Multivariate Time Series - 2

22.05

Vector autoregressive models

The autoregressive moving average models can be extended to the multivariate case.

Then the stochastic process that generates the time series of a *vector* of variables is modelled.

A **vector autoregressive** (**VAR**) model describes the dynamic evolution of a number of variables from their common history.

For two variables, Y_t and X_t , the VAR consists of two equations.

A first order VAR would be given by

AR(1) part

$$Y_t = \delta_1 + \theta_{11} Y_{t-1} + \theta_{12} X_{t-1} + \varepsilon_{1t} \qquad (9.29)$$

$$X_t = \delta_2 + \theta_{21} Y_{t-1} + \theta_{22} X_{t-1} + \varepsilon_{2t}, \qquad (9.30)$$

! Models should be constructed from X and Y which are stationary

Assumption: all the regressors are endogenous (we are not sure if X depends on Y or Y depends on X, there may be several directions of the dependence etc.)

Vector autoregressive models

The most important feature of the model is that ε_{1t} and ε_{2t}, two white noise processes (independent of the history of Y and X), **may be correlated**.

If, for example, $\theta_{12} \neq 0$ it means that the history of X helps explaining Y.

A compact way of writing (9.29)–(9.30) is

AR(p) model in a vector form

$$\begin{pmatrix} Y_t \\ X_t \end{pmatrix} = \begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix} + \begin{pmatrix} \theta_{11} & \theta_{12} \\ \theta_{21} & \theta_{22} \end{pmatrix} \begin{pmatrix} Y_{t-1} \\ X_{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \qquad (9.31)$$

or, with appropriate definitions, as

vector of coefficients

$$\vec{Y}_t = \delta + \Theta_1 \vec{Y}_{t-1} + \vec{\varepsilon}_t, \qquad (9.32)$$

where $\vec{Y}_t = (Y_t, X_t)'$ and $\vec{\varepsilon}_t = (\varepsilon_{1t}, \varepsilon_{2t})'$.

Assumption: all the regressors are endogenous (we are not sure if X depends on Y or Y depends on X, there may be several directions of the dependence etc.)

Vector autoregressive models: examples

$$\begin{pmatrix} Y_t \\ X_t \end{pmatrix} = \begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix} + \begin{pmatrix} \theta_{11} & \theta_{12} \\ \theta_{21} & \theta_{22} \end{pmatrix} \begin{pmatrix} Y_{t-1} \\ X_{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}$$

RW: $\Theta_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

White noise plus a constant: $\Theta_1 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

Two independent time series: $\Theta_1 = \begin{pmatrix} \theta_{11} & 0 \\ 0 & \theta_{22} \end{pmatrix}$

First time series depends on the second, where as there is no reverse causality: $\Theta_1 = \begin{pmatrix} \theta_{11} & \theta_{12} \\ 0 & \theta_{22} \end{pmatrix}$

Vector autoregressive models: examples

VAR estimation for bivariate model with two lags

	Coef.	Std. Err.	z	P>|z|	[95% Conf. Interval]	
dln_inc						
dln_inc						
L1.	-.101226	.1240019	-0.82	0.414	-.3442653	.1418133
L2.	.0137257	.1221872	0.11	0.911	-.2257569	.2532083
dln_consump						
L1.	.3433416	.1362228	2.52	0.012	.0763498	.6103334
L2.	.1476759	.1364602	1.08	0.279	-.1197812	.415133
_cons	.0115223	.0032791	3.51	0.000	.0050955	.0179492
dln_consump						
dln_inc						
L1.	.3231097	.1086299	2.97	0.003	.1101991	.5360203
L2.	.3607425	.1070401	3.37	0.001	.1509477	.5705373
dln_consump						
L1.	-.3030441	.1193358	-2.54	0.011	-.5369379	-.0691502
L2.	-.0599513	.1195437	-0.50	0.616	-.2942527	.1743501
_cons	.0120202	.0028726	4.18	0.000	.0063901	.0176503

Here $\Theta_1 = \begin{pmatrix} -0.10 & 0.34 \\ 0.32 & -0.30 \end{pmatrix}$ and $\Theta_2 = \begin{pmatrix} 0.01 & 0.14 \\ 0.36 & -0.06 \end{pmatrix}$.

Vector autoregressive models: examples

OLS

```
. reg l(0/2).dln_inc l(1/2).dln_consump
```

Source	SS	df	MS
Model	.001118023	4	.000279506
Residual	.011055329	84	.000131611
Total	.012173352	88	.000138334

Number of obs	=	89
F(4, 84)	=	2.12
Prob > F	=	0.0850
R-squared	=	0.0918
Adj R-squared	=	0.0486
Root MSE	=	.01147

dln_inc	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
dln_inc						
L1.	-.101226	.1276391	-0.79	0.430	-.3550504	.1525984
L2.	.0137257	.1257712	0.11	0.913	-.2363841	.2638355
dln_consump						
L1.	.3433416	.1402185	2.45	0.016	.0645018	.6221813
L2.	.1476759	.1404628	1.05	0.296	-.1316498	.4270017
_cons	.0115223	.0033752	3.41	0.001	.0048103	.0182344

VAR

dln_inc		
dln_inc		
L1.	-.101226	.1240019
L2.	.0137257	.1221872
dln_consump		
L1.	.3433416	.1362228
L2.	.1476759	.1364602
_cons	.0115223	.0032791
dln_consump		
dln_inc		
L1.	.3231097	.1086299
L2.	.3607425	.1070401
dln_consump		
L1.	-.3030441	.1193358
L2.	-.0599513	.1195437
_cons	.0120202	.0028726

VAR: Stata

```
*** Vector Error-correction model and multivariate cointegration ***
clear all
use http://www.stata-press.com/data/r13/balance2
describe
*y is ln(GDP)
*i is ln(income)
*c is ln(consumption)
```

Explore the data

```
 obs:            96                    macro data for VECM/balance study
vars:             7                    30 May 2019 09:37
                                       (_dta has notes)
```

