Econometrics

Week 3

Institute of Economic Studies
Faculty of Social Sciences
Charles University in Prague

Fall 2022

Recommended Reading

For today

- Further issues in using OLS with time series data.
- Chapter 11.

For next week

- Serial correlation and heteroskedasticity in time series regressions.
- Chapter 12.

Today's Lecture

- We have learned that OLS has the same desirable finite sample properties in the time-series case as in the cross-sectional case...
- ...if a somewhat altered set of assumptions is satisfied.
- Today, we will study large sample properties of OLS, which are much more problematic in time-series than in cross-sectional case.
- We will introduce key concepts needed to apply the usual large sample approximations in regression analysis with time series.
- We will learn how to test whether asymptotic properties are satisfied and what to do if they are not

Question

Consider the following regression model built to study the relationship between short-term interest rates and GDP growth.

$$gGDP_t = \beta_0 + \beta_1 int_t + u_t$$

Is the OLS estimate of β_1 unbiased in this example?

- Interest rates are adjusted based on the state of the economy $\rightarrow$ strict exogeneity is not satisfied
- Under (only) contemporaneous exogeneity, OLS estimator is biased, but it might still be consistent!

Large Sample Properties of OLS

What do we mean by large sample properties (asymptotic properties)?

- Consistency
- Asymptotic variance
- Asymptotic normality

When deriving large sample properties we rely on:

- Law of Large Numbers (LLN)
- Central Limit Theorem (CLT)

Large Sample Properties of OLS

- Both the LLN and CLT in the form you know them require random sampling.
- But time series data cannot be treated as random samples!
- For the LLN and CLT to hold random sampling is not necessary, but other assumptions have to be made.
- In particular, the data generating process has to be **stationary** and **weakly dependent**.
- We will introduce these two concepts during today's lecture.

Stationary and Nonstationary Time Series

- The concept of stationarity is important for time series analysis.
- Probability distribution function of a stationary process is **stable over time**.
- If we shift the sequence of the collection of random variables, joint probability distribution must remain unchanged.

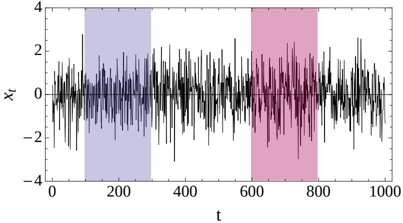
Stationarity

The stochastic process $\{x_t : t = 1, 2, \ldots\}$ is *stationary* if for every collection of time indices $1 \leq t_1 < t_2 < \ldots < t_m$, the joint probability distribution of $(x_{t_1}, x_{t_2}, \ldots, x_{t_m})$ is the same as the joint probability distribution of $(x_{t_1+h}, x_{t_2+h}, \ldots, x_{t_m+h})$ for all integers $h \geq 1$.

- Definition may seem little abstract, but meaning is pretty straightforward, see example on the next slide.

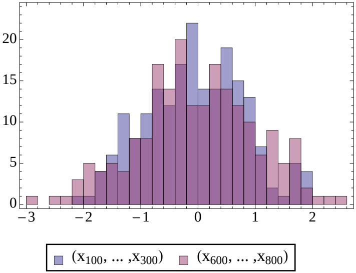
STATIONARY

Time Series 1



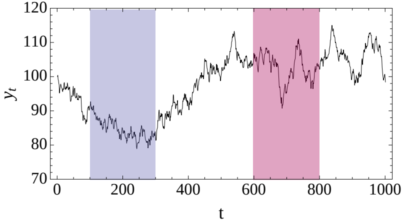
$\Downarrow$

Histograms



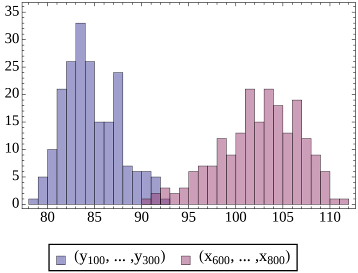
NONSTATIONARY

Time Series 2



$\Downarrow$

Histograms



Covariance Stationary Process

Covariance Stationary Process

A stochastic process $\{x_t : t = 1, 2, \ldots\}$ with finite second moment $[E(x_t^2) < \infty]$ is *covariance stationary* if:

- $E(x_t)$ is constant.
- $Var(x_t)$ is constant.
- for any $t, h \geq 1$, $Cov(x_t, x_{t+h})$ depends only on h and not on t.

