Correcting for Serial Correlation

Tyler Ransom

Univ of Oklahoma

Feb 28, 2019

Today's plan

1. Review reading topics
    1.1 Correcting for Serial Correlation
2. In-class activity: more practice with TS data, correcting for serial correlation

Review

Time series models

- Consider the following general time series regression model:

$$y_t = \beta_0 + \beta_1 x_t + \beta_2 x_{t-1} + \beta_3 z_t + \beta_4 z_{t-1} + \beta_5 y_{t-1} + u_t$$

- The current interest rate depends on:
  - current and previous trade deficit
  - current and previous nominal GDP
  - previous interest rate

Gauss-Markov in time series models

- Same Gauss-Markov assumptions as cross-sectional data
- Some slight tweaks:
- Replace A(2) [random sampling] with *no serial correlation*:

$$E(u_t u_s|\mathbf{X}) = 0 \text{ for all } t \neq s$$

- Homoskedasticity and $E(u|\mathbf{x})$ are same as before:

$$\begin{aligned} Var(u_t|\mathbf{X}) &= \sigma^2 \\ E(u_t|\mathbf{X}) &= 0 \end{aligned}$$

Serial correlation

- when $Corr(u_t, u_s) \neq 0$ for some (t, s) pairs
- say that errors exhibit **serial correlation** or **autocorrelation**
- No need to worry about it in cross-sectional data
- Why? $Corr(u_i, u_j) = 0$ in a random sample

How to test for serial correlation

- Run your regression of interest
- Extract residuals
- Run a regression of $\hat{u}_t$ on $\hat{u}_{t-1}, \ldots$
- Look at heteroskedasticity-robust t- and F-stats
- Practiced this in Lab 11 last time

Why do we care about serial correlation?

- Serial correlation, like heteroskedasticity, will invalidate hypothesis tests

- From a journal editor commenting on a paper I recently submitted:

> *I was surprised by the lack of serial correlation in wage shocks, especially given your analysis at the monthly level. This is almost certainly why you have such small standard errors—probably smaller than what they should be given there is likely a lot of serial correlation*

- Failure to correct for it may give you an incorrect conclusion

Correcting for Serial Correlation

Ways of correcting for serial correlation

1. Use FGLS with a weighting function that nullifies the serial correlation:
    - Cochrane-Orcutt or Prais-Winsten
    - These are now out-of-style because they rule out lagged DV models
    - Also only fix $AR(1)$ serial correlation (not more general forms)
2. Compute heteroskedasticity and autocorrelation robust SEs
    - a.k.a. “HAC standard errors” (most common: Newey-West)
3. Use fancier models of heteroskedasticity and autocorrelation (ARCH)

HAC standard errors

- Nowadays: treat serial correlation in TS models like heterosk. in CS models:

  as a nuisance that causes the usual inference to be incorrect

- For heteroskedasticity, use robust SEs:

  ```
  coeftest(est, vcov=hccm)
  ```

- We can do this same process for making serial correlation robust SEs

- serial correlation robust SEs also robust to heteroskedasticity

How to get HAC standard errors

- The underlying theory is complicated, but it is easy to describe the idea
- For example, allow u_t to be correlated with u_{t-1} and u_{t-2} ...

  but not the errors more than two periods apart
- Resulting SEs are usually called **Newey-West standard errors**
- a.k.a. **HAC (heteroskedasticity and autocorrelation consistent) SEs**

More about Newey-West SEs

- The N-W SEs are not as automated as the adjustment for heteroskedasticity
- We have to choose a lag before doing any estimation
- With annual data, the lag is short—maybe a couple of years, so $lag = 2$
- With quarterly or monthly data we tend to try longer lags, such as $lag = 24$

Newey-West SEs in R

- In R, the code is similar to correcting for heteroskedasticity:

```
coeftest(est, vcov=NeweyWest(est))
```

- R automatically chooses the *lag* for you according to a rule-of-thumb
- You can choose your own by specifying, e.g.

```
coeftest(est, vcov=NeweyWest(est, lag=3))
```