Natural "Natural Experiments" in Economics

Rosenzweig and Wolpin

2000

Natural Experiments

- Natural Experiments: Natural outcomes, which are plausibly random with respect to the two major sources of heterogeneity: tastes and abilities.
- Examples: twin births, human cloning (monozygotic twins), birth date, gener, and weather events.
- They are used as instruments for returns to schooling, labor supply sensitivity to temporary and permanent changes in income, women's labor force participation to fertility change.

- Randomness of the instrument and the explanatory power is not the only condition for the identification of the structural parameters of interest.
- In many studies using natural experiments, implicit model assumptions are made, which are not always plausible.

Natural Experiment Model

$$Y = \alpha + \beta X + \epsilon$$

Instrumental variable estimator

$$\beta_{IV} = \frac{cov(Y, Z)}{cov(X, Z)} = \frac{cov(X, Z)\beta + cov(\epsilon, Z)}{cov(X, Z)}$$

Example: Returns to Schooling, Work Experience

The returns to schooling equation

$$lny_a = f(S, \mu) + g(X_a, \mu)$$

S: schooling
X_a: work experience at age a
y_a: earnings.
μ: parameters

The present value of attending school at age 0:

$$V_1(s_1 = 1|S_0) = exp\left[f(S_0 + 1, \mu)\right] + \sum_{a=0}^{A-1} \beta^{a+1} exp\left[g(a, \mu)\right] - c$$

The present value of not attending school:

$$V_1(s_1 = 0|S_0) = exp\left[f(S_0, \mu)\right] + \sum_{a=0}^{A-1} \beta^{a} exp\left[g(a, \mu)\right]$$

Then,

$$s_1 = 1 \quad \text{if} \quad f(S_0 + 1, \mu) - f(S_0, \mu) \geq r + ln\left[\frac{c}{V_1(s_1 = 0|S_0)} + 1\right]$$

$$s_1 = 0 \quad \text{otherwise}$$

Ability Bias

If ability increases the marginal return to schooling,

$$\frac{\partial \left[f(S_0+1,\mu)-f(S_0,\mu)\right]}{\partial \mu} > 0$$

then, only individuals wigh high ability attend schooling and those with low ability do not. Simple regression where the independent variable is years of schooling overstates the returns to schooling.

Angrist and Krueger:Returns to Education

Variation in date of birth as a natural experiment.

- Fixed school entry date and minimum school leaving age create an instrument. Students born later than the school entry have to stay at school longer.
- No need for the state level variation in school leaving age regulation.
- High ability types stay at school longer than S_0 years.
- Low ability types with birth date that are subject to the regulation stay for $S_0 + 1$ years, while low ability types with birth date that makes them not subject to regulation stay for S_0 years. IV:

$$\frac{E\left[lny_a|A_r\right] - E\left[lny_a|A_n\right]}{Prob(S_0|A_n)} = \frac{f(S_0+1,\mu_1) - f(S_0,\mu_1)}{Prob(\mu_1)} \tag{1}$$

A_r: birth dates subject to regulation.

However, change in years of schooling also changes years of postschooling experience. Assume that individuals either attend school or work every year. For those born earlier than the school start date,

$$
\begin{aligned}
E\left[lny_a | A_n\right] &= \pi_1\left[f(S_0+1,\mu_1)+g(a-a_K-1,\mu_1)\right] \\
&\quad +(1-\pi_1)\left[f(S_0,\mu_2)+g(a-a_K,\mu_2)\right]
\end{aligned}
$$

High ability type 1 attends school one year more and has one year less experience, given age. Low ability type 2 attends school one year less and has one year more experience.
Those born later than the school start date are subject to the regulation. They also attend school at earlier age.

$$
\begin{aligned}
E\left[lny_a | A_r\right] &= \pi_1\left[f(S_0+1,\mu_1)+g(a-a_K,\mu_1)\right] \\
&\quad +(1-\pi_1)\left[f(S_0+1,\mu_2)+g(a-a_K,\mu_2)\right]
\end{aligned}
$$

Then,

$$E\left[lny_a|A_r\right]-E\left[lny_a|A_n\right] = \pi_1\left[g(a-a_K,\mu_1)-g(a-a_K-1,\mu_1)\right] \\ +(1-\pi_1)\left[f(S_0+1,\mu_2)-f(S_0,\mu_2)\right]$$

Hence, the returns to schooling is overstate by the returns to experience of type 1.

Angrist Draft Lottery Example

Use Mincer equation to estimate returns to experience. Does not have schooling data.

$$\begin{aligned} lny_a &= \alpha_0 + \alpha_1 S_a + \alpha_2 X_a + u_a \\ &= \alpha_0' + \alpha_2(a - m) + u_a' \end{aligned}$$

where civilian experience is

$$X_a = a - S_a - m - a_e$$

where a_e is the school entry age. Then, without schooling data,

$$u_a' = (\alpha_1 - \alpha_2)S_a + u_a$$

- Draft lottery determining military service could be correlated with schooling level (GI bill). In that case, Angrist argues that one could interpret the results as the returns to postmilitary.
- There is additional effect of military draft on schooling: some individuals may be drafted before schooling completion. Then, the effect of military service could be due to the reduced schooling, not due to reduced experience. experience and schooling.

Endogeneity of Experience

Suppose in the Angrist Krueger returns to schooling example, experience also is affected by ability.

$$
\begin{aligned}
lny_a &= \alpha_1 S_a + \alpha_2 X_a + \mu + \epsilon_a \\
S &= \beta_1 a_e(q) + \beta_2 \mu + \epsilon_S \\
X_a &= \gamma_1 S + \gamma_2 \mu + \epsilon_X
\end{aligned}
$$

Suppose ϵ_a, ϵ_S, ϵ_X are independent. Then, if experience is also a function of ability $\gamma_2 \neq 0$, the returns to schooling parameter cannot be estimated consistently even with the age of birth instrument q. There is only one instrument even though there are two endogenous variables, S_a and X_a.

The Natural Human Cloning Experiment

Use within pair differences of “identical twins” as instruments. They have the same innate ability μ. Any difference in schooling, etc. between those twince are uncorrelated to their ability, and thus can be used as instruments.

$$\begin{aligned} lny_{ij} &= \alpha_1 S_{ij} + \alpha_2 X_{ij} + \mu_j + \epsilon_{ij} \\ S_{ij} &= b_s \mu_j + f_{Sj} + u_{S,ij} \\ X_{ij} &= b_X \mu_j + f_{Xj} + u_{X,ij} \end{aligned}$$

for identical twin i in family j.
Differencing within twins in a family removes ability term. Any within twins difference in schooling and experience are orthogonal to the ability, and thus are valid instruments.

- Using twins differencing does not remove the bias due to correlation between the error terms $u_{S,ij}$, $u_{X,ij}$ and ϵ_{ij}. Unobservables that affect schooling, experience and earnings may not be solely to genetic origin.
- They may receive different random wage offers. The average (accepted) wage for twins with lower schooling will therefore overstate the average (offered) wage and thus the return to schooling will be understated.
- Even though any post-birth outcome can be used as potential instruments, they are only orthogonal to ability shock.

- In general, decision process of individuals are comples, and thus the outcome varible of interest typicaly depends on variables of which several are endogenous.
- In most cases of interest, the number of plausible natural experiments are not enough to deal with all the potential endogeneities. The conventional approach of natural experiments make very strong model assumptions to make it feasible to estimate the outcome equations with insufficient number of natural experiments.
- Those results, even though they seem to require less parametric assumptions, are in no ways less restrictive than the parameter estimates based on a fully specified dynamic model of individual decisions, and often are less realistic.