UNIVERSITÀ DEGLI STUDI DI MILANO
Dipartimento di Economia, Management e Metodi Quantitativi

Academic Year 2019-2020

Time Series Econometics

Fabrizio Iacone

Chapter 11: Regressions with I(0) and I(1) processes

Topics: Regression with stationary and invertible ARMA processes, Spurious regression for I(1) processes, Cointegration for I(1) processes, Testing for cointegration with the ADF test

Two notes on the notation

♦In

$$Y_t = Y_{t-1} + \varepsilon_t, \;\; \varepsilon_t \; i.i.d.(0,\sigma^2), \text{if } t > 0$$

$$Y_t = 0 \text{ if } t \leq 0,$$

we saw that the initialisation $Y_0 = 0$ is used to have a finite variance (and a finite covariance structure) and also in some limits, like $\sqrt{T}\frac{1}{T^2}\sum_{t=1}^{T} Y_{t-1} \rightarrow_d \sigma\int_0^1 W(r)dr$. However, this is often neglected in many references, and we too will omit the repetition in the next lecture, and write

$$Y_t = Y_{t-1} + \varepsilon_t, \;\; \varepsilon_t \; i.i.d.(0,\sigma^2)$$

instead.

♦ We will discuss the linear model

$$Y_t = X_t\beta + e_t$$

and the estimate $\widehat{\beta}$.

Notice that Hamilton discussed

$$Y_t = X_t\gamma + e_t$$

and then $\widehat{\gamma}$. We used a different symbol because γ and $\widehat{\gamma}$ refers to the autocovariances.

Regressions with time series

$$Y_t = X_t\beta + e_t$$

$$\widehat{\beta} = \left(\sum_{t=1}^{T} X_t^2\right)^{-1} \sum_{t=1}^{T} X_t Y_t.$$

X_t stationary and invertible ARMA(p,q) with i.i.d. innovations, with $E(X_t X_{t+j}) = \Gamma_j^X$ ($\left|\Gamma_j^X\right| < \infty$);
e_t stationary and invertible ARMA(p,q) with i.i.d. innovations, with $E(e_t) = 0$, $E(e_t e_{t+j}) = \gamma_j^e$ ($\left|\gamma_j^e\right| < \infty$);

X_t independent from e_s at all t, s.

Then,

$$\widehat{\beta} \to_p \beta$$

$$\sqrt{T}\left(\widehat{\beta} - \beta\right) \to_d N(0, V)$$

where

$$V = (\Gamma_0^X)^{-2} \sum_{j=-\infty}^{\infty} \Gamma_j^X \gamma_j^e$$

★ i.e. we can generalise the results of the standard regression model to ARMA dependent structures in X_t and e_t

What does it happen if X_t &/or e_t are $I(1)$?

The Granger Newbold experiment

$$Y_{1,t} = Y_{1,t-1} + \varepsilon_{1,t}, \quad \varepsilon_{1,t} \; i.i.d.(0, \sigma_1^2)$$

$$Y_{2,t} = Y_{2,t-1} + \varepsilon_{2,t}, \quad \varepsilon_{2,t} \; i.i.d.(0, \sigma_2^2)$$

$\varepsilon_{1,t}$ independent from $\varepsilon_{2,s}$ for all t, s

Regress $Y_{2,t}$ on $Y_{1,t}$:

$$\widehat{\beta} = \frac{\sum_{t=1}^{T} Y_{1,t} Y_{2,t}}{\sum_{t=1}^{T} Y_{2,t}^2}$$

Since $\varepsilon_{1,t}$ is independent from $\varepsilon_{2,s}$ for all t, s, then $Y_{1,t}$ independent from $Y_{2,s}$ for all t, s, so the parameter that $\widehat{\beta}$ is estimating is 0. However, $\widehat{\beta}$ does not converge to 0 as $T \to \infty$.

On the contrary, let $W_1(.)$, $W_2(.)$ two Brownian motions, such that

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{[rT]}\varepsilon_{1,t} \rightarrow_d \sigma_1 W_1(r)$$

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{[rT]}\varepsilon_{2,t} \rightarrow_d \sigma_2 W_2(r)$$

Then, $W_1(.)$ and $W_2(.)$ are independent, i.e. $E(W_1(r)W_2(s)) = 0$ for any $r, s\,(r \in [0,1], s \in [0,1])$;

Rewriting

$$\widehat{\beta} = \frac{\frac{1}{T^2}\sum_{t=1}^{T} Y_{1,t}Y_{2,t}}{\frac{1}{T^2}\sum_{t=1}^{T} Y_{2,t}^2}$$

$$= \frac{\frac{1}{T}\sum_{t=1}^{T}\frac{1}{\sqrt{T}}Y_{1,t}\frac{1}{\sqrt{T}}Y_{2,t}}{\frac{1}{T}\sum_{t=1}^{T}\left(\frac{1}{\sqrt{T}}Y_{2,t}\right)^2},$$

$$\widehat{\beta} \rightarrow_d \frac{\sigma_1}{\sigma_2}\frac{\int_0^1 W_1(r)W_2(r)dr}{\int_0^1 W_2(r)^2 dr}$$

★ $\widehat{\beta}$ does not converge to 0 as $T \to \infty$

★ if we test $H_0 : \{\beta = 0\}$ vs. $H_A : \{\beta \neq 0\}$, we reject H_0 as $T \to \infty$ ("spurious evidence of a significant regression parameter")

★ including a constant in the regression changes the limit distribution of $\widehat{\beta}$ but not the essence of the results

★ These results can be generalised even further, to

$$Y_t = X_t\beta + e_t,$$

where X_t and e_t are generic $I(1)$ processes with no deterministic trend: even when $\beta \neq 0$,

(i) the OLS estimate $\widehat{\beta}$ is still inconsistent, and

(ii) the correct null hypothesis is rejected as $T \to \infty$ (in this sense, it still is a "spurious regression").

