

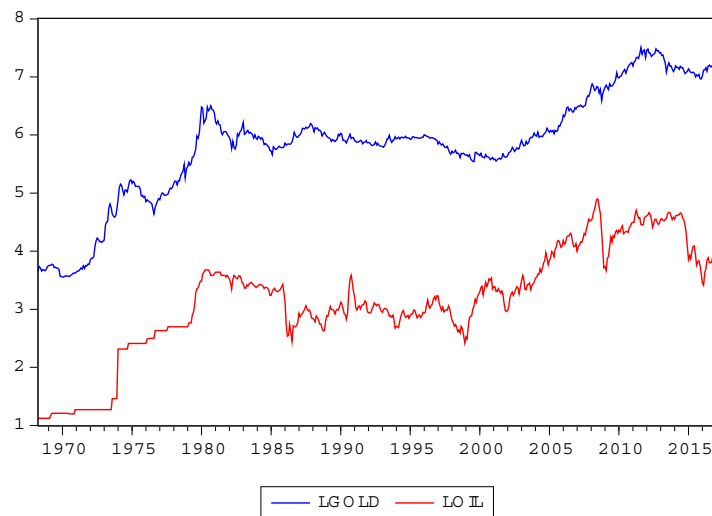
Solution to Exercise 7

Simon A. Broda

1. (a) If the relative price of oil expressed in units of gold, $\text{oil}_t/\text{gold}_t$, is stationary, then this implies that $\log(\text{oil}_t/\text{gold}_t) = \log(\text{oil}_t) - \log(\text{gold}_t)$ is also stationary, so $\log(\text{oil}_t)$ and $\log(\text{gold}_t)$ must be cointegrated with cointegrating vector (1, -1) if the individual series are integrated.
- (b) We begin by transforming the data to logs and making a plot:

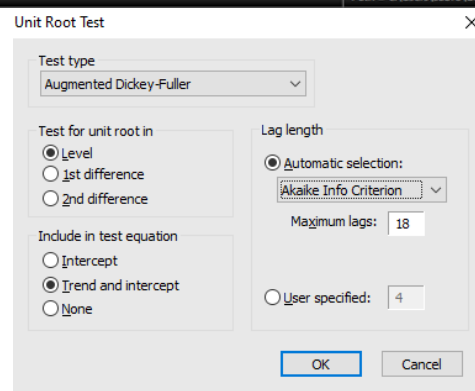
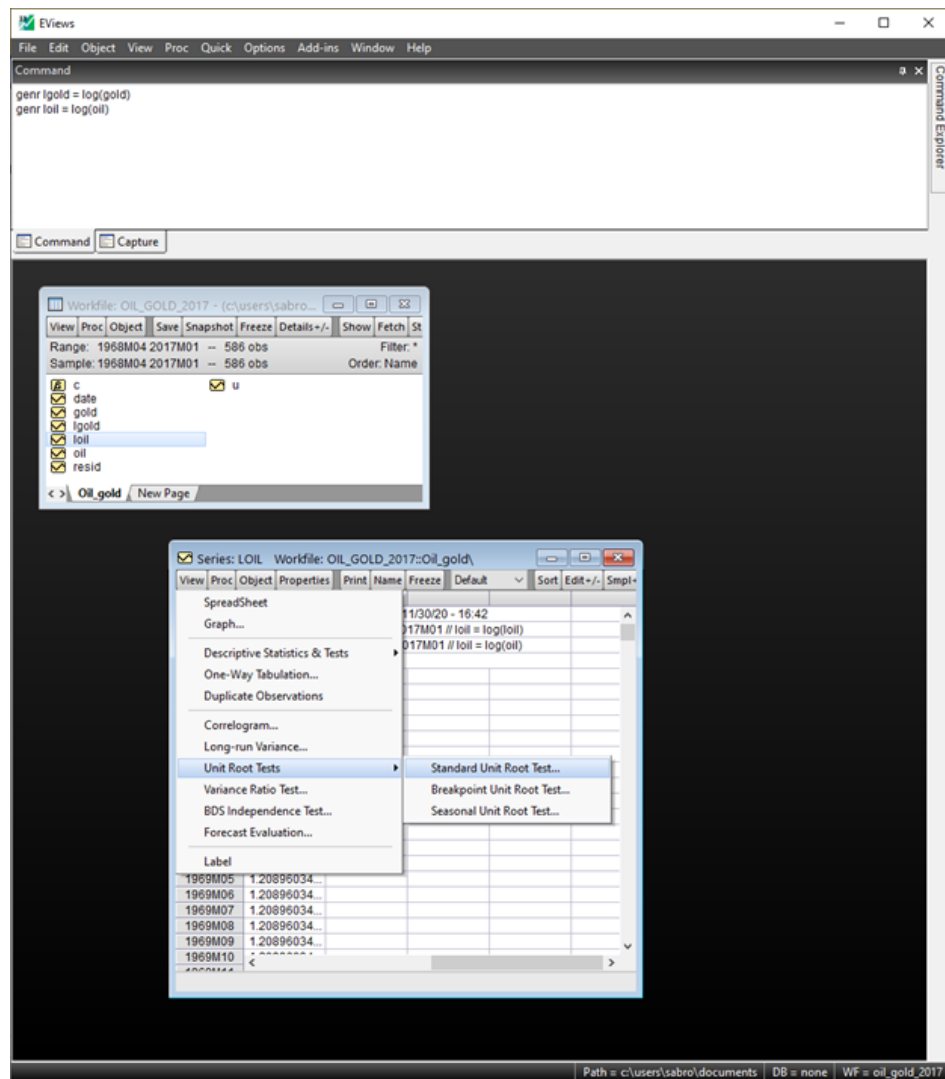
```
genr lgold = log(gold)
genr loil = log(oil)
```

Plotting the data requires opening the two series as a group. The resulting plot is given below.



