# Returns to Education

January 10, 2019

# Markets in Education

Education is a unique marketplace:

- It is not perfectly competitive (e.g. large fixed costs, imperfect information);
- Spillovers or externalities are quite common;
- Supply is often local(ish);
- Education is highly differentiated, both vertically (age) and horizontally (different quality/classes for individuals who are the same age);
- The quality of your education (might) depend on the characteristics of other consumers (i.e. your peers).

That means that we can use economic tools to answer questions about education policy, but we need to go beyond the simplest models of perfect competition. Although plenty of policy-makers and commentators focus on trying to get the market more toward their perfect competition ideal.

# Education as an Investment

- One of the earliest (and most important) questions in the economic of eduction: what are the returns to eduction?
  - For now, we will focus on the individual returns to eduction as opposed to the social returns to eduction.
- What do we mean by the "returns to eduction?"
- Why do we care about the individual returns to eduction?

# Education as an Investment

- One of the earliest (and most important) questions in the economic of eduction: what are the returns to eduction?
  - For now, we will focus on the individual returns to eduction as opposed to the social returns to eduction.
- What do we mean by the "returns to eduction?"
- Why do we care about the individual returns to eduction?

# Education as an Investment

- One of the earliest (and most important) questions in the economic of eduction: what are the returns to eduction?
  - For now, we will focus on the individual returns to eduction as opposed to the social returns to eduction.
- What do we mean by the "returns to eduction?"
- Why do we care about the individual returns to eduction?

# Education as an Investment

- One of the earliest (and most important) questions in the economic of eduction: what are the returns to eduction?
  - For now, we will focus on the individual returns to eduction as opposed to the social returns to eduction.
- What do we mean by the "returns to eduction?"
- Why do we care about the individual returns to eduction?

# Estimating the Individual Returns to Eduction

- Let's suppose that you ran an OLS regression:

$$ln(Y_i) = \beta_0 + \beta_1 Education_i + \epsilon_i \qquad (1)$$

- Does $\beta_1$ define the individual returns to education?
- If not, is $\beta_1$ larger or smaller than the returns to eduction?

# Estimating the Individual Returns to Eduction

- Let's suppose that you ran an OLS regression:

$$ln(Y_i) = \beta_0 + \beta_1 Education_i + \epsilon_i \quad (1)$$

- Does $\beta_1$ define the individual returns to education?

- If not, is $\beta_1$ larger or smaller than the returns to eduction?

# Estimating the Individual Returns to Eduction

- Let's suppose that you ran an OLS regression:

$$ln(Y_i) = \beta_0 + \beta_1 Education_i + \epsilon_i \tag{1}$$

- Does $\beta_1$ define the individual returns to education?
- If not, is $\beta_1$ larger or smaller than the returns to eduction?

# Omitted Variable Bias

- Let's assume that an individual's outcome $Y_i$ is a function of her years of education ($Ed_i$) and ability ($Ability_i$)

$$Y_i = \beta_0 + \beta_1 Ed_i + \beta_2 Ability_i + \epsilon \tag{2}$$

- What happens if we just estimate our original model, omitting ability?

$$Y_i = \gamma_0 + \gamma_1 Ed_i + \mu \tag{3}$$

- Using our formula for a regression coefficient $E(\hat{\gamma_1}) = \frac{cov(Y, Ed_i)}{var(Ed_i)}$, and substituting $\alpha + \beta_1 Ed_i + \beta_2 Ability_i + \epsilon$ for $Y_i$, we get:

$$E(\hat{\gamma_1}) = \frac{cov(\alpha + \beta_1 Ed_i + \beta_2 Ability_i + \epsilon, Ed_i)}{var(Ed_i)} \tag{4}$$

# Omitted Variable Bias

- Let's assume that an individual's outcome $Y_i$ is a function of her years of education ($Ed_i$) and ability ($Ability_i$)

$$Y_i = \beta_0 + \beta_1 Ed_i + \beta_2 Ability_i + \epsilon \tag{2}$$

- What happens if we just estimate our original model, omitting ability?

$$Y_i = \gamma_0 + \gamma_1 Ed_i + \mu \tag{3}$$

- Using our formula for a regression coefficient $E(\hat{\gamma_1}) = \frac{cov(Y, Ed_i)}{var(Ed_i)}$, and substituting $\alpha + \beta_1 Ed_i + \beta_2 Ability_i + \epsilon$ for $Y_i$, we get:

$$E(\hat{\gamma_1}) = \frac{cov(\alpha + \beta_1 Ed_i + \beta_2 Ability_i + \epsilon, Ed_i)}{var(Ed_i)} \tag{4}$$

# Omitted Variable Bias

- Let's assume that an individual's outcome $Y_i$ is a function of her years of education ($Ed_i$) and ability ($Ability_i$)

$$Y_i = \beta_0 + \beta_1 Ed_i + \beta_2 Ability_i + \epsilon \tag{2}$$

- What happens if we just estimate our original model, omitting ability?

