AFRE 835: Introductory Econometrics

Chapter 11: Further Issues in Using OLS with Time Series Data

Spring 2017

Introduction

- In chapter 10, we looked at basic properties of time series models and how they differ from the pure cross-sectional setting.
- We also examined what assumptions were needed to restore the Gauss-Markov Theorem in that setting.
- Of particular importance was the need to assert the **strict exogeneity** assumption (imbedded in TS.3)
- This chapter explores an alternative set of assumptions that lead to consistency of OLS, but not unbiasedness.

Outline

1. Stationarity and Weakly Dependent Time Series
2. Asymptotic Properties of OLS
3. Highly Persistent Time Series
4. Dynamically Complete Models/Absence of Serial Correlation

Stationary and Nonstationary Time Series

- In order to characterize the asymptotic properties of OLS estimators in a time series setting (properties such as consistency and asymptotic normality) we need to use laws of large numbers and the central limit theorem.
- These, in turn, will require that the time series itself satisfy certain characteristics.
- One important characteristic is (strict) **stationarity**, which basically says that the distributional properties of the time series do not change over time.
- Formally: A stochastic process $\{x_t : t = 1, 2, \ldots\}$ is *stationary* if for every collection of time indices $1 \leq t_1 < t_2 < \cdots < t_m$, the joint distribution of $(x_{t_1}, x_{t_2}, \ldots, x_{t_m})$ is the same as the joint distribution of $(x_{t_1+h}, x_{t_2+h}, \ldots, x_{t_m+h})$.
- A stochastic process that is not stationary is said to be *nonstationary*.

Stationarity (cont'd)

- A weaker form of stationarity is Covariance Stationarity:

  ...A stochastic process $\{x_t : t = 1, 2, \ldots\}$ with finite second moments $[E(x_t^2) < \infty]$ is **covariance stationary** if (i) $E(x_t)$ is constant, (ii) $Var(x_t)$ is constant, and (iii) for any $t, h \geq 1$, $Cov(x_t, x_{t+h})$ depends only on h and not on t.
- The following stochastic sequence is neither strictly nor covariance stationary:

$$y_t = \delta_0 + \delta_1 t + e_t \tag{1}$$

  where $\delta_1 \neq 0$ and $\{e_t : t = 1, 2, \ldots\}$ is an *iid* sequence with mean of zero and variance σ_e^2.

Weak Dependence

- Another useful concept is one of weak dependence, which can defined loosely in the case of a stationary process as follows:

  ...A stationary stochastic 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.
- A related assumption for a covariance stationary process is:

  ...A covariance stationary sequence is said to be **asymptotically uncorrelated** if $Corr(x_t, x_{t+h}) \to 0$ as $h \to \infty$.
- These assumptions are important in that they place enough assumptions on the stochastic process so that the law of large numbers (LLN) and the central limit theorem (CLT) apply.

Example #1 of a Weakly Dependent Series: MA(1)

- Moving Average process of order one (MA1) is one example of a weakly dependent stochastic process, where

$$x_t = e_t + \alpha_1 e_{t-1}, \quad t = 1, 2, \ldots \tag{2}$$

with $\{e_t : t = 0, 2, \ldots\}$ being an *iid* sequence with mean of zero and variance σ_e^2.
- Note that this sequence is stationary and weakly dependent.
- In particular, for $h \geq 1$:

$$\begin{aligned} Cov(x_t, x_{t-h}) &= E(x_t x_{t-h}) \\ &= E([e_t + \alpha_1 e_{t-1}][e_{t-h} + \alpha_1 e_{t-h-1}]) \\ &= E(e_t e_{t-h} + \alpha_1 e_t e_{t-h-1} + \alpha_1 e_{t-1} e_{t-h} + \alpha_1^2 e_{t-1} e_{t-h-1}) \end{aligned}$$

$$= \begin{cases} \alpha_1 \sigma_e^2 & h = 1 \\ 0 & h > 1 \end{cases} \tag{3}$$

- MAq ($q < \infty$) would also be weakly dependent and stationary.

Example #2 of a Weakly Dependent Series: Stable AR(1)

- A second example is a special case of the **autoregressive process of order 1 (AR1)**, which (given a sequence starting point y_0 at time $t = 0$) evolves recursively according to

$$y_t = \rho_1 y_{t-1} + e_t, \quad t = 1, 2, \ldots \tag{4}$$

with $\{e_t : t = 0, 2, \ldots\}$ being an *iid* sequence with mean of zero and variance σ_e^2.
- An AR1 process is a **stable AR1 process** if $|\rho_1| < 1$

  ... and a stable AR1 process is weakly stationary.

