Econ 312

Monday, March 30

Estimation of time-series regression models with stationary variables

Reading: Online time-series Chapter 2

Class notes: Pages 101 to 105

Daily problem: Breusch-Godfrey test (assigned for 3/16)

Today's Far Side offering

Context and overview

- **Last class**: We talked in detail about stationarity and its implications for time-series regression
- **Today**: We discuss how to proceed with estimation if all variables are stationary
  - **Assumptions** underlying use of OLS
  - **Autocorrelation** in the error
    - Implications
    - Testing
    - Correcting

Exogeneity and strict exogeneity

- To use OLS, we need the independent variable to be **stationary** and **exogenous**
- Exogeneity is the standard OLS assumption that the error is uncorrelated with x: $E(u_t \mid x_t) = 0$
- In a time-series context, we also need it to be uncorrelated with past and perhaps future x
- **Exogeneity**: $E(u_t \mid x_t, x_{t-1}, \ldots) = 0$
- **Strict exogeneity**: $E(u_t \mid \ldots, x_{t+2}, x_{t+1}, x_t, x_{t-1}, x_{t-2}, \ldots) = 0$

TS assumptions for OLS

- **TS1**: **Linear model** with additive error
- **TS2**: No perfect **multicollinearity**
- **TS3**: Regressor is **exogenous** to u
- **TS4**: **Homoskedasticity**
- **TS5**: **No autocorrelation** between one u_t and another u_{t-s}
- **TS6**: Error term is **normally distributed** (not totally necessary)

Autocorrelation in *x* and *u*

- What's the distinction between x and u?
  - We observe x but not u, but both affect y
- Nearly **all time-series variables are autocorrelated** over time
  - That includes the unobserved variables that we put in u
  - It also includes the right-hand variables x
- Autocorrelation in $\boldsymbol{x}$ **does not violate** the TS assumptions that underlie OLS estimation
- Autocorrelation in $\boldsymbol{u}$ **violates** assumption TS-5, so it requires attention in choosing and evaluating estimators

Serial autocorrelation

- Assumption TS-5 is almost never satisfied
- **Implications** ~ heteroskedasticity:
  - OLS is unbiased and consistent
  - OLS is not fully efficient
  - OLS standard errors are biased, so we can't use usual *t* statistics
- **Solutions**
  - Estimate by OLS but use HAC robust standard errors
  - Estimate by efficient generalized least squares
- In practice, can often reduce autocorrelation by **adding more lags** to achieve a **dynamically complete model**

Detecting serial correlation

- OLS estimator is consistent, so can use **OLS residuals** as basis for testing for autocorrelation
- We would be looking to see if $\hat{u}_t$ is correlated with $\hat{u}_{t-s}$ (for $s = 1$ in particular)
  - Recall that we defined $\text{corr}(u_t, u_{t-s}) = \rho_s$
  - We want to test $H_0 : \rho_s = 0$
  - An obvious choice would be to look at the sample autocorrelations r_s
  - **Three common tests** are tests of correlation in OLS residuals

Breusch-Godfrey test

- **Regress** either y_t or $\hat{u}_t$ on all of your x variables, **plus one or more lags of the residuals:** $\hat{u}_{t-1}, \hat{u}_{t-2}, \ldots, \hat{u}_{t-s}$
- If there is no autocorrelation, the coefficients on the lagged residuals are zero
- Test the significance of the lagged residuals with
  - **Joint *F* test**, if using y as the dependent variable (t statistic works if $s = 1$)
  - **Lagrange multiplier test:** $T \times R^2 \sim \chi^2$ with s degrees of freedom, if using residuals, where T is the number of observations

Other tests

- Page 102 of the notes describes two other tests that are less common:
- **Durbin-Watson test** was developed decades ago, but is awkward to use
  - You'll see it in older papers
  - The critical values for the test depend on the sample, so only upper and lower bounds are tabulated
- **Box-Lyung *Q* test** (sometimes called the "portmanteau test") is sometimes used
  - Formula is on page 102 of the class notes