Then we follow the Engle-Granger procedure.

Step 1: Conduct individual unit root tests to make sure both series are integrated. We include a time trend for both because the data look trending, and choose the lag length automatically by the AIC.



The results is shown below.

Null Hypothesis: LOIL has a unit root
 Exogenous: Constant, Linear Trend
 Lag Length: 1 (Automatic - based on SIC, maxlag=18)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-2.552422	0.3027
Test critical values: 1% level	-3.973847	
5% level	-3.417533	
10% level	-3.131184	

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

Null Hypothesis: LGOLD has a unit root
 Exogenous: Constant, Linear Trend
 Lag Length: 0 (Automatic - based on SIC, maxlag=18)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-1.820648	0.6937
Test critical values: 1% level	-3.973820	
5% level	-3.417519	
10% level	-3.131176	

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

EViews chose to include one lagged difference. Neither test rejects, so the series are I(1).

Step 2: Estimate the long-run relationship (cointegrating relationship)

$$\text{loil}_t = \beta_1 + \beta_2 \text{lgold}_t + U_t.$$

Equation Estimation

Specification Options

Equation specification

Dependent variable followed by list of regressors including ARMA and PDL terms, OR an explicit equation like $Y=c(1)+c(2)*X$.

loil c lgold

Estimation settings

Method: LS - Least Squares (NLS and ARMA)

Sample: 1968M04 2017M01

OK Cancel

The result is

Dependent Variable: LOIL
Method: Least Squares
Date: 11/30/20 Time: 16:50
Sample: 1968M04 2017M01
Included observations: 586

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-2.258963	0.086329	-26.16698	0.0000
LGOLD	0.926232	0.014618	63.36354	0.0000
R-squared	0.873014	Mean dependent var		3.141848
Adjusted R-squared	0.872797	S.D. dependent var		0.929639
S.E. of regression	0.331561	Akaike info criterion		0.633398
Sum squared resid	64.20072	Schwarz criterion		0.648324
Log likelihood	-183.5855	Hannan-Quinn criter.		0.639214
F-statistic	4014.938	Durbin-Watson stat		0.074317
Prob(F-statistic)	0.000000			

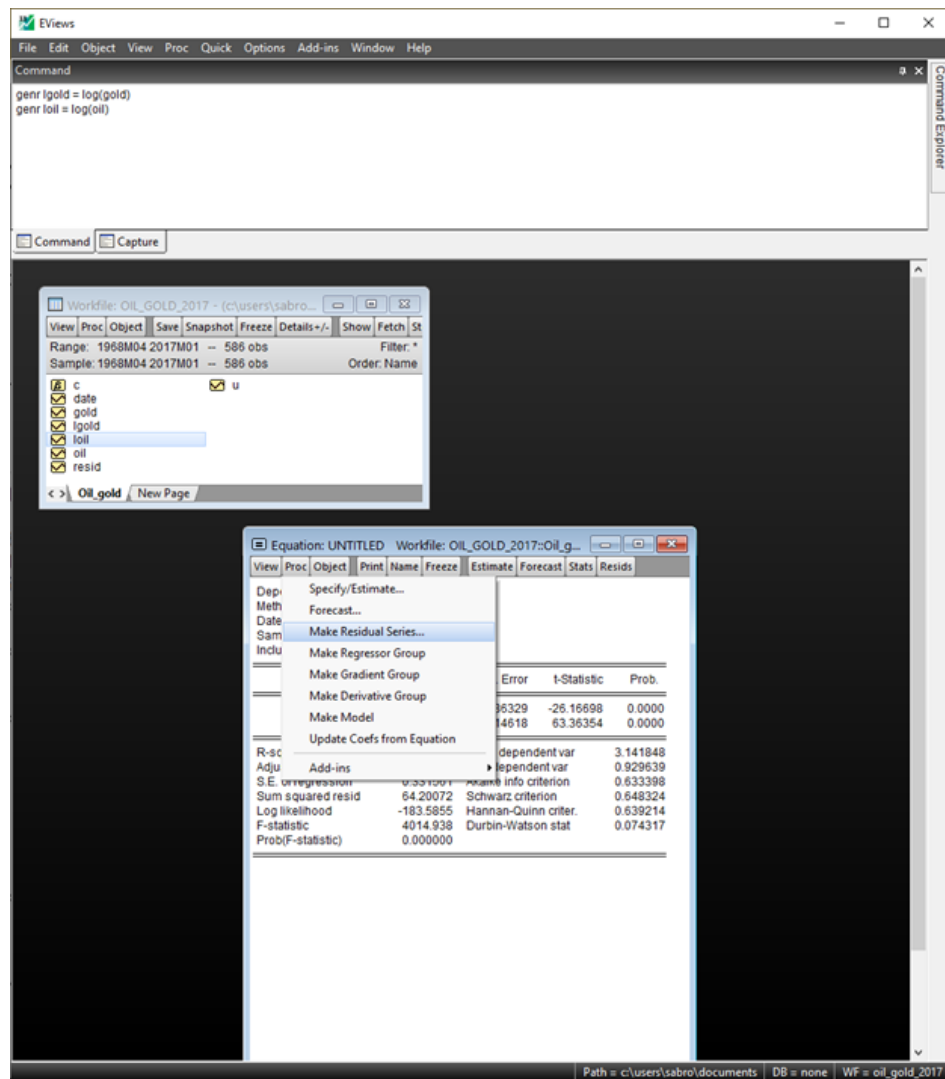
The estimated long-run relationship is

$$\text{loil}_t = -2.26 + 0.926\text{lgold}_t + U_t.$$

The estimated cointegrating vector (provided we find cointegration) is $(1, -0.926)$, i.e.,

$$\text{loil}_t - 0.926\text{lgold}_t = -2.26 + U_t$$

is stationary. We save the residuals for later use and plot them:



