Econometrics

Chapter 11: Regressions with Serially Correlated Errors

In Choi

Sogang University

Testing for AR(1) Serial Correlation

- Want to be able to test for whether the regression errors are serially correlated or not.
- Want to test the null that $\rho = 0$ in

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

  where u_t is the model error term and e_t is iid.
- With strictly exogenous regressors, the test is very straightforward – simply regress the residuals on lagged residuals and use a t-test. the t-test is defined by

$$\frac{\hat{\rho}}{\sqrt{\hat{\sigma}^2 \left(\sum_{t=2}^{n} \hat{u}_{t-1}^2 \right)^{-1}}},$$

  where $\hat{\sigma}^2 = \frac{1}{n-1} \sum_{t=2}^{n} (\hat{u}_t - \hat{\rho}\hat{u}_{t-1})^2$. This test has a standard normal distribution in the limit.

Testing for AR(1) Serial Correlation

Heuristics We may write

$$\frac{\hat{\rho}}{\sqrt{\hat{\sigma}^2\left(\sum_{t=2}^n \hat{u}_{t-1}^2\right)^{-1}}} = \left(\frac{\sum_{t=2}^n \hat{u}_{t-1}^2}{n}\right)^{-1/2} \frac{\sum_{t=2}^n \hat{u}_t \hat{u}_{t-1}}{\sqrt{n}\sqrt{\hat{\sigma}^2}}.$$

But since $\hat{u}_t = u_t +$ a negligible term,
$\frac{\sum_{t=2}^n \hat{u}_{t-1}^2}{n} \simeq E(u_{t-1}^2) = \sigma^2$ by the law of large numbers and $\frac{\sum_{t=2}^n \hat{u}_t \hat{u}_{t-1}}{\sqrt{n}} \simeq \frac{\sum_{t=2}^n u_t u_{t-1}}{\sqrt{n}} \simeq N(0, \sigma^4)$ by the central limit theorem. Thus, $\frac{\hat{\rho}}{\sqrt{\hat{\sigma}^2\left(\sum_{t=2}^n \hat{u}_{t-1}^2\right)^{-1}}} \simeq N(0,1)$.

Testing for AR(1) Serial Correlation

- An alternative is the Durbin-Watson (DW) statistic, which is calculated by many packages. This depends on the assumption of normality.

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

- If the DW statistic is around 2, then we can reject serial correlation, while if it is significantly < 2 we cannot reject.
- Critical values are difficult to calculate, making the t- test easier to work with.

Testing for AR(1) Serial Correlation

- If the regressors are not strictly exogenous (e.g., lagged dependent variables), then the t- and DW tests will not work.
- Suppose that the model is $y_t = \beta y_{t-1} + u_t \ (|\beta| < 1)$ and we want to test whether or not $\{u_t\}$ is serially correlated. Let $\hat{u}_t = u_t - (\hat{\beta} - \beta) y_{t-1}$. It is required that $\frac{\sum_{t=2}^{n} \hat{u}_t \hat{u}_{t-1}}{\sqrt{n}}$ has a normal distribution in the limit under the null of $\rho = 0$ to use the t-ratio as before. But it doesn't.

Testing for AR(1) Serial Correlation

Heuristics Consider

$$\frac{1}{\sqrt{n}}\sum_{t=2}^{n}\hat{u}_t\hat{u}_{t-1}$$
$$= \frac{1}{\sqrt{n}}\sum_{t=2}^{n}u_tu_{t-1} + n(\hat{\beta}-\beta)^2\frac{1}{n^{1.5}}\sum_{t=2}^{n}y_{t-1}y_{t-2}$$
$$-\sqrt{n}(\hat{\beta}-\beta)\frac{1}{n}\sum_{t=2}^{n}u_ty_{t-2} - \sqrt{n}(\hat{\beta}-\beta)\frac{1}{n}\sum_{t=2}^{n}u_{t-1}y_{t-1}.$$

When $\rho = 0$, $\frac{1}{\sqrt{n}}\sum_{t=2}^{n}u_tu_{t-1}$ and $\sqrt{n}(\hat{\beta}-\beta)$ are normally distributed in the limit, and $\frac{1}{n}\sum_{t=2}^{n}u_ty_{t-2} \xrightarrow{p} 0$. But the probability limit of $\frac{1}{n}\sum_{t=2}^{n}u_{t-1}y_{t-1}$ is not zero, and this causes the problem.

