Applied (Financial) Econometrics

Time series

Monique de Haan
(monique.dehaan@uva.nl)

Data

- The data file USMacro_Quarterly.dta contains quarterly data on several macroeconomic variables.
- Data are from the first quarter of 1957 to the fourth quarter of 2013.
- All data series are from the Federal Reserve Economic Data (FRED) database maintained by the Federal Reserve Bank of St. Louis.
- We will construct forecasting models for the rate of inflation.
- In the exercise we only consider the time period 1963:Q1-2012:Q4, but we use data from before 1963 to create lagged values.

Data

Variable Name	Description
GDPC96	Real Gross Domestic Product
JAPAN_IP	Production of Total Industry in Japan (FRED series name is JPNPROINDQISMEI)
PCECTPI	Personal Consumption Expenditures: Chain-type Price Index
GS10	10-Year Treasury Constant Maturity Rate (Quarterly Average of Monthly Values)
GS1	1-Year Treasury Constant Maturity Rate (Quarterly Average of Monthly Values)
TB3MS	3-Month Treasury Bill: Secondary Market Rate (Quarterly Average of Monthly Values)
UNRATE	Civilian Unemployment Rate (Quarterly Average of Monthly Values)
EXUSUK	U.S. / U.K. Foreign Exchange Rate (Quarterly Average of Daily Values)
CPIAUCSL	Consumer Price Index for All Urban Consumers: All Items (Quarterly Average of Monthly Values)

Part 1 and part 2

We have to.....

open the datafile and tell stata that we are using time series data

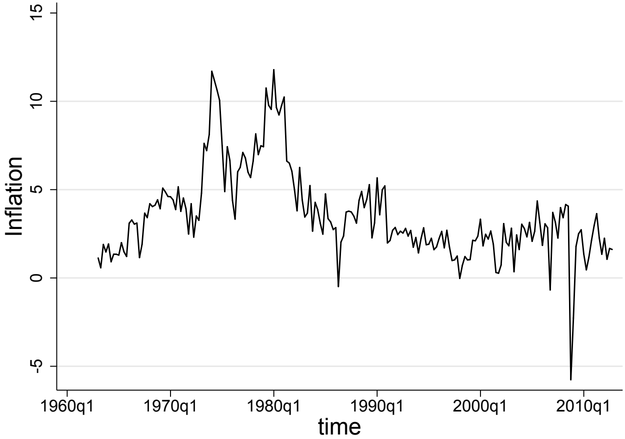
```
. *Part 1
. use USMacro_Quarterly.dta, clear

. ** Tell stata that we are using time series data
. tsset time
        time variable:  time, 1957q1 to 2013q4
                delta:  1 quarter
```

create the dependent variable

```
.
. *Part 2
. gen infl=400*(ln(PCECTPI)-ln(L1.PCECTPI))
(1 missing value generated)
```

What are the units of `infl`?

Part 3 and part 4



- Inflation increased over the 20-year period 1960-1980, then declined for a decade and has been reasonably stable since then.
- It appears to have a stochastic trend.

Part 5

The j^{th}autocorrelation$= \rho_j = \frac{Cov(infl_t, iflt-j)}{Var(ifl_t)}$ can be estimated by

$$\widehat{\rho}_j = \frac{\widehat{Cov(infl_t, infl_{t-j})}}{\widehat{Var\,(infl_t)}}$$

```
. corrgram infl if tin(1963q1,2012q4), noplot lags(4)
```

LAG	AC	PAC	Q	Prob>Q
1	**0.8260**	**0.8284**	**138.51**	**0.0000**
2	**0.7373**	**0.1824**	**249.43**	**0.0000**
3	**0.7124**	**0.2199**	**353.5**	**0.0000**
4	**0.6432**	**-0.0451**	**438.79**	**0.0000**

The first 4 autocorrelations are positive, reflecting positive serial correlation.

If inflation is high in one quarter it tends to be high in the next (four) quarter(s).

