Elements of Econometrics. Lecture 22. Modeling with Nonstationary Time Series.

FCS, 2022-2023

Time Series Data: Order of Integration

1. Weakly dependent (stationary) time series are integrated of order zero (= I(0))
2. If a time series has to be differenced one time in order to obtain a weakly dependent series, it is called integrated of order one (= I(1))
3. If a time series has to be differenced *m* times in order to obtain a weakly dependent series, it is called integrated of order *m* (= I(m))

Examples for I(1) processes

$$y_t = y_{t-1} + e_t \;\Rightarrow\; \Delta y_t = y_t - y_{t-1} = e_t \leftarrow$$

After differencing, the resulting series are weakly dependent (because e_t is weakly dependent).

$$\Delta \log(y_t) \approx (y_t - y_{t-1})/y_{t-1}$$

Differencing of the log function provides the growth rates (approx.) which are often weakly dependent (stationary).

Differencing is often a way to achieve weak dependence of economic time series with clear interpretation of the indicators.

Cointegration

The order of integration of linear combination of series $X_1, . . . , X_n$ is usually equal to the maximum order of integration among the series.

Example: if series X_1 and X_2 are integrated of the order 1 and 2 respectively, then their linear combination will be integrated of the order 2.

A linear combination of series with the same order of integration k will be integrated of the order k.

However, if the series have some **long-run relationship**, the order of integration of their linear combination can be lower.

Two or more series (with the same order of integration $k \geq 1$) are called **cointegrated** if there exists their **stationary** linear combination.

COINTEGRATION

$$\boldsymbol{Y_t = \beta_1 + \beta_2 X_{2t} + \ldots + \beta_k X_{kt} + u_t}$$

$$\boldsymbol{u_t = Y_t - \beta_1 - \beta_2 X_{2t} - \ldots - \beta_k X_{kt}}$$

$$e_t = Y_t - b_1 - b_2 X_{2t} - \ldots - b_k X_{kt} - test\ for\ unit\ root$$

If there exists linear relationship between variables Y_t, X_{2t}, …, X_{kt}, the disturbance term u_t is measuring the deviation between the terms in the model. The test is indirect since done with e_i which are more close to stationary than u_i. Hence the critical values are lower.

Asymptotic Critical Values of the Dickey-Fuller Statistic for a Cointegrating Relationship with Two Variables

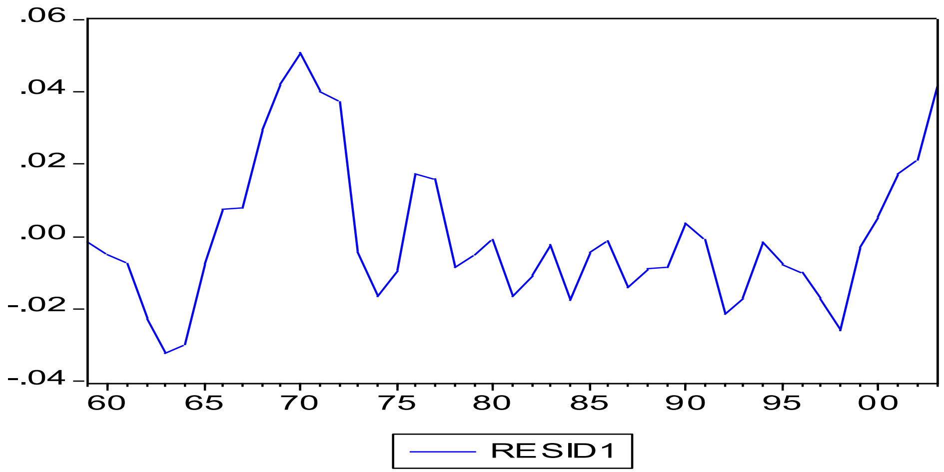
Regression equation contains:	5%	1%
Constant, but no trend	–3.34	–3.90
Constant and trend	–3.78	–4.32

Cointegration: Example. LGFOOD (I(1)), LGDPI (I(1)), LPRFOOD (I(0)).