Example #2 (cont'd)

- Weak stationarity argument: note by repeated substitution that

$$\begin{aligned} y_t &= \rho_1 y_{t-1} + e_t \\ &= \rho_1(\rho_1 y_{t-2} + e_{t-1}) + e_t \\ &= \rho_1^2 y_{t-2} + e_t + \rho_1 e_{t-1} \\ &= \rho_1^2(\rho_1 y_{t-3} + e_{t-2}) + e_t + \rho_1 e_{t-1} \\ &= \rho_1^3 y_{t-3} + e_t + \rho_1 e_{t-1} + \rho_1^2 e_{t-2} = \cdots = \rho_1^h y_{t-h} + \sum_{s=0}^{h-1} \rho_1^s e_{t-s} \end{aligned} \tag{5}$$

- Assuming covariance stationarity (which one can show), then

$$Cov(y_t y_{t-h}) = E(y_t y_{t-h}) = \rho_1^h E(y_{t-h}^2) + \sum_{s=0}^{h-1} \rho_1^s E(y_{t-h} e_{t-s}) = \rho_1^h \sigma_y^2$$

which shrinks to zero as $h \to \infty$, as long as $|\rho_1| < 1$.

Weak Dependence and Stationarity

- These two concepts are not equivalent.
- In particular, one can have a non-stationary process that is weakly dependent.
- This is true, for example, for a trending series.
- Specifically, a series that is stationary about a time trend *and* weakly dependent is known as a **trend stationary process.**

Weakening the Strict Exogeneity Assumptions - at a Price

- With the concepts from above, we can now consider relaxing the strict exogeneity assumption required for the Gauss-Markov Theorem in chapter 10.
- In doing so, however, we lose unbiasedness and the finite sample properties of the OLS estimators.

New Time Series Assumptions

- Assumption TS.1' (linearity and weak dependence): The stochastic process $\{(\boldsymbol{x}_t, y_t) : t = 1, \ldots, n\}$ is stationary and weakly dependent and follows the linear model:

$$y_t = \beta_0 + \beta_1 x_{t1} + \cdots + \beta_k x_{tk} + u_t \tag{6}$$

  where $\boldsymbol{x}_t = (x_{t1}, \ldots, x_{tk})'$ is a row vector of the regressors for time period t and $u_t : t = 1, \ldots, n$ denotes a sequence of errors.

  ...What's new here is stationarity and weak dependence - this is what allows the LLN and CLT to apply to sample averages.
- Assumption TS.2' (No Perfect Collinearity): Same as TS.2.
- Assumption TS.3' (Zero Conditional Mean): The explanatory variables are **contemporaneously exogenous**; i.e., $E(u_t|\boldsymbol{x}_t) = 0$.

Theorem 11.1: Unbiasedness

- **Theorem 11.1**: Under TS.1' through TS.3', the OLS estimators are consistent, with $plim\hat{\beta}_j = \beta_j \ \forall \ j = 0, 1, \ldots, n$.
- The key gain here relative to Theorem 10.1 is that we have relaxed the strict exogeneity assumption (TS.3 vs. TS.3').

  ... This allows for feedback effects.
- The cost here is that we now only have consistency. OLS need no longer be unbiased.
- The problem of having to rely on asymptotics in the context of time series data is that we often have relatively short time series.

The Static Model

- In the static model, with two explanatory variables, we have

$$y_t = \beta_0 + \beta_1 z_{t1} + \beta_2 z_{t2} + u_t \tag{7}$$

  ... the condition for consistency under weak dependence is

$$E(u_t | z_{t1}, z_{t2}) = 0. \tag{8}$$

- This *does not* rule out possible feedback effects from y_{t-1} to z_t; e.g.

$$z_{t1} = \delta_0 + \delta_1 y_{t-1} + v_t, \tag{9}$$

  which is ruled out by strict exogeneity.
- It does, however, rule out classical measurement errors, omitted variables bias, or errors in functional form.

The AR(1) Model

- Suppose that we have an AR(1) model,

$$y_t = \beta_0 + \beta_1 y_{t-1} + u_t \tag{10}$$

where

$$E(u_t | y_{t-1}, y_{t-2}, \ldots) = 0. \tag{11}$$

- This model satisfies TS.3', since

$$E(y_t | y_{t-1}, y_{t-2}, \ldots) = E(y_t | y_{t-1}) = \beta_0 + \beta_1 y_{t-1} \tag{12}$$

- The model, however, would clearly not satisfy TS.3.
- In this case, OLS estimators are consistent, but not unbiased.

