Lecture 10 - Economics 113

Professor Spearot

- **Agenda**
  1. Bias Examples
  2. Other Examples

- **Exam Reminders**
  1. No graphing calculators or notes
  2. Be on time!!!
Multivariate Regression

Omitted variable bias

- What happens when we omit an important variable?
- Need to conjecture regarding the relationship between the omitted variable and included $x$ and $y$ variables

**The Table**

<table>
<thead>
<tr>
<th>Corr(omitted variable, $x$)</th>
<th>Corr(omitted variable, $y$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>positive</td>
<td>positive</td>
</tr>
<tr>
<td>negative</td>
<td>negative</td>
</tr>
<tr>
<td>upward bias</td>
<td>downward bias</td>
</tr>
<tr>
<td>downward bias</td>
<td>upward bias</td>
</tr>
</tbody>
</table>

- **Upward bias:**
  - Estimate is *higher* than the true parameter: $\beta < \hat{\beta}$
- **Downward bias:**
  - Estimate is *lower* than the true parameter: $\hat{\beta} < \beta$
Multivariate Regression

Omitted variable bias - Examples

- Example: Effect of class attendance on grades
- Population follows:

\[
\text{final} = \beta_0 + \beta_1 \text{attend} + \beta_2 \text{study} + u
\]

- We instead forget about \textit{study} and estimate:

\[
\hat{\text{final}} = \hat{\beta}_0 + \hat{\beta}_1 \text{attend}
\]

- Suppose we estimate \(\hat{\beta}_1 > 0\), and conclude that attendance increases your grade (\(\beta_1 > 0\)). Is this right?
- Positive correlation between \textit{study} and \textit{final}
- Positive correlation between \textit{study} and \textit{attend}
- \(\hat{\beta}_1\) suffers from an upward bias. \(\beta_1 < \hat{\beta}_1\)
Multivariate Regression

Omitted variable bias - Examples

- Intuition
  - $\hat{\beta}_1 > 0$ suggests that higher attendance improves your grade
  - However, students who attend class often tend to study more
  - Thus, attend may actually be accounting for the effects of studying, and not attendance.

- Overall, given $\beta_1 < \hat{\beta}_1$, the result $\hat{\beta}_1 > 0$ is insufficient to guarantee that $\beta_1 > 0$. 
Multivariate Regression

Omitted variable bias - Examples

- Example: Effect of drugs on crime
- Population follows:

\[
\text{crime} = \beta_0 + \beta_1 \text{educ} + \beta_2 \text{drugs} + u
\]

- We instead forget about drugs and estimate:

\[
\hat{\text{crime}} = \hat{\beta}_0 + \hat{\beta}_1 \text{educ}
\]

- Suppose we estimate \(\hat{\beta}_1 < 0\), and conclude education reduces your likelihood of committing a crime \((\beta_1 \leq 0)\)
- Positive correlation between drugs and crime
- Negative correlation between drugs and educ
- \(\hat{\beta}_1\) suffers from an downward bias. \(\hat{\beta}_1 < \beta_1\)
Multivariate Regression

Omitted variable bias - Examples

- Intuition
  - \( \hat{\beta}_1 < 0 \) suggests that education reduces your likelihood of committing a crime
  - However, people who go to school are less likely to abuse drugs
  - Thus, educ may actually be accounting for the propensity of drug use, not the effects of education

- Overall, given \( \hat{\beta}_1 < \beta_1 \), the result \( \hat{\beta}_1 < 0 \) is insufficient to guarantee that \( \beta_1 < 0 \).
Example: Effect of graduate education on wages

Population follows:

\[ \log(wage) = \beta_0 + \beta_1 geduc + \beta_2 Exper + u \]

We instead forget about \( Exper \) and estimate:

\[ \log(wage) = \hat{\beta}_0 + \hat{\beta}_1 \log(geduc) \]

Suppose we estimate \( \hat{\beta}_1 > 0 \), and conclude that graduate education increases your wage \( (\beta_1 > 0) \)

Positive correlation between \( Exper \) and \( \log(wage) \)

Negative correlation between \( Exper \) and \( geduc \) (by construction)

\( \hat{\beta}_1 \) suffers from a downward bias. \( \hat{\beta}_1 < \beta_1 \)
Multivariate Regression

Omitted variable bias - Examples

- **Intuition**
  - $\hat{\beta}_1 > 0$ suggests that graduate education of some sort increases your wage
  - However, people who pursue graduate education have lower levels of experience
  - Thus, people with no graduate education may earn relatively high wages since they have lots of experience.
  - Overall, given $\hat{\beta}_1 < \beta_1$, the result $\hat{\beta}_1 > 0$ is *sufficient* to guarantee that $\beta_1 > 0$. 