- Covariance stationarity focuses only on the first two moments of a stochastic process and the covariance between x_t and x_{t+h}.
- If a stationary process has a finite second moment, **it must be covariance stationary**, BUT NOT VICE VERSA.
- $\Rightarrow$ strict stationarity is stronger than covariance stationarity (for processes with finite second moments).

(Covariance) Stationary Process

Are these time series stationary?

- Restaurants price index See data plot
  - Not stationary.
- Time series with a trend
  - Not stationary - its mean changes in time.
  - May be trend-stationary.
- A moving average process?

$$x_t = \epsilon_t + \alpha_1 \epsilon_{t-1}, \qquad t = 1, 2, \dots,$$

where $\{\epsilon_t : t = 0, 1, \dots\}$ is i.i.d.

  - Stationary - because $\{\epsilon_t\}$ is i.i.d.

Weakly Dependent Time Series

- A very different concept from stationarity.

Weakly Dependent Process

A stationary time series process $\{x_t : t = 1, 2, \ldots\}$ is said to be *weakly dependent* if x_t and x_{t+h} are "almost independent" as h increases without bound.

- Very vague definition, as there are many cases of weak dependence $\Rightarrow$ Not covered by this course.
- If for a covariance stationary process, $Corr(x_t, x_{t+h}) \to 0$ as $h \to \infty$, we will say this process is weakly dependent.
- In other words, as the variables get farther apart in time, the correlation between them becomes smaller and smaller.
- Technically, we assume that the correlation converges to zero fast enough.

Weakly Dependent Time Series

- Asymptotically uncorrelated sequences are characterized by weak dependence.
- Why we need it? It replaces assumption of random sampling in implying that the Law of Large Numbers (LLN) and Central Limit Theorem (CLT) holds.

The simplest example of Weakly Dependent Time Series

- Independent, identically distributed sequence.
- A sequence that is independent is trivially weakly dependent.

Moving Average – MA(1) – Process

- More interesting example is the Moving Average Process of order one – MA(1).

  $$x_t = \epsilon_t + \alpha_1 \epsilon_{t-1}, \qquad t = 1, 2, \ldots,$$

  where $\{\epsilon_t : t = 0, 1, \ldots\}$ is an i.i.d. sequence with zero mean and variance σ_ϵ^2.
- Is MA(1) stationary? YES!
  $\Leftarrow$ because $\{\epsilon_t\}$ is i.i.d.
- Is MA(1) weakly dependent? YES!
  $\Leftarrow$ its consecutive terms x_{t+1} and x_t are correlated, with $Corr(x_t, x_{t+1}) = \alpha_1/(1 + \alpha_1^2)$, but x_{t+2} and x_t are independent.

Autoregressive – AR(1) – Process

- The process $\{x_t\}$ is called Autoregressive process of order one – AR(1), when it follows:

  $$x_t = \rho x_{t-1} + \epsilon_t, \qquad t = 1, 2, \ldots$$

  The starting point in the sequence is x_0 at time $t = 0$, and $\{\epsilon_t : t = 1, 2, \ldots\}$ is an i.i.d. sequence with zero mean and variance σ_ϵ^2.
- This process is weakly dependent and stationary in case when $|\rho| < 1$ (in other words, we assume the so-called *stability condition*).

Autoregressive – AR(1) – Process

- To show that AR(1) is weakly dependent, we will assume that it is covariance stationary.
- Then, $E(x_t) = E(x_{t-1}) \Rightarrow E(x_t) = 0$.
- $Var(x_t) = \rho^2 Var(x_{t-1}) + Var(\epsilon_t)$, and under covariance stationarity, we must have:

$$\sigma_x^2 = \rho^2 \sigma_x^2 + \sigma_\epsilon^2$$

- For $\rho^2 < 1$, we can easily solve for σ_x^2:

$$\sigma_x^2 = \frac{\sigma_\epsilon^2}{1 - \rho^2}$$

Autoregressive – AR(1) – Process

- The covariance between x_t and x_{t+h} for $h \geq 1$ can be found by repeated substitution:
- $x_{t+h} = \rho x_{t+h-1} + \epsilon_{t+h} = \rho(\rho x_{t+h-2} + \epsilon_{t+h-1}) + \epsilon_{t+h}$
  $\ldots$
  $x_{t+h} = \rho^h x_t + \rho^{h-1}\epsilon_{t+1} + \ldots + \rho\epsilon_{t+h-1} + \epsilon_{t+h}$
- As $E(x_t) = 0$, we can multiply it by x_t and take expectations to obtain $Cov(x_t, x_{t+h})$.
- Using the fact that ϵ_{t+j} is uncorrelated with x_t for all integers $j \geq 1$ we get:
  $Cov(x_t, x_{t+h}) = E(x_t x_{t+h}) = \rho^h E(x_t^2) + \rho^{h-1} E(x_t \epsilon_{t+1}) + \ldots + E(x_t \epsilon_{t+h}) = \rho^h E(x_t^2) = \rho^h \sigma_x^2$.
- Finally, $Corr(x_t, x_{t+h}) = \frac{Cov(x_t, x_{t+h})}{\sigma_x \sigma_x} = \rho^h$
- So ρ is correlation between any 2 adjacent terms, and for large h, correlation gets very low ($\rho^h \to 0$ as $h \to \infty$).

