ECON 2007: Quant Econ and Econometrics

Serial Correlation and Heteroskedasticity

Dr. Áureo de Paula

Department of Economics
University College London

What of OLS with Serially Correlated Errors?

- Unbiasedness: TS.1-TS.3 regardless of serial correlation!
- Consistency: TS.1'-TS.3' regardless of serial correlation!
- Lack of serial correlation TS.5 and TS.5' was invoked to obtain standard errors and necessary to show that OLS was B(est) L(inear) U(nbiased) E(stimator).
- With serial correlation, OLS is no longer BLUE.

Variance of OLS under Serial Correlation

- To see how the variance of OLS is affected, imagine that

$$u_t = \rho u_{t-1} + e_t, \quad t = 1, \ldots, n$$

  with $|\rho| < 1$ and e_t, iid with mean zero and variance σ^2.
- Now consider the simple static linear regression:

$$y_t = \beta_0 + \beta_1 x_t + u_t.$$

- In this case, remember that

$$\hat{\beta}_1 = \beta_1 + \frac{\sum_{t=1}^{n}(x_t - \overline{x})u_t}{\sum_{t=1}^{n}(x_t - \overline{x})^2} = \beta_1 + \frac{\sum_{t=1}^{n}(x_t - \overline{x})u_t}{SST_x}$$

Variance of OLS under Serial Correlation

- The variance of the OLS estimator (conditional on $x_1, \ldots, x_T$) is then given by

$$Var(\hat{\beta}_1) = \frac{Var(\sum_{t=1}^{T}(x_t - \overline{x})u_t)}{SST_x^2}$$

- If $n = 2$ and $Var(u_t) = \sigma_u^2$, the above becomes

$$\frac{\sum_{t=1}^{2}(x_t - \overline{x})^2\sigma_u^2 + 2(x_1 - \overline{x})(x_2 - \overline{x})Cov(u_1, u_2)}{SST_x^2}$$

Variance of OLS under Serial Correlation

- Since $Cov(u_1, u_2) = \rho\sigma_u^2$, we have that

$$Var(\hat{\beta}_1) = \frac{\sigma_u^2}{SST_x} + 2\frac{\sigma_u^2 \rho (x_1 - \overline{x})(x_2 - \overline{x})}{SST_x^2}$$

- If $n > 2$ the formula can be shown to equal

$$Var(\hat{\beta}_1) = \frac{\sigma_u^2}{SST_x} + 2\frac{\sigma_u^2 \sum_{t=1}^{T-1} \sum_{j=1}^{T-t} \rho^j (x_t - \overline{x})(x_{t+j} - \overline{x})}{SST_x^2}$$

Variance of OLS under Serial Correlation

- When $\rho = 0$ (no serial correlation), the previous formula coincides with the usual formula for the variance of the OLS estimator variance.
- Otherwise, the usual formula may over- or under-estimate the variance of the coefficient estimator.
- If $\rho > 0$ and x_t is positively correlated through time, the usual formula will underestimate the variance.

OLS with Lag Dependent Variables

- If $\mathbf{x}_t$ contains a lag dependent variable, TS.3 no longer holds (even when there is no serial correlation) and OLS is not unbiased.
- When there is no serial correlation, OLS may still be consistent provided TS.3' holds.
- Even when there is serial correlation, as long as TS.1'-TS.3' hold, OLS is consistent.

OLS with Lag Dependent Variables

- For example, take

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

with $|\beta_1| < 1$ (TS.1' holds). Let

$$u_t = y_t - E(y_t|y_{t-1})$$

Because $E(u_t|y_{t-1}) = 0$, TS.3' holds and OLS is consistent (provided TS.2' also holds). In this case,

$$Cov(u_t, u_{t-1}) = Cov(u_t, y_{t-1} - \beta_0 - \beta_1 y_{t-2}) = -\beta_1 Cov(u_t, y_{t-2})$$

which needs not be zero.

