Further Issues in Time Series Regressions: Large Sample Properties of OLS Estimators

Hüseyin Taştan[1]

[1]Yıldız Technical University
Department of Economics

Econometrics II

Asymptotic Properties of OLS

- In Chapter 10, we saw some cases where the classical linear model assumptions are not satisfied for certain time series problems.
- In such cases, we must appeal to large sample properties of OLS.
- In this section, we state and soften the assumptions and main results that justify OLS more generally.

TS.1' Linearity and Weak Dependence

- TS.1 assumed that the model is linear in β parameters.
- If x variables contain lagged dependent variables, such as y_{t-1}, y_{t-2}, TS.1 is replaced with:

Assumption TS.1': Linearity and Weak Dependence

Assumption TS.1' is the same as TS.1, except we must also assume that $\{(x_t, y_t) : t = 1, 2, ..\}$ is weakly dependent. In other words, the law of large numbers and the central limit theorem can be applied to sample averages.

TS.2' No Perfect Collinearity

Assumption TS.2': No Perfect Collinearity

Same as Assumption TS.2

This is the familiar full rank condition (see the matrix algebra notes).

TS.3' Zero Conditional Mean

- Instead of TS.3 (u_t is uncorrelated with past, present and future values of independent variables), we'll use

Assumption TS.3': Zero Conditional Mean

For each t, $E(u_t|x_t) = 0$.

- u_t and the explanatory variables are contemporaneously uncorrelated.

$$E(u_t) = 0, \quad Cov(x_{tj}, u_t) = 0, \quad j = 1, ..., k.$$

Consistency

- **Under these assumptions, OLS estimators are consistent.**

Theorem 11.1 Consistency of OLS Estimators

Under TS.1', TS.2' and TS.3', OLS estimators are consistent: $\text{plim}\, \hat{\beta}_j = \beta_j$, j=0,1,...,k.

Consistency of OLS Estimators

- There are some key practical differences between Theorems 10.1 and 11.1.
- Theorem 11.1, we conclude that the OLS estimators are consistent, but not unbiased.
- In Theorem 11.1, we have weakened the sense in which the explanatory variables (xs) must be exogenous.
- But weak dependence is required in the underlying time series.

Example

- In our example, z_{t1} is monthly percentage change in the money supply and y_t is monthly inflation rate.

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

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

- Suppose that the change in money supply depends on last month's rate of inflation (y_{t-1}).

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

- This mechanism generally causes z_{t1} and u_t to be correlated (as can be seen by plugging in for y_t.)
- This kind of feedback is allowed under Assumption TS.3'. We can use z_{t1} as an explanatory variable.

Example

- In the previous model, the consistency of OLS is

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

- This assumption does not allow the other factors contained in u_t correlated with z_{t1} and z_{t2}.
- But, the lagged values of random error term may be correlated with explanatory variables. For example, u_{t-1} and z_{t1} is correlated.
- Just as in cross-sectional regression, misspecified functional form and measurement errors in explanatory variables cause TS.3' to be invalid.

Example: Finite Distributed Lag Model

- In the finite distributed lag model:

$$y_t = \alpha_0 + \delta_0 z_t + \delta_1 z_{t-1} + \delta_2 z_{t-2} + u_t$$

- the expected value of u_t, given current and all past values of z, is zero:

$$E(u_t|z_t, z_{t-1}, z_{t-2}, z_{t-3}, ...) = 0$$

- This means that, once z_t,z_{t-1} ve z_{t-2} are included, no further lags of z affect $E(u_t|z_t, z_{t-1}, z_{t-2}, z_{t-3}, ...)$); if this were not true, we would put further lags into the equation.
- When we set $X_t = (z_t, z_{t-1}, z_{t-2})$, Assumption TS.3' is then satisfied: OLS will be consistent.
- As in the previous example, TS.3' does not rule out feedback from y to future values of z.

Example: First Order Autoregressive Model

- Consider the AR(1) model:

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

- where the error u_t is zero expected value, given all past values of y:

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

- Combining these two equations

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

- It means that, once y lagged one period has been controlled for, no further lags of y affect the expected value of y_t. (This is where the name "first order" originates). And the relationship is assumed to be linear.

