

ECON-C4200 - Econometrics II: Capstone

Lecture 10: Time series III

Otto Toivanen

Learning outcomes

- At the end of lecture 10, you know
 - 1 what dynamic causal effects are
 - 2 what the assumptions behind such a model are
 - 3 how to estimate a dynamic causal model
 - 4 how to interpret the results of a dynamic causal model
 - 5 what **H**eteroskedasticity and **A**utoregression **C**onsistent (HAC) standard errors are
 - 6 what a **V**ector **A**utoregressive **R**egression (VAR) model is
 - 7 how to estimate a VAR
 - 8 how to estimate a VAR how to and interpret the results of a VAR model

Dynamic causal effect

- Think of macroeconomic data and questions.
- Example: The effect of a change in the ECB interest rate on inflation next month, in 6 months, in 12, in 24,... months.
- Often, we only have one "subject", say the country affected by ECB decisions.
- Difficult to think of an RCT.

Dynamic causal effect

- Substitute thought experiment: different treatments to the same (subject / observation unit) at different points in time.
- Treatment = e.g. level of interest rate t periods ago.
- If the different times are drawn from the same distribution (i.e., Y_t, X_t is stationary), then we can estimate the dynamic causal effects from an OLS regression of Y_t on X_t and its lags.
- Such an estimator is called a **distributed lag** estimator.

The distributed lag model

$$Y_t = \beta_0 + \beta_1 X_t + \beta_2 X_{t-1} + \dots + \beta_{t-p+1} X_{t-p} + u_t$$

- Interpretation of β_1 : The **impact effect of change in X_t**
- Notice that we are keeping the **past** constant!
- β_2 : The **1-period dynamic multiplier**. The effect of a change in X_{t-1} on Y_t , keeping $X_t, X_{t-2}, \dots, X_{t-p}$ constant.
- β_3 : The **2-period dynamic multiplier**. The effect of a change in X_{t-2} on Y_t , keeping $X_t, X_{t-1}, X_{t-3}, \dots, X_{t-p}$ constant.
- **j-period cumulative dynamic multiplier** = $\sum_{k=1}^{j-1} \beta_k$
- Example: 3-period cumulative dynamic multiplier = $\beta_1 + \beta_2 + \beta_3 + \beta_4$
...that is, the effect of all X_j for $j = t, t-1, t-2, t-3$.

The distributed lag model: assumptions

$$Y_t = \beta_0 + \beta_1 X_t + \beta_2 X_{t-1} \dots + \beta_{t-p+1} X_{t-p} + u_t$$

- **Exogeneity:** $\mathbb{E}[u_t | X_t, X_{t-1}, \dots] = 0$
- **Strict exogeneity:** $\mathbb{E}[u_t | \dots, X_{t+1}, X_t, X_{t-1}, \dots] = 0$
- Notice that if X is strictly exogenous, then it is also exogenous.
- We will proceed with assuming exogeneity.

The distributed lag model: assumptions

A1 $\mathbb{E}[u_t | X_t, X_{t-1}, \dots] = 0$

A2 i) Y, X are stationary and ii) (Y_t, X_t) and (Y_{t-j}, X_{t-j}) become independent as j grows large.

A3 (Y_t, X_t) have 8 nonzero finite moments; large outliers are unlikely.

A4 There is no perfect multicollinearity.

The distributed lag model: assumptions

- Assumptions 1 and 4 are familiar.
- Assumption 2 is different. The difference is due to us studying time series.
- Assumption 3 is familiar otherwise, but now need 8 finite moments. This is to ensure "good" (=HAC) standard errors.

The distributed lag model: assumptions

- Assumption 2.i yields
 - ① internal validity
 - ② external validity
 - ③ This is the time series equivalent of the "identically distributed" of the cross-section i.i.d assumption
- Assumption 2.ii
 - ① enables the use of a version of CLT
 - ② This is the time series equivalent of the "independently distributed" of the cross-section i.i.d assumption
- The intuition is that when the time periods are sufficiently separated, we have independent experiments.