OLS with Lag Dependent Variables

- In other circumstances, serial correlation and a lagged dependent variable may invalidate TS.3'.
- For instance, let

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

and

$$u_t = \rho u_{t-1} + e_t$$

where $E(e_t | u_{t-1}, u_{t-2}, \ldots) = E(e_t | y_{t-1}, y_{t-2}, \ldots) = 0$.

- In this case,

$$Cov(u_t, y_{t-1}) = \rho Cov(u_{t-1}, y_{t-1}) \neq 0$$

unless $\rho = 0$ (because $y_{t-1} = \beta_0 + \beta_1 y_{t-2} + u_{t-1}$).

- OLS would not be consistent then.

OLS with Lag Dependent Variables

- But in this case, notice that

$$
\begin{aligned}
y_t &= \beta_0 + \beta_1 y_{t-1} + u_t \\
&= \beta_0 + \beta_1 y_{t-1} + \rho u_{t-1} + e_t \\
&= \beta_0 + \beta_1 y_{t-1} + \rho(y_{t-1} - \beta_0 - \beta_1 y_{t-2}) + e_t \\
&= \alpha_0 + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + e_t
\end{aligned}
$$

where $\alpha_0 = \beta_0(1-\rho)$, $\alpha_1 = \beta_1 + \rho$ and $\alpha_2 = \rho\beta_1$.

- Since e_t is uncorrelated with y_{t-1} and y_{t-2} the OLS estimator for the above model would be consistent!

OLS with Lag Dependent Variables

"You need a good reason for having both a lagged dependent variable in a model and a particular model of serial correlation in the errors. Often serial correlation in the errors of a dynamic model simply indicates that the dynamic regression function has not been completely specified." (p.412)

Testing for Serial Correlation

- How can we test for serial correlation in the residual of

$$y_t = \beta_0 + \beta_1 x_{1t} + \cdots + \beta_k x_{kt} + u_t$$

- The most common form of serial correlation in the literature is an AR(1) structure: $u_t = \rho u_{t-1} + e_t$. We will investigate a few tests of the null hypothesis:

$$H_0 : \rho = 0$$

- We discuss tests when $x_{1t}, \ldots, x_{kt}$ are (1) strictly exogenous and (2) when they are no strictly exogenous.

Testing for Serial Correlation: Strictly Exogenous Regressors

- Here we assume that $E(e_t|u_{t-1}, u_{t-2}, \dots) = 0$ and $Var(e_t|u_{t-1}) = Var(e_t) = \sigma^2$.
- If we observed u_t one could simply test H_0 by regressing u_t on u_{t-1}. Given our assumptions this test would be valid in large samples (i.e., asymptotically).
- We do not observe u_t, but we can estimate it!

Testing for Serial Correlation with Strictly Exogenous Regressors

- So, to test for AR(1) serial correlation we
  1. Regress y_t on $x_{1t}, \ldots, x_{kt}$ and save the estimated residuals $\hat{u}_t$.
  2. Regress $\hat{u}_t$ on $\hat{u}_{t-1}$ and obtain a coefficient estimate $\hat{\rho}$ and its t-stat $t_{\hat{\rho}}$.
  3. Use $t_{\hat{\rho}}$ to test H_0.
- Any source of serial correlation that causes adjacent errors u_t to be correlated can be detected using this test.
- Beware of numerical versus statistical significance when sample size is large (not likely in many applications).

Testing for Serial Correlation with Strictly Exogenous Regressors

- Another test for AR(1) serial correlation relies on the Durbin-Watson statistic:

$$DW = \frac{\sum_{t=2}^{T}(\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^{T}\hat{u}_t^2} \approx 2(1-\hat{\rho})$$

- The distribution of this test statistic (conditional on regressors) depends on the explanatory variables, the sample size, the number of regressors and whether the regression contains an intercept.