Dependent Variable: LGFOOD
Method: Least Squares
Sample: 1959 2003
Included observations: 45

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2.236	0.388	5.76	0.000
LGDPI	0.5002	0.0088	56.87	0.000
LPRFOOD	-0.075	0.073	-1.025	0.311

R-squared	0.992	Mean dependent var	6.021
S.D. dependent var	0.223	Sum squared resid	0.017
F-statistic	2606.860	Durbin-Watson stat	0.479



Cointegration: Example. LGFOOD (I(1)), LGDPI (I(1)), LPRFOOD (I(0)).

Augmented Dickey-Fuller Unit Root Test on RESID1
Lag Length: 0 (Automatic based on SIC, MAXLAG=9)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-1.927	0.318
Test critical values	1% level	-3.589	
	5% level	-2.930	
	10% level	-2.603	

*MacKinnon (1996) one-sided p-values.

Asymptotic Critical Values of the Dickey-Fuller Statistic for a Cointegrating Relationship with Two Variables

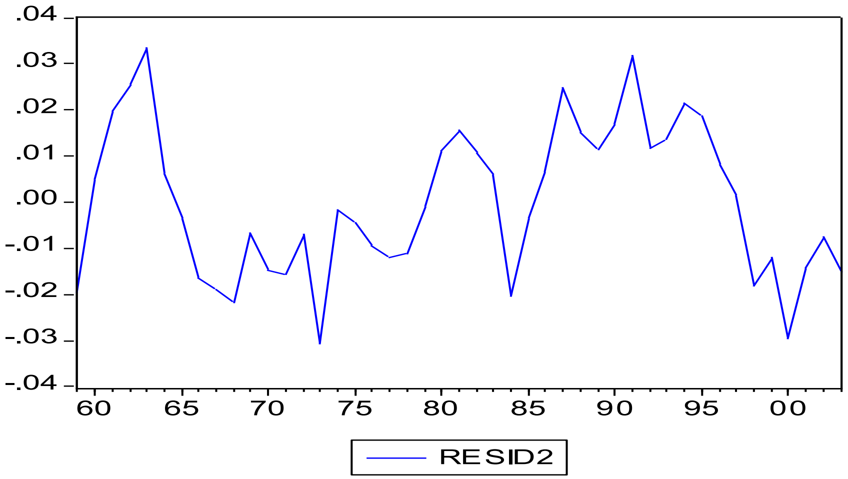
Regression equation contains:	5%	1%
Constant, but no trend	–3.34	–3.90
Constant and trend	–3.78	–4.32

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(RESID1) Method: Least Squares Sample (adjusted): 1960 2003
Included observations: 44 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
RESID1(-1)	-0.208	0.108	-1.927	0.061
C	0.0008	0.002	0.389	0.700

The test statistic is –1.93, not significant at the 5 percent level. The null hypothesis of nonstationarity is not rejected (the variables are not cointegrated). But this may be due to its low power against the alternative hypothesis that the disturbance term is a highly autocorrelated stationary process.

Cointegration, Example 2. LGHOUS, LGDPI, LPRHOUS.



Null Hypothesis: RESID2 has a unit root
Exogenous: Constant   Lag Length: 0 (Fixed)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-2.911003	0.0521
Test critical values:			
1% level	-3.588509		
5% level	-2.929734		
10% level	-2.603064		

The residuals are either nonstationary, or the power of the test is insufficient.

Asymptotic Critical Values of the Dickey-Fuller Statistic for a Cointegrating Relationship with Two Variables

Regression equation contains:	5%	1%
Constant, but no trend	–3.34	–3.90
Constant and trend	–3.78	–4.32

Fitting Models with Nonstationary Time Series: Detrending

In the models with variables which include time trends, removal of the trends, or detrending, allows to avoid getting spurious regressions.