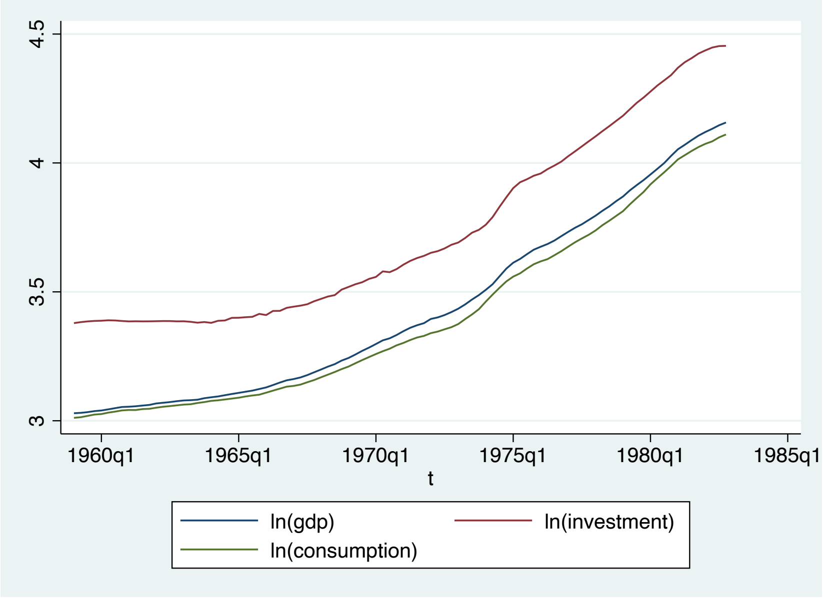
variable name	storage type	display format	value label	variable label
gdp	float	%9.0g		
t	int	%tq		
inv	float	%9.0g		
consump	float	%9.0g		
y	double	%10.0g		**ln(gdp)**
i	double	%10.0g		**ln(investment)**
c	double	%10.0g		**ln(consumption)**

VAR: Stata

```
tsset t
tsline y i c
dfuller d.y
dfuller d.i
dfuller d.c

dfuller d2.y
dfuller d2.i
dfuller d2.c
```

```
tsset t
        time variable:  t, 1959q1 to 1982q4
                delta:  1 quarter
```



Stationarity testing

```
. dfuller d.y

Dickey-Fuller test for unit root                   Number of obs   =        94

                               ---------- Interpolated Dickey-Fuller ---------
                  Test         1% Critical       5% Critical      10% Critical
               Statistic           Value             Value             Value
------------------------------------------------------------------------------
 Z(t)             -2.472            -3.518            -2.895            -2.582
------------------------------------------------------------------------------
MacKinnon approximate p-value for Z(t) = 0.1225

. dfuller d.i

Dickey-Fuller test for unit root                   Number of obs   =        94

                               ---------- Interpolated Dickey-Fuller ---------
                  Test         1% Critical       5% Critical      10% Critical
               Statistic           Value             Value             Value
------------------------------------------------------------------------------
 Z(t)             -3.587            -3.518            -2.895            -2.582
------------------------------------------------------------------------------
MacKinnon approximate p-value for Z(t) = 0.0060

. dfuller d.c

Dickey-Fuller test for unit root                   Number of obs   =        94

                               ---------- Interpolated Dickey-Fuller ---------
                  Test         1% Critical       5% Critical      10% Critical
               Statistic           Value             Value             Value
------------------------------------------------------------------------------
 Z(t)             -2.279            -3.518            -2.895            -2.582
------------------------------------------------------------------------------
MacKinnon approximate p-value for Z(t) = 0.1788
```

VAR: Stata

```
tsset t
tsline y i c
dfuller d.y
dfuller d.i
dfuller d.c

dfuller d2.y
dfuller d2.i
dfuller d2.c
```

What is the order of integration?

Stationarity testing

```
. dfuller d2.y

Dickey-Fuller test for unit root                   Number of obs   =        93

                               ---------- Interpolated Dickey-Fuller ---------
                  Test         1% Critical       5% Critical      10% Critical
               Statistic           Value             Value             Value
------------------------------------------------------------------------------
 Z(t)           -11.508           -3.520            -2.896            -2.583
------------------------------------------------------------------------------
MacKinnon approximate p-value for Z(t) = 0.0000

. dfuller d2.i

Dickey-Fuller test for unit root                   Number of obs   =        93

                               ---------- Interpolated Dickey-Fuller ---------
                  Test         1% Critical       5% Critical      10% Critical
               Statistic           Value             Value             Value
------------------------------------------------------------------------------
 Z(t)           -17.118           -3.520            -2.896            -2.583
------------------------------------------------------------------------------
MacKinnon approximate p-value for Z(t) = 0.0000

. dfuller d2.c

Dickey-Fuller test for unit root                   Number of obs   =        93

                               ---------- Interpolated Dickey-Fuller ---------
                  Test         1% Critical       5% Critical      10% Critical
               Statistic           Value             Value             Value
------------------------------------------------------------------------------
 Z(t)           -11.824           -3.520            -2.896            -2.583
------------------------------------------------------------------------------
MacKinnon approximate p-value for Z(t) = 0.0000
```

Reminder

1. Testing for stationarity. Dickey – Fuller test.

$$X_t = \theta X_{t-1} + \varepsilon_t$$
$$\theta = 1 \Rightarrow unit\ root$$
$$(X_t - X_{t-1}) = (\theta - 1)X_{t-1} + \varepsilon_t$$
$$H_o: (\theta - 1) = 0 \Rightarrow unit\ root$$
$$H_a: |\theta - 1| < 0 \Rightarrow stationarity$$

$$AR(1): Y_t = \delta + \theta Y_{t-1} + \varepsilon_t$$
$$H_0: \theta = 1 \quad (a \quad unit \quad root)$$
$$\mathrm{H}_1: |\theta| < 1 \quad (stationarity)$$
$$Dickey - Fuller \quad test:$$
$$DF = \frac{\hat{\theta} - 1}{se(\hat{\theta})}$$

$$More\ convenient \quad procedure$$
$$\Delta Y_t = \delta + (\theta - 1)Y_{t-1} + \varepsilon_t$$
$$\theta = 1 \quad corresponds$$
$$\Delta Y_t = \delta + \varepsilon_t - \quad random \quad walk \quad with \quad drift.$$
$$Y_t - difference \quad stationary.$$

$$It \quad is \quad possible \quad that \quad nonstationary \quad is \quad caused \quad by$$
$$the \quad presence \quad of \quad the \quad deterministic\ time\ trend:$$
$$Y_t = \delta + \theta Y_{t-1} + \gamma t + \varepsilon_t$$
$$with \quad |\theta| < 1 \quad and \quad \gamma \neq 0.$$
$$Y_t - trend \quad stationary.$$
$$Usual \quad DF \quad test:$$
$$\Delta Y_t = \delta + (\theta - 1)Y_{t-1} + \gamma t + \varepsilon_t,$$
$$different \quad critical \quad values \quad of \quad DF \quad statistics: \tau_0, \tau_\mu, \tau_\tau.$$

