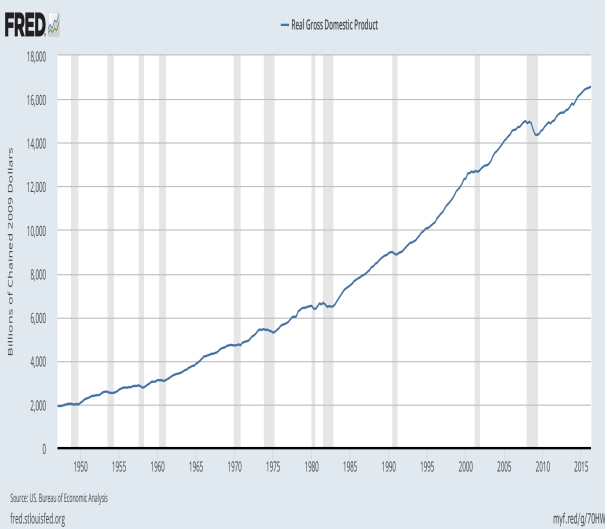
Error-Correction Models
Lecture

Laurent Ferrara

CEF, Ljubljana
Nov. 2019

Plan

PIB US : 1950 - 2015



Introduction: Long-term growth

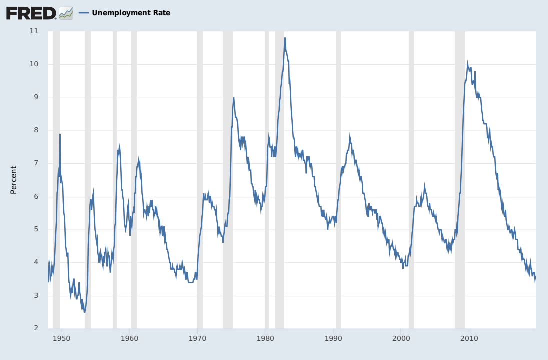
3 factors are responsible for this long-term growth

1. Increase of the population : more people can produce a greater quantity of goods and services
2. Stock of equipment and facilities has increased overtime
3. Techniques of production have led to increases in the productivity

Introduction: Long-term growth

- In spite of this long-term growth, evidence on periods with negative growth rates = economic recessions. This corresponds to business cycles.
- Not related to long-term factors.
- Clearly visible on unemployment rate, with asymmetric behaviour.
- Is this series stationary ?

US unemployment rate

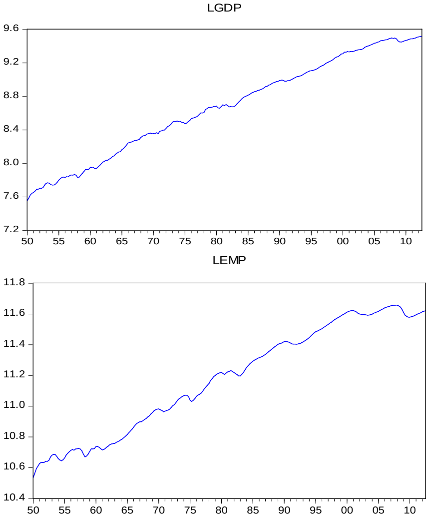


Shaded areas indicate U.S. recessions Source: U.S. Bureau of Labor Statistics fred.stlouisfed.org

Introduction

- How to deal with integrated $I(1)$ time series?
- Option 1: Stationarize by differentiation (eg: $\Delta \log$) or detrending (eg: HP filter, linear trend ...) to get $I(0)$ series
- Option 2: Keep the information and put forward a model that accounts for common trends = Error-Correction Models (ECM hereafter)

Example: US GDP and US Employment (in logs)



Introduction

- From an economic point of view, the presence of equilibrium relations justifies the presence of cointegration
- Examples:
  1. Consumption and income
  2. Money, interest rate, output and prices
  3. Output and employment
  4. Purchasing power parity
- The stationary cointegration relationship of the integrated variables can be considered as a long-run equilibrium
- Any short-run deviation from the long-run equilibrium will dissipate after some periods, depending on the dynamics of the model (more or less persistent)