Heteroskedasticity and Autocorrelation Consistent (HAC) standard errors

- Under the above assumptions a distributed lag model, estimated via OLS, yields consistent estimates of the coefficients β_k .
- What about the standard errors? The variance of the sampling distribution is different as X_t and u_t may be autocorrelated (think of what the equivalent would mean in a cross-section context).
- \rightarrow need standard errors that are robust not only to heteroskedasticity but also to autocorrelation.

Heteroskedasticity and Autocorrelation Consistent (HAC) standard errors

- In (well-behaving) cross-section data, the variance of β_1 in a univariate regression is

$$\text{var}(\hat{\beta}_1) = \frac{\sigma_u^2}{T(\sigma_X^2)^2}$$

- Let's consider a time-series with $T = 2$, with $\text{cov}(u_t, u_{t-1}) \neq 0$:

$$\begin{aligned}\text{var}\left(\frac{1}{T} \sum_{t=1}^T u_t\right) &= \text{var}[0.5(u_1 + u_2)] \\ &= 0.25[\text{var}(u_1) + \text{var}(u_2) + 2\text{cov}(u_1, u_2)] \\ &= 0.5(\sigma_u^2 + \rho)\end{aligned}$$

- Autocorrelation means that the usual variance formula for the error term does not apply.

Example: A (sort-of) Phillips curve

$$\text{infl}_t = \beta_0 + \beta_1 \delta UE_t + \beta_2 \delta UE_{t-1} + u_t$$

- Let's estimate the model with regular and HAC s.e.

Sort-of Phillips curve

```
. regress dlnpind_euro d.lnue_euro d2.lnue_euro, robust
```

```
Linear regression                Number of obs   =      187
                                F(2, 184)       =      0.23
                                Prob > F            =     0.7956
                                R-squared           =     0.0030
                                Root MSE        =     .0041
```

dlnpind_euro	Robust					
	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
lnue_euro						
D1.	-.0234946	.0368771	-0.64	0.525	-.0962509	.0492617
D2.	.0059077	.0289012	0.20	0.838	-.0511127	.062928
_cons	.001453	.0002997	4.85	0.000	.0008616	.0020444

```
. newey dlnpind_euro d.lnue_euro d2.lnue_euro, lag(1)
```

```
Regression with Newey-West standard errors    Number of obs   =      187
maximum lag: 1                               F( 2, 184)     =      0.22
                                              Prob > F       =     0.8056
```

dlnpind_euro	Newey-West					
	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
lnue_euro						
D1.	-.0234946	.037917	-0.62	0.536	-.0983027	.0513134
D2.	.0059077	.0276402	0.21	0.831	-.0486248	.0604402
_cons	.001453	.0003067	4.74	0.000	.0008479	.0020581

Estimation of the distributed lag model

$$Y_t = \beta_0 + \beta_1 X_t + \beta_2 X_{t-1} \dots + \beta_{t-p+1} X_{t-p} + u_t$$

- You can estimate the model with OLS.
- You directly get the impact effect of X and the k -period dynamic multipliers & their standard errors
- You do not get the cumulative multipliers directly, nor their standard errors.

Estimating the cumulative dynamic multipliers

$$Y_t = \beta_0 + \beta_1 X_t + \beta_2 X_{t-1} \dots + \beta_{t-p+1} X_{t-p} + u_t$$

- Option 1: you estimate the above specification and then use post-estimation commands to calculate the cumulative dynamic multipliers and their standard errors.

Cumulative dynamic multiplier

```
. newey dlnpind_euro d.lnue_euro d2.lnue_euro, lag(3)
```

```
Regression with Newey-West standard errors      Number of obs   =       187  
maximum lag: 3                                F( 2,          184) =       0.23  
                                              Prob > F        =       0.7916
```

dlnpind_euro	Newey-West		t	P> t	[95% Conf. Interval]	
	Coef.	Std. Err.				
lnue_euro						
D1.	-.0234946	.0356168	-0.66	0.510	-.0937645	.0467753
D2.	.0059077	.0245427	0.24	0.810	-.0425137	.054329
_cons	.001453	.0002694	5.39	0.000	.0009215	.0019845

```
. scalar cdm = _b[d.lnue_euro] + _b[d2.lnue_euro]
```

```
. scalar list cdm  
      cdm = -.01758694
```

```
. testnl _b[d.lnue_euro] + _b[d2.lnue_euro] = 0
```

```
(1) _b[d.lnue_euro] + _b[d2.lnue_euro] = 0
```

```
      chi2(1) =       0.37  
      Prob > chi2 =     0.5450
```