Reminder

Dolado, Jenkinson, Sosvilla-Rivero testing procedure for unit root

$$Step1. \quad \Delta Y_t = \delta + \gamma t + \pi Y_{t-1} + c_1 \Delta Y_{t-1} + \ldots + c_{p-1} \Delta Y_{t-p+1} + \varepsilon_t.$$

$$H_0 : \pi = 0 \quad (unit \quad root, \quad DS),$$

$$H_1 : \pi < 0 \quad (TS).$$

$$If \quad t < \tau_\tau (H_0 \quad is \quad rejected) \Rightarrow TS, \quad no\, unit \quad root.$$

$$If \quad t > \tau_\tau \quad (H_0 \quad is \quad not \quad rejected) \quad go \quad to \quad step \quad 2.$$

$$Step2. \quad \Delta Y_t = \delta + \gamma t + c_1 \Delta Y_{t-1} + \ldots + c_{p-1} \Delta Y_{t-p+1} + \varepsilon_t.$$

$$H_0 : \quad \gamma = 0,$$

$$H_1 : \gamma \neq 0.$$

$$If \quad H_0 \quad is \quad rejected, \quad stop.$$

$$If \quad H_0 \quad is \quad not \quad rejected, \quad go \quad to \quad step \quad 3.$$

$$Test \quad statistics \quad \tau_0, \tau_\mu, \tau_\tau.$$

Demidova: https://www.hse.ru/mirror/pubs/share/219723293

Reminder

Dolado, Jenkinson, Sosvilla-Rivero testing procedure for unit root

$Step3.\quad \Delta Y_t = \delta + \pi Y_{t-1} + c_1 \Delta Y_{t-1} + \ldots + c_{p-1}\Delta Y_{t-p+1} + \varepsilon_t.$

$H_0 : \pi = 0 \quad (unit \quad root, \quad DS),$

$H_1 : \pi < 0 \quad (const).$

$If \quad t < \tau_\mu (H_0 \quad is \quad rejected) \Rightarrow no\, unit \quad root, \quad no \quad time \quad trend.$

$If \quad t > \tau_\mu \quad (H_0 \quad is \quad not \quad rejected) \quad go \quad to \quad step \quad 4.$

$Step4.\quad \Delta Y_t = \delta + c_1 \Delta Y_{t-1} + \ldots + c_{p-1}\Delta Y_{t-p+1} + \varepsilon_t.$

$H_0 : \quad \delta = 0,$

$H_1 : \delta \neq 0.$

$If \quad H_0 \quad is \quad rejected, \quad stop.$

$If \quad H_0 \quad is \quad not \quad rejected, \quad go \quad to \quad step \quad 5.$

Reminder

Dolado, Jenkinson, Sosvilla-Rivero testing procedure for unit root

$$Step5. \quad \Delta Y_t = \pi Y_{t-1} + c_1 \Delta Y_{t-1} + \ldots + c_{p-1} \Delta Y_{t-p+1} + \varepsilon_t .$$

$$H_0 : \pi = 0 \quad (unit \quad root, \quad DS),$$

$$H_1 : \pi < 0 \quad (stationarity).$$

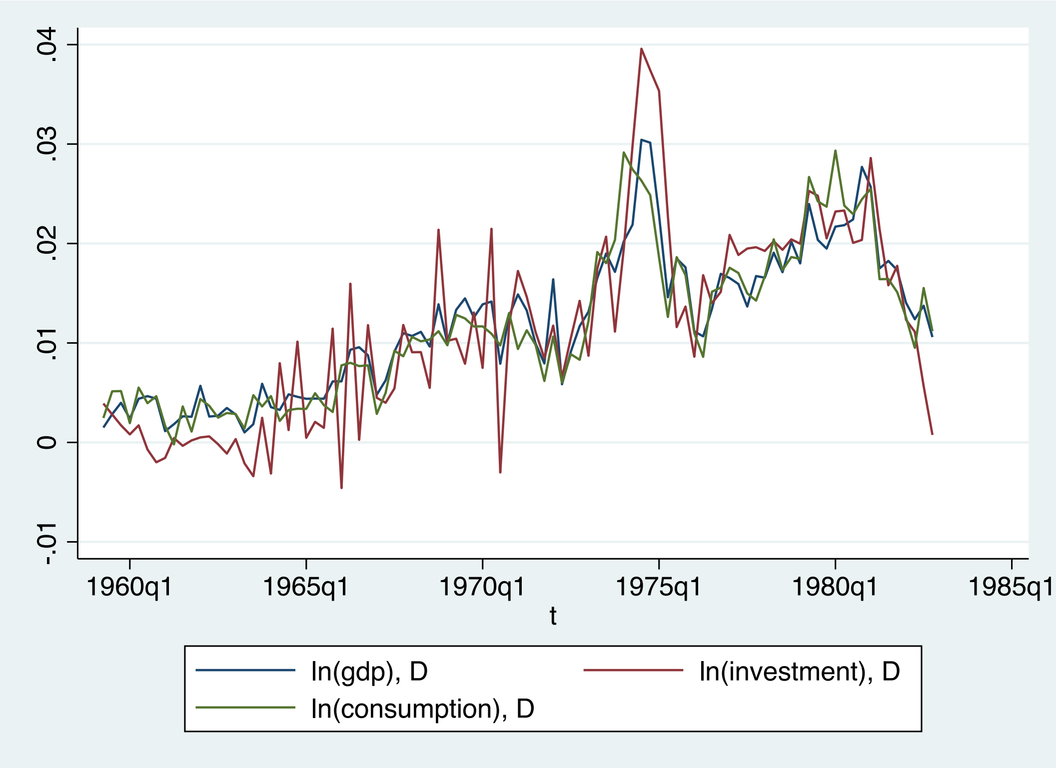
$$If \quad t < \tau_0 (H_0 \quad is \quad rejected) \Rightarrow no\, unit \quad root.$$

$$If \quad t > \tau_0 \quad (H_0 \quad is \quad not \quad rejected) \Rightarrow unit \quad root..$$

