Econometrics

Week 4

Institute of Economic Studies
Faculty of Social Sciences
Charles University in Prague

Fall 2022

Recommended Reading

For today

- Serial correlation and heteroskedasticity in time series regressions
- Chapter 12

For the next week

- Pooling Cross Sections across Time: Simple Panel Data Methods
- Chapter 13

What have we learned so far?

- With time series data OLS has the same finite-sample properties as with cross-sectional data as long as a **stricter** set of CLM assumptions is satisfied
- The most restrictive assumption is **strict exogeneity**
- Strict exogeneity is often not satisfied, what results is biased OLS estimates
- OLS estimates on time series data can be **consistent** if all the involved stochastic processes are stationary and weakly dependent and we have at least **contemporaneous exogeneity** in the model.

OLS with Serially Correlated Errors: Unbiasedness and Consistency

- Note that when proving unbiasedness and consistency we do not assume anything about the error structure (the variance-covariance matrix of disturbances).
- As long as explanatory variables are strictly exogenous (and the model is linear, and there is no perfect colinearity between explanatory variables), $\hat{\beta}_{OLS}$'s are unbiased:
  - regardless the degree of serial correlation in disturbances,
  - regardless the degree of heteroskedasticity.
- As long as explanatory variables are contemporaneously exogenous ($E(u_t|\mathbf{x}_t) = 0$) and the involved variables are stationary and weakly dependent, $\hat{\beta}_{OLS}$'s are consistent (although not necessarily unbiased).
- So why do we need the assumption of no serial correlation?

OLS with Serially Correlated Errors: Efficiency and Inference

- Gauss-Markov Theorem requires both homoskedasticity and serially uncorrelated errors.
  - OLS is not BLUE under the presence of serial correlation.
  - Specifically, OLS is not efficient (is not the best among linear estimators)

- Moreover, variance of OLS estmates is misscalculated under the presence of serial correlation.
  - Both the small-sample variance and asymptotic variance.
  - We cannot use t-test or F-test for hypothesis testing.

Variance of OLS estimator

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

$$\begin{aligned} Var[\hat{\beta}_1] &= Var\left[\frac{\sum_{t=1}^{T}(x_t-\overline{x})u_t}{\sum_{t=1}^{T}(x_t-\overline{x})^2}\right] = \frac{Var\left[\sum_{t=1}^{T}(x_t-\overline{x})u_t\right]}{\left(\sum_{t=1}^{T}(x_t-\overline{x})^2\right)^2} = \\ &= \frac{1}{{SST_x}^2}Var\left[\sum_{t=1}^{T}(x_t-\overline{x})u_t\right] \end{aligned}$$

Variance of OLS estimator

For simplicity of notation, assume that $\overline{x} = 0$. Then:

$$Var(\hat{\beta}_1) = \frac{1}{{SST_x}^2} Var\left[\sum_{t=1}^{T} x_t u_t\right] = \frac{1}{{SST_x}^2} Var\left[x_1 u_1 + ... + x_T u_T\right]$$

$$= \frac{1}{{SST_x}^2}\left[\sum_{t=1}^{T} x_t^2 Var(u_t) + 2\sum_{t=1}^{T-1}\sum_{j=1}^{T-t} x_t x_{t+j} E[u_t u_{t+j}]\right]$$

$$= \underbrace{\sigma^2/SST_x}_{\text{est. variance of } \hat{\beta}_1} \quad \underbrace{+\, 2/SST_x^2 \sum_{t=1}^{T-1}\sum_{j=1}^{T-t} x_t x_{t+j} Cov[u_t u_{t+j}]}_{\text{bias}},$$

where $\sigma^2 = Var(u_t)$. If we ignore serial correlation (the fact that $Cov[u_t u_{t+j}] \neq 0$) and estimate the variance in the usual way, variance estimator is biased.