Estimating the cumulative dynamic multipliers

- Option 2: Transform the model to yield direct estimates of the cumulative dynamic multipliers.
- Example: A 1-lag distributed lag model:

$$\begin{aligned} Y_t &= \beta_0 + \beta_1 X_t + \beta_2 X_{t-1} + u_t \\ &= \beta_0 + \beta_1 X_t - \beta_1 X_{t-1} + \beta_2 X_{t-1} + \beta_1 X_{t-1} + u_t \\ &= \beta_0 + \beta_1 (X_t - X_{t-1}) + (\beta_2 + \beta_1) X_{t-1} + u_t \\ &= \beta_0 + \beta_1 \Delta X_t + (\beta_2 + \beta_1) X_{t-1} + u_t \end{aligned}$$

- Same trick works for higher order distributed lag models.

Cumulative dynamic multiplier

```
. gen diff = d.lnue_euro - d2.lnue_euro  
(6 missing values generated)  
  
. newey dlnpind_euro diff d2.lnue_euro, lag(3)
```

```
Regression with Newey-West standard errors      Number of obs   =      187  
maximum lag: 3                                F( 2,          184) =      0.23  
                                                Prob > F         =      0.7916
```

dlnpind_euro	Newey-West		t	P> t	[95% Conf. Interval]	
	Coef.	Std. Err.				
diff	-.0234946	.0356168	-0.66	0.510	-.0937645	.0467753
lnue_euro D2.	-.0175869	.0290564	-0.61	0.546	-.0749134	.0397396
_cons	.001453	.0002694	5.39	0.000	.0009215	.0019845

Distributed lag model with autocorrelation

$$Y_t = \beta_0 + \beta_1 X_t + \beta_2 X_{t-1} \dots + \beta_{t-p+1} X_{t-p} + u_t$$

- Now assume that

$$u_t = \tilde{u}_t + \phi u_{t-1}$$

in other words, u_t is autocorrelated (AR(1)).

- Besides using HAC standard errors, we could consider transforming the regression equation so that the error term in the transformed equation is i.i.d.
- As an example, let's work with the following distributed lag model:

$$Y_t = \beta_0 + \beta_1 X_t + \beta_2 X_{t-1} + u_t$$

Estimating the cumulative dynamic multipliers

- We can proceed as follows:

$$\begin{aligned} Y_t &= \beta_0 + \beta_1 X_t + \beta_2 X_{t-1} + u_t \\ Y_t - \phi Y_{t-1} &= \beta_0 + \beta_1 X_t + \beta_2 X_{t-1} + u_t - \phi Y_{t-1} \\ &= \beta_0 + \beta_1 X_t + \beta_2 X_{t-1} + u_t \\ &\quad - \underbrace{\phi(\beta_0 + \beta_1 X_{t-1} + \beta_2 X_{t-2} + u_{t-1})}_{= Y_{t-1}} \\ &= (\beta_0 - \phi\beta_0) + \beta_1 X_t + (\beta_2 - \phi\beta_1) X_{t-1} - \phi\beta_2 X_{t-2} + \tilde{u}_t \\ &\quad + \phi u_{t-1} - \phi u_{t-1} \\ Y_t &= \delta_0 + \phi Y_{t-1} + \delta_1 X_t + \delta_2 X_{t-1} + \delta_3 X_{t-2} + \tilde{u}_t \end{aligned}$$

- The outcome is an ADL(1,2) model with i.i.d error term.