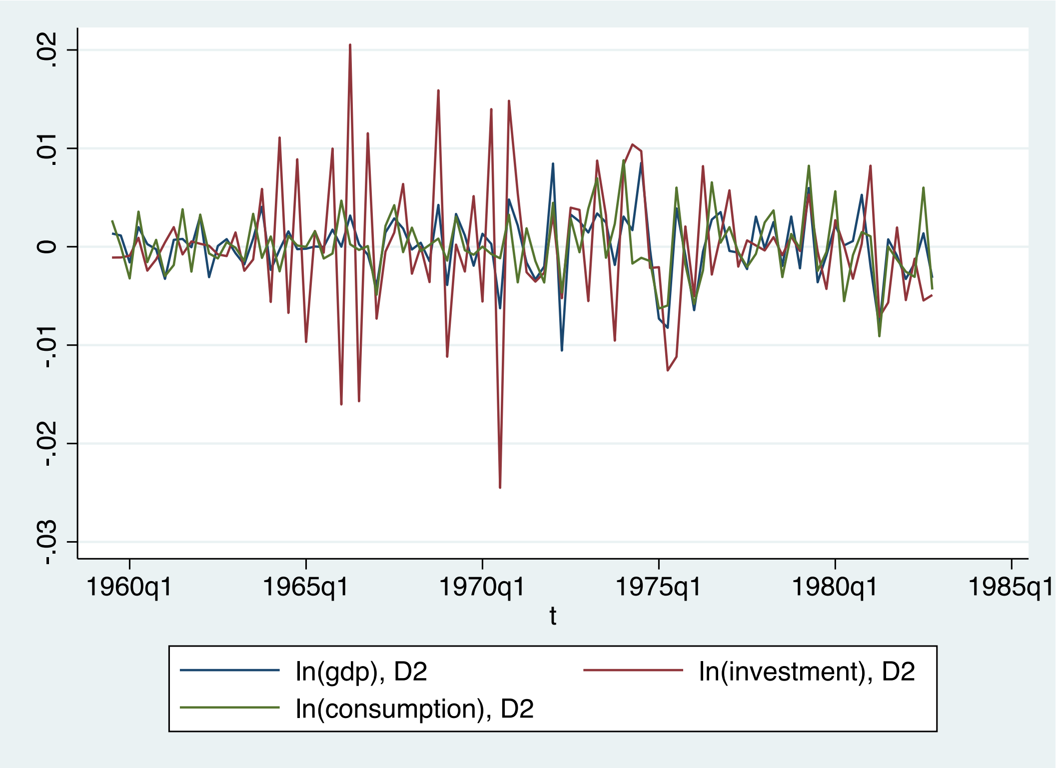
VAR: Stata

```
tsline d.(y i c)
tsline d2.(y i c)
```

Stationarity testing



First difference



Second difference

VAR: Stata

```
reg y i c
estat dwatson
reg i y c
estat dwatson
reg c y i
estat dwatson
```

```
reg d.(y i c)
estat dwatson
reg d.(i y c)
estat dwatson
reg d.(c y i)
estat dwatson
```

y, i, c – the same order of integration -> dw test for cointegration after model estimation

If y, i, c are cointegrated -> residuals are stationary

```
. reg d.(y i c)

      Source |       SS           df       MS      Number of obs   =        95
-------------+----------------------------------   F(2, 92)        =    723.14
       Model |  .004550013         2  .002275007   Prob > F        =    0.0000
    Residual |  .000289432        92  3.1460e-06   R-squared       =    0.9402
-------------+----------------------------------   Adj R-squared   =    0.9389
       Total |  .004839446        94  .000051483   Root MSE        =    .00177

------------------------------------------------------------------------------
         D.y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           i |
         D1. |   .2618364   .0334642     7.82   0.000     .1953736    .3282993
             |
           c |
         D1. |   .6176699   .0435336    14.19   0.000     .5312084    .7041314
             |
       _cons |   .0017598   .0003363     5.23   0.000     .0010918    .0024278
------------------------------------------------------------------------------

. estat dwatson

Durbin-Watson d-statistic(  3,    95) =  1.550373
```

```
. reg d.(i y c)

      Source |       SS           df       MS      Number of obs   =        95
-------------+----------------------------------   F(2, 92)        =    201.83
       Model |  .007401153         2  .003700576   Prob > F        =    0.0000
    Residual |  .001686816        92  .000018335   R-squared       =    0.8144
-------------+----------------------------------   Adj R-squared   =    0.8104
       Total |  .009087969        94  .000096681   Root MSE        =    .00428

------------------------------------------------------------------------------
         D.i |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           y |
         D1. |   1.525986   .1950301     7.82   0.000      1.13864    1.913333
             |
           c |
         D1. |  -.2933017   .1851441    -1.58   0.117    -.6610138    .0744104
             |
       _cons |  -.0033995   .0008543    -3.98   0.000    -.0050962   -.0017029
------------------------------------------------------------------------------

. estat dwatson

Durbin-Watson d-statistic(  3,    95) =  2.052595
```

Cointegration

I(d) time series

$$X_t \sim I(0) - stationary \quad time \quad series$$

$$X_t \sim I(d) \quad if \quad X_t, \quad \Delta X_t, \quad \Delta^2 X_t, \ldots, \Delta^{d-1} X_t$$
$$are \quad nonstationary, \quad \Delta^d X_t \quad is \quad stationary.$$

Tests for cointegration

$$\text{Idea}: \quad Y_t - \beta X_t = \varepsilon_t.$$
$$If \quad \varepsilon_t \sim I(0) \Rightarrow cointegration.$$
$$But \quad we \quad don't \quad know \quad \beta.$$
$$OLS: \quad \hat{Y}_t = \hat{\alpha} + \hat{\beta} X_t,$$
$$e_t = Y_t - \hat{\alpha} - \hat{\beta} X_t,$$
$$e_t \sim I(0)???, \quad DF, ADF \quad are \quad not \quad valid.$$
$$DF2$$

Properties of I(d) time series:

If $X_t \sim I(0)$, $Y_t \sim I(1)$ then $Z_t = X_t + Y_t \sim I(1)$

If $X_t \sim I(d)$ then $Z_t = a + bX_t \sim I(d)$

If $X_t \sim I(d_1)$, $Y_t \sim I(d_2)$ then $Z_t = aX_t + bY_t \sim I(d\max(d_1,d_2))$

If $X_t \sim I(d)$, $Y_t \sim I(d)$ then $Z_t = aX_t + bY_t \sim I(d^*)$,

$d^* \leq d$.