Example: AR(1) Model

- In AR(1) model, u_t is uncorrelated with y_{t-1}, but u_t and y_t are correlated.
- Therefore, **a model with a lagged dependent variable cannot satisfy the strict exogeneity assumption TS.3**
- For the weak dependence condition to hold, we must assume that $|\rho_1| < 1$, as we mentioned. If this condition holds, then Theorem 11.1 implies that the OLS estimator of AR(1) model produces consistent estimators.
- Unfortunately, the OLS estimator of AR(1) model is biased and this bias can be large if the sample size is small or if β_1 is near one (if β_1 is near one, there is a severe downward bias).

Assumptions TS.4' Homoscedasticity and TS.5' No Serial Correlation

Assumption TS.4' Homoscedasticity

For all t, $Var(u_t|X_t) = \sigma^2$.

Assumption TS.5' No Serial Correlation

For all $t \neq s$, $E(u_t u_s|X_t, X_s) = 0$.

- In TS.4', note how we condition only on the explanatory variables at time t (compare to TS.4).
- In TS.5', we condition only on the explanatory variables in the time periods coinciding with u_t and u_s.

Asymptotic Normality of OLS

Theorem 11.2: Asymptotic Normality of OLS

Under TS.1' through TS.5', the OLS estimators are asymptotically normally distributed. Further, the usual OLS standard errors, t statistics, F statistics, and LM statistics are asymptotically valid.

- Under TS.1' through TS.5', we now obtain an asymptotic result that is practically identical to the cross sectional case.
- Even if the classical linear model assumptions do not hold, OLS is still consistent for large samples, and the usual inference procedures are valid. Of course, this hinges on TS.1' through TS.5' being true.
- In this section, we discuss ways in which the weak dependence assumption can fail. The problems of serial correlation and heteroskedasticity are treated in Chapter 12.

Example 11.4: Efficient Market Hypothesis

- A strict form of the efficient markets hypothesis (EMH) states that information observable to the market prior to week t should not help to predict the return during week t. If we use only past information on y, the EMH is stated as

$$E(y_t|y_{t-1}, y_{t-2}, ...) = E(y_t)$$

- If this information is false, then we could use information on past weekly returns to predict the current return. The EMH presumes that such investment opportunities will be noticed and will disappear almost instantaneously.

$$\widehat{\text{return}} = \underset{(0.081)}{0.180} + \underset{(0.038)}{0.059}\, return_{t-1}$$

$$n = 689 \quad R^2 = 0.0035 \; \bar{R}^2 = 0.0020$$

Efficient Market Hypothesis: AR(2) Model

- For example, an autoregressive model of order two, or AR(2) model, is:

$$y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + u_t$$
$$E(u_t|y_{t-1}, y_{t-2}, ...) = 0$$

- To test whether β_1 and β_2 are jointly significant, our null hypothesis is

$$H_0 : \beta_1 = \beta_2 = 0$$

- Adding homoskedasticity assumption $Var(u_t|y_{t-1}, y_{t-2}) = \sigma^2$, we can use F statistic to test the null. If we do not reject H_0, we will have evidence in favor of EMH.
- If we estimate an AR(2) model for $return_t$, the two lags are individually insignificant at the $\%10$ level. They are also jointly insignificant: the F statistic is approximately $F = 1.65$; the p-value for this F statistic (with 2 and 685 degrees of freedom) is about 0.193. Thus, we do no reject at even the $\%15$ significance level.