Distributional Properties of OLS

- Assumption TS.4' (Homoskedasticity): The errors are **contemporaneously homoskedastic**; i.e., $Var(u_t|\boldsymbol{x}_t) = \sigma^2$.
- Assumption TS.5' (No Serial Correlation): For all $t \neq s, E(u_t u_s | \boldsymbol{x}_t, \boldsymbol{x}_s) = 0$
- **Theorem 11.2** (Asymptotic Normality of OLS): Under assumptions TS.1' through TS.5', the OLS estimators are asymptotically normally distributed. Further, the usual OLS standard errors, t-statistics, F-statistics, an LM statistics are asymptotically valid.

Efficient Market Hypotheses

- Under the efficient markets hypothesis (EMH),

$$E(y_t|y_{t-1}, y_{t-2}, \ldots) = E(y_t) \tag{13}$$

  ...Translation: Knowing the past does not help predict subsequent returns.
- Using an AR(q) model, and letting y_t denote the weekly returns on the NY stock exchange, we would have

$$y_t = \beta_0 + \beta_1 y_{t-1} + \cdots + \beta_q y_{t-q} + e_t \tag{14}$$

- The EMH would correspond to the hypothesis $\delta_1 = \cdots = \delta_q = 0$, with $H_1 : H0$ is not true.
- The following tables tests the EMH using $q = 1$, 2, and 4.
- In each case we fail to reject the null hypothesis using a joint F-statistic test (e.g., p-value = 0.22 when q=4).

Efficient Markets Hypothesis

	AR(1)	AR(2)	AR(4)
L.return	**0.059**	**0.060**	**0.063**
	(0.038)	**(0.038)**	**(0.038)**
L2.return		**-0.038**	**-0.043**
		(0.038)	**(0.038)**
L3.return			**0.030**
			(0.038)
L4.return			**-0.052**
			(0.038)
_cons	**0.180**	**0.186**	**0.185**
	(0.081)*	**(0.081)***	**(0.082)***
R2	**0.00**	**0.00**	**0.01**
N	**689**	**688**	**686**

* p<0.05; ** p<0.01

Highly Persistent Time Series

- Weak Dependence essentially requires that the influence on any one innovation or factor eventually dies out over time.
- But there are many examples of economic time series that are not weakly dependent.
- These series are referred to as **highly persistent** or **strongly dependent** time series.
- In their raw form, we cannot then use the standard LLN and CLT.
- However, they can sometimes be transformed into series that are weakly dependent.

Examples of Highly Persistent Time Series

- One highly persistent time series is an AR(1) process with $\rho_1 = 1$; i.e.,

$$y_t = y_{t-1} + e_t, \quad t = 1, 2, \ldots \tag{15}$$

  with $\{e_t : t = 0, 2, \ldots\}$ being an *iid* sequence with mean of zero and variance σ_e^2 and y_0 is independent of $e_t \ \ \forall t$.
- This series is also known as a **random walk.**
- Using $h = t$ and $\rho_1 = 1$ in (5), we have