Definition. If $X_t \sim I(1)$, $Y_t \sim I(1)$, $Z_t = Y_t - \beta X_t \sim I(0)$

then X_t and Y_t are cointegrated,

$(1, -\beta)$ is being called the cointegrating vector.

Cointegration

An important exception to the results in the previous subsection arises **when the two nonstationary series have the same stochastic trend in common**.

Consider two series, **integrated of order one, Y_t and X_t**, and suppose that **a linear relationship exists** between them.

In other words, there exists some value β such that **$Y_t - \beta X_t$ is $I(0)$**, although Y_t and X_t are both $I(1)$.

In such a case it is said that Y_t and X_t are **cointegrated**, and that they share a **common trend**.

Although **the relevant asymptotic theory is nonstandard**, it can be shown that **one can consistently estimate β** from an OLS regression of Y_t on X_t as

$$Y_t = \alpha + \beta X_t + \varepsilon_t$$

In this case, the **OLS estimator b** is said to be **super consistent** for β, because it **converges to β at a much faster rate than with conventional asymptotics**.

In the standard case, we have that $\sqrt{T}(b - \beta)$ is asymptotically normal and we say that b is **$\sqrt{T}$ –consistent for β**.

In the cointegration case, the appropriate asymptotic distribution is the one of **$T(b - \beta)$**.

The intuition behind the super consistency result is quite straightforward.

Since **OLS** chooses a and b to minimize the sample variance of e_t, it is **extremely good in finding an estimate close to β**.

Test for cointegration: residuals, DW

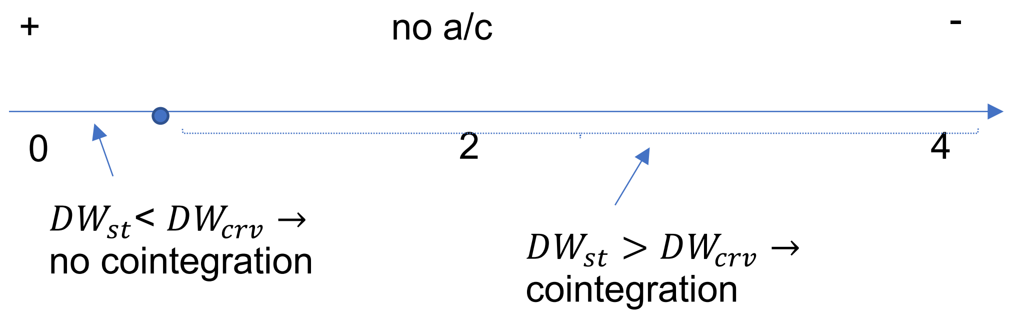


Table 9.3 5% Critical values CRDW tests for co-integration (Banerjee *et al.*, 1993)

Number of variables (incl. Y_t)	Number of observations 50	100	200
2	0.72	0.38	0.20
3	0.89	0.48	0.25
4	1.05	0.58	0.30
5	1.19	0.68	0.35

VAR: Stata

```
// to choose number of lags in VAR
varsoc d.(y i c)
varsoc d2.(y i c)
```

What should be the order of the VAR model?

If y, i, c are cointegrated -> residuals are stationary

```
. varsoc d.(y i c)

   Selection-order criteria
   Sample:  1960q2 - 1982q4                     Number of obs      =        91
```

lag	LL	LR	df	p	FPE	AIC	HQIC	SBIC
0	1110.04				5.4e-15	-24.3306	-24.2972	-24.2478
1	1206.45	192.82	9	0.000	8.0e-16	-26.2516	-26.1181*	-25.9205*
2	1218.68	24.454*	9	0.004	7.4e-16*	-26.3226*	-26.0888	-25.7431
3	1223.82	10.287	9	0.328	8.1e-16	-26.2378	-25.9039	-25.4101
4	1231.1	14.568	9	0.104	8.4e-16	-26.2001	-25.766	-25.124

```
   Endogenous:  D.y D.i D.c
    Exogenous:  _cons
```

```
. varsoc d2.(y i c)

   Selection-order criteria
   Sample:  1960q3 - 1982q4                     Number of obs      =        90
```

lag	LL	LR	df	p	FPE	AIC	HQIC	SBIC
0	1131.9				2.6e-15	-25.0866	-25.053	-25.0032
1	1166.84	69.887	9	0.000	1.4e-15	-25.6631	-25.5287	-25.3298*
2	1178.11	22.545	9	0.007	1.4e-15	-25.7136	-25.4784	-25.1303
3	1194.85	33.482*	9	0.000	1.2e-15*	-25.8856*	-25.5496*	-25.0523
4	1199.93	10.162	9	0.338	1.3e-15	-25.7985	-25.3617	-24.7153

```
   Endogenous:  D2.y D2.i D2.c
    Exogenous:  _cons
```

VAR: Stata

```
41
42   var d2.(y i c), lags(1/3)
```

VAR model, second differences + 3 lags

```
Vector autoregression

Sample:  1960q2 - 1982q4                  Number of obs     =         91
Log likelihood =   1208.928               AIC               =   -25.9105
FPE            =   1.12e-15               HQIC              =  -25.57655
Det(Sigma_ml)  =   5.80e-16               SBIC              =  -25.08275
```

Equation	Parms	RMSE	R-sq	chi2	P>chi2
D2_y	10	.002859	0.3200	42.82699	0.0000
D2_i	10	.00541	0.4749	82.29336	0.0000
D2_c	10	.003061	0.2058	23.57714	0.0050