Detrending of each variable in the model is equivalent to including the time trend as an explanatory variable.

Economic indicators often behave not as series including time trends, but as random walks.

If you detrend a series which is a random walk with drift, then its variance increases proportionally to time, the series does not become stationary, and hence the problem of spurious regressions is not resolved.

Fitting Models with Nonstationary Time Series: Differencing

If having random walk time series, differencing is a procedure which can be applied for making it stationary:

subtracting $Y_{t-1} = \beta_1 + \beta_2 X_{t-1} + u_{t-1}$ from $Y_t = \beta_1 + \beta_2 X_t + u_t$, we get $\Delta Y_t = \beta_2 \Delta X_t + \Delta u_t$

The series ΔY_t and ΔX_t are stationary, and the coefficient β_2 can be estimated from this model.

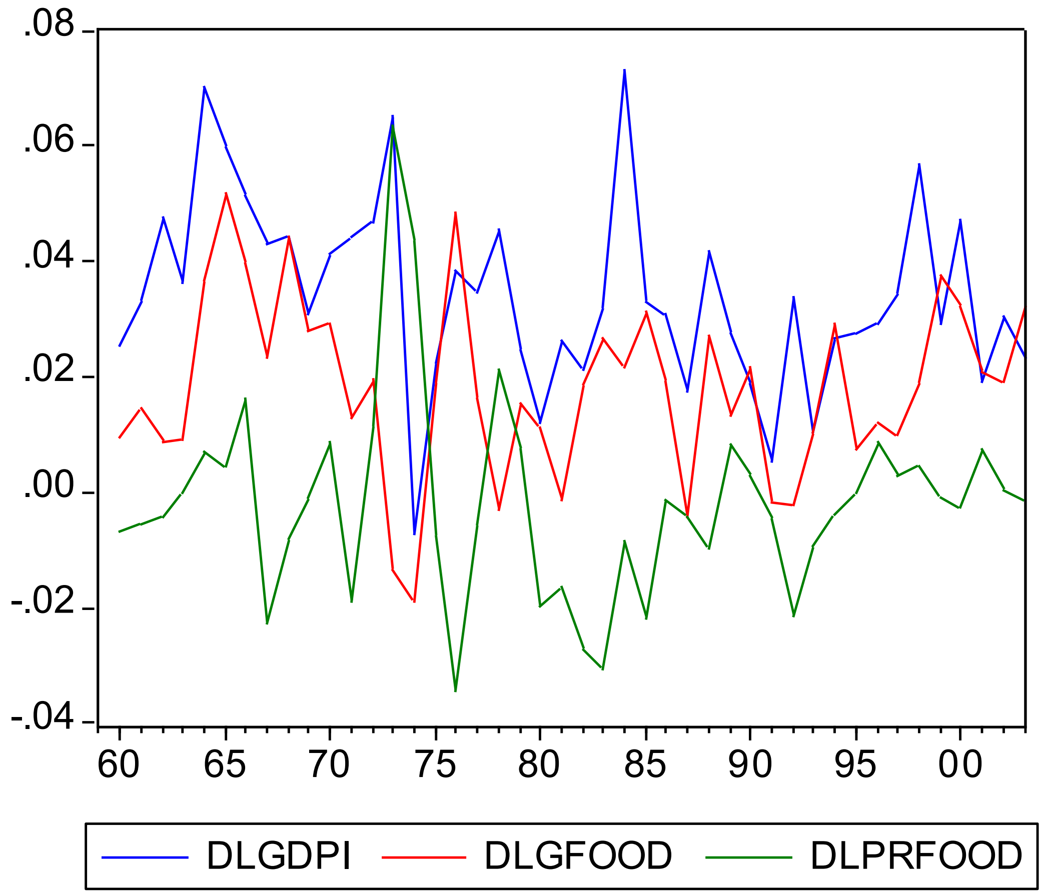
The new disturbance term Δu_t is subject to autocorrelation, and appropriate remedial measures should be applied. Only in the case of severe autocorrelation of u_t in the initial model (ρ is close to 1) differencing helps to reduce autocorrelation.