- In this case, regress the residual on the lagged residual and all of the regressors and use t-test as above.

Correcting for Serial Correlation

- Start with case of strictly exogenous regressors, and maintain all G-M assumptions except no serial correlation.
- Prais-Winsten procedure
  Assume errors follow AR(1). We need to try and transform the equation so we have no serial correlation in the errors.

$$y_t = \beta_0 + \beta_1 x_{1t} + \ldots + \beta x_{kt} + u_t$$

$$u_t = \rho u_{t-1} + e_t, \;\; e_t \sim iid\left(0, \sigma^2\right), \;\; |\rho| < 1.$$

Correcting for Serial Correlation

The transformed model is

$$\begin{aligned}
\sqrt{1-\rho^2}\, y_1 &= \sqrt{1-\rho^2} \sum_{j=1}^{k} x_{j1}\beta_j + e_1^* \\
y_t - \rho y_{t-1} &= \sum_{j=1}^{k} \left(x_{jt} - \rho x_{j,t-1}\right) \beta_j + e_t
\end{aligned}$$

where $e_1^* = \sqrt{1-\rho^2}\, u_1$. Since
(i) $Ee_i^* = Ee_t = 0$
(ii) $Var\left(e_1^*\right) = \left(1-\rho^2\right) Var\left(u_1\right) = \left(1-\rho^2\right) \cdot \frac{\sigma^2}{1-\rho^2} = \sigma^2$
(iii) $E\left(e_1^* e_t\right) = \sqrt{1-\rho^2} E\left(u_1 e_t\right) = 0,\ t = 2, \cdots, n,$
The error terms of the transformed model satisfy the condition of the standard linear regression model.

Note Since $u_1 = \sum_{j=0}^{\infty} \rho^j e_{1-j}$,

$$
\begin{aligned}
Var(u_1) &= E(u_1^2) \\
&= E(\sum_{j=0}^{\infty} \rho^j e_{1-j})^2 \\
&= \sum_{j=0}^{\infty} \rho^{2j} E(e_{1-j})^2 = \frac{\sigma^2}{1-\rho^2}.
\end{aligned}
$$

In addition,

$$
\begin{aligned}
E(u_1 e_t) &= E\left(\sum_{j=0}^{\infty} \rho^j e_{1-j} e_t\right) \\
&= \sum_{j=0}^{\infty} \rho^j E\left(e_{1-j} e_t\right) = 0 \text{ for } t \geq 2.
\end{aligned}
$$

Correcting for Serial Correlation

- When

$$y_t - \rho y_{t-1} = \sum_{j=1}^{k} (x_{jt} - \rho x_{j,t-1}) \beta_j + e_t$$

is used ignoring the first observation, it is called the Cochrane-Orcutt procedure.

Feasible GLS Estimation

- Feasible Prais-Winsten procedure
  1. Run OLS and get $\hat{u}_t$.
  2. Run $AR(1)$ regression using $\hat{u}_t$. This gives $\hat{\rho}$.
  3. Transform the model using $\hat{\rho}$ and run OLS.
- Eviews tip: y c x ar(1)

Robust standard error

Consider the linear regression model

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

where $\{x_t, u_t\}$ is a zero-mean, stationary process of unknown structure. Write

$$\sqrt{n}\left(\hat{\beta}_1 - \beta_1\right) = \frac{\sum (x_t - \bar{x}) u_t / \sqrt{n}}{\sum (x_t - \bar{x})^2 / n}.$$

Robust standard error

Under some additional conditions on $\{x_t, u_t\}$, we have

$$\frac{1}{\sqrt{n}}\sum(x_t-\bar{x})u_t \xrightarrow{d} N(0, \lim \frac{1}{n}E\left(\sum(x_t-\bar{x})u_t\right)^2).$$

Thus, if $\sum(x_t-\bar{x})^2/n \xrightarrow{p} M$,

$$\sqrt{n}\left(\hat{\beta}_1-\beta_1\right) \xrightarrow{d} N(0, M^{-2}\lim \frac{1}{n}E\left(\sum(x_t-\bar{x})u_t\right)^2).$$

The Newey-West procedure estimates $\sqrt{M^{-2}\lim \frac{1}{n}E\left(\sum(x_t-\bar{x})u_t\right)^2}$, which is the robust standard error in the presence of serial correlation in the errors.