

## **Estimating the returns to education**

**Earnings regression – Jacob Mincer, reduced form equation**

**Log (real wage) = f (age, age<sup>2</sup>, years of education, years of education<sup>2</sup>, degrees attained, gender, race, marital status, # of dependent children, region, industry)**

**Estimated relationship – increased earnings over the life-cycle associated with education**

**Distinct increase associated with earning a degree**

## **Interpretations**

- 1) productivity effect – years coefficient reflects productivity**
- 2) omitted variable effect – years is correlated to inherent productivity which is omitted, coefficient on years is biased upwards**

## Omitted Variable Bias

Suppose that the TRUE model is

$$Y = a + bX + b_1ED + b_2AB + e$$

And that there is a relationship between ED and AB

$$AB = q + j ED + u$$

If we estimate

$$Y = a + b_1X + b_2ED + v$$

It is equivalent to estimating:

$$Y = a + b_1X + b_2ED + b_3(q + j ED + u) + v$$

Or

$$Y = (a + b_3q) + b_1X + (b_2 + b_3j) ED + (b_3v + u)$$

Or

$$Y = d_0 + d_1X + d_2ED + W$$

**In general  $E(d_2) = b_1 + b_3(\text{SEDAB}/\text{SED}^2)$**

**In other words unless  $\text{Corr}(\text{AB}, \text{ED}) = 0$  the estimated coefficient is biased. In effect, the estimated coefficient on education is partly an effect of education and partly an effect of ability.**

## **How do you control for inherent productivity?**

### **1) Direct controls measuring “productivity”**

**IQ**

**GCSE maths, english results**

**Beauty**

**Leadership skills**

**Results – productivity variables always significant, education still has a positive (though reduced effect)**

**Problems – may not fully control for productivity, availability of data**

**2) twins studies (fixed effects) – inherent productivity is a function of genetics and family environment which are unobservable, but the same for twins.**

**Econometric strategy - Use difference-in-difference to eliminate effect of genetics and family.**

**Results – returns to education reduced by about 50%**

**Problem – estimates identified by cases where there are large differences between twins (e.g. accidents)**

**3) Natural experiments – one group of individuals receives more education than an otherwise similar group, and participates in the same labour market.**

**Examples: Draft lotteries, changes in length of the school year**

**Evidence: A substantial proportion of the return to schooling is due to inherent productivity (perhaps half)**

**Problem: estimates are very “local” (one group averages 11.7 years ED the other 12.0 years)**

**How do you test for signalling?**

**1) returns to degree – if completion of the third year of University has a much higher rate of return than the second**

**But – need control for selection in completion of the degree (worst students drop out)**

**Compare degree holders to exogenous drop outs  
drafted  
degree program was cancelled  
change in family finances**

**2) Compare the employed to the self employed –  
don't need to signal to yourself that you are  
qualified**