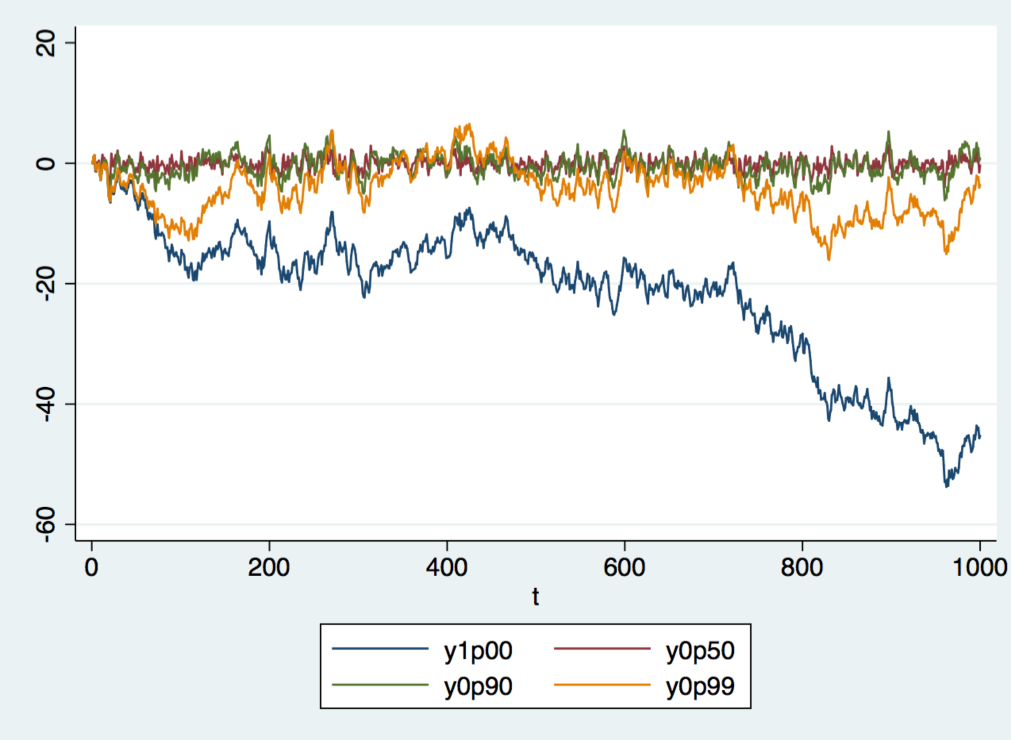
$$y_t = y_0 + \sum_{s=0}^{t-1} e_{t-s} = y_0 + \sum_{s=1}^{t} e_s \tag{16}$$

- Assuming $E(y_0) = 0$ and $Var(y_0)$, we the have $E(y_t) = 0 \ \ \forall t$ and

$$Var(y_t) = Var\left(\sum_{s=1}^{t} e_s\right) = \sum_{s=1}^{t} Var\left(e_s\right) = \sigma_e^2 t \tag{17}$$

  so the series is nonstationary.

The graph below plots an AR(1) process where $y_0 = 0$, $e_t \overset{iid}{\sim} \mathcal{N}(0,1)$ and $\rho_1 \in \{1.00, 0.99, 0.90, 0.50\}$.



The Random Walk (cont'd)

- The random walk is not only nonstationary, but it is *not* asymptotically uncorrelated.
- Wooldridge notes that $Corr(y_t, y_{t+h}) = \sqrt{t(t+h)}$.
- The influence of y_t on future values of y_{t+h} is very persistent.
- Using $r = h - s$, (5) implies that for the random walk:

$$y_{t+h} = y_t + \sum_{s=0}^{h-1} e_{t+h-s} = y_t + \sum_{r=1}^{h} e_{t+r}$$

- For the random walk, this in turn implies that

$$E(y_{t+h}|y_t) = E\left(y_t + \sum_{r=1}^{h} e_{t+r}\middle| y_t\right) = y_t. \tag{18}$$

- The random walk is a special case of the **unit root process**.

Trending versus Persistent Behavior

- It is important to distinguish trending versus persistent behavior.
- One illustration of this is an AR(1) model with drift; i.e.,

$$y_t = \alpha_0 + \rho_1 y_{t-1} + e_t, \quad t = 1, 2, \ldots \tag{19}$$

  with $\{e_t : t = 0, 2, \ldots\}$ being an *iid* sequence with mean of zero and variance σ_e^2 and y_0 is independent of $e_t \ \ \forall t$.
- If ρ_1 is small, the process is still trending, but the impact of a given y_t on future periods still diminishes over time.
- If $\rho_1 = 1$, then the stochastic process is known as a **random walk with drift**, and exhibits both a trend effect and highly persistent behavior., with

$$y_t = \alpha_0 + y_{t-1} + e_t, \quad t = 1, 2, \ldots \tag{20}$$

- The extent of persistence can have implications for the long term importance of any policy shift.

Transforming Highly Persistent Time Series

- Highly persistent time series violate assumption TS.1'.
- However, one can often transform them into sequences that are weakly dependent.
- A weakly dependent process is said to be **integrated of order zero** or **I(0)** and can be analyzed using standard regressions tools.
- Unit root processes are said to be **integrated of order one** or **I(1)**.
- First differencing these processes (also known as **difference-stationary processes**) will yield a weakly dependent series.
- This is easy to see in the case of a random walk, since

$$\Delta y_t = y_t - y_{t-1} = e_t, \tag{21}$$

  which is clearly stationary and weakly dependent.