If Y_t and X_t are unrelated I(1) processes, absence of their relationship will be revealed in the differenced model, so the problem of spurious regressions will be resolved.

Shortcomings of the differenced model:

- constant disappears
- only short-run relationships can be investigated since in the long-run equilibrium $\Delta Y = \Delta X = 0$

Differencing: Regress DLGFOOD on DLGDPI and DLPRFOOD.



LGFOOD, LGDPI and LPRFOOD are I(1) processes since their first differences are stationary. DLGFOOD, DLGDPI and DLPRFOOD are the Growth Rates of FOOD, DPI and PRFOOD respectively.

Differencing: Regress DLGFOOD on DLGDPI and DLPRFOOD.

Dependent Variable: DLGFOOD
Method: Least Squares Sample (adjusted): 1960 2003
Included observations: 44 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.0036	0.0046	0.7827	0.438
DLGDPI	0.407	0.122	3.338	0.002
DLPRFOOD	-0.379	0.115	-3.286	0.002

R-squared	0.321	Mean dependent var	0.018
S.D. dependent var	0.015	S.E. of regression	0.013
Sum squared resid	0.0069	F-statistic	9.712
Durbin-Watson stat	1.582	Prob(F-statistic)	0.00035

Constant is insignificant: no time trend in LGFOOD, LGDPI and LPRFOOD. Coefficients are short run income and price elasticities.

Error Correction Model

$$Y_t^* = \alpha_1 + \alpha_2 X_t$$

$$\Delta Y_t = \left(Y_t - Y_{t-1}\right) = \lambda\left(Y_{t-1}^* - Y_{t-1}\right) + \delta \Delta X_t + u_t$$

$$= \lambda\left(\alpha_1 + \alpha_2 X_{t-1} - Y_{t-1}\right) + \delta\left(X_t - X_{t-1}\right) + u_t$$

$$= \lambda\alpha_1 + \delta X_t + \left(\lambda\alpha_2 - \delta\right) X_{t-1} - \lambda Y_{t-1} + u_t$$

$$Y_t = \beta_1 + \beta_2 Y_{t-1} + \beta_3 X_t + \beta_4 X_{t-1} + u_t$$

$$\beta_1 = \lambda\alpha_1 \qquad \beta_3 = \delta$$

$$\beta_2 = 1 - \lambda \qquad \beta_4 = \lambda\alpha_2 - \delta$$

Y* is a desirable (appropriate) unobserved value of Y. In the short run, $\Delta Y_t = Y_t - Y_{t-1}$, is determined by two components: closing the discrepancy between its previous "appropriate" and actual values, $Y^*_{t-1} - Y_{t-1}$, and a straightforward response to ΔX_t. This is ADL(1,1) model.

Fitting Models with Nonstationary Time Series:

Error Correction Model.

ADL(1,1) model

$$Y_t = \beta_1 + \beta_2 Y_{t-1} + \beta_3 X_t + \beta_4 X_{t-1} + \varepsilon_t$$

In equilibrium

$$\bar{Y} = \beta_1 + \beta_2 \bar{Y} + \beta_3 \bar{X} + \beta_4 \bar{X}.$$

Hence

$$\bar{Y} = \frac{\beta_1}{1-\beta_2} + \frac{\beta_3 + \beta_4}{1-\beta_2} \bar{X}$$

Cointegrating relationship

$$Y_t = \frac{\beta_1}{1-\beta_2} + \frac{\beta_3 + \beta_4}{1-\beta_2} X_t$$

Cointegrating relationship describes the long-run effects. We will construct a model combining short-run and long-run dynamics.

Fitting Models with Nonstationary Time Series: Error Correction Model.