Dickey-Fuller tests

- Let's consider a given variable (X_t), the ADF test is based on the following regression:

$$\Delta X_t = C + \delta t + \rho X_{t-1} + \sum_{i=1}^{p} a_i \Delta X_{t-i} + u_t,$$

where u_t is a weak WN, p in the AR order for ΔX_t

- Constant C and linear trend δt may or not be included in the regression, leading to various possible tests:
    1. $C = 0$ and $\delta = 0$
    2. $C = 0$ and $\delta \neq 0$
    3. $C \neq 0$ and $\delta = 0$
    4. $C \neq 0$ and $\delta \neq 0$

Dickey-Fuller tests

- Under the null hypothesis $H_0 : \rho = 0$ the series X_t is assumed to be weakly stationary ($I(0)$)
- The null hypothesis $H_0 : \rho = 0$ is tested using the Student statistics for ρ, that is $\hat{\rho}/\sqrt{Var(\hat{\rho})}$.
- Standard critical values are not theoretically available but have been tabulated by Dickey-Fuller and many others

Definition of cointegrated series

- Let's first consider a bi-variate case with 2 variables of interest x_t and y_t
- We say that the two time series x_t and y_t are supposed to be cointegrated if the following conditions are verified:
  1. x_t is $I(1)$ and y_t is $I(1)$
  2. There exist (α, β) such as $\alpha x_t + \beta y_t$ is $I(0)$
- We note (x_t, y_t) is $CI(1)$ and (α, β) the cointegration vector

Definition of cointegrated series

- The previous definition can be generalized to n series $(y_t^1, \ldots, y_t^n) = y_t$ supposed to be integrated of order 1.
- The cointegration vector is then $\beta = (\beta_1, \ldots, \beta_n)$ such $\beta' x_t$ is I(0)
- β_1 is often assumed to be equal to 1

Basic ECM with 2 variables

- A model able to account for cointegration in the bivariate case (with variables in logs) is:

$$\Delta y_t = c + \alpha \Delta x_t + \gamma(y_{t-1} - \beta x_{t-1}) + \varepsilon_t$$

  where
  α : short-term elasticity
  β: long-term elasticity
  γ: speed of adjustment to the long-run equilibrium, $\gamma < 0$

- The ECM has 2 components: short-term (with $I(0)$ variables) and long-term (with lagged $I(1)$ variables)

Basic ECM with n variables

- Let's consider n variables $(y_{1t}, \ldots, y_{nt})$, supposed to be co-integrated
- A model able to account for cointegration in the n-variate case (with variables in logs) is:

$$\Delta y_{1t} = c + \sum_{i=2}^{n} \alpha_i \Delta y_{it} + \gamma(y_{1,t-1} - \sum_{i=2}^{n} \beta_i y_{i,t-1}) + \varepsilon_t$$

where
$\alpha = (\alpha_2, \ldots, \alpha_n)$: short-term elasticities
$\beta = (\beta_2, \ldots, \beta_n)$: long-term elasticities
γ: speed of adjustment to the long-run equilibrium, $\gamma < 0$

Extended ECM with n variables

- Let's consider a given $I(1)$ variable y_t to explain
- We can put forward variables that are likely to explain y_t
  1. in the short-run (m-vector x_t, supposed to be $I(0)$)
  2. in the the long-run (q-vector z_t, supposed to be $I(1)$)
- Thus the model becomes:

$$\Delta y_t = c + \sum_{i=1}^{m} \alpha_i x_{it} + \gamma(y_{t-1} - \sum_{i=1}^{q} \beta_i z_{i,t-1}) + \varepsilon_t$$

- Note that past values of Δy_t can enter into the short-run vector x_t (AR components)

Vector-ECM with n variables

- Let's consider a VAR(p) in level

$$\Phi(B)y_t = \varepsilon_t,$$

as

$$\Delta y_t = -\Pi y_{t-1} + \Gamma_1 \Delta y_{t-1} + \ldots + \Gamma_p \Delta y_{t-p+1} + \varepsilon_t$$

where

$$\Pi = (I - \Phi_1 - \ldots - \Phi_p)$$

and

$$\Gamma_i = -(\Phi_{i+1} - \ldots - \Phi_p)$$

VECM with n variables

- It is well known that y_t is weakly stationary if $|\Phi(z)| = 0$ has all the roots outside the unit circle
- In the presence of unit roots, it is $|\Phi(1)| = 0$ which implies that $\Phi(1)$ has a reduced rank, i.e. smaller than n
- Since $\Phi(1) = I - \Phi_1 - \ldots - \Phi_p$, we have $\Phi(1) = \Pi$. Thus the rank of Π is associated to the presence of unit roots in the vector y_t

