Module 9.3: Time Series Analysis
Fall Term 2022

Week 7:

Cointegration

Lucerne University of Applied Sciences and Arts

Outline in Weeks

1. Introduction; Descriptive Modelling
2. Returns; Autocorrelation; Stationarity
3. ARMA Models
4. Unit Roots; Regressions between Time Series
5. Volatility Modelling
6. Value at Risk
7. Cointegration

Outline

Cointegration and Common Trends

Suppose we have two time series Y_t and X_t, which are both $I(1)$, and we analyze a regression model of the form

$$Y_t = \beta_1 + \beta_2 X_t + U_t.$$

Here U_t has mean zero but may display autocorrelation. Two cases:

- $U_t \sim I(1)$: if U_t displays no mean-reversion, then Y_t does not revert to the explained part $\beta_1 + \beta_2 X_t$. Even if $\beta_2 = 0$, its t-statistic and R^2 will often seem significant (*spurious regressions*). To avoid this, one should estimate a model in differences, i.e., $\Delta Y_t = a_1 + a_2 \Delta X_t + \Delta U_t$.
- $U_t \sim I(0)$: now Y_t and X_t have a *common* stochastic trend, such that the linear combination $Y_t - \beta_2 X_t$ does not have a trend. This is called *cointegration*.

Example

- Consider the model

$$
\begin{aligned}
Y_t &= \beta_1 + \beta_2 X_t + U_{1,t} \\
X_t &= X_{t-1} + U_{2,t}
\end{aligned}
$$

  where $\beta_2 \neq 0$, $U_{1,t}, U_{2,t} \overset{\text{iid}}{\sim} (0, \sigma^2)$ independently of each other.
- X_t is a random walk and thus nonstationary. Y_t contains X_t and is thus also nonstationary. But

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

  is stationary: the RHS is white noise plus a constant.
- $(1, -\beta_2)$ is called the *cointegrating vector*.

Cointegration and Common Trends, contd.

- The concept is easily extended to more than two series: if $X_{2t}, \ldots, X_{kt}$ are all $I(1)$ variables, and

$$Y_t = \beta_1 + \beta_2 X_{2t} + \ldots + \beta_k X_{kt} + U_t,$$

  then this is a spurious regression if $U_t \sim I(1)$ (and $\beta_i = 0$), and a cointegrating relation if U_t is stationary.
- In other words, cointegration between k integrated series means that there exists a *linear combination*[1] of them which is stationary.
- Examples of possibly cointegrated time series:
  - exchange rates and relative prices (*purchasing power parity*);
  - spot and futures prices of assets or exchange rates;
  - short- and long-term interest rates (*term structure models*);
  - stock prices and dividends (*present value relations*).

---

[1] i.e., a weighted sum

Outline

Testing for Cointegration

- For cointegrated series, one should exploit the long-run equilibrium relationship between variables for estimation rather than differencing. Differencing would *remove* that structure.
- Engle and Granger proposed the following procedure:
  - Conduct individual unit root tests to ensure all series are $I(1)$.
  - Estimate the regression model

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

    by ordinary least-squares. Estimates are (super)consistent, but standard errors are wrong because series are $I(1)$.
  - Apply an ADF unit root test (with constant) to the residuals $\hat{u}_t$ from this regression. This yields a test for $H_0 : U_t \sim I(1)$ (spurious regression) against $H_1 : U_t \sim I(0)$ (cointegration). The critical values depend on k. E.g., for $k = 2$ (X_{2t} and an intercept), the 5% c.v. is -3.41.
  - If H_0 is rejected, estimate an *error correction model*.

Engle-Granger critical values

Number of series (w/o constant)	2	3	4	5	6
Critical value	-3.41	-3.80	-4.16	-4.49	-4.74

Error Correction Models

- Cointegration between Y_t and X_t implies that deviations $(Y_{t-1} - \beta_1 - \beta_2 X_{t-1})$ from the equilibrium level should be (partially) corrected in the next period, by Y_t, X_t, or both.
- This leads to a *vector error correction model* (VECM), which in the simplest form is

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

  where e_{1t} and e_{2t} are two white noise errors (possibly correlated), and where we expect $\alpha_1 < 0$ and/or $\alpha_2 \beta_2 > 0$.
- We might need to add lags of ΔY_t and/or ΔX_t on RHS to combat autocorrelation.
- The *Granger representation theorem* states that cointegration implies an error correction model (possibly with more lags), and vice versa; see exercises.