Part 6

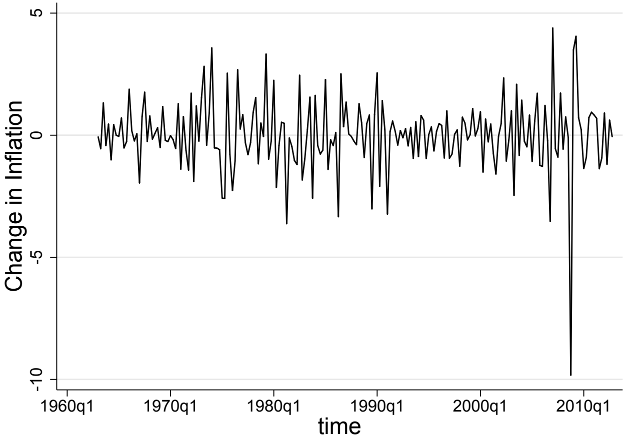
The j^{th}autocorrelation$= \rho_j = \frac{Cov(\triangle infl_t, \triangle iflt-j)}{Var(\triangle ifl_t)}$ can be estimated by

$$\widehat{\rho}_j = \frac{\widehat{Cov(\triangle infl_t, \triangle infl_{t-j})}}{\widehat{Var\left(\triangle infl_t\right)}}$$

```
. corrgram D.infl if tin(1963q1,2012q4), noplot lags(4)
```

LAG	AC	PAC	Q	Prob>Q
1	**-0.2531**	**-0.2531**	**13.005**	**0.0003**
2	**-0.1829**	**-0.2640**	**19.827**	**0.0000**
3	**0.1326**	**0.0094**	**23.435**	**0.0000**
4	**-0.0924**	**-0.1110**	**25.194**	**0.0000**

Part 7



- The change in inflation is slightly negatively serially correlated (the first autocorrelation is -0.25) so that values above the mean tend to be followed by values below the mean.

Part 8

$$AR(1): \quad \triangle infl_t = \beta_0 + \beta_1 \triangle infl_{t-1} + u_t$$

```
. regress D.infl L1.D.infl if tin(1963q1, 2012q4), robust

Linear regression                               Number of obs     =        200
                                                F(1, 198)         =      12.67
                                                Prob > F          =     0.0005
                                                R-squared         =     0.0641
                                                Root MSE          =     1.4549
```

D.infl	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
infl						
LD.	-.2531082	.0710979	-3.56	0.000	-.3933145	-.1129019
_cons	.0028786	.1028678	0.03	0.978	-.1999785	.2057357

Part 8

$$AR(1): \quad \triangle infl_t = \beta_0 + \beta_1 \triangle infl_{t-1} + u_t$$

```
. regress D.infl L1.D.infl if tin(1963q1, 2012q4), robust

Linear regression                               Number of obs     =        200
                                                F(1, 198)         =      12.67
                                                Prob > F          =     0.0005
                                                R-squared         =     0.0641
                                                Root MSE          =     1.4549
```

D.infl	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
infl LD.	-.2531082	.0710979	-3.56	0.000	-.3933145	-.1129019
_cons	.0028786	.1028678	0.03	0.978	-.1999785	.2057357

- The coefficient on $\triangle ifl_{t-1}$ is statistically significant at a 1% level, so that the lagged change in inflation helps predict the current change in inflation.

Part 9

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

```
. gen BIC_1=ln(e(rss)/e(N))+e(rank)*(ln(e(N))/e(N)) if tin(1963q1, 2012q4)
(28 missing values generated)

. gen AIC_1=ln(e(rss)/e(N))+e(rank)*(2/e(N)) if tin(1963q1, 2012q4)
(28 missing values generated)

. sum BIC_1 AIC_1
```

Variable	Obs	Mean	Std. Dev.	Min	Max
BIC_1	200	.7927663	0	.7927663	.7927663
AIC_1	200	.7597831	0	.7597831	.7597831