VECM with n variables

- If all the n variables are stationary, then Π will have full rank n. That is the VAR(p) in level is a VAR(p) for stationary variables
- If all the n variables are integrated and not cointegrated, we have $\Pi = 0$ and the model is a VAR($p-1$) in first differences such as:

$$\Delta y_t = \Gamma_1 \Delta y_{t-1} + \ldots + \Gamma_p \Delta y_{t-p+1} + \varepsilon_t$$

- If the variables are integrated and cointegrated , Π has rank r with $0 < r < n$, r being the number of cointegration relations, namely the number of independent stationary linear combinations of the n integrated variables.

VECM with n variables

- In that case:

$$\Pi = \alpha \times \beta'$$

α is $(n \times r)$ and β' is $(r \times n)$

- Matrix β contains the coefficients of the r independent stationary combinations of the n integrated variables,

$$\beta' y_{t-1} \sim I(0)$$

- The model becomes a Vector ECM (VECM)

$$\Delta y_t = -\alpha\beta' y_{t-1} + \Gamma_1 \Delta y_{t-1} + \ldots + \Gamma_p \Delta y_{t-p+1} + \varepsilon_t$$

where matrix α contains for each equation the loadings of the r cointegrating relationships $\beta' y_{t-1}$

Engle-Granger cointegration test

- Engle and Granger (1987) suggest the following 2-step procedure to test for the presence of cointegration among $I(1)$ variables $(y_{1t}, \ldots, y_{nt})$
    1. Estimate by OLS the regression:

$$y_{1t} = \beta_2 y_{2t} + \ldots + \beta_n y_{nt} + u_t \tag{1}$$

    2. Get estimated residuals $\hat{u}_t$ and test for unit root. It there is one, then no cointegration. Otherwise, if $\hat{u}_t \sim I(0)$, then variables are cointegrated
- Critical values for unit root tests based on ADF tests are different from those of standard ADF tests. Indeed, $\hat{u}_t$ are OLS residuals, thus have a minimized variance which could biased the test toward rejecting a unit root when using ADF values.

Engle-Granger cointegration test

- Engle and Granger (1987) provide proper critical values using simulations as the asymptotic distribution of the test is non-standard.
- Standard econometric softwares report those tabulated critical values
- If estimated residuals remain $I(1)$, it means that the variables are not cointegrated, we are in the case of a *spurious regression*, i.e. linear regression of $I(1)$ variables that are not cointegrated.
- Another sign of a spurious regression is when the explanatory power (e.g. R^2) is very high.

Engle-Granger cointegration test

- Engle-Granger test is intuitive and easy to implement
- But the distribution of the OLS estimator $\hat{\beta}$ is non-standard, which complicates statistical inference. Many papers have tried to provide solutions for this issue.
- Another caveat is that the Engle-Granger does not allow for to determine the number of cointegrated relationships, r.

Engle-Granger cointegration test

- Once the test has been carried out, the full model can be estimated
  1. Get the estimated stationary residuals from the long-run equation:
  $$\hat{u}_t = y_{1t} - \hat{\beta}_2 y_{2t} - \dots \hat{\beta}_n y_{nt}$$
  2. Run the short-run equation:
  $$\Delta y_{1t} = c + \sum_{i=2}^{n} \alpha_i \Delta y_{it} + \gamma \hat{u}_{t-1} + \varepsilon_t$$

Engle-Granger in the bivariate case

- 2-step estimation Ã la Engle-Granger:
  1/ Long-run estimate by OLS : $\hat{\beta}$
  $\hat{u}_t = y_t - \hat{\beta} x_t$
  Check if $\hat{u}_t$ is $I(0)$.
  2/ Short-run estimate by OLS:

$$\Delta \hat{y}_t = \hat{c} + \hat{\alpha} \Delta x_t + \hat{\gamma} \hat{u}_{t-1}$$

