労働経済 I 期末レポート

経済学研究科 23A18014 丹治伶峰

1.
2.
3. (1)
   - Ability bias
   - Selection bias

   (2)
   - Ability bias

     Optimal wage depending on the worker's ability θ is :

$$
\begin{aligned}
w(\theta) &= Ah_0 e^{\theta x(\theta)} \\
&\text{where} \\
h_0 &: \text{initial human capital} \\
x(.) &: \text{optimal length of education as a function of } \theta
\end{aligned}
$$

     When θ goes up, there are two different effects : direct effect on $w(\theta)$ and indirect effect, that raises $x(\theta)$, and then affect on $w(\theta)$.
     Then, assume simplified form regression on loggarithm wage w_i by education length t_i,

$$\ln w_i = \beta_0 + \beta_1 t_i + \epsilon_i$$

     t_i and error term ϵ_i are positively correlated, which causes bias by OLS.
   - Selection bias

     Consider decision-making whether to go to college or not.
     Define B^i be the education effect for type $i \in \{C, H\}$ individual, then we would like to specify

$$
\begin{aligned}
B^C &= E^C_C - E^C_H \\
B^H &= E^H_C - E^H_H
\end{aligned}
$$

Note that E is return of education, superscript standing for the type of individual, and subscript denoting her/his actual choice (: C is to go to college, while H not).

By observed sample, however, the "appeared" effect B is :

$$B = E_C^C - E_H^H$$

since type C individual usually go to school and vice versa.

Again, assume the wage regression

$$\ln w_i = \beta_0 + \beta_1 t_i + \epsilon_i$$

If $E_H^C < E_H^H$, where the earnings of job for high school would have been lower for type C than for type H, then t_i and ϵ_i are negatively correlated, and so B^C is underestimated.

Similarly, if $E_C^H < E_C^C$, then t_i and ϵ_i are positively correlated, so B^H is overestimated.

(3)

Exclusive instruments should have explanatory power to the independent variable you are interested in, but be independent of the error term of the original regression model.

Consider the regression model

$$\ln y_i = \beta_0 + \beta_1 t_i + \beta_2 + x_i + \epsilon$$

Then, excluded variable Z_i is introduced to the following regression :

$$t_i = \gamma_0 + \gamma_1 Z_i + \gamma_2 + e_i$$

For the excluded variable being valid, two conditions below are required :

- γ_1 is statistically significant.
- Z_i and ϵ_i are independent