Most frequent form of serial correlation in errors is an autoregressive process

Let us assume a model with AR(1) errors

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

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

for $t = 1, 2, \ldots, T$, where $|\rho| < 1$ and $\{\epsilon_t\}$ is i.i.d. with zero mean and variance σ_ϵ^2.

$$\begin{aligned} Var(\hat{\beta}_1) &= \sigma^2/SST_x + 2/SST_x^2 \sum_{t=1}^{T-1} \sum_{j=1}^{T-t} x_t x_{t+j} Cov[u_t u_{t+j}] \\ &= \sigma^2/SST_x + 2(\sigma^2/SST_x^2) \sum_{t=1}^{T-1} \sum_{j=1}^{T-t} \rho^j x_t x_{t+j}, \end{aligned}$$

where $Cov[u_t, u_{t+j}] = \rho^j \sigma^2$ for integer $j \geq 1$ (see Lecture 3).

Properties of OLS with Serially Correlated Errors: Efficiency and Inference

Consequences:

- In most economic applications, $\rho > 0$ (and therefore $\rho^j > 0$) and the usual formula to estimate the variance of OLS estimate *underestimates* its true variance.

- We tend to think that OLS slope estimator is more precise than it actually is.

- Main consequence is that standard errors are invalid $\Rightarrow$ t-statistics for testing single hypotheses are invalid $\Rightarrow$ statistical inference is invalid.

Serial Correlation in Errors (Disturbances)

- It is important to know whether errors are serially correlated.
  - If they are - OLS is not BLUE (but still might be unbiased and consistent).
  - If they are - standard errors of $\hat{\beta}$'s computed in the usual way are invalid $\Rightarrow$ statistical inference is invalid.

- Before interpreting regression coefficients, test for serial correlation in errors!

- When serial correlation is detected:
  - Transform the model to get rid of serial correlation, OR
  - Keep the original model, but use different formula to compute standard errors of $\hat{\beta}$'s.

Example - Phillips Curve

$$inf_t = \beta_0 + \beta_1 unem_t + u_t$$

- There is AR(1) serial correlation in disturbances.
- Disturbances are heteroskedastic.

	OLS (usual st.err)	OLS (robust st. err)	FGLS
unem	0.468	0.468	0.716
	(0.076)	(0.112)	(0.072)
const	1.424	1.424	8.296
	(0.412)	(0.345)	(0.001)

p-values in parentheses

Asymptotic Test for AR(1) Serial Correlation

- It is useful to test for serial correlation in error terms:

$$y_t = \beta_0 + \beta_1 x_{t1} +, \ldots + \beta_k x_{tk} + u_t,$$

  with $u_t = \rho u_{t-1} + \epsilon_t$, $t = 2, \ldots T$.

- We would like to know whether $\rho = 0$ (no serial correlation) or $\rho \neq 0$ (AR(1) serial correlation).

- Remember, even with serial correlation, OLS estimates may be consistent!
  $\Rightarrow$ Residuals from the regression can be used to test the behavior of random disturbance.

Asymptotic Test for AR(1) Serial Correlation

- With **strictly exogenous** regressors, the test works as follows:

Testing for AR(1) Serial Correlation with Strictly Exogenous Regressors

(i) Run the OLS regression of y_t on $x_{t1}, \dots x_{tk}$ and obtain the OLS residuals, $\hat{u}_t$, for all $t = 1, \dots, T$.
(ii) Run the regression of $\hat{u}_t$ on $\hat{u}_{t-1}$, for all $t = 2, \dots, T$, obtaining the coefficient $\hat{\rho}$ on $\hat{u}_{t-1}$ and its t statistic $t_{\hat{\rho}}$.
(iii) Use $t_{\hat{\rho}}$ to test $H_0 : \rho = 0$ against $H_A : \rho \neq 0$ in the usual way.

Testing for AR(1) Serial Correlation

- In case we do not have strictly exogenous regressors (one or more x_{tj} is correlated with u_{t-1}), the t-test does not work.
- Why?
- We have to include all explanatory variables in the test:

Testing for AR(1) Serial Correlation with General Regressors

(i) Run the OLS regression of y_t on $x_{t1}, \ldots x_{tk}$ and obtain the OLS residuals, $\hat{u}_t$, for all $t = 1, \ldots, T$.
(ii) Run the regression of $\hat{u}_t$ on $x_{t1}, \ldots x_{tk}, \hat{u}_{t-1}$, for all $t = 2, \ldots, T$, obtaining the coefficient $\hat{\rho}$ on $\hat{u}_{t-1}$ and its t statistic $t_{\hat{\rho}}$.
(iii) Use $t_{\hat{\rho}}$ to test $H_0 : \rho = 0$ against $H_A : \rho \neq 0$ in the usual way.