Part 10

$$AR(2): \quad \triangle infl_t = \beta_0 + \beta_1 \triangle infl_{t-1} + \beta_2 \triangle infl_{t-2} + u_t$$

```
. regress D.infl L1.D.infl L2D.infl if tin(1963q1, 2012q4), robust

Linear regression                               Number of obs     =        200
                                                F(2, 197)         =      12.65
                                                Prob > F          =     0.0000
                                                R-squared         =     0.1293
                                                Root MSE          =     1.4068
```

D.infl	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
infl						
LD.	-.3199827	.0668392	-4.79	0.000	-.4517949	-.1881704
L2D.	-.2641477	.0825592	-3.20	0.002	-.426961	-.1013344
_cons	.0026297	.0994816	0.03	0.979	-.193556	.1988153

Part 10

$$AR(2): \quad \triangle infl_t = \beta_0 + \beta_1 \triangle infl_{t-1} + \beta_2 \triangle infl_{t-2} + u_t$$

```
. regress D.infl L1.D.infl L2D.infl if tin(1963q1, 2012q4), robust

Linear regression                               Number of obs     =        200
                                                F(2, 197)         =      12.65
                                                Prob > F          =     0.0000
                                                R-squared         =     0.1293
                                                Root MSE          =     1.4068
```

D.infl	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
infl						
LD.	-.3199827	.0668392	-4.79	0.000	-.4517949	-.1881704
L2D.	-.2641477	.0825592	-3.20	0.002	-.426961	-.1013344
_cons	.0026297	.0994816	0.03	0.979	-.193556	.1988153

- The estimated coefficient on $\triangle infl_{t-2}$ is statistically significant, so the AR(2) model seems to be preferred to the AR(1) model.

Part 11

```
. gen BIC_2=ln(e(rss)/e(N))+e(rank)*(ln(e(N))/e(N)) if tin(1963q1, 2012q4)
(28 missing values generated)

. gen AIC_2=ln(e(rss)/e(N))+e(rank)*(2/e(N)) if tin(1963q1, 2012q4)
(28 missing values generated)

. sum AIC* BIC*
```

Variable	Obs	Mean	Std. Dev.	Min	Max
AIC_1	**200**	**.7597831**	**0**	**.7597831**	**.7597831**
AIC_2	**200**	**.6974945**	**0**	**.6974945**	**.6974945**
BIC_1	**200**	**.7927663**	**0**	**.7927663**	**.7927663**
BIC_2	**200**	**.7469693**	**0**	**.7469693**	**.7469693**

Part 12 & part 13

Variable	Obs	Mean	Std. Dev.	Min	Max
BIC_1	200	.7927663	0	.7927663	.7927663
BIC_2	200	.7469693	0	.7469693	.7469693
BIC_3	200	.7733617	0	.7733617	.7733617
BIC_4	200	.7874006	0	.7874006	.7874006
BIC_5	200	.8005706	0	.8005706	.8005706
BIC_6	200	.818479	0	.818479	.818479
AIC_1	200	.7597831	0	.7597831	.7597831
AIC_2	200	.6974945	0	.6974945	.6974945
AIC_3	200	.7073953	0	.7073953	.7073953
AIC_4	200	.7049426	0	.7049426	.7049426
AIC_5	200	.7016211	1.11e-16	.7016211	.7016211
AIC_6	200	.7030379	0	.7030379	.7030379

Part 12 & part 13

Variable	Obs	Mean	Std. Dev.	Min	Max
BIC_1	200	.7927663	0	.7927663	.7927663
BIC_2	200	.7469693	0	.7469693	.7469693
BIC_3	200	.7733617	0	.7733617	.7733617
BIC_4	200	.7874006	0	.7874006	.7874006
BIC_5	200	.8005706	0	.8005706	.8005706
BIC_6	200	.818479	0	.818479	.818479
AIC_1	200	.7597831	0	.7597831	.7597831
AIC_2	200	.6974945	0	.6974945	.6974945
AIC_3	200	.7073953	0	.7073953	.7073953
AIC_4	200	.7049426	0	.7049426	.7049426
AIC_5	200	.7016211	1.11e-16	.7016211	.7016211
AIC_6	200	.7030379	0	.7030379	.7030379