Make Residuals



Residual type

☒ Ordinary

☐ Standardized

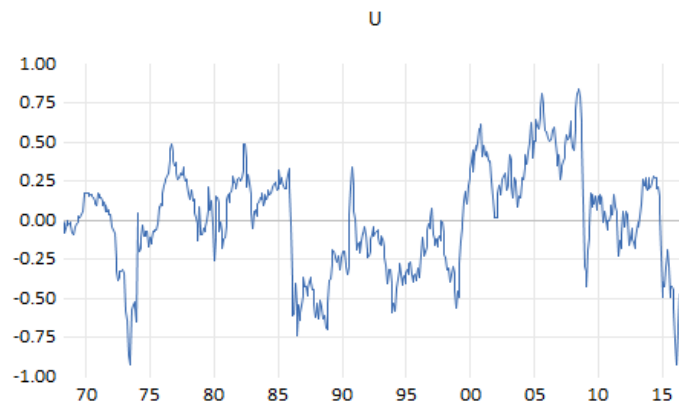
☐ Generalized

OK

Name for resid series

u

Cancel



Step 3: Conduct an ADF test (with just an intercept, no trend!) for the residuals to test H_0 : No Cointegration. Result:

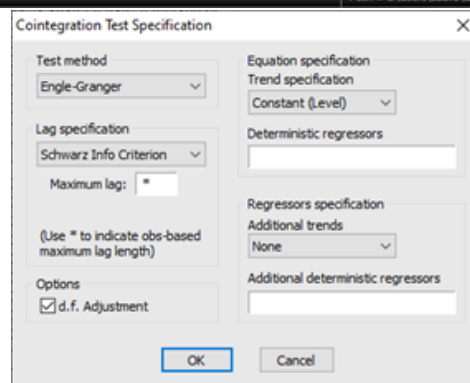
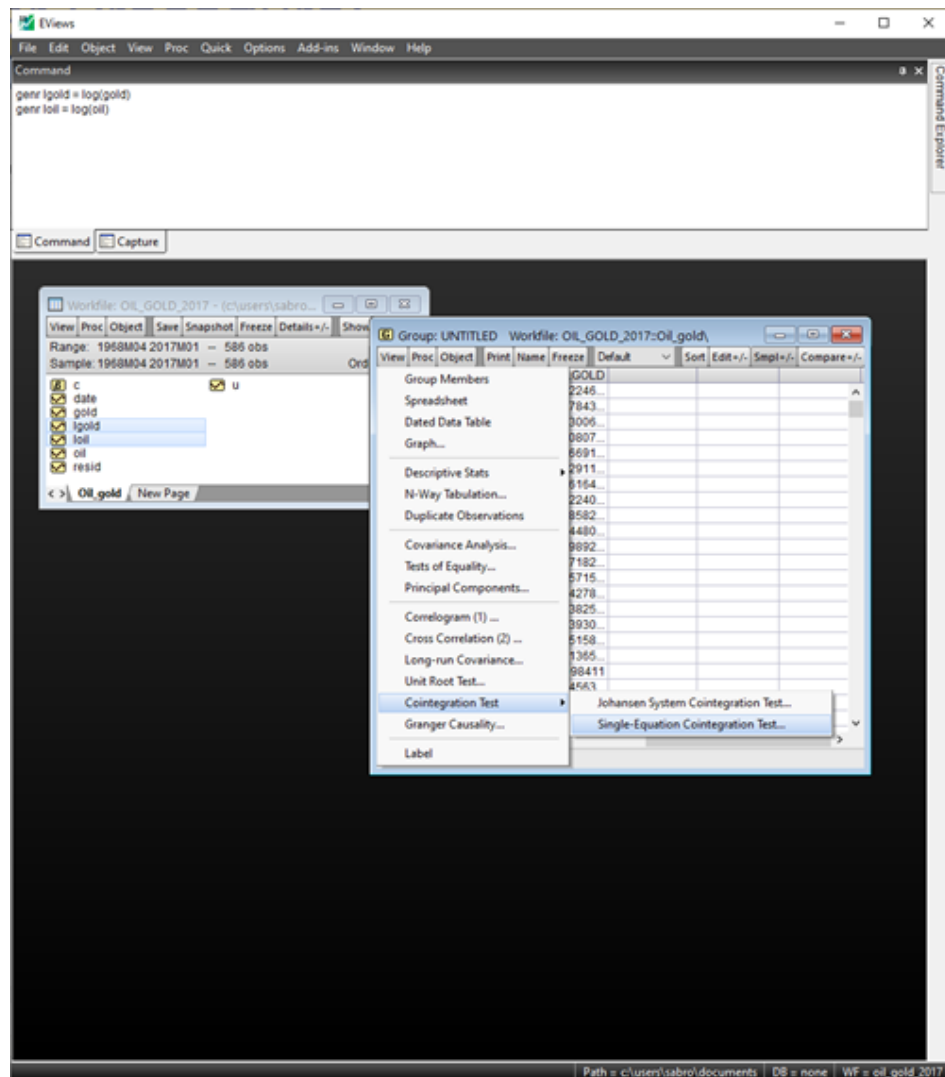
Null Hypothesis: U has a unit root
Exogenous: Constant
Lag Length: 1 (Automatic - based on SIC, maxlag=18)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-3.724771	0.0040
Test critical values: 1% level	-3.441318	
5% level	-2.866270	
10% level	-2.569348	

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

Be careful to use the Engle-Granger critical value of -3.41 at the 5% level. The test rejects the null. Conclusion: there is indeed cointegration.