ADL(1,1) model $$Y_t = \beta_1 + \beta_2 Y_{t-1} + \beta_3 X_t + \beta_4 X_{t-1} + \varepsilon_t$$

Cointegrating relationship $$Y_t = \frac{\beta_1}{1-\beta_2} + \frac{\beta_3 + \beta_4}{1-\beta_2} X_t$$

$$Y_t - Y_{t-1} = \beta_1 + (\beta_2 - 1)Y_{t-1} + \beta_3 X_t + \beta_4 X_{t-1} + \varepsilon_t =$$

$$= \beta_1 + (\beta_2 - 1)Y_{t-1} + \beta_3 X_t - \beta_3 X_{t-1} + \beta_3 X_{t-1} + \beta_4 X_{t-1} + \varepsilon_t =$$

$$= (\beta_2 - 1)\left(Y_{t-1} - \frac{\beta_1}{1-\beta_2} - \frac{\beta_3 + \beta_4}{1-\beta_2} X_{t-1} \right) + \beta_3 (X_t - X_{t-1}) + \varepsilon_t$$

$$\Delta Y_t = (\beta_2 - 1)\left(Y_{t-1} - \frac{\beta_1}{1-\beta_2} - \frac{\beta_3 + \beta_4}{1-\beta_2} X_{t-1} \right) + \beta_3 \Delta X_t + \varepsilon_t .$$

The ADL(1,1) relationship may be rewritten to incorporate the cointegrating relationship by subtracting Y_{t-1} from both sides, subtracting $\beta_3 X_{t-1}$ from the right side and adding it back again, and rearranging. If the series X_t and Y_t are cointegrated, and the first differences are stationary, we get all the series as *I(0)*.
The Error Correction Model has been obtained.

Fitting Models with Nonstationary Time Series: Error Correction Model.

$$\Delta Y_t = (\beta_2 - 1)\left(Y_{t-1} - \frac{\beta_1}{1-\beta_2} - \frac{\beta_3 + \beta_4}{1-\beta_2} X_{t-1} \right) + \beta_3 \Delta X_t + \varepsilon_t .$$

The Error Correction Model states that the change in *Y* in any period will be governed by the change in *X* and the discrepancy between Y_{t-1} and the value predicted by the cointegrating relationship.

If *Y* and *X* are I(1) and the cointegrating relationship found, then ΔY_t, ΔX_t, and the error correction term are I(0).

Problem of estimation: the coefficients for calculating the error correction term are not known. Solution: Engle-Granger two step procedure.

Error Correction Model Estimation: Engle-Granger Two Step Procedure.

$$\Delta Y_t = (\beta_2 - 1)\left(Y_{t-1} - \frac{\beta_1}{1-\beta_2} - \frac{\beta_3 + \beta_4}{1-\beta_2} X_{t-1} \right) + \beta_3 \Delta X_t + \varepsilon_t .$$

If *Y* and *X* are I(1), ΔY_t, ΔX_t to be used in further modeling.

Step 1: the cointegrating relationship is estimated, its residuals tested for stationarity. If the hypothesis of nonstationarity is rejected, go to Step 2.

Step 2: estimate the Error Correction Model using the residuals for cointegrating relationship.

Engle and Granger: asymptotically, in the cointegrating term the estimates of βs can be used instead of true values, and hence the residuals from the cointegrating regression can be further used.

Error Correction Model, Example.

ADL(1,1) model

$$LGHOUS_t = \beta_1 + \beta_2 LGHOUS_{t-1} + \beta_3 LGDPI_t + \beta_4 LGDPI_{t-1} +$$
$$+ \beta_5 LPRHOUS_t + \beta_6 LPRHOUS_{t-1} + \varepsilon_t$$

Cointegrating relationship:

$$LGHOUS_t = \frac{\beta_1}{1-\beta_2} + \frac{\beta_3 + \beta_4}{1-\beta_2} LGDPI_t + \frac{\beta_5 + \beta_6}{1-\beta_2} LPRHOUS_t$$

Error Correction Model:

$$\Delta LGHOUS_t = (\beta_2 - 1)\left(LGHOUS_{t-1} - \frac{\beta_1}{1-\beta_2} - \frac{\beta_3 + \beta_4}{1-\beta_2} LGDPI_{t-1} - \frac{\beta_5 + \beta_6}{1-\beta_2} LPRHOUS_{t-1} \right) +$$
$$+ \beta_3 \Delta LGDPI_t + \beta_5 \Delta LPRHOUS_t + \varepsilon_t .$$

Engle-Granger Two Step Procedure: Example.