- Lag length on basis of BIC: 2
- Lag length on basis of AIC: 2

Part 14

$$\triangle infl_t = \beta_0 + \delta infl_{t-1} + \gamma_1 \triangle infl_{t-1} + \gamma_1 \triangle infl_{t-2} + u_t$$

$$H_0 : \delta = 0 \qquad vs \quad H_1 : \delta < 0$$

```
. regress D.infl L1.infl L1.D.infl L2.D.infl if tin(1963q1, 2012q4)
```

Source	SS	df	MS
Model	71.4633738	3	23.8211246
Residual	376.321249	196	1.92000637
Total	447.784622	199	2.25017398

Number of obs	=	200
F(3, 196)	=	12.41
Prob > F	=	0.0000
R-squared	=	0.1596
Adj R-squared	=	0.1467
Root MSE	=	1.3856

D.infl	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
infl						
L1.	-.109079	.0410631	-2.66	0.009	-.1900612	-.0280967
LD.	-.2547434	.0720035	-3.54	0.001	-.3967445	-.1127422
L2D.	-.2205139	.0696671	-3.17	0.002	-.3579073	-.0831204
_cons	.3901379	.1757289	2.22	0.028	.0435757	.7367002

Part 14

$$\triangle infl_t = \beta_0 + \delta infl_{t-1} + \gamma_1 \triangle infl_{t-1} + \gamma_1 \triangle infl_{t-2} + u_t$$

$$H_0 : \delta = 0 \qquad vs \quad H_1 : \delta < 0$$

```
. regress D.infl L1.infl L1.D.infl L2.D.infl if tin(1963q1, 2012q4)
```

Source	SS	df	MS
Model	71.4633738	3	23.8211246
Residual	376.321249	196	1.92000637
Total	447.784622	199	2.25017398

Number of obs	=	200
F(3, 196)	=	12.41
Prob > F	=	0.0000
R-squared	=	0.1596
Adj R-squared	=	0.1467
Root MSE	=	1.3856

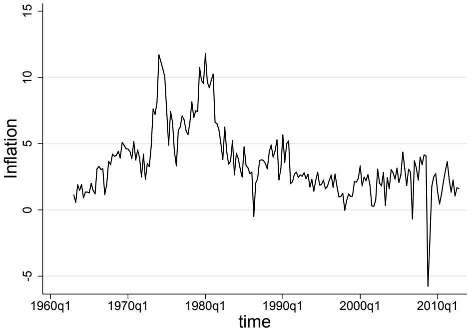
D.infl	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
infl						
L1.	-.109079	.0410631	-2.66	0.009	-.1900612	-.0280967
LD.	-.2547434	.0720035	-3.54	0.001	-.3967445	-.1127422
L2D.	-.2205139	.0696671	-3.17	0.002	-.3579073	-.0831204
_cons	.3901379	.1757289	2.22	0.028	.0435757	.7367002

$DF - statistic = -2.66$ The 10% critical value is -2.57 and the 5% critical value is -2.86; thus the unit root null hypothesis can be rejected at the 10% but not the 5% significance level.

Part 15

The Dickey-Fuller regression with a deterministic trend

$$\triangle Y_t = \beta_0 + {\color{blue}\alpha t} + \delta Y_{t-1} + \gamma_1 \triangle Y_{t-1} + u_t$$



The inflation rate does not exhibit a linear trend, so that the specification that includes an intercept, but no time trend is appropriate.

Part 16

Both the AIC and the BIC choose two lags, so the specification with 2 lags is appropriate.

Part 17

- Inflation is highly persistent.
- The null hypothesis of a unit root cannot be rejected at the 5% significance level, and given the precision of the estimate, this suggests that inflation indeed has a unit root.
- However we can never accept a null hypothesis!