Alternative to this manual approach: select `l oil` and `lgold` (click on one, press control, click on the other), and open them as a group. Then do



Result:

Date: 11/30/20 Time: 20:28
Series: LOIL LGOLD
Sample: 1968M04 2017M01
Included observations: 586
Null hypothesis: Series are not cointegrated
Cointegrating equation deterministics: C
Automatic lags specification based on Schwarz criterion (maxlag=18)

Dependent	tau-statistic	Prob.*	z-statistic	Prob.*
LOIL	-3.728235	0.0177	-28.50078	0.0087
LGOLD	-3.488501	0.0346	-25.44978	0.0171

*Mackinnon (1996) p-values.

Same result (first row), and we even get a p -value! Note that EViews does the test “in both directions”, once with Y_t as dependent variable and once with X_t . As mentioned in the slides, this doesn’t matter asymptotically, but in finite samples it might matter which of the variables we consider endogenous and which one exogenous. In this particular case though, they both give the same answer. Still, it would be better not to have to make a choice here. That’s what the Johansen procedure accomplishes.

Step 4: Estimate the VECM

$$\begin{aligned}\Delta Y_t &= c_1 + \alpha_1(Y_{t-1} - \beta_1 - \beta_2 X_{t-1}) + e_{1t}, \\ \Delta X_t &= c_2 + \alpha_2(Y_{t-1} - \beta_1 - \beta_2 X_{t-1}) + e_{2t}.\end{aligned}$$

replacing $Y_{t-1} - \beta_1 - \beta_2 X_{t-1}$ by the OLS residual $\hat{u}_{t-1} = Y_{t-1} - \hat{\beta}_1 - \hat{\beta}_2 X_{t-1}$ we saved earlier. Then, we can estimate α_1 and α_2 by OLS. We start with the equation for LOIL:

Equation Estimation

Specification Options

Equation specification

Dependent variable followed by list of regressors including ARMA and PDL terms, OR an explicit equation like $Y=c(1)+c(2)*X$.

$d(oil)c u(-1)$

Estimation settings

Method: **LS - Least Squares (NLS and ARMA)**

Sample: 1968m04 2017m01

OK Cancel

Dependent Variable: D(LOIL)
Method: Least Squares
Date: 11/30/20 Time: 20:16
Sample (adjusted): 1968M05 2017M01
Included observations: 585 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.004870	0.003377	1.442012	0.1498
U(-1)	-0.027903	0.010205	-2.734171	0.0064
<hr/>				
R-squared	0.012660	Mean dependent var		0.004853
Adjusted R-squared	0.010967	S.D. dependent var		0.082140
S.E. of regression	0.081688	Akaike info criterion		-2.168397
Sum squared resid	3.890354	Schwarz criterion		-2.153451
Log likelihood	636.2562	Hannan-Quinn criter.		-2.162573
F-statistic	7.475692	Durbin-Watson stat		1.498327
Prob(F-statistic)	0.006444			

We notice that there is autocorrelation (look at the DW stat). We can cure this by adding a lagged difference (like the ADF test did automatically when it selected the lag length).

Equation Estimation

Specification Options

Equation specification

Dependent variable followed by list of regressors including ARMA and PDL terms, OR an explicit equation like $Y=c(1)+c(2)*X$.

$d(loil) c u(-1) d(loil(-1))$

Estimation settings

Method: LS - Least Squares (NLS and ARMA)

Sample: 1968m04 2017m01

OK Cancel

Dependent Variable: D(LOIL)
Method: Least Squares
Date: 11/30/20 Time: 17:09
Sample (adjusted): 1968M06 2017M01
Included observations: 584 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.003641	0.003278	1.110606	0.2672
U(-1)	-0.034971	0.009942	-3.517509	0.0005
D(LOIL(-1))	0.256547	0.040092	6.399038	0.0000
R-squared	0.077666	Mean dependent var		0.004862
Adjusted R-squared	0.074491	S.D. dependent var		0.082210
S.E. of regression	0.079089	Akaike info criterion		-2.231362
Sum squared resid	3.634195	Schwarz criterion		-2.208914
Log likelihood	654.5578	Hannan-Quinn criter.		-2.222613
F-statistic	24.46179	Durbin-Watson stat		2.006622
Prob(F-statistic)	0.000000			

Now it's fine. Let's repeat for the other variable:

Equation Estimation ×

Specification Options

Equation specification

Dependent variable followed by list of regressors including ARMA and PDL terms, OR an explicit equation like $Y=c(1)+c(2)*X$.