$$Y_i = \gamma_0 + \gamma_1 Ed_i + \mu \tag{3}$$

- Using our formula for a regression coefficient $E(\hat{\gamma_1}) = \frac{cov(Y, Ed_i)}{var(Ed_i)}$, and substituting $\alpha + \beta_1 Ed_i + \beta_2 Ability_i + \epsilon$ for $Y_i$, we get:

$$E(\hat{\gamma_1}) = \frac{cov(\alpha + \beta_1 Ed_i + \beta_2 Ability_i + \epsilon, Ed_i)}{var(Ed_i)} \tag{4}$$

## Omitted Variable Bias

$$= \frac{cov(\beta_0, Ed_i)}{var(Ed_i)} + \beta_1 + \beta_2 \frac{cov(Ability_i, Ed_i)}{var(Ed_i)} + \frac{cov(\epsilon, Ed_i)}{var(Ed_i)} \tag{5}$$

- The first term is zero since $\beta_0$ is a constant and the last term is also zero given the main OLS assumptions ($cov(\epsilon, X_i) = 0$)
- This leaves us with the omitted variable bias formula:

$$E(\hat{\gamma_1}) = +\beta_1 + \beta_2 \frac{cov(Ability_i, Ed_i)}{var(Ed_i)} \tag{6}$$

- Note that the last term $\frac{cov(Ability_i, Ed_i)}{var(Ed_i)}$ is the coefficient from a bivariate regression of $Ability_i$ on $Ed_i$
- OVB is driven by factor(s) that are more correlated with your independent variables of interest and your dependent variable.

# Omitted Variable Bias

$$= \frac{cov(\beta_0, Ed_i)}{var(Ed_i)} + \beta_1 + \beta_2 \frac{cov(Ability_i, Ed_i)}{var(Ed_i)} + \frac{cov(\epsilon, Ed_i)}{var(Ed_i)} \quad (5)$$

- The first term is zero since $\beta_0$ is a constant and the last term is also zero given the main OLS assumptions ($cov(\epsilon, X_i) = 0$)
- This leaves us with the omitted variable bias formula:

$$E(\hat{\gamma_1}) = +\beta_1 + \beta_2 \frac{cov(Ability_i, Ed_i)}{var(Ed_i)} \quad (6)$$

- Note that the last term $\frac{cov(Ability_i, Ed_i)}{var(Ed_i)}$ is the coefficient from a bivariate regression of $Ability_i$ on $Ed_i$
- OVB is driven by factor(s) that are more correlated with your independent variables of interest and your dependent variable.

# Omitted Variable Bias

$$= \frac{cov(\beta_0, Ed_i)}{var(Ed_i)} + \beta_1 + \beta_2 \frac{cov(Ability_i, Ed_i)}{var(Ed_i)} + \frac{cov(\epsilon, Ed_i)}{var(Ed_i)} \tag{5}$$

- The first term is zero since $\beta_0$ is a constant and the last term is also zero given the main OLS assumptions ($cov(\epsilon, X_i) = 0$)
- This leaves us with the omitted variable bias formula:

$$E(\hat{\gamma_1}) = +\beta_1 + \beta_2 \frac{cov(Ability_i, Ed_i)}{var(Ed_i)} \tag{6}$$

- Note that the last term $\frac{cov(Ability_i, Ed_i)}{var(Ed_i)}$ is the coefficient from a bivariate regression of $Ability_i$ on $Ed_i$
- OVB is driven by factor(s) that are more correlated with your independent variables of interest and your dependent variable.

# Approaches to Estimating the Returns to Education

- There are (roughly) six approaches to estimating the returns:

1. Bound the returns to education
2. Selection-on-observables
3. Structural models
4. Twin studies
5. Quasi-experimental evidence
6. Random Control Trials (RCTs)

# Approaches to Estimating the Returns to Education

- There are (roughly) six approaches to estimating the returns:

1. Bound the returns to education
2. Selection-on-observables
3. Structural models
4. Twin studies
5. Quasi-experimental evidence
6. Random Control Trials (RCTs)

# Approaches to Estimating the Returns to Education

- There are (roughly) six approaches to estimating the returns:

1. Bound the returns to education
2. Selection-on-observables
3. Structural models
4. Twin studies
5. Quasi-experimental evidence
6. Random Control Trials (RCTs)

# Approaches to Estimating the Returns to Education

- There are (roughly) six approaches to estimating the returns:

1. Bound the returns to education
2. Selection-on-observables
3. Structural models
4. Twin studies
5. Quasi-experimental evidence
6. Random Control Trials (RCTs)

# Approaches to Estimating the Returns to Education

- There are (roughly) six approaches to estimating the returns:

1. Bound the returns to education
2. Selection-on-observables
3. Structural models
4. Twin studies
5. Quasi-experimental evidence
6. Random Control Trials (RCTs)

# Approaches to Estimating the Returns to Education

- There are (roughly) six approaches to estimating the returns:

  1. Bound the returns to education
  2. Selection-on-observables
  3. Structural models
  4. Twin studies
  5. Quasi-experimental evidence
  6. Random Control Trials (RCTs)

# Approaches to Estimating the Returns to Education

- There are (roughly) six approaches to estimating the returns:

1. Bound the returns to education
2. Selection-on-observables
3. Structural models
4. Twin studies
5. Quasi-experimental evidence
6. Random Control Trials (RCTs)