Testing for Serial Correlation with Strictly Exogenous Regressors

- Instead of tabulating critical values for all these scenarios, two thresholds (d_U and d_L) are usually presented for a test of H_0 agains $H_1 : \rho > 0$ at a given signicance level.
- For example, when $n = 45$, $k = 4$ and the significance level is 5%, $d_L = 1.336$ and $d_U = 1.720$.
- When H_0 is true, $DW \approx 2$. If $DW > d_U$, we fail to reject H_0. If $DW < d_L$, we reject H_0 in favor of the alternative. When $d_L < DW < d_U$, the test is inconclusive.

Testing for Serial Correlation without Strictly Exogenous Regressors

- Without strict exogeneity, at least one x_t is correlated with u_{t-1}.
- When this is the case, Durbin (1970) suggested an alternative testing procedure that proceeds as follows:
  1. Run OLS of y_t on $x_{1t}, \ldots, x_{kt}$ and obtain the residuals $\hat{u}_t$.
  2. Run the regression of $\hat{u}_t$ on $x_{1t}, \ldots, x_{kt}$ and $\hat{u}_{t-1}$ and obtain the coefficient $\hat{\rho}$ on $\hat{u}_{t-1}$ and its t statistic, $t_{\hat{\rho}}$.
  3. Use $t_{\hat{\rho}}$ to test H_0.
- $x_{1t}, \ldots, x_{kt}$ may contain lagged dependent variables and other nonstrictly exogenous explanatory variables.

Testing for Serial Correlation without Strictly Exogenous Regressors

- Because $\hat{u}_t = y_t - \hat{\beta}_0 - \hat{\beta}_1 x_{1t} - \cdots - \hat{\beta}_k x_{kt}$, the t-stat is the same if we replace $\hat{u}_t$ with y_t in step 2 above.
- Heteroskedasticity of unknown form can be accomodated by using heteroskedasticity-robust t stats.
- The test can be modified to test for higher order serial correlation (e.g., $u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2}$).

Serial Correlation-Robust Inference with OLS

- OLS can be more robust than estimators that allow for serial correlation (see book). We need nevertheless to use appropriate standard errors.

- In the simple linear regression model

$$y_t = \beta_0 + \beta_1 x_t + u_t,$$

remember that we have

$$\hat{\beta}_1 = \beta_1 + \frac{\sum_{t=1}^{T}(x_t - \overline{x})u_t}{\sum_{t=1}^{T}(x_t - \overline{x})^2} = \beta_1 + \frac{\sum_{t=1}^{T} r_t u_t}{\sum_{t=1}^{T} r_t^2}$$

and the variance of $\hat{\beta}_1$ (given $x_t, t = 1, \ldots, T$) is

$$\frac{Var(\sum_{t=1}^{T} a_t)}{\left(\sum_{t=1}^{T} r_t^2\right)^2} = \frac{TVar(a_t) + 2\sum_{j=1}^{T-1}(T-j)Cov(a_t, a_{t-j})}{\left(\sum_{t=1}^{T} r_t^2\right)^2}$$

where $r_t = x_t - \overline{x}$ and $a_t = r_t u_t$.

Serial Correlation-Robust Inference with OLS

- This can be further simplified to:

$$\frac{TVar(a_t)}{\left(\sum_{t=1}^{T} r_t^2\right)^2} \times \frac{v}{Var(a_t)}$$

where $v = \left(Var(a_t) + 2\sum_{j=1}^{T-1} \frac{T-j}{T} Cov(a_t, a_{t-j}) \right)$.

- Notice that the first term is the usual *robust* variance for the estimator.