Vector autoregressive (VAR) models

- What if you want to model two or more series simultaneously?
- What if they interact, i.e., the past values of Y affect current value of X and vice versa?
- VAR models are designed for these situations. Below an example:

$$Y_t = \beta_{10} + \beta_{11} Y_{t-1} + \beta_{12} Y_{t-2} + \gamma_{11} X_{t-1} + \gamma_{12} X_{t-2} + u_{1t}$$

$$X_t = \beta_{20} + \beta_{21} X_{t-1} + \beta_{22} X_{t-2} + \gamma_{21} Y_{t-1} + \gamma_{22} Y_{t-2} + u_{2t}$$

Vector autoregressive (VAR) models

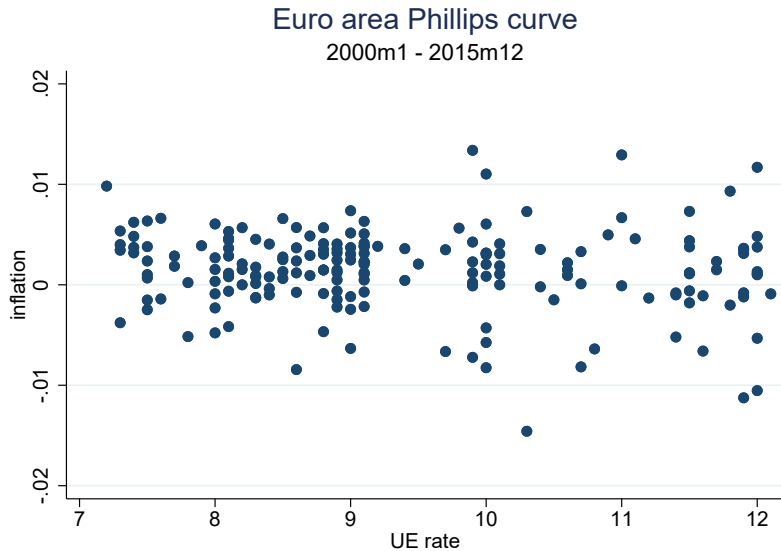
- Without any constraints on the coefficients, one could estimate the VAR equation by equation.
- Often, theory suggests some constraints on the coefficients, e.g., in macroeconomic models.
- Example of VAR: The Euro-area Phillips curve.
- We use data from 2000m1 - 2015m12.

Plotting the Phillips curve

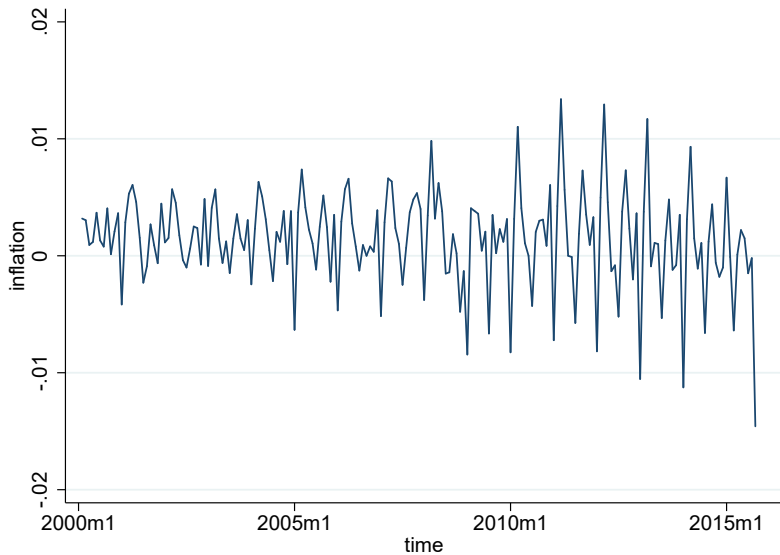
Stata code

```
1 use "mytimeseries.dta", clear
2 tsset time_ind
3 label var dlnpind_euro "inflation"
4 label var ue_euro "UE rate"
5 scatter dlnpind_euro ue_euro , ///
6     title("Euro area Phillips curve") ///
7     subtitle("2000m1 - 2015m12") ///
8     graphregion(color(white)) bgcolor(white)
9
10 graph export "L_11_10_4_phillips.pdf", replace
```