★ Modelling strategy: model ΔY_t and ΔX_t

Cointegration

$$Y_{1,t} = \beta Y_{2,t} + \varepsilon_{1,t}, \quad \varepsilon_{1,t} \; i.i.d.(0, \sigma_1^2)$$

$$Y_{2,t} = Y_{2,t-1} + \varepsilon_{2,t}, \quad \varepsilon_{2,t} \; i.i.d.(0, \sigma_2^2)$$

$$\varepsilon_{1,t} \text{ independent from } \varepsilon_{2,s} \text{ for all } t, s$$

Letting again $W_1(.)$, $W_2(.)$ such that

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{[rT]} \varepsilon_{1,t} \to_d \sigma_1 W_1(r), \; \frac{1}{\sqrt{T}} \sum_{t=1}^{[rT]} \varepsilon_{2,t} \to_d \sigma_2 W_2(r)$$

($W_1(.)$ and $W_2(.)$ are independent) then

$$T\left(\widehat{\beta} - \beta\right) \to_d \frac{\sigma_1}{\sigma_2} \frac{\int_0^1 W_2(r) dW_1(r)}{\int_0^1 W_2(r)^2 dr}$$

★ $\widehat{\beta}$ is "superconsistent" (see rate T rather than $\sqrt{T}$)

★ the limit distribution of $T\left(\widehat{\beta} - \beta\right)$ is known; it is possible to test $H_0 : \{\beta = \beta_0\}$ vs. $H_A : \{\beta \neq \beta_0\}$ (for any β_0 except $\beta_0 = 0$).

★ including a constant in the regression changes the limit distribution of $\widehat{\beta}$ but not the essence of the results.

★ Modelling strategy: Error Correction Mechanism:

The first equation can be rewritten as

$$Y_{1,t} - Y_{1,t-1} = \beta Y_{2,t} - \beta Y_{2,t-1} + \beta Y_{2,t-1} - Y_{1,t-1} + \varepsilon_{1,t}$$
$$\Delta Y_{1,t} = \beta \Delta Y_{2,t} - (Y_{1,t-1} - \beta Y_{2,t-1}) + \varepsilon_{1,t}$$

Here, $\beta \Delta Y_{2,t}$ gives the effect of changes of $Y_{2,t}$ on $Y_{1,t}$ ("short term dynamics"), and $(Y_{1,t-1} - \beta Y_{2,t-1})$ gives the "adjustment to the long run equilibrium".

These results can be generalised even further, to

$$Y_t = X_t\beta + e_t,$$

where X_t is a generic $I(1)$ process with no deterministic trends and e_t is a generic $I(0)$ process with $E(e_t) = 0$: then, there is a non-degenerate limit distribution ϖ such that

$$T\left(\widehat{\beta} - \beta\right) \to_d \varpi$$

★ the estimate $\widehat{\beta}$ can be rearranged, so it is possible to test $H_0 : \{\beta = \beta_0\}$ vs. $H_A : \{\beta \neq \beta_0\}$ (for any β_0 except $\beta_0 = 0$)

★ $\widehat{\beta}$ is "superconsistent"

★X_t does not need to be independent from e_s

★ the example can also be generalised to multidimensional X_t, and to X_t with linear deterministic trends (however, some additional conditions may have to be specified, for these cases)

★ The ECM may be generalised to allow for stationary AR processes in ΔX_t and/or e_t

Testing for cointegration

Consider the generic model

$$Y_t = \alpha + X_t^{'}\beta + e_t$$

where X_t and Y_t are $I(1)$ (possibly with deterministic trends) and e_t may either be $I(1)$ or $I(0)$.

In order to know wheter to go for the ECM modelling or for the differencing, and in order to know if the estimates of α and β are reliable, we must find out if e_t is $I(1)$.

If we don't know e_t:

✠ estimate $\widehat{\alpha}$, $\widehat{\beta}$ by regressing Y_t on a constant and on X_t, compute the residuals

$$\widehat{e}_t = Y_t - \widehat{\alpha} - X_t^{'}\widehat{\beta}$$

and test applying the ADF test statistic (Case 1) to $\widehat{e}_t$.

However,

★the limit distribution of the t statistic of the ADF test statistic for $\widehat{e}_t$ depends on: the number of regressors included; the type of deterministic components included.

★The limit distribution is even more skewed to the left, and it gets more and more so the more regressors and the more deterministic terms are considered (for example, when X_t is a scalar and it has no linear deterministic trend, α is included in the regression, the 5% critical value is -3.37 as opposed to -1.95, the one we would use if e_t was observable).

If we think we know β, then, it would then be advisable to test $Y_t - X_t^{'}\beta$ instead;
If we think we know α and β, then, it would then be advisable to test $Y_t - \alpha - X_t^{'}\beta$ instead.