Serial Correlation-Robust Inference with OLS

- This can be generalized to a multiple linear regression.
- In the linear regression model

$$y_t = \beta_0 + \beta_1 x_{1t} + \cdots + \beta_k x_{kt} + u_t,$$

the standard error for the coefficient estimator $\hat{\beta}_1$ can be shown to be equal to

$$Avar(\hat{\beta}_1) = \left(\sum_{t=1}^{T} E(r_t^2) \right)^{-2} Var\left(\sum_{t=1}^{T} r_t u_t \right)$$

where r_t is the residual in

$$x_{1t} = \delta_0 + \delta_2 x_{2t} + \cdots + \delta_k x_{kt} + r_t$$

Serial Correlation-Robust Inference with OLS

- This variance can be estimated as follows:
  1. Regress x_{1t} on $x_{2t}, \ldots, x_{kt}$ and save the residuals $\hat{r}_t$.
  2. For some integer g, compute
  $$\hat{v} = \sum_{t=1}^{T} \hat{a}_t^2 + 2 \sum_{h=1}^{g} [1 - h/(g+1)] \left(\sum_{t=1}^{T} \hat{a}_t \hat{a}_{t-h} \right)$$
  where $\hat{a}_t = \hat{r}_t \hat{u}_t$.
  3. The heteroskedasticity and autocorrelation robust standard error is then
  $$[\text{“}se(\hat{\beta}_1)\text{”}/\hat{\sigma}]^2 \sqrt{\hat{v}}.$$

  where “$se(\hat{\beta}_1)$” denote the usual OLS standard error.
- This obtains as the variance formula (in the simple regression) can also be written as:

$$\left(\frac{\sum u_t^2 / T / \left(\sum_{t=1}^{T} r_t^2 \right)}{\sum u_t^2 / T} \right)^2 \times Tv$$

Serial Correlation-Robust Inference with OLS

- The corrected standard errors will perform better in large samples.
- One needs to choose g! Newey and West (1987) recommend (the integer part of) $4(n/100)^{2/9}$.
- With strong serial correlation OLS can be very inefficient. In this case it might make sense to first difference the data before estimation:

$$\Delta y_t = \beta_1 \Delta x_t + \Delta u_t.$$

  (With AR(1) residuals, this will typically eliminate most of the serial correlation when ρ is large and positive.)

Heteroskedasticity

- Heteroskedasticity would not cause bias or inconsistency and HAC standard errors can be used for inference.
- Some forms of heteroskedasticity are nevertheless interesting in their own. Especially that implied by autoregressive conditional heteroskedasticity (ARCH) models and their generalizations.
- These models postulate something like

$$E(u_t^2 | u_{t-1}, u_{t-2}, \ldots) = \alpha_0 + \alpha_1 u_{t-1}^2$$

  whereby the variance of the residuals changes dynamically.
- Notice that conditions TS.1'-TS.5' may still hold and inference with OLS is still correct.

Heteroskedasticity

- We might nonetheless be able to get asymptotically more efficient estimators if we exploit the particular nature of the heteroskedasticity.
- This particular form of heteroskedasticity is also of great interest for its connection with the analysis of volatility dynamics.
- It was originally used to investigate the behavior of inflation in the UK for example.

Heteroskedasticity

For the United States (Engle, 1983):

$$h_t = \alpha_0 + \alpha_1 \sum_{j=1}^{8} (9 - j)\epsilon_{t-j}^2/36.$$

$$\begin{aligned}\dot{P} = &+ \underset{(4.1)}{0.33\dot{P}_{-1}} + \underset{(2.6)}{0.20\dot{P}_{-2}} + \underset{(3.8)}{0.06P\dot{M}_{-1}} + \underset{(3.5)}{0.16\dot{W}_{-1}} + \underset{(0.9)}{0.05\dot{M}_{-1}} \\ &+ \underset{(1.4)}{0.00002t} + \underset{(0.07)}{0.00006}\end{aligned}$$

$$\alpha_0 = \underset{(2.7)}{0.000006} \quad \alpha_1 = \underset{(2.7)}{0.56.}$$

where $h = E(\epsilon_t^2 | \epsilon_{t-1}, \epsilon_{t-2}, \dots)$, $\dot{\text{P}}$ = inflation, $\dot{\text{W}}$ = wage change, $\dot{\text{M}}$ = money supply change.

These slides covered:

Wooldridge 12, Stock and Watson 14 and 15.