Newey-West (HAC consistent) standard errors

- If inefficient OLS estimators are acceptable, we can conduct valid tests using **standard errors that are robust** to autocorrelation (and heteroskedasticity) developed by **Newey and West**
- The formula is complicated (see pages 103 and 104) and there is one parameter that must be specified: the **number of lags** to use in approximating the standard errors
  - This is independent of the number of lags in the actual model
  - Stock and Watson suggest using $m = 0.75T^{1/3}$
  - Implement in Stata with the post-estimation command **newey , lags(m)**

GLS estimation with AR(1) error

- Suppose that

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

$$u_t = \rho u_{t-1} + \varepsilon_t, \text{ with } \varepsilon \text{ being white noise and } -1 < \rho < 1$$

- As when we used weighted least squares to solve heteroskedasticity, we seek to transform the model into one with a classical error term
- In this case, we can "**quasi-difference**" y and x by subtracting their lagged values multiplied by ρ (assumed to be known for now)

Quasi-differencing

$$\tilde{y}_t = \begin{cases} y_t\sqrt{\left(1-\rho^2\right)}, & t=1, \\ y_t - \rho y_{t-1}, & t=2,3,\ldots,T, \end{cases} \quad \tilde{x}_t = \begin{cases} x_t\sqrt{\left(1-\rho^2\right)}, & t=1, \\ x_t - \rho x_{t-1}, & t=2,3,\ldots,T, \end{cases} \quad \tilde{u}_t = \begin{cases} u_t\sqrt{\left(1-\rho^2\right)}, & t=1, \\ u_t - \rho u_{t-1}, & t=2,3,\ldots,T. \end{cases}$$

- We can then estimate the model $\tilde{y}_t = (1-\rho)\beta_0 + \beta_1\tilde{x}_t + \tilde{u}_t$
  - This model has an error term that is not serially correlated
  - If TS-5 was the only OLS assumption violated, then OLS is fully efficient for the quasi-differenced model because all assumptions are satisfied
- Note the asymmetric treatment of the first observation: some methods simply delete this observation
- **What is ρ**? How about using $\hat{\rho} = \text{corr}\left(\hat{u}_t, \hat{u}_{t-1}\right)$?

Estimating ρ with OLS residuals

- **Two-step** procedure
  1. Run OLS regression and get residuals; estimate $\hat{\rho} = \text{corr}\left(\hat{u}_t, \hat{u}_{t-1}\right)$
  2. Quasi-difference the model and run OLS to get efficient estimator

- **Prais-Winsten** method uses all T observations
- **Cochrane-Orcutt** method omits the first observation

- In Stata, use the **prais** command (option **corc** to get Cochrane-Orcutt)

A problematic special case

- OLS residuals are consistent provided the error is uncorrelated with x, but there is an important case where this is violated
- **Lagged dependent variable**: If y_{t-1} is a regressor, that regressor is affected by lagged error u_{t-1}
  - If u is autocorrelated, then u_{t-1} also affects u_t, making the regressor y_{t-1} correlated with u_t
  - TS-3 is violated, making OLS biased and inconsistent
  - Can't rely on OLS residuals to provide estimate of ρ
  - **Hildreth-Lu** procedure searches directly for the $\hat{\rho}$ that minimizes the model SSR rather than using (inconsistent) OLS residuals
  - Use **ssesearch** option in **prais** to do use Hildreth-Lu

Review and summary

- Exogeneity and strict exogeneity are important in time series regressions
- The assumptions underlying OLS are slightly modified
- The error term is nearly always serially correlated in time-series regressions
- This leads OLS to be inefficient and its standard errors to be biased
- We can use Newey-West HAC robust standard errors
- We can use GLS to get efficient estimators that are better than OLS

From *The Devil's Dictionary*

Distance, *n*. The only thing that the rich are willing for the poor to call theirs, and keep.

[Note that for Ambrose Bierce, as well as for me, "distance" is a NOUN, not a VERB. I refuse to engage in "social distancing," though I am perfectly willing to be "keeping social distance"!]

What’s next?

- This session has examined the problem of a serially correlated error term in a time-series regression model
- In the next class (April 1), we consider how to estimate models in which y responds to x slowly over time, so that lagged values of x affect the current value of y
  - These are called **distributed-lag models**
  - The lags here relate to the **deterministic** part of the model, not to the error term
  - These models are covered in Chapter 3 of the online readings