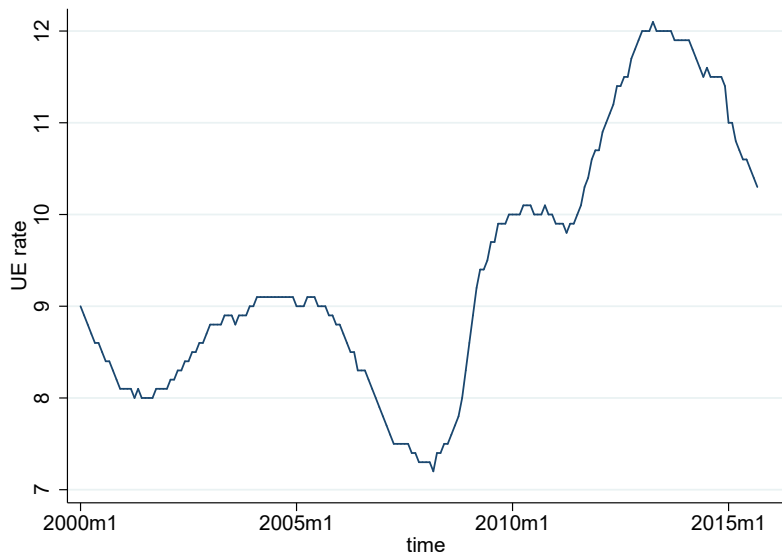
Plotting the Phillips curve



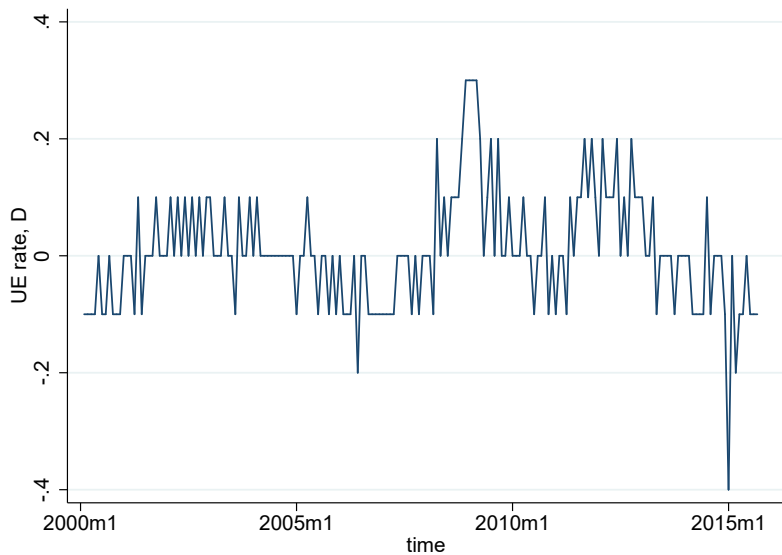
Inflation



Unemployment



Change in unemployment



VAR on inflation and UE

Stata code

```
1 use "mytimeseries.dta", clear
2 forvalues t = 1/10{
3     dfuller ue_euro, lags('t')
4 }
5 forvalues t = 1/10{
6     dfuller d.ue_euro, lags('t')
7 }
8 forvalues t = 1/10{
9     dfuller dlnpind_euro, lags('t')
10 }
```

DF tests

```
. forvalues t = 1/10{  
2.     dfuller ue_euro, lags(`t')  
3. }
```

Augmented Dickey-Fuller test for unit root Number of obs = 187

Test Statistic	Interpolated Dickey-Fuller			
	1% Critical Value	5% Critical Value	10% Critical Value	
Z(t)	-0.724	-3.481	-2.884	-2.574

MacKinnon approximate p-value for Z(t) = 0.8405

```
. forvalues t = 1/10{  
2.     dfuller d.ue_euro, lags(`t')  
3. }
```

Augmented Dickey-Fuller test for unit root Number of obs = 186

Test Statistic	Interpolated Dickey-Fuller			
	1% Critical Value	5% Critical Value	10% Critical Value	
Z(t)	-4.188	-3.481	-2.884	-2.574

MacKinnon approximate p-value for Z(t) = 0.0007

DF tests

```
. forvalues t = 1/10{  
2.     dfuller dlnpind_euro, lags(`t')  
3.     }
```

Augmented Dickey-Fuller test for unit root Number of obs = 186

Test Statistic	Interpolated Dickey-Fuller			
	1% Critical Value	5% Critical Value	10% Critical Value	
Z(t)	-11.302	-3.481	-2.884	-2.574