- The inclusion of $x_{t1}, x_{t2}, \ldots x_{tk}$ explicitly allows each x_{tj} to be correlated with u_{t-1} $\Rightarrow$ no need for strict exogeneity.

Testing for Higher Order Serial Correlation

- We can easily extend the test for *second order* (AR(2)) serial correlation.
- In the model $u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \epsilon_t$, we test the

$$H_0 : \rho_1 = 0, \rho_2 = 0$$

- We regress $\hat{u}_t$ on $x_{t1}, x_{t2}, \ldots x_{tk}, \hat{u}_{t-1}, \hat{u}_{t-2}$ for all $t = 3, \ldots, T$.
- ...and obtain F test for *joint* significance of $\hat{u}_{t-1}$ and $\hat{u}_{t-2}$. If they are jointly significant, we reject the null $\Rightarrow$ errors are serially correlated of order two.

Testing for Higher Order Serial Correlation

- We can generally include q lags to test for high order serial correlation.

Testing for AR(q) Serial Correlation:

(i) Run the OLS regression of y_t on $x_{t1}, \ldots x_{tk}$ and obtain the OLS residuals, $\hat{u}_t$, for all $t = 1, \ldots, T$.
(ii) Run the regression of $\hat{u}_t$ on $x_{t1}, \ldots x_{tk}, \hat{u}_{t-1}, \hat{u}_{t-2}, \ldots, \hat{u}_{t-q}$, for all $t = (g+1), \ldots, T$, obtaining the coefficient $\hat{\rho}$ on $\hat{u}_{t-1}$ and its t statistic $t_{\hat{\rho}}$.
(iii) Compute the F test for joint significance of $\hat{u}_{t-1}, \hat{u}_{t-2}, \ldots, \hat{u}_{t-q}$.

- Or use LM version of test – **Breusch-Godfrey test**:

$$LM = (T-q)R^2_{\hat{u}},$$

where $R^2_{\hat{u}}$ is the usual R-squared from the regression above.

- Under the null hypothesis (which still is no serial correlation), $LM \overset{a}{\sim} \chi^2_q$.

Correcting for Serial Correlation

- When autocorrelation is detected, we need to treat it.
- We know that OLS is inefficient.
- So what is the efficient estimator in the AR(1) setting?
- We assume Gauss-Markov Assumptions TS1-TS4, but we relax TS5 and assume errors to follow AR(1): $u_t = \rho u_{t-1} + \epsilon_t$, for all $t = 1, 2, \ldots$
- $Var(u_t) = \sigma_\epsilon^2/(1-\rho^2)$.
- We need to transform the regression equation so that there is no serial correlation in the errors.
- How?

Correcting for Serial Correlation

Transform the data to eliminate the autoregressive component from the error.

$$y_t = \beta_0 + \beta_1 \cdot x_t + u_t \quad (1)$$
$$\text{with: } u_t = \rho u_{t-1} + \epsilon_t$$

$$y_{t-1} = \beta_0 + \beta_1 \cdot x_{t-1} + u_{t-1}$$
$$\rho y_{t-1} = \rho\beta_0 + \rho\beta_1 \cdot x_{t-1} + \rho u_{t-1} \quad (2)$$

Subtracting (2) from (1), we get for $t \geq 2$:

$$\underbrace{y_t - \rho y_{t-1}}_{\tilde{y_t}} = \beta_0 (1-\rho) + \beta_1 \underbrace{(x_t - \rho x_{t-1})}_{\tilde{x_t}} + \underbrace{u_t - \rho u_{t-1}}_{\epsilon_t \text{ - not autocorrelated}}$$

To make the estimates of the above regression BLUE, we need to add the lost observation for $t = 1$:

$$\left(1-\rho^2\right)^{1/2} y_1,\ \left(1-\rho^2\right)^{1/2} x_1$$

Correcting for Serial Correlation

- To make the estimates of the transformed regression BLUE, we need to add the lost observation for $t = 1$:

$$\left(1-\rho^2\right)^{1/2} y_1, \left(1-\rho^2\right)^{1/2} x_1$$

- Why is the first observation multiplied by $\left(1-\rho^2\right)^{1/2}$?