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
D2_y						
y						
LD2.	-.5379116	.1571339	-3.42	0.001	-.8458884	-.2299349
L2D2.	-.7881112	.1600753	-4.92	0.000	-1.101853	-.4743694
L3D2.	-.6245419	.15392	-4.06	0.000	-.9262195	-.3228643
i						
LD2.	-.0109498	.0677212	-0.16	0.872	-.1436809	.1217812
L2D2.	.1740143	.0800009	2.18	0.030	.0172154	.3308132
L3D2.	.1563892	.0645166	2.42	0.015	.0299391	.2828393
c						
LD2.	.3047081	.1337558	2.28	0.023	.0425514	.5668647
L2D2.	.6129672	.1429175	4.29	0.000	.332854	.8930804
L3D2.	.5384643	.1359958	3.96	0.000	.2719175	.8050112
_cons	.0001627	.0002838	0.57	0.567	-.0003936	.0007189
D2_i						
y						
LD2.	.3014075	.297329	1.01	0.311	-.2813465	.8841616
L2D2.	-.007382	.3028947	-0.02	0.981	-.6010447	.5862808
L3D2.	.1377462	.2912476	0.47	0.636	-.4330885	.7085809
i						
LD2.	-.7333108	.1281421	-5.72	0.000	-.9844647	-.4821568
L2D2.	-.1740623	.1513778	-1.15	0.250	-.4707574	.1226328
L3D2.	-.1839026	.1220783	-1.51	0.132	-.4231717	.0553665
c						
LD2.	.7041148	.253093	2.78	0.005	.2080617	1.200168
L2D2.	.5647705	.2704288	2.09	0.037	.0347399	1.094801
L3D2.	.4534464	.2573314	1.76	0.078	-.0509139	.9578068
_cons	-.0001443	.000537	-0.27	0.788	-.0011968	.0009083
D2_c						
y						
LD2.	-.0729829	.1682315	-0.43	0.664	-.4027105	.2567448
L2D2.	-.4034248	.1713806	-2.35	0.019	-.7393247	-.067525
L3D2.	-.4644299	.1647906	-2.82	0.005	-.7874135	-.1414463
i						
LD2.	-.0595044	.072504	-0.82	0.412	-.2016096	.0826008
L2D2.	.1458614	.085651	1.70	0.089	-.0220115	.3137342
L3D2.	.1539156	.069073	2.23	0.026	.018535	.2892963
c						
LD2.	-.1695081	.1432024	-1.18	0.237	-.4501795	.1111634
L2D2.	.2298018	.1530111	1.50	0.133	-.0700944	.5296981
L3D2.	.3563617	.1456005	2.45	0.014	.07099	.6417335
_cons	.0001594	.0003039	0.52	0.600	-.0004362	.0007549

no inf of y on investments

Granger causality

Granger causality test is used to determine if one time series will be useful to forecast another variable by investigating causality between two variables in a time series. The method is a probabilistic account of causality; it uses observed data sets to find patterns of *correlation*. One good thing about time series vector autoregression (VAR) is that we could test 'causality' in some sense. This test is first proposed by Granger (1969), and therefore we refer to it as the Granger causality.

Simple Mechanism to define Granger Causality:
It is based on the **idea that if X causes Y, then the forecast of Y based on previous values of Y AND the previous values of X should best result in the forecast of Y based on previous values of Y alone.**
Granger causality should not be used to test if a lag of Y causes Y. Instead, it is generally used on exogenous (not Y lag) variables only. In simple terms '*X* is said to Granger-cause *Y* if *Y* can be better predicted using the histories of both *X* and *Y* than it can by using the history of *Y* alone'
When performing Granger Causality Test we need to consider two assumptions:
1. Future values cannot **cause** the past values.
2. A notably distinct information is contained in **cause** about effect which will not be available elsewhere

$$Y_t = \gamma_1 + \dots + \theta_1 X_{t-1} + \theta_2 X_{t-2} + \theta_3 X_{t-3}$$

$H_0: \theta_1 = \theta_2 = \theta_3 = 0$
if H_0 rejected $\Rightarrow$ X influences Y

Granger causality

- Testing for Granger-causality using F-statistics when one or both time series are non-stationary can lead to *nearly* false causality. If both the time series are **NOT** stationary then differencing, de-trending or other techniques must first be employed before using the Granger Causality test.

- We say that *x* Granger-causes y when the **null hypothesis** is rejected.

•The **null hypothesis** for the test is that lagged X-values do not explain the variation in Y. Put simple, it assumes that X(t) doesn't Granger-cause Y(t).

- Cointegration analysis is a useful tool in order to examine if there exists a long run equilibrium relationship between two or more time series. (*For example: nominal interest rate is connected with an increase on expected inflation in the long run.*)

- Causation is can be One-Direction, Both-Direction or NO-Direction

After fitting a VAR, we may want to know whether one variable "Granger-causes" another (Granger 1969). A variable x is said to Granger-cause a variable y if, given the past values of y, past values of x are useful for predicting y. A common method for testing Granger causality is to regress y on its own lagged values and on lagged values of x and test the null hypothesis that the estimated coefficients on the lagged values of x are jointly zero. Failure to reject the null hypothesis is equivalent to failing to reject the hypothesis that x does not Granger-cause y.

For each equation and each endogenous variable that is not the dependent variable in that equation, `vargranger` computes and reports Wald tests that the coefficients on all the lags of an endogenous variable are jointly zero. For each equation in a VAR, `vargranger` tests the hypotheses that each of the other endogenous variables does not Granger-cause the dependent variable in that equation.

VAR: Stata

```
// to choose number of lags in VAR
varsoc d.(y i c)
varsoc d2.(y i c)
```

0.035 also signif.

Granger causality testing

Null hypothesis: estimated coefficients on the lagged x are equal to zero

Wald Test

reject Ho at 5% sign. level

```
. vargranger

   Granger causality Wald tests
```

Equation	Excluded	chi2	df	Prob > chi2
D2_y	D2.i	8.5987	3	0.035
D2_y	D2.c	27.454	3	0.000
D2_y	ALL	32.439	6	0.000
D2_i	D2.y	1.3423	3	[0.719]
D2_i	D2.c	10.826	3	0.013
D2_i	ALL	32.891	6	0.000
D2_c	D2.y	10.813	3	0.013
D2_c	D2.i	8.4446	3	0.038
D2_c	ALL	14.47	6	0.025

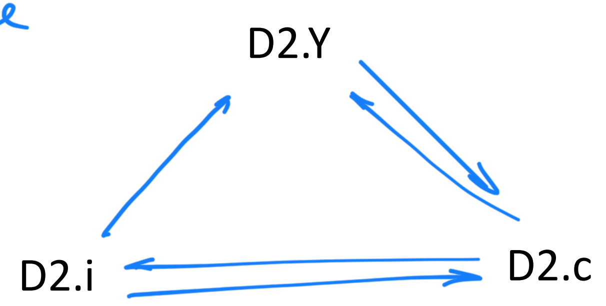
if reject Ho → these variables have influence

Test if [i] is a cause for [y]

if [c] is a cause for y

Test if [c+i] are jointly a cause for [y]

accepted only here



Johansen Test

moving to ECM enables us to identify short run effect

allows us to analyze whether two or more time series can form a cointegrating relationship

(in our case: if there is a long-term relationship between GDP, investments and consumption?)

