- The Akaike Information Criterion (AIC)

The F/t statistic approach

- Choose the order of an AR(p) based on whether or not the coefficients are statistically significant at, say, the 5 percent level
- This technique can often choose a model that is too large
- For example, suppose the true model is an AR(5); at the 5 percent level of significance, the t-statistic will incorrectly reject the null that the coefficient on the sixth lag is zero 5 percent of the time

R^2 and Adjusted R^2

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{SSR}{TSS}$$

$ESS = \sum_{i=1}^{T} (\hat{Y} - \bar{Y})^2$
$TSS = \sum_{i=1}^{T} (Y - \bar{Y})^2$
The R^2 is always increasing in the number of lags

$$\bar{R}^2 = 1 - \frac{T-1}{T-p-1} \frac{SSR}{TSS}$$

If T is large and p is small (as it often is for financial data), $\bar{R}^2$ will not penalize the addition of extra lag terms enough.

The Akaike Information Criterion (AIC)

$$AIC(p) = ln\left[\frac{SSR(p)}{T}\right] + (p+1)\frac{2}{T}$$

- The first term (the sum of squared residuals) measures how close the model is to the actual data and does not increase as you add more lags
- The second term is a penalty term that is increasing in p
- Akaike recommended selecting forecasting models by finding the one model with the smallest AIC.

The Bayes Information Criterion

$$BIC(p) = ln\left[\frac{SSR(p)}{T}\right] + (p+1)\frac{ln(T)}{T}$$

The BIC is similar to the AIC, but it adds a higher penalty as you add more lags.
As with the AIC, a lower score is better.

F/t statistics vs AIC vs BIC

In general, the model chosen by F/t statistics will be greater than or equal to that chosen by the AIC which will be greater than or equal to that chosen by the BIC.

AIC vs BIC vs R^2

Table: AIC vs BIC vs R^2

p	AIC	BIC	$\bar{R}^2$
1	1055.2	1065.7	0.074
2	1020.3	1034.4	0.191
3	1019.2	1036.8	0.190
4	1010.3	<u>1031.4</u>	0.211
5	1007.2	1031.9	0.215
6	1005.5	1033.6	0.217
7	1003.3	1035.0	0.221
8	<u>1001.5</u>	1036.7	<u>0.223</u>