- To make sure that the variance of disturbance for the first observation is the same as the variance of disturbances for other observations! (see derivations on board)

Feasible GLS Estimation with AR(1) Errors

- The problem with the GLS estimation is that we do not know the value of ρ.
- But we already know how to obtain the estimate of ρ: Simply regress the OLS residuals on their lagged values and get $\hat{\rho}$.

Feasible GLS (FGLS) Estimation with AR(1) Errors

(i) Run the OLS regression of y_t on $x_{t1}, \ldots, x_{tk}$ and obtain the OLS residuals $\hat{u}_t$, $t = 1, 2, \ldots, T$.
(ii) Run the regression of $\hat{u}_t$ on $\hat{u}_{t-1}$ to obtain $\hat{\rho}$.
(iii) Run OLS equation:

$$\tilde{y}_t = \beta_0 \tilde{x}_{t0} + \beta_1 \tilde{x}_{t1} + \ldots + \beta_k \tilde{x}_{tk} + \epsilon_t,$$

where $\tilde{y}_t = y_t - \hat{\rho} y_{t-1}$, $\tilde{x}_t = x_t - \rho\hat{\rho} x_{t-1}$ for $t \geq 2$, and $\tilde{x}_1 = (1 - \hat{\rho}^2)^{1/2} x_1$, $\tilde{y}_1 = (1 - \hat{\rho}^2)^{1/2} y_1$ for $t = 1$.

Feasible GLS Estimation with AR(1) Errors

- GLS is BLUE under TS1 - TS5 and we can use t and F tests from the transformed equation for inference.
- Distributions conditional on $\mathbf{X}$ are exact (with minimum variance) if TS6 (TS6a) holds for ϵ_t.

- FGLS estimators are **not** unbiased, but are **consistent** when the regressors are strictly exogenous.
- Asymptotically, FGLS is more efficient than OLS.
- FGLS is close to first-differencing and thus can approximately eliminate unit roots.
- This method can be extended for higher order serial correlation, AR(q) in the error term.

- FGLS estimator is called the **Prais-Winsten estimator**.
- If we omit the first equation (for $t = 1$), it is called the **Cochrane-Orcutt estimator**.

Serial Correlation-Robust Standard Errors

- Problem: If the regressors are not strictly exogenous, FGLS is no longer consistent.
- Problem: If we do not model the AR process correctly, FGLS is no longer efficient.

- It's possible to calculate serial correlation (and heteroskedasticity) robust standard errors of OLS estimates. OLS will be consistent, but inefficient.

- The idea is to scale OLS standard errors to take into account serial correlation.
- It can be shown that the asymptotic variance of OLS estimator $\hat{\beta}_1$ is:

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

- Serial correlation robust standard errors can poorly behave in small samples in presence of large serial correlation.

Heteroskedasticity in Time Series Regressions

- Even if OLS estimators are **unbiased** and/or **consistent**, the OLS inference is **invalid**, if TS4 (or TS4a) (homoskedasticity) fails.
- Heteroskedasticity-robust statistics can be easily derived in the same manner as in cross-sectional data (if TS1a, TS2a, TS3a and TS5a hold).
- However, in small samples these robust standard errors may be large. $\Rightarrow$ We want to test for heteroskedasticity.
- We can use the same tests as in the cross-sectional case, but we need to have no serial correlation in the errors.
- Also for the Breusch-Pagan test where we specify $\hat{u_t}^2 = \delta_0 + \delta_1 x_{t1} + \ldots + \delta_k x_{tk} + \nu_t$ and test $H_0 : \delta_1 = \delta_2 = \ldots = \delta_k = 0$, we need $\{\nu_t\}$ to be **homoskedastic and serially uncorrelated**.
- If we find heteroskedasticity, we can use heteroskedasticity robust statistics.

Thank you

Thank you for your attention!

... and do not forget to read Chapter 13 for the next week!

Remember about the Home Assignment - Due November 2, 2022