In particular we need to consider Vector Autoregressive Models (VAR)

A general vector autoregressive model is similar to the AR(p) model except that each quantity is vector valued and matrices are used as the coefficients. The general form of the VAR(p) model, without drift, is given by:

$$\mathbf{x_t} = \mu + A_1\mathbf{x_{t-1}} + \ldots + A_p\underbrace{\mathbf{x_{t-p}}}_{\text{vectors}} + \mathbf{w_t}$$

Where μ is the vector-valued mean of the series, A_i are the coefficient matrices for each lag and $\mathbf{w_t}$ is a multivariate Gaussian noise term with mean zero.

At this stage we can form a Vector Error Correction Model (VECM) by differencing the series:

L1 in stata = short run effect

$$\Delta\mathbf{x_t} = \mu + A\mathbf{x_{t-1}} + \Gamma_1\Delta\mathbf{x_{t-1}} + \ldots + \Gamma_p\Delta\mathbf{x_{t-p}} + \mathbf{w_t}$$

$H_0: A=0$, looks @ rank A | How many factors are cointegrated

Where $\Delta\mathbf{x_t} := \mathbf{x_t} - \mathbf{x_{t-1}}$ is the differencing operator, A is the coefficient matrix for the first lag and Γ_i are the matrices for each differenced lag.

The test checks for the situation of no cointegration, which occurs when the matrix $A = 0$.

https://www.stata.com/manuals/tsvecrank.pdf

Johansen Test

allows us to analyze whether two or more time series can form a cointegrating relationship

2. Johansen test for cointegration

An alternative approach that does not suffer from these drawbacks was proposed by Johansen (1988), who developed a maximum likelihood estimation procedure, which also **allows one to test for the number of cointegrating relations**.

The details of the Johansen procedure are **very complex** and we shall only focus on a few aspects.

The starting point of the Johansen procedure is the VAR representation of $\vec{Y}_t$ given in (9.44):

$$\Delta\vec{Y}_t = \delta + \Gamma_1 \Delta\vec{Y}_{t-1} + \cdots + \Gamma_{p-1}\Delta\vec{Y}_{t-p+1} + \Pi\vec{Y}_{t-1} + \vec{\varepsilon}_t,$$

where $\vec{\varepsilon}_t$ is $NID(0,\Sigma)$. Note that the use of maximum likelihood requires us to impose a **particular distribution** for the white noise terms.

Johansen Test

allows us to analyze whether two or more time series can form a cointegrating relationship

2. Johansen test for cointegration (continued)

H0: $\vec{Y}_t$ is a vector of $I(1)$ variables, while r linear combinations of $\vec{Y}_t$ are stationary.

Trace test

H0: $r \leq r_0$ versus the alternative $H1$: $r_0 < r \leq k$, where is r_0 is the rank tested at the current step.

Maximum eigenvalue test

H0: $r \leq r_0$ versus the more restrictive alternative

The two tests are **likelihood ratio tests**, but **do not have the usual Chi-squared distributions**. Instead, the appropriate distributions are **multivariate extensions of the Dickey–Fuller distributions**.

As with the unit root tests, the percentiles of the distributions depend on the fact whether a constant (and a time trend) are included.

Johansen Test

allows us to analyze whether two or more time series can form a cointegrating relationship

Table **9.9** Critical values Johansen's LR tests for cointegration

	λ_{trace}-statistic $H_0: r \leq r_0$ vs $H_1: r > r_0$		λ_{max}-statistic $H_0: r \leq r_0$ vs $H_1: r = r_0 + 1$	
$k - r_0$	5%	10%	5%	10%
Case 1: restricted intercepts in VAR (in cointegrating relations only)				
1	9.16	7.53	9.16	7.53
2	20.18	17.88	15.87	13.81
3	34.87	31.93	22.04	19.86
4	53.48	49.95	28.27	25.80
5	75.98	71.81	34.40	31.73
Case 2: unrestricted intercepts in VAR				
1	8.07	6.50	8.07	6.50
2	17.86	15.75	14.88	12.98
3	31.54	28.78	21.12	19.02
4	48.88	45.70	27.42	24.99
5	70.49	66.23	33.64	31.02

Johansen Test

allows us to analyze whether two or more time series can form a cointegrating relationship

Example. Testing cointegration rank

Income and consumption

```
. vecrank ln_inc ln_consump

                        Johansen tests for cointegration
Trend: constant                                        Number of obs =      90
Sample:  1960q3 - 1982q4                                        Lags =       2
-------------------------------------------------------------------------------
                                                         5%
maximum                                      trace    critical
  rank    parms       LL       eigenvalue  statistic    value
    0       6     574.20894         .       20.4454     15.41
    1       9      582.7739     0.17332      3.3155*     3.76
    2      10     584.43167     0.03617
-------------------------------------------------------------------------------
```

Let’s include income, consumption and investment in the model

```
. vecrank ln_inc ln_consump ln_inv

                        Johansen tests for cointegration
Trend: constant                                        Number of obs =      90
Sample:  1960q3 - 1982q4                                        Lags =       2
-------------------------------------------------------------------------------
                                                         5%
maximum                                      trace    critical
  rank    parms       LL       eigenvalue  statistic    value
    0      12     735.12264         .       32.6777     29.68
    1      17     746.03476     0.21533     10.8534*    15.41
    2      20     749.73178     0.07887      3.4594      3.76
    3      21     751.46148     0.03771
-------------------------------------------------------------------------------
```

VAR: Stata

```
//Johansen test is implemented to choose number of cointegrating equations with the given above number of lags
vecrank d.(y i c), lags(2)
```

Johansen test

rank = # cointegrated equations

$$\Delta\vec{Y}_t = \delta + \Gamma_1\Delta\vec{Y}_{t-1} + \cdots + \Gamma_{p-1}\Delta\vec{Y}_{t-p+1} + \Pi\vec{Y}_{t-1} + \vec{\varepsilon}_t, \qquad \Pi = \gamma\beta'$$

Where β is a matrix of cointegration vectors and γ represents the matrix of weights with which each cointegrating vector enters each of the ΔY_t equations.