`d(lgold) c u(-1)`

Estimation settings

Method: LS - Least Squares (NLS and ARMA)

Sample: 1968m04 2017m01

OK
Cancel

Dependent Variable: D(LGOLD)
Method: Least Squares
Date: 11/30/20 Time: 17:17
Sample (adjusted): 1968M05 2017M01
Included observations: 585 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.005866	0.002342	2.504458	0.0125
U(-1)	0.008997	0.007077	1.271298	0.2041
R-squared	0.002765	Mean dependent var		0.005871
Adjusted R-squared	0.001054	S.D. dependent var		0.056676
S.E. of regression	0.056646	Akaike info criterion		-2.900562
Sum squared resid	1.870743	Schwarz criterion		-2.885616
Log likelihood	850.4143	Hannan-Quinn criter.		-2.894737
F-statistic	1.616198	Durbin-Watson stat		1.932704
Prob(F-statistic)	0.204130			

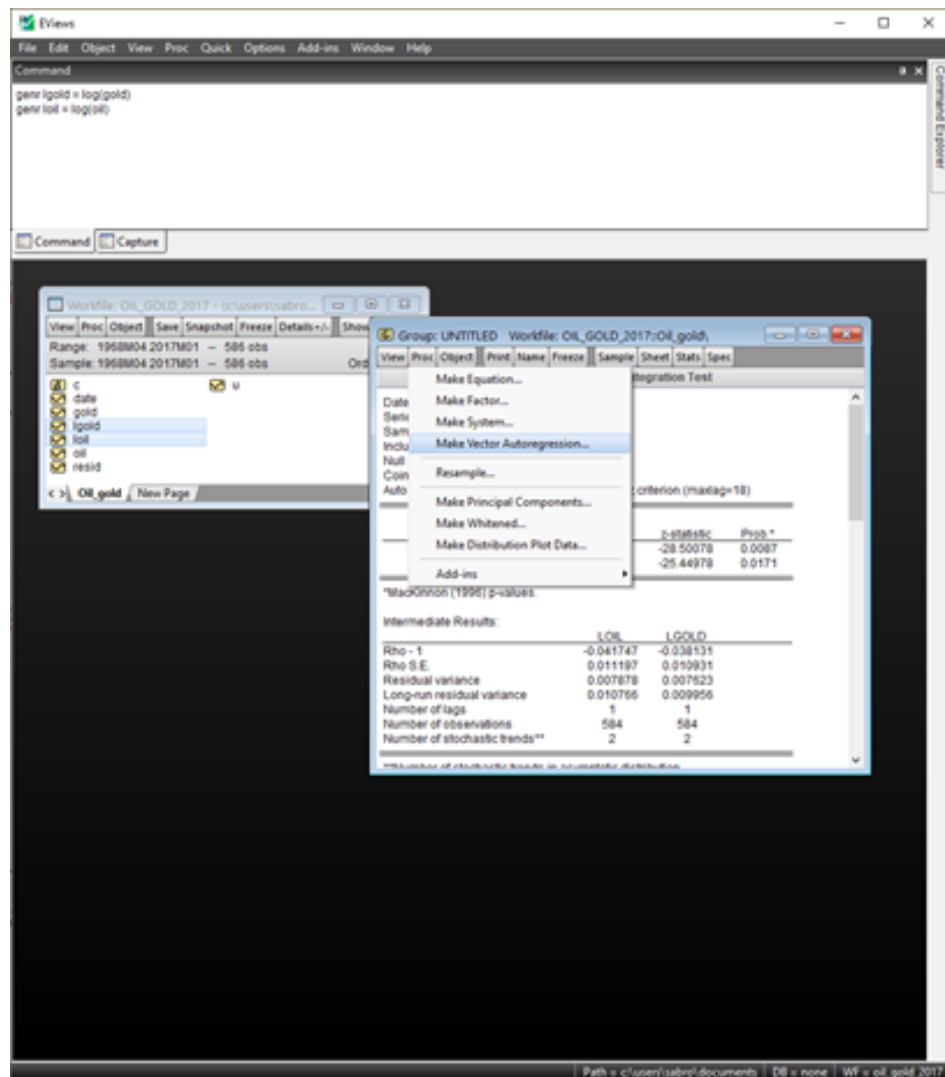
This one seems fine without any lagged differences. We notice that u_{t-1} is insignificant, so we can stick with a single-equation ECM. The final model is

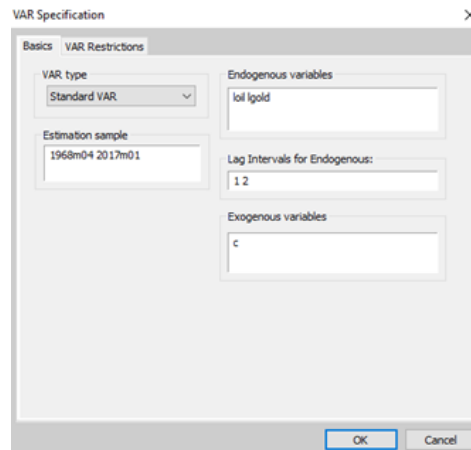
$$\Delta \text{loil}_t = 0.0036 - 0.035(\text{loil}_{t-1} - 0.926 \text{lgold}_{t-1} + 2.26) + 0.25 \Delta \text{loil}_{t-1} + e_{1t}$$

Interpretation: there is an equilibrium relationship between `loil` and `lgold`, with cointegrating vector $(1, -0.926)$. In case of a disequilibrium, `loil` adjusts towards the equilibrium. The adjustment amounts to 3.5% of the disequilibrium per period.