Determining Order of Integration

- Determining whether a process is I(0) or I(1) (or...) requires more complex methods, covered later in the course.
- An informal tools is to simple estimate an AR(1) using OLS and see if ρ_1 lies close to one.
- A formal test is harder to do, since the resulting standard errors are incorrect under the null $H_0 : \rho_1 = 1$.
- Moreover, $\hat{\rho}_1$ can be severely biased downward when ρ_1 is close to one.

Fertility Equation

- In chapter 10, Wooldridge looked at the link between the general fertility rate *gfr* and the value of the personal tax exemption *pe*.
- We saw that *pe* was not statistically significant in a FDL specification.
- However, the variables do exhibit high degrees of autocorrelation (0.977 and 0.964, respectively).
- Using first differences, *cgfr* and *cpe* respectively, we still don't see a significant effect.
- But using lagged differences, there is a significant effect.
- Moreover, the results are robust to the inclusion of auxiliary variables and a time trend.

General Fertility Rate

	First Diff	FDL3 First Diff	FDL3 and Time	FDL3 and aux	FDL3 and gen
cpe	-0.043	-0.036	-0.035	-0.075	-0.062
	(0.028)	(0.027)	(0.027)	(0.032)*	(0.032)
L.cpe		-0.014	-0.013	-0.051	-0.039
		(0.028)	(0.028)	(0.033)	(0.032)
L2.cpe		0.110	0.111	0.088	0.095
		(0.027)**	(0.027)**	(0.028)**	(0.027)**
t			0.008		0.094
			(0.024)		(0.038)*
ww2				4.839	3.251
				(2.832)	(2.797)
pill				-1.676	-4.888
				(1.005)	(1.616)**
_cons	-0.785	-0.964	-1.267	-0.650	-3.141
	(0.502)	(0.468)*	(1.046)	(0.582)	(1.150)**
R2	0.03	0.23	0.23	0.30	0.36
N	71	69	69	69	69

* p<0.05; ** p<0.01

Dynamically Complete Models

- A **Dynamically Complete Model** is a time series model where no further lags of either the dependent variable or the explanatory variables helps to explain the mean of the dependent variable.
- Effectively, it means you are assuming that you have model correctly specified.
- In general, if we have:

$$y_t = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + u_t \tag{22}$$

then

$$E(y_t | \boldsymbol{x}_t, y_{t-1}, \boldsymbol{x}_{t-1}, y_{t-2}, \boldsymbol{x}_{t-2}, y_{t-3} \cdots) = E(y_t | \boldsymbol{x}_t) \tag{23}$$

or, given TS.3',

$$E(u_t | \boldsymbol{x}_t, y_{t-1}, \boldsymbol{x}_{t-1}, y_{t-2}, \boldsymbol{x}_{t-2}, y_{t-3} \cdots) = E(u_t | \boldsymbol{x}_t) = 0 \tag{24}$$

- Note: $\boldsymbol{x}_t$ in (22) can include lagged values of the dependent variable or of any regressor.

Example #1: Fertility FDL Model

The *gfr* model using differences and an FDL3 specification is apparently not dynamically complete.

General Fertility Rate

	First Diff	FDL3 First Diff	FDL3 and Time	FDL3 and aux	FDL3 and gen	FDL3 and gen	Dyn Comp?
cpe	-0.043	-0.036	-0.035	-0.075	-0.062	-0.062	-0.045
	(0.028)	(0.027)	(0.027)	(0.032)*	(0.032)	(0.032)	(0.026)
L.cpe		-0.014	-0.013	-0.051	-0.039	-0.039	0.002
		(0.028)	(0.028)	(0.033)	(0.032)	(0.032)	(0.027)
L2.cpe		0.110	0.111	0.088	0.095	0.095	0.105
		(0.027)**	(0.027)**	(0.028)**	(0.027)**	(0.027)**	(0.026)**
t			0.008		0.094	0.094	
			(0.024)		(0.038)*	(0.038)*	
ww2				4.839	3.251	3.251	
				(2.832)	(2.797)	(2.797)	
pill				-1.676	-4.888	-4.888	
				(1.005)	(1.616)**	(1.616)**	
L.cgfr							0.300
							(0.106)**
_cons	-0.785	-0.964	-1.267	-0.650	-3.141	-3.141	-0.702
	(0.502)	(0.468)*	(1.046)	(0.582)	(1.150)**	(1.150)**	(0.454)
R2	0.03	0.23	0.23	0.30	0.36	0.36	0.32
N	71	69	69	69	69	69	69

* p<0.05; ** p<0.01