- An extended model by acocunting for short-run dynamics

$$\Delta y_t = c + \sum_{i=1}^{p} \phi_i \Delta y_{t-i} + \sum_{i=0}^{q} \alpha_i \Delta x_{t-i} + \gamma (y_{t-1} - \beta x_{t-1}) + \varepsilon_t$$

Johansen cointegration test

- Johansen (1995) developed a Maximum Likelihood approach to test for the cointegration rank r, obtain ML estimators of α and β and test specific hypotheses on parameters
- The procedure sequentially test the folllwing hypotheses:
  1. H0: $r = 0$ vs H1: $r = 1$
  2. H0: $r \leq 1$ vs H1: $r = 2$
  3. H0: $r \leq 2$ vs H1: $r = 3$

     ...
- If at step 1, the test does not reject H0, we set $r = 0$.
  If at step i, the test rejects H0, we set $r = i$

Johansen cointegration test

- Johansen (1995) suggested 2 statistics to be used in each step of the sequential procedure: the trace test and the maximum eigenvalue test. In practice, it is common to compute both statistics and compare their outcomes.
- Both have asymptotic distributions that depend on the deterministic components of the model. In practice, it is common to compute for various model specifications.

Johansen cointegration test

- Once r is determined by the test, we need to identify the coefficients of the cointegration vectors.
- But for any matrix Q of dimension $r \times r$, we get:

$$\Pi = \alpha\beta' = \alpha Q Q^{-1}\beta' = \gamma\delta'$$

- This means that the cointegration coefficients and the loadings are not uniquely identified. Some restrictions have to be imposed.
- Once r has been determined and α and β identified, estimation and inference in the VECM is standard

Forecasting with VECM with n variables

- The VECM forecast for $h = 1$ is:

$$\Delta\hat{y}_t(1) = -\hat{\alpha}\hat{\beta}' y_t + \hat{\Gamma}_1 \Delta y_t + \ldots + \hat{\Gamma}_{p-1} \Delta y_{t-p}$$

- Forecasts for the level of y_{t+1} is thus given by:

$$\hat{y}_t(1) = y_t + \Delta\hat{y}_t(1)$$

- For $h > 0$, we iterate the forecasts such that

$$\Delta\hat{y}_t(h) = -\hat{\alpha}\hat{\beta}'\hat{y}_t(h-1) + \hat{\Gamma}_1 \Delta\hat{y}_t(h-1) + \ldots + \hat{\Gamma}_{p-1} \Delta\hat{y}_t(h-p)$$

  where forecasts on the rhs are replaced by true values when available

Forecasting with VECM for $h = 1$

- As for the specific bivariate case with (y_t, x_t), the 1-step-ahead predictor for the level of y_{t+1} is

$$\hat{y}_t(1) = y_t + \hat{\alpha}(\hat{x}_t(1) - x_t) + \hat{\gamma}(y_t - \hat{\beta} x_t)$$

- Note that in that case, we need to forecast the 1-step-ahead value of x_t

Alternative forecasts with n variables

- The VAR model in differences is a first alternative model
- In the presence of cointegration, the VAR in differences is misspecified due to the omission of the cointegration relationships $\beta' y_{t-1}$.
- Hence forecasts are suboptimal
- However, in the case of unaccounted changes in the cointegrating vectors β or in their loadings α, VAR forecasts could be more robust.
- Indeed, the VECM constraints the long-run forecasts to satisfy to the relationship $\beta' \hat{y} t(h) \sim I(0)$, i.e. forecasts go back to their long-run equilibrium.

Alternative forecasts with n variables

- The VAR model in levels is a second alternative model
- The model is correctly specified but parametes are not efficiently estimated as the cointegration restriction is not imposed
- But if the sample is long enough, the OLS estimates remain consistent and will reflect at some mpoint the cointegrating restrictions.
- Moreover, this model does not require the specification of the cointegration rank r, which maybe an advantage when there is uncertainty.
- Empirical comparisons between forecast stemming from VAR in differences, VAR in levels and VECM make sense

References

- Hamilton, J.D. (1994), *Time Series Analysis*, Princeton University Press.
- Ghysels, E. and M. Marcellino (2018), *Applied Economic Forecasting using Time Series Methods*, Oxford University Press
- Johansen, S. (1995), *Likelihood-based Inference in Cointegrated Vector Autoregressive Models*, Oxford University Press