Trends Revisited

- A trending series is not stationary, but
- A trending series can be weakly dependent.

Trend-stationarity

A series that is stationary about its time trend, as well as weakly dependent, is called **trend-stationary process**.

- As long as a trend is included in the regression, everything is OK.

Asymptotic Properties of OLS

- Proofs are quite technical.
- If interested, consult Wooldrige (1994) Estimation and inference for dependent processes, in *Handbook of Econometrics*, 4, 2639-2738.

Asymptotic Properties of OLS

The following assumptions will alter the assumptions for unbiasedness introduced last week (TS1, TS2 and TS3).

TS1a: Linearity and Weak Dependence

In a linear model:

$$y_t = \beta_0 + \beta_1 x_{t1} + \ldots + \beta_k x_{tk} + u_t,$$

we assume that $\{(\mathbf{x}_t, y_t) : t = 1, 2, \ldots\}$ is stationary and weakly dependent.

Thus the Law of Large Numbers and Central Limit Theorem can be applied to sample averages.

- Note that $\{(\mathbf{x}_t, y_t) : t = 1, 2, \ldots\}$ is stationary and weakly dependent
- All involved variables, explanatory and dependent, need to be stationary and weakly dependent!

Asymptotic Properties of OLS

TS2a: No Perfect Collinearity

No independent variable is constant or a perfect linear combination of the others. Same as TS2.

TS3a: Zero Conditional Mean

For each t, $E(u_t|\mathbf{x}_t) = 0$.

TS3a: is much weaker than TS3 $\Leftarrow$ no restriction on the relationship between u_t and x_s when $t \neq s$.

Asymptotic Properties of OLS

Theorem 1: Consistency of OLS

Under TS1a, TS2a and TS3a, the OLS estimators are consistent: plim $\hat{\beta}_j = \beta_j$, $j = 0, 1, \ldots, k$.

The theorem is similar to the one we used in cross-sectional data, but we have omitted random sampling assumption (thanks to TS3a.).

Large-Sample Inference

Assumptions TS4a and TS5a are weaker than their CLM counterparts:

TS4a: Homoskedasticity

For all t, $Var(u_t|\mathbf{x}_t) = \sigma^2$.

TS5a: No Serial Correlation

For all $t \neq s$, $E(u_t u_s|\mathbf{x}_t, \mathbf{x}_s) = 0$.

Theorem 2: Asymptotic Normality of OLS

Under the TS1a–TS5a for time series, the OLS estimators are asymptotically normally distributed, and the usual OLS standard errors, t statistics, F statistics, and LM statistics are asymptotically valid.

Large-Sample Inference

Why is the Theorem 2 so important?

Even if strict exogeneity is not satisfied (one of the CLM assumptions), OLS may still be consistent, and the usual inference may still be valid!
Of course, only if TS1a–TS5a are met !!!

- In other words: Provided the time series are stationary and weakly dependent, usual OLS inference is valid under assumptions weaker than CLM assumptions.
- Under TS1a–TS5a, we can show that the OLS estimators are asymptotically efficient (as in the cross-sectional case).
- Finally, when a regression model with trend-stationary explanatory variable satisfies TS1a – TS5a, and we add time trend to the regression, usual inference is valid.

Highly Persistent Time Series in Regression Analysis

- Highly persistent = strongly dependent.
- Unfortunately, many economic time series have strong dependence. Weak dependence assumption violated.
- This is not a problem, if the CLM assumptions (TS1 – TS6) from the previous lecture hold.
- But, usual inference is very sensitive to violation of these assumptions.