Engle-Granger Procedure

- The VECM

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

  is estimated by replacing $Y_{t-1} - \beta_1 - \beta_2 X_{t-1}$ by OLS residual $\hat{u}_{t-1} = Y_{t-1} - \hat{\beta}_1 - \hat{\beta}_2 X_{t-1}$, and estimating α_1 and α_2 by OLS.
- Note that $\hat{u}$ is stationary, so this is a valid regression!
- If $\alpha_2 = 0$, then all correction is done by Y_t, and not by X_t. In that case it makes sense to treat X_t as exogenous and Y_t as endogenous, and consider the "single-equation" error correction model

$$\Delta Y_t = c + \alpha_1(Y_{t-1} - \beta_1 - \beta_2 X_{t-1}) + e_t.$$

- In general, both Y_t and X_t are endogenous.

Outline

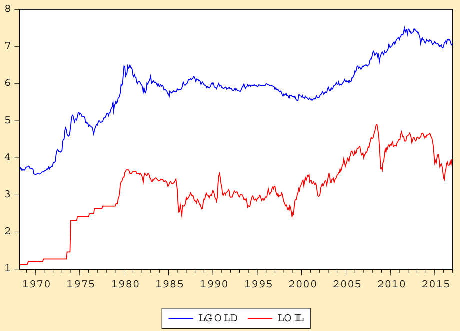
Example

- Until 1971, as part of the Bretton-Woods system of fixed exchange rates, the US dollar was convertible to gold, i.e., it was possible for foreign central banks to redeem US dollars for gold at a fixed rate of 35$ per troy ounce, so that the price of gold was fixed.
- In 1971, US president Nixon unilaterally cancelled the direct convertibility, ultimately ending the Bretton-Woods agreement.
- Gold became a floating asset, and its price increased sharply; in other words, the US$ was massively devalued.

Example continued

- We want to analyze the hypothesis that the increasing price (in US$) of oil is not a consequence of an increased demand for (or a reduced supply of) oil, but rather of a continued devaluation of the US$.
- We have at our disposal monthly data from April 1968 to January 2017 (586 observations) on the following variables:
  - gold_t, the spot price of one troy ounce of gold in US$;
  - oil_t, the spot price of one barrel of WTI crude oil in US$.
- Idea: if the relative price of oil expressed in units of gold $\text{oil}_t/\text{gold}_t$ is stationary, then this implies that $\log(\text{oil}_t) - \log(\text{gold}_t)$ is stationary, so that $\log(\text{oil}_t)$ and $\log(\text{gold}_t)$ must be cointegrated if the individual series are integrated.

$\log(\text{gold}_t)$ and $\log(\text{oil}_t)$



Example continued

Step 0 Take logs: `genr lgold = log(gold), genr loil = log(oil).`

Step 1 Test that the variables are integrated (ADF test with constant and trend)

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.

Neither test rejects, so the series are I(1).

Example continued

Step 2 Estimate long-run relationship

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

and save the residuals (`Proc→Make Residual Series→Ordinary`) as `u`.

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

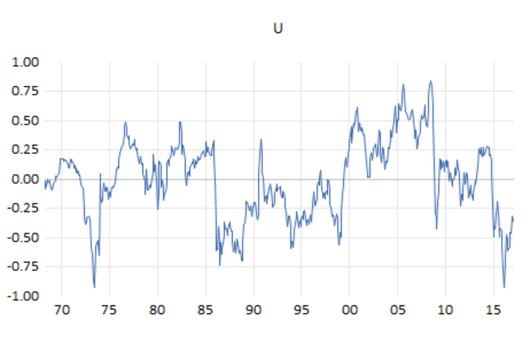
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-2.258963	0.086329	-26.16698	0.0000
LGOLD	0.926232	0.014618	63.36354	0.0000

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 cointegrating vector is $(1, -\beta_2) = (1, -0.926)$ (if the Engle-Granger test rejects).
Careful: standard errors are wrong, because variables are $I(1)$.

Example continued

Residuals (=equilibrium error) $\hat{u}_t$



Example continued

Step 3 Apply ADF test (with intercept) to `u`.

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.

Careful: we need to use Engle-Granger 5% critical value of -3.41, not the ones given by EViews. Conclusion: test rejects the null of "no cointegration".

Example continued

Step 4 Estimate VECM. Note: I threw in a lag of the dependent variable in the equation for `d(loil)`, as there was autocorrelation without it (cf. selected lag length in ADF test).

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		

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		

The adjustment coefficient α_2 in the equation for `d(lgold)` is insignificant. So all the adjustment is done by `loil` $\rightarrow$ single-equation ECM.

Example continued

- The final model is the single-equation ECM

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

  In our earlier notation,

$$\Delta Y_t = c + \alpha_1 (Y_{t-1} - \beta_1 - \beta_2 X_{t-1}) + \gamma \Delta Y_{t-1} + e_{1t},$$

  with $c = 0.0036$, $\alpha_1 = -0.035 < 0$ as desired, $\beta_1 = -2.26$, $\beta_2 = 0.926$, and $\gamma = 0.25$.
- Interpretation: there is an equilibrium relationship between loil and lgold. In case of a disequilibrium, loil adjusts towards the equilibrium. The adjustment amounts to 3.5% per period.

