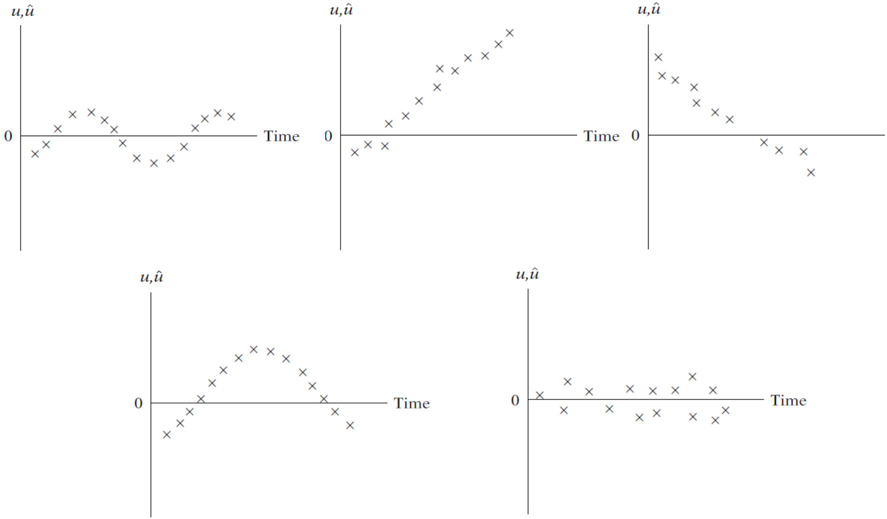
L9: Autocorrelation

Feng Li
feng.li@cufe.edu.cn

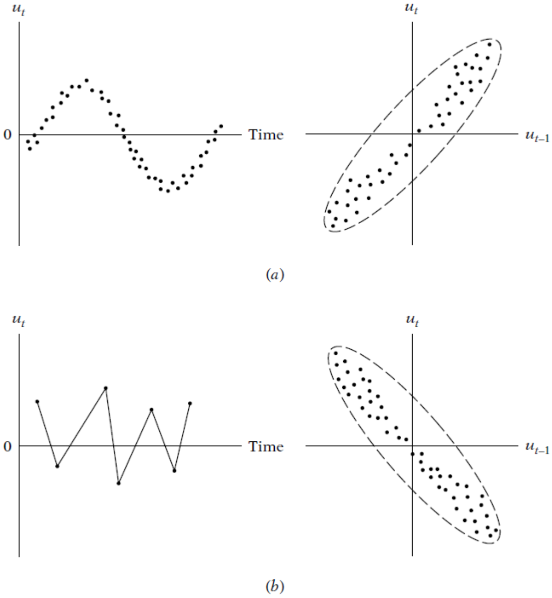
School of Statistics and Mathematics
Central University of Finance and Economics

Introduction

- In the classic regression model we assume $cov(u_i, u_j | x_i, x_k) = E(u_i, u_j) = 0$
- What if we break the assumption? Look at these patterns



Positive and negative autocorrelation



(*a*) Positive and (*b*) negative autocorrelation.

What happens to OLS/GLS when autocorrelation exists?

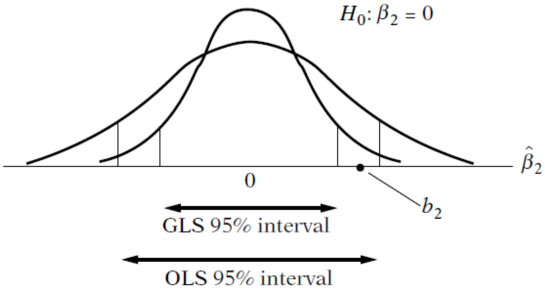
- Consider a simple model $Y_t = \beta_1 + \beta_2 X_t + u_t$ where the error terms is autocorrelated as $u_t = \rho u_{t-1} + \epsilon_t$
- The OLS estimator will not change (still linear unbiased)
- But is not efficient since

$$Var\left(\hat{\beta}_2\right) = \frac{\sigma^2}{\sum x_t^2}\left(1 + 2\rho\frac{\sum x_t x_{t-1}}{\sum x_t^2} + 2\rho^2\frac{\sum x_t x_{t-2}}{\sum x_t^2} + ...\right) > \frac{\sigma^2}{\sum x_t^2}$$

  where $\sigma^2/\sum x_t^2$ is the variance of $\hat{\beta}_2$ when no autocorrelation presents.
- The GLS estimators under autocorrelation is BLUE.

Consequences of using OLS when autocorrelation

- The confidence intervals are likely wider than those from GLS.

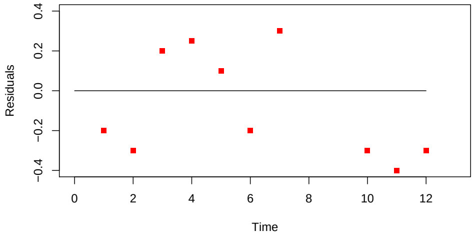


- The usual t and F tests are not valid.
- The residual variance $\hat{\sigma}^2$ likely to underestimate the true σ^2
- Likely to overestimate R^2

Detection of autocorrelation

↪ The runs test (nonparametric test)

- **step 1:** Plot the residuals as follows



- **step 2:** Count the runs as: (- -) (+ + +) (-) (+) (- - -)
- **step 3:** N_1 = number of "+" residuals = 4,
  N_2 = number of "-" residuals = 6, $N = N_1 + N_2 = 10$,
  R = number of runs = 5
- **step 3:** The number of runs is normally distributed

$$R \sim N\left(\frac{2N_1N_2}{N} + 1,\ \frac{2N_1N_2(2N_1N_2 - N)}{N^2(N+1)}\right)$$

Detection of autocorrelation

↪ The runs test – example

- Carry out the runs test with the residuals give in the previous slides.
- What critical value are you look for?
- Why this test is called nonparametric test?