Random Walk

- A simple example of highly persistent time series often found in economics is AR(1) with $\rho = 1$:

$$x_t = x_{t-1} + \epsilon_t, \qquad t = 1, 2, \ldots,$$

  where $\{\epsilon_t : t = 1, 2, \ldots\}$, is an i.i.d. with zero mean and variance σ_ϵ^2.
- We assume that the initial value, x_0, is independent of e_t for all $t \geq 1$.
- Expected value of a random walk does not depend on t:

$$x_t = \epsilon_t + \epsilon_{t-1} + \ldots + \epsilon_1 + x_0$$

$$E(x_t) = E(\epsilon_t) + E(\epsilon_{t-1}) + \ldots + E(\epsilon_1) + E(x_0) = x_0$$

- If $x_0 = 0$, then $E(x_t) = 0$ for all t.
- $Var(x_t) =$ $Var(\epsilon_t) + Var(\epsilon_{t-1}) + \ldots + Var(\epsilon_1) + Var(x_0) = \sigma_\epsilon^2 t$. Process is not stationary, as its variance increases in time.

Random Walk

- Random walk is not stationay, as its variance increases in time.
- Random walk is highly persistent, as value of x_t today influences all the future values (all x_{t+h})
  $E(x_{t+h}|x_t) = x_t$ for all $h \geq 1$.
- Do you remember the AR(1) with $|\rho| < 1$?
  $E(x_{t+h}|x_t) = \rho^h x_t$ for all $h \geq 1 \Rightarrow$
  with $h \to \infty$, $E(x_{t+h}|x_t) = 0$.
- Random walk is a special case of a **unit root process**.

Random Walk

- It is important to distinguish between highly persistent time series and trending time series.
- Interest rates, inflation, unemployment rate, etc. can be highly persistent, but they do not have a trend.
- Finally, we can have highly persistent time series with trend: **random walk with drift**:

$$x_t = \alpha + x_{t-1} + \epsilon_t, \qquad t = 1, 2, \ldots,$$

  where α is the **drift term** and $\{\epsilon_t : t = 1, 2, \ldots\}$ and x_0 satisfy the same properties as in the random walk model.
- $E(x_{t+h}|x_t) = \alpha h + x_t$.

Transforming Highly Persistent Time Series

- Highly persistent time series can lead to very misleading results if CLM are violated.
  (Do you remember spurious regression problem?)
- We need to transform them into weekly dependent processes.
- Weakly dependent process will be referred to as **integrated of order zero, I(0)**.
- Random Walk is integrated of order one, I(1).
- It's first difference is integrated of order zero, I(0):

$$\Delta x_t = x_t - x_{t-1} = \epsilon_t, \qquad t = 2, 3, \ldots$$

- Therefore $\{\Delta x_t : t = 2, 3, \ldots\}$ is an i.i.d. sequence and we can use it for analysis.

Further issues

- It is important to distinguish between unit root processes and stable processes.
  - In random walk process its history has no relevance for the future.
  - If stock returns follow random walk, examining recent movements in stock history is worthless.
  - If GDP follows random walk, the effects of shocks will be permanent.
- If a stochastic trend process (random walk with drift) is misclassified as a deterministic trend, we run into risk of spurious regression.
- How to find out whether a time series is I(0) or I(1)?
  - Detrend first!
  - Informal test: examine correlation between x_t and x_{t-1}.
  - Formal tests: Dickey-Fuller and augmented Dickey-Fuller test (not covered by this course - Wooldrige chapter 18).

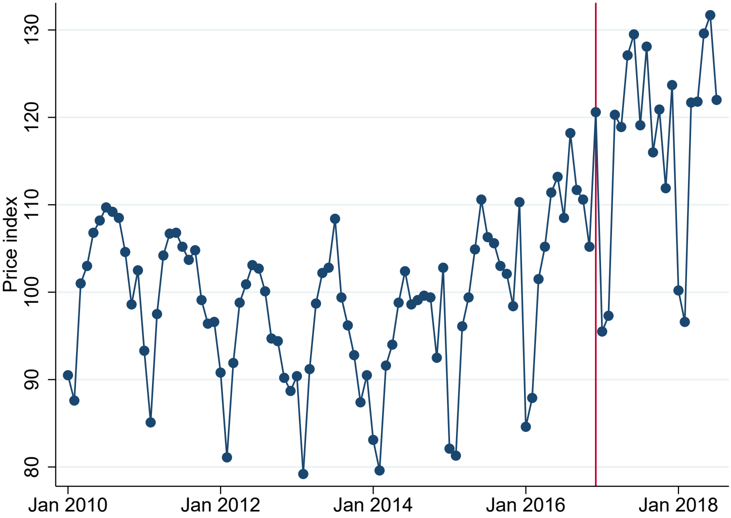
Thank you

Thank you for your attention!

... and do not forget to read Chapter 12 for the next week!

Please, revise also: Variance of OLS estimator and its derivation

Restaurants price index



◂ Back