Outline

Limitations of Engle-Granger Procedure:

- Choice of the dependent variable Y_t in Engle-Granger test is arbitrary. Selecting another X_{it} as dependent variable should not matter asymptotically (as number of observations $T \to \infty$), but will make a difference in practice.
- Method can only be used if there is only a single unique cointegrating relation, which involves Y_t. So $(X_{2t}, \ldots, X_{kt})$ are not allowed to be cointegrated without Y_t.
- Although the OLS estimators $\hat{\beta}_i$ are consistent under cointegration, they do not have an asymptotic normal distribution, so standard inference (t-tests) fails.

Johansen Procedure

- An alternative is to use *Johansen's procedure*.
- Johansen has derived the MLE of the general VECM with normally distributed errors $\mathbf{u}_t$ and k variables.
- This facilitates likelihood ratio tests for H_{0r}: "r or fewer cointegrating relationships" against the alternative "more than r"
- The test is known as the *Johansen trace test*. The test statistics $\lambda_{trace}(r)$ can be expressed in terms of particular eigenvalues $\hat{\lambda}_i$. Their asymptotic distribution under the null is a multivariate version of the Dickey-Fuller distribution.
- We reject for large positive values of the test statistic; critical values and p-values are built into EViews and similar programs.

Johansen's cointegration test

These tests may be used to estimate the cointegrating rank r in the following way:

1. Start with $r = 0$;
2. Test H_{0r} with $\lambda_{trace}(r)$;
3. If H_{0r} is not rejected, then $\hat{r} = r$; if it is rejected, replace r by $r + 1$ and go back to step 2;
4. If H_{0r} is rejected for all $r = 0, 1, \ldots, k - 1$, then conclude $\hat{r} = k$ (this corresponds to a stationary system).

Treatment of constant and linear trend

Just like with the Dickey-Fuller test, we have to allow for a constant and possibly a linear trend in the VECM. Eviews distinguishes five cases:

1. No constant, no trend. This implies that the variables have mean zero. Rarely applicable.
2. Constant in the cointegrating relations, no trend. Applicable when the data have no trend (interest rates, inflation, real exchange rates).
3. Constant in the VECM equation. This leads to a drift, so applicable for trending series (real gdp, stock prices). This is the default, for good reasons.
4. Constant in VECM, trend in cointegration relations. Similar to 3, but hard to interpret.
5. Trend in the VECM equation. Leads to a quadratic trend. Rarely applicable.

Outline

Example continued

Step 1 Apply Johansen integration test.
Choose Option 3 (unrestricted constant) and one lagged Δ on RHS (lag specification `1 1`), because that's what we found in the Engle-Granger approach.

Johansen test

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

H_{00}: "No Cointegration" is rejected. H_{01}: "One cointegrating relationship" is not.
Conclusion: there is one cointegrating (=equilibrium) relationship.

Example continued

Step 2 Estimate VECM. Choose "1 cointegrating relationship", and lag specification `1 1`.

Vector Error Correction Estimates
Date: 11/30/20 Time: 17:32
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]

Example continued

- Final model is

$$\Delta \mathrm{loil}_t = 0.0034 - 0.034(\mathrm{loil}_{t-1} - 0.914\mathrm{lgold}_{t-1} + 2.19) + 0.25\Delta\mathrm{loil}_{t-1} + 0.05\Delta\mathrm{lgold}_{t-1} + e_{1t}$$
$$\Delta \mathrm{lgold}_t = 0.0054 - 0.007(\mathrm{loil}_{t-1} - 0.914\mathrm{lgold}_{t-1} + 2.19) + 0.049\Delta\mathrm{loil}_{t-1} + 0.019\Delta\mathrm{lgold}_{t-1} + e_{2t}$$

- Note how similar the first equation is to what we found with Engle-Granger! Here, too, $\alpha_2 = -0.007$ is insignificant, so we could ignore the 2nd equation.
- The cointegrating vector is $(1, -0.914)$.
- Interpretation: there is an equilibrium relationship between loil and lgold. In case of a disequilibrium, loil adjusts towards the equilibrium. The adjustment amounts to 3.4% per period.

Outline

Learning Goals

Students

- Understand the concept of cointegration,
- are able to test for cointegration using both the Engle-Granger procedure and the Johansen test,
- and are able to estimate an error correction model.

Homework

- Exercise 7
- Problem 6 from Chapter 8 of Brooks (2019)