Cointegrating relationship:

$$LGHOUS_t = \frac{\beta_1}{1-\beta_2} + \frac{\beta_3+\beta_4}{1-\beta_2} LGDPI_t + \frac{\beta_5+\beta_6}{1-\beta_2} LPRHOUS_t$$

Dependent Variable: LGHOUS Method: Least Squares
Sample: 1959 2003 Included observations: 45

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.0056	0.168	0.034	0.97
LGDPI	1.032	0.0066	155.2	0.000
LPRHOUS	-0.483	0.042	-11.57	0.000

The residuals are marginally stationary. Suppose they are stationary, and we have got cointegrating relationship:

Engle-Granger Two Step Procedure: Example.

Error Correction Model:

Dependent Variable: DLGHOUS Method: Least Squares
Sample (adjusted): 1960 2003 Included observations: 44

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLGDPI	0.938	0.049	19.22	0.000
DLPRHOUS	-0.498	0.122	-4.070	0.0002
RESID(-1)	-0.311	0.111	-2.80	0.0077

R-squared	0.250	Mean dependent var	0.034
S.D. dependent var	0.013	S.E. of regression	0.012
Sum squared resid	0.0057	Durbin-Watson stat	1.626

The coefficients of *DLGDPI* and *DLPRHOUS* provide estimates of the short-run income and price elasticities of demand for HOUSING. They are slightly lower than the long-run elasticities estimated in the cointegrating relationship.

The coefficient of the cointegrating term *RESID(-1)* indicates that about 31 per cent of the gap is eliminated in one year.

Granger Causality

Granger test for causality (1969): regress current value of *X* on the past values of *X* and *Y*, and the current value of *Y* on the past values of *X* and *Y*:

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \ldots + \alpha_m Y_{t-m} + \beta_1 X_{t-1} + \ldots + \beta_m X_{t-m} + \varepsilon_t \quad (1)$$

$$X_t = \alpha_0 + \alpha_1 X_{t-1} + \ldots + \alpha_m X_{t-m} + \beta_1 Y_{t-1} + \ldots + \beta_m Y_{t-m} + \varepsilon_t \quad (2)$$

Test statistic: F-statistic for H_0: $\beta_1 = \beta_2 = \ldots = \beta_m = 0$. If in (1) *Ho* is rejected, then *X* Granger causes *Y*.

If in (2) *Ho* is rejected, then *Y* Granger causes *X*.

Independence, or Unidirectional or bilateral Granger causality is possible.

EViews: select the group of series, indicate “Granger Causality” option in “View” menu. Then EViews runs bivariate regressions for all possible pairs of series in the group selected. The reported *F*-statistics are the Wald test statistics for H_0 for each equation.

Granger Causality Tests: Example, PTPE, PFOOD and PHOUS (EViews)

Pairwise Granger Causality Tests

Sample: 1959 2003
Lags: 2

Null Hypothesis:	Obs	F-Statistic	Probability
PFOOD does not Granger Cause PTPE	43	4.78473	0.01402
PTPE does not Granger Cause PFOOD		0.02229	0.97797
PHOUS does not Granger Cause PTPE	43	0.18886	0.82868
PTPE does not Granger Cause PHOUS		6.44592	0.00389
PHOUS does not Granger Cause PFOOD	43	0.13250	0.87630
PFOOD does not Granger Cause PHOUS		7.16765	0.00228

Number of lags is fixed by the user. For Lag=3 conclusions are the same.