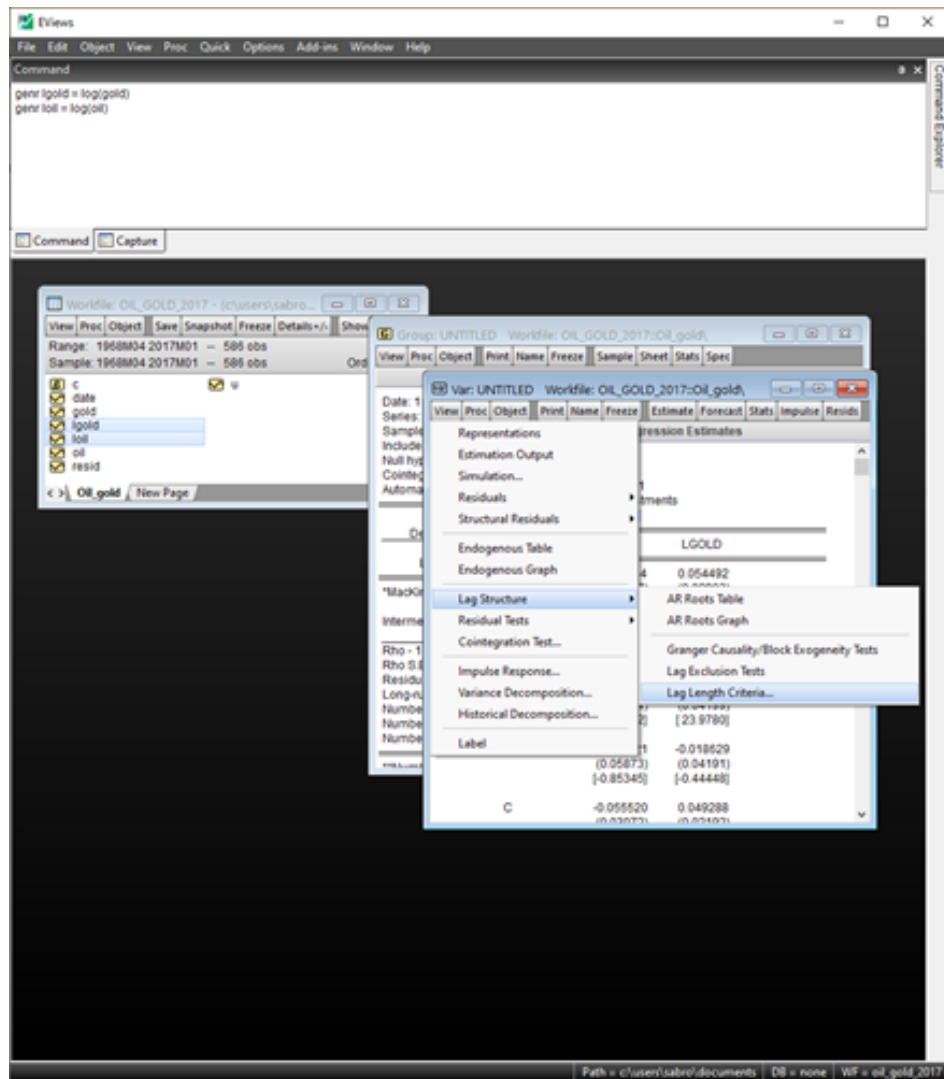
- (c) We now repeat the analysis, but using Johansen's procedure, rather than Engle and Granger's.

Step 0 The first step is to pick a lag length. In slides, we didn't do this: we just picked a lag length of 1 because that's what we used in the Engle-Granger procedure. If we hadn't done Engle-Granger and started with Johansen right away, we'd need a way to do this. The easiest way is the following: open the variables as a group like before, then click on Proc→Make Vector Autoregression.





Keep the defaults, hit enter, and ignore the output. We're just doing this to get to the multivariate information criteria, by clicking View→Lag Structure→Lag length criteria:



This yields

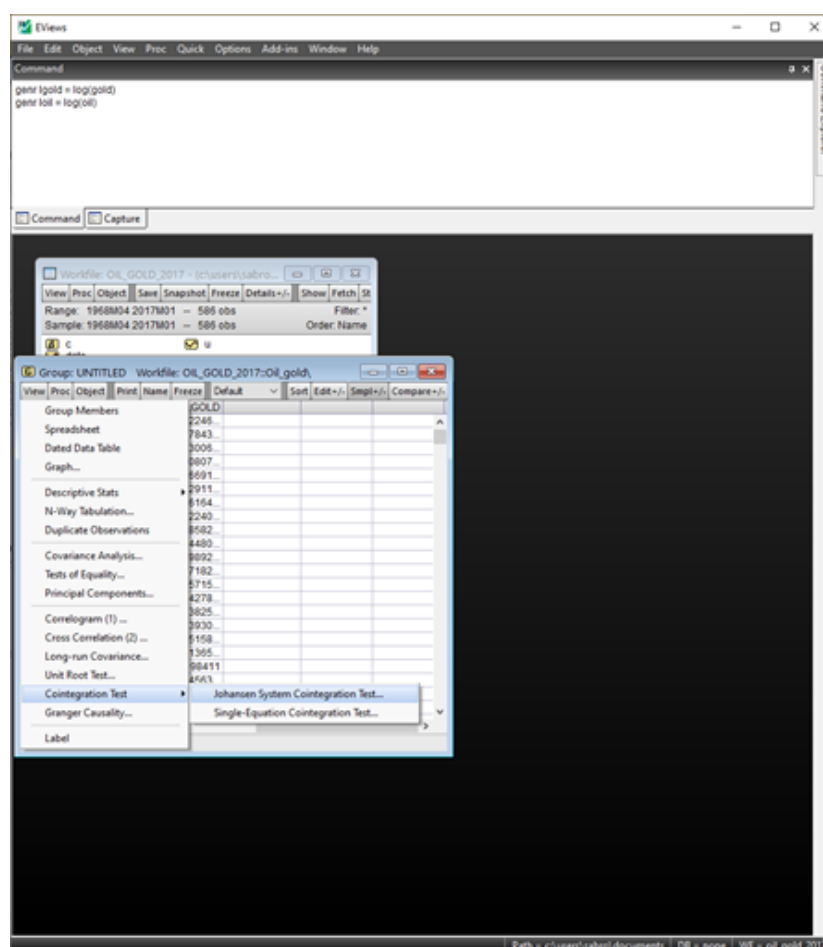
VAR Lag Order Selection Criteria
Endogenous variables: LOIL LGOLD
Exogenous variables: C
Date: 11/30/20 Time: 20:40
Sample: 1968M04 2017M01
Included observations: 578