MacKinnon approximate p-value for Z(t) = 0.0000

VAR on inflation and UE

Stata code

```
1 var dlnpind_euro d.ue_euro
2 var dlnpind_euro d.ue_euro , lags(1 2 3)
3 var dlnpind_euro d.ue_euro
4 fcast compute f_ , step(12)
5 fcast graph f_dlnpind_euro , ///
6     title("Forecast for Euro area inflation") ///
7     graphregion(color(white)) bgcolor(white)
8     graph export "L-II-10-8-fcast-inflation.pdf" , replace
9
10 fcast graph f_D_ue_euro , ///
11     title("Forecast for Euro area UE") ///
12     graphregion(color(white)) bgcolor(white)
13     graph export "L-II-10-9-fcast_ue.pdf" , replace
```

VAR estimation results

- HQIC = Hannan-Quick information criterion

```
. var dlnpind_euro d.ue_euro
```

```
Vector autoregression
```

```
Sample: 2000m4 - 2015m9           Number of obs   =       186
Log likelihood =    978.682         AIC              =   -10.41594
FPE           =    1.03e-07        HQIC             =   -10.34566
Det(Sigma_ml) =    9.22e-08        SBIC            =   -10.24251
```

Equation	Parms	RMSE	R-sq	chi2	P>chi2
dlnpind_euro	5	.004021	0.0567	11.18402	0.0246
D_ue_euro	5	.077844	0.3899	118.8916	0.0000

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
dlnpind_euro						
dlnpind_euro						
L1.	.0607082	.0745786	0.81	0.416	-.0854631	.2068795
L2.	-.2318106	.0746343	-3.11	0.002	-.3780911	-.08553
ue_euro						
L1.	-.0019017	.0033073	-0.57	0.565	-.008384	.0045806
L2D.	.0024801	.0033123	0.75	0.454	-.0040119	.0089722
_cons	.0016858	.0003333	5.06	0.000	.0010326	.002339
D_ue_euro						
dlnpind_euro						
L1.	1.173706	1.443759	0.81	0.416	-1.65601	4.003421
L2.	-1.1060152	1.444838	-0.07	0.942	-2.937846	2.725816
ue_euro						
L1.	.2300881	.0640266	3.59	0.000	.1045982	.355578
L2D.	.4883861	.064123	7.62	0.000	.3627074	.6140648
_cons	.0006486	.0064515	0.10	0.920	-.0119962	.0132933

VAR estimation results

```
. var dlnpind_euro d_ue_euro, lags(1 2 3)
```

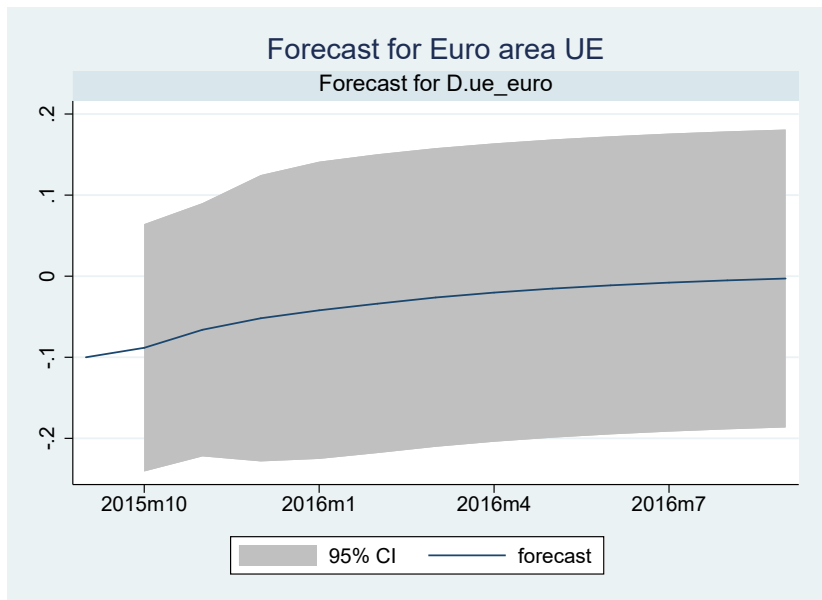
Vector autoregression

```
Sample: 2000m5 - 2015m9           Number of obs   =       185
Log likelihood = 975.7829          AIC              = -10.39765
FFE           = 1.05e-07           HQIC            = -10.29889
Det(Sigma_ml) = 8.99e-08          SBIC           = -10.15395
```