Granger Causality Tests, Example (Eviews):
COVID 19 cases in Moscow, Moscow oblast and Tver oblast, 25.03-02.08, 2020

Pairwise Granger Causality Tests
Date: 02/24/21 Time: 16:43
Sample: 1 131
Lags: 5

Null Hypothesis:	Obs	F-Statistic	Prob.
MO does not Granger Cause MOSCOW	126	5.54762	0.0001
MOSCOW does not Granger Cause MO		1.52675	0.1869
TVERO does not Granger Cause MOSCOW	126	1.13723	0.3448
MOSCOW does not Granger Cause TVERO		4.54270	0.0008
TVERO does not Granger Cause MO	126	3.63247	0.0044
MO does not Granger Cause TVERO		0.45441	0.8093

Vector Autoregression (VAR)

Vector autoregression (VAR) is used for systems of interrelated time series. In the VAR model every endogenous variable in the system is a function of the lagged values of all endogenous variables in the system, and of the values of exogenous variables.

VAR model:

$$Y_t = A_1 Y_{t-1} + \ldots + A_m Y_{t-m} + BX_t + \varepsilon_t$$

where Y is a vector of endogenous variables, X is a vector of exogenous variables, A and B are matrices of coefficients to be estimated, and ε is a vector of innovations (uncorrelated with their own lagged values and with all of the right-hand side variables).

Vector Autoregression (VAR), Example (EViews): PTPE, PFOOD, PHOUS

EViews, example: VAR(PTPE,PFOOD,PHOUS)
Endogenous variables: PTPE PFOOD PHOUS
Lag Intervals for Endogenous: 1 1

Vector Autoregression Estimates

Sample (adjusted): 1960 2003

Included observations: 44 after adjustments

t-statistics in [ ]

	PTPE	PHOUS	PFOOD
PTPE(-1)	0.742	-0.176	-0.307
	[8.846]	[-2.94]	[-1.82]
PHOUS(-1)	-0.065	0.952	0.003
	[-1.736]	[35.42]	[0.040]
PFOOD(-1)	0.348	0.265	1.324
	[6.230]	[6.647]	[11.75]
C	-0.626	-1.057	0.041
	[-2.32]	[-5.50]	[0.076]

Vector Autoregression (VAR), Example (Eviews): COVID 19 cases in Moscow, Moscow oblast and Tver oblast, 25.03-02.08, 2020

Vector Autoregression Estimates
Date: 02/24/21 Time: 16:54
Sample (adjusted): 3 131
Included observations: 129 after adjustments
Standard errors in ( ) & t-statistics in [ ]

	MOSCOW	MO	TVERO
MOSCOW(-1)	0.578232	0.006031	0.002788
	(0.08222)	(0.00738)	(0.00129)
	[7.03292]	[0.81768]	[2.16738]
MOSCOW(-2)	0.229729	0.004296	-0.005486
	(0.08294)	(0.00744)	(0.00130)
	[2.76991]	[0.57736]	[-4.22749]
MO(-1)	4.644605	0.743408	0.014723
	(0.99020)	(0.08883)	(0.01549)
	[4.69056]	[8.36934]	[0.95022]
MO(-2)	-3.804878	0.207676	0.003171
	(0.99186)	(0.08897)	(0.01552)
	[-3.83611]	[2.33413]	[0.20434]
TVERO(-1)	-1.714575	0.702538	0.409965
	(4.94875)	(0.44392)	(0.07744)
	[-0.34647]	[1.58257]	[5.29419]
TVERO(-2)	-4.338766	-1.033738	0.401060
	(4.92489)	(0.44178)	(0.07706)
	[-0.88099]	[-2.33992]	[5.20428]
C	137.5249	19.37508	2.972403
	(117.517)	(10.5417)	(1.83887)
	[1.17026]	[1.83794]	[1.61643]
R-squared	0.860812	0.969853	0.825001