Lag	LogL	LR	FPE	AIC	SC	HQ
0	-949.7189	NA	0.092309	3.293145	3.308230	3.299027
1	1473.535	4821.353	2.14e-05	-5.077976	-5.032721	-5.060329
2	1493.683	39.94747*	2.02e-05*	-5.133851*	-5.058426*	-5.104441*
3	1494.055	0.735532	2.05e-05	-5.121298	-5.015703	-5.080124
4	1494.302	0.485776	2.07e-05	-5.108311	-4.972546	-5.055373
5	1496.013	3.357075	2.09e-05	-5.100391	-4.934456	-5.035688
6	1498.043	3.969130	2.10e-05	-5.093575	-4.897470	-5.017108
7	1501.239	6.225229	2.11e-05	-5.090792	-4.864516	-5.002561
8	1502.698	2.832684	2.13e-05	-5.082000	-4.825555	-4.982005

* indicates lag order selected by the criterion
LR: sequential modified LR test statistic (each test at 5% level)
FPE: Final prediction error
AIC: Akaike information criterion
SC: Schwarz information criterion
HQ: Hannan-Quinn information criterion

Most criteria agree on 2 lags. For technical reasons (this is a VAR, not a VECM), this means that we need 1 lag in the Johansen test and in the VECM.

Step 1 Conduct the Johansen test. Open the variables as a group as before, then do



Choose Option 3 to allow for a trend, and specify the lag interval as 1 1 (this means

all lagged differences from lag 1 to lag 1), because that is what the information criteria suggested.

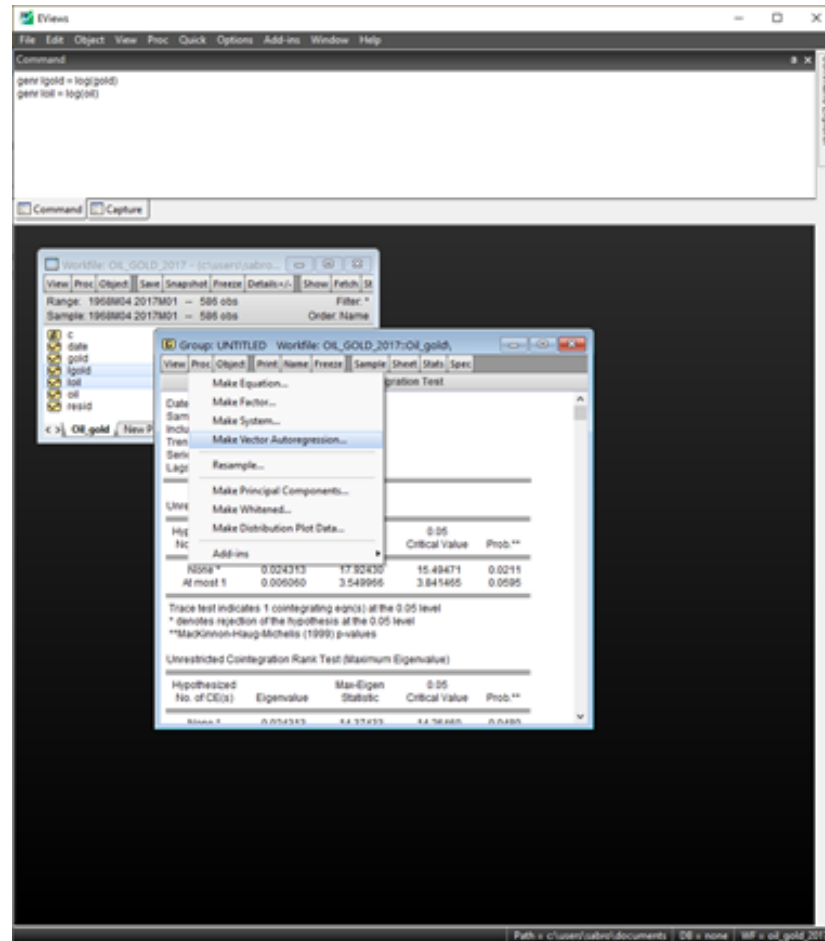
Result:

Unrestricted Cointegration Rank Test (Trace)				
Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Prob.**
None *	0.024313	17.92430	15.49471	0.0211
At most 1	0.006060	3.549966	3.841465	0.0595

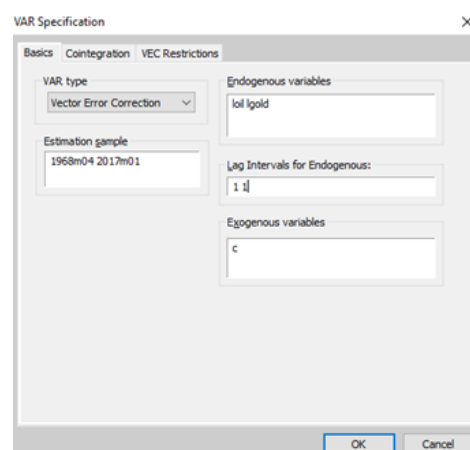
Trace test indicates 1 cointegrating eqn(s) at the 0.05 level
 * denotes rejection of the hypothesis at the 0.05 level
 **Mackinnon-Haug-Michelis (1999) p-values

The test for H_{00} : “No cointegration” rejects. The test for H_{01} : “1 cointegration relationship” accepts. Conclusion: there is one cointegrating relationship. If the latter test had rejected as well, then that would mean that the model is stationary; i.e., the variables aren’t integrated in the first place.