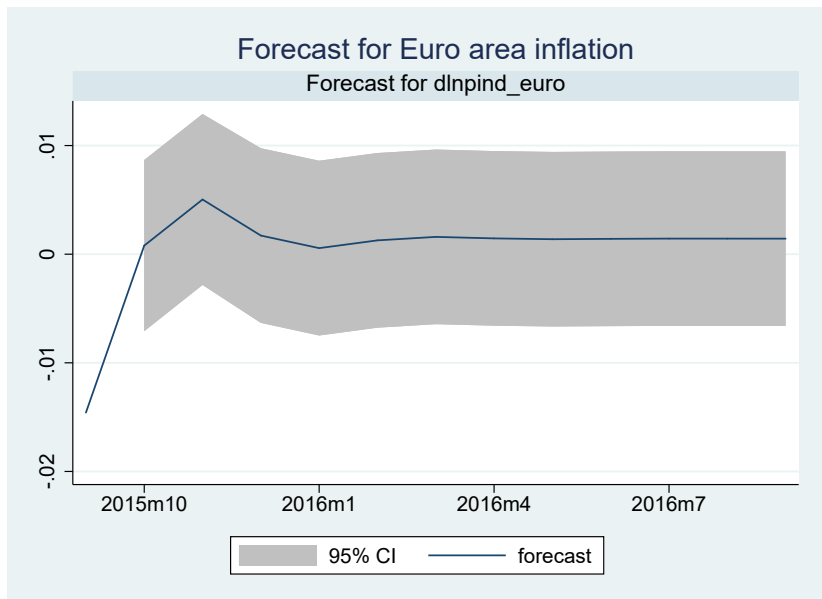
Equation	Farms	RMSE	R-sq	chi2	P>chi2
dlnpind_euro	7	.004039	0.0641	12.67015	0.0486
D_ue_euro	7	.077409	0.4029	124.8098	0.0000

	Coeff.	Std. Err.	z	P> z	[95% Conf. Interval]
dlnpind_euro					
dlnpind_euro					
L1.	.03993011	.0773214	0.51	0.611	-.112246 .1908483
L2.	-.2281727	.0747801	-3.05	0.002	-.3747389 -.0816064
L3.	-.0917675	.0771614	-1.19	0.234	-.2430011 .059466
ue_euro					
L1.	-.0015937	.0038003	-0.42	0.675	-.0090422 .0058548
L2.	.0027843	.0034209	0.81	0.416	-.0039206 .0094892
L3.	-.0011178	.0038034	-0.29	0.769	-.0085723 .0063366
_cons	.0018605	.0003637	5.12	0.000	.0011478 .0025733
D_ue_euro					
dlnpind_euro					
L1.	.6617554	1.481956	0.45	0.655	-2.242825 3.566336
L2.	.1285749	1.433248	0.09	0.929	-2.68054 2.93769
L3.	-.9287684	1.478889	-0.63	0.530	-3.827338 1.969801
ue_euro					
L1.	.1499964	.0728378	2.06	0.039	.0072368 .2927559
L2.	.452871	.0655663	6.91	0.000	.3243634 .5813785
L3.	.1552676	.0728961	2.13	0.033	.0123939 .2981414
_cons	.0023387	.0069703	0.34	0.737	-.0113229 .0160004

VAR estimation results



VAR estimation results



Granger causality

- Do lagged values of X predict Y_t , or the other way round?
- H_0 : the coefficients on all the lagged values of X in the Y regression are jointly insignificant.
- Granger causality \neq causality.
- Does it make sense to assume the future does not affect the past?

VAR on inflation and UE

Stata code

```
1 var dlnpind_euro d.ue_euro  
2 vargranger
```

Testing Granger causality

```
. vargranger
```

```
Granger causality Wald tests
```

Equation	Excluded	chi2	df	Prob > chi2
dlnpind_euro	D.ue_euro	.63417	2	0.728
dlnpind_euro	ALL	.63417	2	0.728
D_ue_euro	dlnpind_euro	.66253	2	0.718
D_ue_euro	ALL	.66253	2	0.718