Example 11.5: Expectations Augmented Phillips Curve

- A linear version of the expectations augmented Phillips curve can be written as:

$$inf_t - inf_t^e = \beta_1(unem_t - \mu_0) + e_t$$

- Where μ_0 is the natural rate of unemployment and inf_t^e is the expected rate of inflation formed in year t. This model assumes that the natural rate of unemployment is constant.
- Under adaptive expectations, the expected value of current inflation depends on recently observed inflation:

$$inf_t - inf_{t-1} = \beta_0 + \beta_1 unem_t + e_t$$

$$\Delta inf_t = \beta_0 + \beta_1 unem_t + e_t$$

- Where $\Delta inf_t = inf_t - inf_{t-1}$ and $\beta_0 = -\beta_1\mu_0$. Therefore, under adaptive expectations, the expectations augmented Phillips curve relates the change in inflation to the level of unemployment and a supply shock, e_t.

Example 11.5: Expectations Augmented Phillips Curve

$$\widehat{\Delta \mathrm{inf}_t} = \underset{(1.38)}{3.03} - \underset{(0.230)}{0.543}\, unem_t$$

$$n = 48 \quad R^2 = 0.108\ \bar{R}^2 = 0.088$$

- The trade-off between cyclical unemployment and unanticipated inflation is estimated as a one-point increase in $unem$ lowers unanticipated inflation by over one half of a point. The effect is statistically significant (two-sided p-value 0.023).
- An estimate of natural unemployment rate can be obtained from this regression, $\mu_0 = \beta_0/(-\beta_1)$: $\hat{\mu}_0 = \hat{\beta}_0/(-\hat{\beta}_1) = 3.03/0.543 = 5.58$
- We estimate the natural rate to be about 5.6, which is well within the range suggested by macroeconomists $(5 - 6\%)$.

Using Highly Persistent Time Series: Example 11.6 (Fertility Equation)

- gfr: fertility rate, pe: dollar value of personal exemption. The first order autocorrelations for these series are very large: $\hat{\rho_1} = 0.977$ for gfr and $\hat{\rho_1} = 0.964$ for pe.
- These are suggestive of unit root behavior, and they raise questions about the use of the usual OLS t statistics in Chapter 10. We now estimate the equations using the first differences (and dropping the dummy variables for simplicity)

$$\widehat{\Delta \mathrm{gfr}_t} = -\underset{(0.502)}{0.785} - \underset{(0.028)}{0.043}\, \Delta pe_t$$

$$n = 71 \quad R^2 = 0.032\ \bar{R}^2 = 0.018$$

$$\widehat{\Delta \mathrm{gfr}_t} = -\underset{(0.468)}{0.964} - \underset{(0.027)}{0.036}\, \Delta pe_t - \underset{(0.028)}{0.014}\, \Delta pe_{t-1} + \underset{(0.027)}{0.110}\, \Delta pe_{t-2}$$

$$n = 69 \quad R^2 = 0.233\ \bar{R}^2 = 0.197$$

Example 11.7: Wages and Productivity

- The variable $hrwage$ is average hourly wage in the U.S. economy, and $outphr$ is output per hour. One way to estimate the elasticity of hourly wage with respect to output per hour is to estimate the equation:

$$\log(hrwage_t) = \beta_0 + \beta_1 \log(outphr_t) + \beta_2 t + u_t$$

- where the time trend is included because $log(hrwage)$ and $log(outphrt)$ both display clear, upward, linear trends. The estimated model with trend:

$$\widehat{\log(\mathrm{hrwage}_t)} = -\underset{(0.37)}{5.33} + \underset{(0.09)}{1.64}\, \log(outphr_t) - \underset{(0.002)}{0.018}\, \mathrm{t}$$

$$n = 41 \quad R^2 = 0.971\ \bar{R}^2 = 0.970$$

Example 11.7: Wages and Productivity

- The regression results must be viewed with caution. Even after linearly detrending $log(hrwage)$, the first order autocorrelation is 0.967, and for detrended $log(outphr)$, $\hat{\rho}_1 = 0.945$. These suggest that both series have unit roots, so we reestimate the equation in first differences (and we no longer need a time trend)

$$\Delta \log(\widehat{\text{hrwage}_t}) = \underset{(0.0042)}{-0.0036} + \underset{(0.173)}{0.809}\, \Delta \log(outphr_t)$$
$$n = 40 \quad R^2 = 0.364 \;\; \bar{R}^2 = 0.348$$