Step 2: Make a vector autoregression. In the group window, click on Proc→Make Vector Autoregression.



Choose “Vector Error Correction”, and set the lag interval to match what the criteria found.



Go to the Cointegration tab, choose the same model as for the Johansen test (so 3 in this case), and set the number of cointegrating relationships to what the test found (so 1 in this case).

VAR Specification

Rank: 1

Number of cointegrating: 1

Deterministic Trend Specification

No trend in data

☐ 1) No intercept or trend in CE or VAR

☐ 2) Intercept (no trend) in CE - no intercept in VAR

Linear trend in data

☒ 3) Intercept (no trend) in CE and VAR

☐ 4) Intercept and trend in CE - no trend in VAR

Quadratic trend in data

☐ 5) Intercept and trend in CE - linear trend in VAR

OK Cancel

Result:

Vector Error Correction Estimates
Date: 11/30/20 Time: 20:59
Sample (adjusted): 1968M06 2017M01
Included observations: 584 after adjustments
Standard errors in () & t-statistics in []

Cointegrating Eq:	CointEq1
LOIL(-1)	1.000000
LGOLD(-1)	-0.914126 (0.09263) [-9.86821]
C	2.187722

Error Correction:	D(LOIL)	D(LGOLD)
CointEq1	-0.034498 (0.00997) [-3.46000]	0.007463 (0.00713) [1.04641]
D(LOIL(-1))	0.250758 (0.04063) [6.17135]	0.048515 (0.02906) [1.66924]
D(LGOLD(-1))	0.050704 (0.05870) [0.86374]	0.019596 (0.04199) [0.46668]
C	0.003354 (0.00329) [1.01839]	0.005412 (0.00236) [2.29777]

R-squared	0.079176	0.008236
Adj. R-squared	0.074413	0.003106
Sum sq. resids	3.628244	1.856348
S.E. equation	0.079092	0.056574
F-statistic	16.62361	1.605456
Log likelihood	655.0364	850.7165
Akaike AIC	-2.229577	-2.899714
Schwarz SC	-2.199646	-2.869783
Mean dependent	0.004862	0.005761
S.D. dependent	0.082210	0.056662

Determinant resid covariance (dof adj.)	1.96E-05
Determinant resid covariance	1.93E-05
Log likelihood	1512.206
Akaike information criterion	-5.144541
Schwarz criterion	-5.069714
Number of coefficients	10

The red part is the VECM equation for LOIL, the green part is that for LGOLD, and the blue part is the cointegrating relationship. The cointegrating vector is $(1, -0.914)$. The results are remarkably similar to what we found with the Engle-Granger procedure. We could again ignore the equation for LGOLD, because the adjustment coefficient (α_2 , in yellow) is insignificant (t -statistic 1.04). The final model is

$$\Delta \text{loil}_t = 0.0034 - 0.034(\text{loil}_{t-1} - 0.914\text{lgold}_{t-1} + 2.19) + 0.25\Delta \text{loil}_{t-1} + 0.05\Delta \text{lgold}_{t-1} + e_{1t}$$

$$\Delta \text{lgold}_t = 0.0054 - 0.007(\text{loil}_{t-1} - 0.914\text{lgold}_{t-1} + 2.19) + 0.049\Delta \text{loil}_{t-1} + 0.019\Delta \text{lgold}_{t-1} + e_{2t}$$

2. (a) X_t is a random walk, hence $I(1)$. So no, it is not stationary.
 (b) Y_t depends on X_t if $\beta_2 \neq 0$, so it cannot be stationary.
 (c) Yes, because there exists a linear combination of them that is stationary:

$$Y_t - \beta_2 X_t = \beta_1 + U_{1,t}.$$

The cointegrating vector is $(1, -\beta_2)$.

- (d) The goal is to find two equations, one with ΔY_t on the LHS, and one with ΔX_t . Both should have the equilibrium error $Y_{t-1} - \beta_1 - \beta_2 X_{t-1}$ on the RHS.

For Y_t , we find

$$\begin{aligned} Y_t &= \beta_1 + \beta_2 X_t + U_{1,t} & | - Y_{t-1} \\ \Delta Y_t &= -Y_{t-1} + \beta_1 + \beta_2 X_t + U_{1,t} & | \pm \beta_2 X_{t-1} \\ \Delta Y_t &= -(Y_{t-1} - \beta_1 - \beta_2 X_{t-1}) + \beta_2 \Delta X_t + U_{1,t} \\ \Delta Y_t &= \alpha_1 (Y_{t-1} - \beta_1 - \beta_2 X_{t-1}) + \beta_2 \Delta X_t + U_{1,t}, \end{aligned}$$

where $\alpha_1 = -1$. For X_t ,

$$\begin{aligned} X_t &= X_{t-1} + U_{2,t} & | - X_{t-1} \\ \Delta X_t &= U_{2,t} \\ \Delta X_t &= 0(Y_{t-1} - \beta_1 - \beta_2 X_{t-1}) + U_{2,t} \\ \Delta X_t &= \alpha_2 (Y_{t-1} - \beta_1 - \beta_2 X_{t-1}) + U_{2,t} \end{aligned}$$

where $\alpha_2 = 0$. This means that we can treat this as a single-equation ECM.