$r_k(\Pi) = 0$ if $\Pi=0$, i.e. all elements of Y_t are I(1) processes without cointegration

$r_k(\Pi) = (max\ rank\ if\ \Pi)$ is invertible, i.e. all elements of Y_t are I(0) processes

the hypothesis H_0: $r \le r_0$ versus the alternative H_1: $r_0 < r \le k$, can be tested using the statistic

$$\lambda_{trace}(r_0) = -T \sum_{j=r_0+1}^{k} \log(1 - \hat{\lambda}_j). \tag{9.52}$$

This test is the so-called **trace test**. It checks whether the smallest $k - r_0$ eigenvalues are significantly different from zero. Furthermore, we can test H_0: $r \le r_0$ versus the more restrictive alternative H_1: $r = r_0 + 1$ using

$$\lambda_{max}(r_0) = -T \log(1 - \hat{\lambda}_{r_0+1}). \tag{9.53}$$

no stars here

This alternative test is called the **maximum eigenvalue test**, as it is based on the estimated $(r_0 + 1)$th largest eigenvalue.

H0: real rank is less than a value that is assumed

H1: real rank is higher than a value that is assumed

Trace statistics > critical value -> H0 is rejected

Read:

https://www.stata.com/manuals/tsvecrank.pdf

VAR: Stata

```
//Johansen test is implemented to choose number of cointegrating equations with the given
above number of lags
vecrank d.(y i c), lags(2)
```

Johansen test

H0: real rank is less than a value that is assumed

H1: real rank is higher than a value that is assumed

Johansen tests for cointegration

Trend: constant Number of obs = 93

Sample: 1959q4 - 1982q4 Lags = 2

	maximum rank	parms	LL	eigenvalue	trace statistic		5% critical value
H0: rank of matrix A is 0 or less	0	12	1208.2565	.	75.4462	>	29.68
H0: rank of matrix A is 1 or less	1	17	1230.9018	0.38553	30.1556	>	15.41
H0: rank of matrix A is 2 or less	2	20	1244.6437	0.25586	2.6718*	<	3.76
H0: rank of matrix A is 3 or less	3	21	1245.9796	0.02832			

→ **Rank is 2 (we have 2 cointegrating equations)**

if the real rank = 2 or less

Trace statistics > critical value -> H0 is rejected

if everything is not signif in J. test → we have initially stationary TS

If there is a long-term relationship between GDP, investments and consumption? Yes
What form of our VECM should we use? It should include 2 cointegration equations

If the rank is 0 - > there is no cointegration (linear combination)
If the rank is max -> we can say nothing about cointegration, we have stationary time series initially

Vector error-correction model (VECM)

Substituting this produces the model in error-correction form

$$\Delta \vec{Y}_t = \delta + \Gamma_1 \Delta \vec{Y}_{t-1} + \cdots + \Gamma_{p-1} \Delta \vec{Y}_{t-p+1} + \gamma \beta' \vec{Y}_{t-1} + \vec{\varepsilon}_t. \qquad (9.46)$$

The linear combinations $\beta' \vec{Y}_{t\text{-}1}$ present the ***r* cointegrating relationships**.

The coefficients in γ measure how the elements in $\beta' \vec{Y}_{t\text{-}1}$ are adjusted to the *r* 'equilibrium errors' $\vec{Z}_t - \beta' \vec{Y}$.

Thus, (9.46) is a generalization of (9.24) and is referred to as a **vector error-correction model (VECM)**.

It is important to realize that the parameters γ **and** β **are not uniquely** identified in the sense that different combinations of γ and β can produce the same matrix $\Pi = \gamma \beta'$. This is because

$\gamma \beta' = \gamma P P^{-1} \beta'$ for any invertible *r* × *r* matrix *P*.

Example. VECM for income and consumption

Cointegrating equations

Equation	Parms	chi2	P>chi2
_ce1	1	24260.05	0.0000

Identification: beta is exactly identified

Johansen normalization restriction imposed

beta	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
_ce1						
ln_consump	1	.	.	.	.	.
ln_inc	-.9608795	.0061691	-155.76	0.000	-.9729708	-.9487883
_cons	-.2078207	.	.	.	.	.

cointegration equation

The cointegration equation is

$$\ln(consum)-0.21-0.96\ln(income)$$

Compare with OLS estimation of the long-run relationship

$$\ln(consum)=0.86+0.97\ln(income)$$

VECM for income and consumption (continued)

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
D_ln_consump						
_ce1						
L1.	-.2966839	.0762145	-3.89	0.000	-.4460615	-.1473062
ln_consump						
LD.	-.1446509	.111148	-1.30	0.193	-.3624969	.0731951
ln_inc						
LD.	.100926	.1141641	0.88	0.377	-.1228316	.3246836
_cons	-.0024593	.005403	-0.46	0.649	-.013049	.0081303
D_ln_inc						
_ce1						
L1.	-.0779136	.0903667	-0.86	0.389	-.2550292	.0992019
ln_consump						
LD.	.299662	.131787	2.27	0.023	.0413642	.5579597
ln_inc						
LD.	-.0720283	.1353632	-0.53	0.595	-.3373354	.1932788
_cons	.0093648	.0064063	1.46	0.144	-.0031912	.0219208

$\Delta Y = -0.002 + \theta Y_{t-1} + \Delta X_{t-1} + \Delta Y_{t-1} \rightarrow$ as we have coint. equation $\rightarrow Y_{t-1} - 0.21 - 0.96X_t$

Eq.1: $\Delta\ln(consum)_t = -0.002 - 0.14\Delta\ln(consum)_{t-1} + 0.10\Delta\ln(income)_{t-1} + (-0.30[\ln(consum)_{t-1} - 0.21 - 0.96\ln(income)_{t-1}]$

Testing for the number of lags in VECM

LM test for residual autocorrelation after vec

```
. veclmar
```

Lagrange-multiplier test

lag	chi2	df	Prob > chi2
1	9.1166	4	0.05825
2	3.8221	4	0.43061

H0: no autocorrelation at lag order