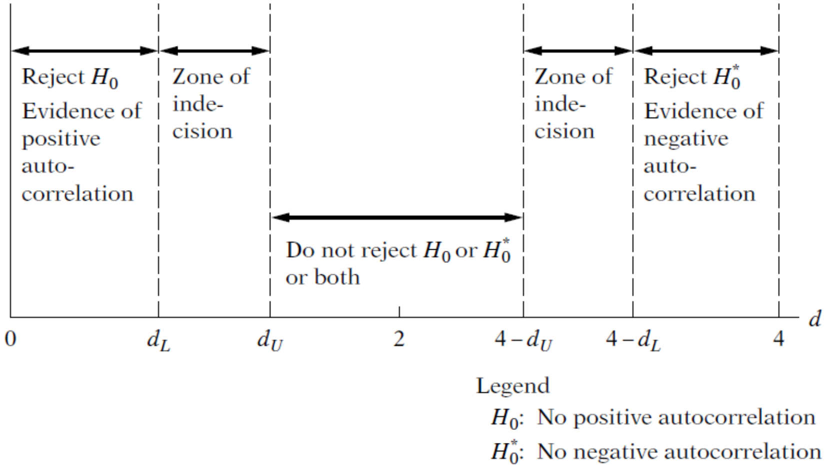
Detection of autocorrelation
↪ Durbin–Watson *d* test (1)

- The Durbin–Watson *d* statistic

$$d = \frac{\sum_{t=2}^{n} (\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^{n} \hat{u}_t^2}$$

- Assumptions underlying the *d* statistic
  - The regression model includes the intercept. If you model does not have the intercept, rerun the model include intercept to obtain $\hat{u}_i$.
  - It can only determine first-order autoregressive scheme.
  - The error term are assumed normally distributed.
  - The explanatory variables do not contains lagged values (we will talk about it more in the time series part).

Detection of autocorrelation
↪ Durbin–Watson *d* test (2)



Durbin–Watson *d* statistic.

Detection of autocorrelation
↬ Durbin–Watson *d* test (3)

- Approximation of *d* statistic

$$d \approx 2(1 - \hat{\rho})$$

where $\hat{\rho} = \frac{\sum \hat{u}_t \hat{u}_{t-1}}{\sum \hat{u}_t^2}$ is the sample first-order coefficient of autocorrelation of $\hat{u}_t$, i.e.,

$$\hat{u}_t = \hat{\rho} u_{t-1}.$$

- When $\rho \to 1$, positive autocorrelation;
- When $\rho \to -1$, negative autocorrelation;
- When $\rho \to 0$, no autocorrelation.

Detection of autocorrelation
↛ Durbin–Watson *d* test – example

- Given a sample of 100 observations and 5 explanatory variables, what can you say about autocorrelation if $\mathrm{d} = 1.2$?
- Can handle without looking at the table?

The comparison between Durbin–Watson test and runs test

- Runs test does not require and probability distribution of the error term.
- **Warning**: The *d* test is not valid if u_i is not iid.
- When n is large,

$$\sqrt{n}(1 - d/2) \sim N(0, 1)$$

- We can use the normality approximation when n is large regardless of iid.
- Durbin–Watson statistic requires the covariates be non stochastic which is difficult to meet in econometrics.
- In this case, try the test on next slides.

Detection of autocorrelation

↪ The Breusch–Godfrey test (Lagrange Multiplier test)

- Consider the model $Y_t = \beta_1 + \beta_2 X_t + u_t$,
- Assume $u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + ... + \rho_p u_{t-p} + \epsilon_t$
- $H_0 : \rho_1 = \rho_2 = ... = \rho_p = 0$, i.e. no autocorrelation.
- Run the auxiliary model

$$\hat{u}_t = \alpha_1 + \alpha_2 X_t + \rho_1 \hat{u}_{t-1} + \rho_2 \hat{u}_{t-2} + ... + \rho_p \hat{u}_{t-p} + \epsilon_t$$

  and obtain R^2.
- When n large,

$$(n - p)R^2 \sim \chi^2(p).$$

- Reject H_0 if $\chi^2_{obs}(p) > \chi^2_{crit}(p)$.
- Question: Have you seen similar another test that has the similar way of constructions as this one?

Model misspecification and pure autocorrelation

- Some variables that were supposed to be in the model but not.
- This is the case of *excluding variable*, which is a type of model specification bias (will talk more in next chapter).
- This may also show patterns in the residuals plot.
- Try to find out if it is pure autocorrelation or misspecification.
- See **example on p. 441**.

Use GLS to correct pure autocorrelation

- Assume you have model $Y_t = \beta_1 + \beta_2 X_t + u_t$ and
- there is pure autocorrelation $u_t = \rho u_{t-1} + \epsilon_t$.
- **When ρ known**
  - From the model we also have $Y_{t-1} = \beta_1 + \beta_2 X_{t-1} + u_{t-1}$
  - Then $Y_t - \rho Y_{t-1} = \beta_1(1-\rho) + \beta_2(X_t - \rho X_{t-1}) + (u_t - \rho u_{t-1})$. Why?
  - The above model can be written as $Y_t^* = \beta_1^* + \beta_2^* X_t^* + \epsilon_t$ which removes autocorrelation. Why?
- **When ρ not known**
  - One may run the model $Y_t - Y_{t-1} = \beta_2(X_t - X_{t-1}) + (u_t - u_{t-1})$
  - We will talk more at second part of this course.

Take home questions

- **12.2, 12.11, 12.12**
- How do you implement GLS in matrix form [Hint: think about the Ω matrix]?