Lectures on Identification 3

Andrew Chesher

CeMMAP & UCL

April 21st 2008

Topics

1. Monday April 14th. Motivation, history, definitions, types of model, parametric and nonparametric IV models.

2. Wednesday April 16th. Quantiles and a non-additive IV model. Continuous endogenous variables: triangular models and control function methods.

3. Today.
    1. Models with excess heterogeneity and index restrictions.
    2. Discrete endogenous variables, set identification and control function methods.
    3. Set identification in IV models for binary data.

4. Wednesday April 23rd. Seminar on "Endogeneity and Discrete Outcomes" - set identification in IV models for discrete data.

Non-additive latent variable model: excess heterogeneity

- Scalar outcome Y is determined by

$$Y = h(\theta(X, Z_1), U)$$

U is a latent vector , $\theta(X, Z_1)$ is a scalar function of endogenous scalar X and vector Z_1.

- The continuous endogenous X is determined by

$$X = g(Z, V)$$

where h is strictly increasing in scalar continuous latent variate V and $Z = (Z_1, Z_2)$.

- (U, V) and Z independent
- Study identification and estimation of **index relative sensitivity**:

$$\frac{\partial\theta(x, z_1)/\partial x}{\partial\theta(x, z_1)/\partial z_{1j}}.$$

Non-additive latent variable model: excess heterogeneity

- Outcome Y is determined by

$$Y = h(\theta(X, Z_1), U)$$

- Continuous *endogenous* X given by

$$X = g(Z, V) \qquad V = r(X, Z)$$

where r is the inverse function such that

$$X = g(Z, r(X, Z)).$$

- $(U, V) \perp\!\!\!\perp Z$. Normalize $V \sim \mathit{Unif}(0, 1)$. Then

$$r(x, z) = F_{X|Z}(x|z).$$

- The conditional distribution function $F_{Y|XZ}(y|x, z) \equiv P[Y \leq y|X = x, Z = z]$:

$$F_{Y|XZ}(y|x, z) = \int \cdots \int_{h(\theta(x,z_1),u) \leq y} dF_{U|V}(u|r(x, z)) \equiv s(\theta(x, z_1), r(x, z), y)$$

Non-additive latent variable model: excess heterogeneity

$$F_{Y|XZ}(y|x,z) = \int \cdots \int_{h(\theta(x,z_1),u) \leq y} dF_{U|V}(u|r(x,z)) \equiv s(\theta(x,z_1), r(x,z), y)$$

- To study identification of $\nabla_x \theta / \nabla_{z_1} \theta$ consider derivatives of $F_{Y|XZ}$.

$$\begin{aligned}
\nabla_x F_{Y|XZ} &= \boldsymbol{\nabla_\theta s \times \nabla_x \theta} + \nabla_r s \times \nabla_x r \\
\nabla_{z_1} F_{Y|XZ} &= \boldsymbol{\nabla_\theta s \times \nabla_{z_1} \theta} + \nabla_r s \times \nabla_{z_1} r \\
\nabla_{z_2} F_{Y|XZ} &= \nabla_r s \times \nabla_{z_2} r
\end{aligned}$$

$$\nabla_x F_{X|Z} = \nabla_x r \qquad \nabla_{z_1} F_{X|Z} = \nabla_{z_1} r \qquad \nabla_{z_2} F_{X|Z} = \nabla_{z_2} r$$

- There is identification of $\nabla_x \theta / \nabla_{z_1} \theta$ if $\nabla_{z_2} F_{X|Z} \neq 0$ because

$$\begin{aligned}
\boldsymbol{\nabla_\theta s \times \nabla_x \theta} &= \nabla_x F_{Y|XZ} - \nabla_x F_{X|Z} \times \frac{\nabla_{z_2} F_{Y|XZ}}{\nabla_{z_2} F_{X|Z}} \\
\boldsymbol{\nabla_\theta s \times \nabla_{z_1} \theta} &= \nabla_{z_1} F_{Y|XZ} - \nabla_{z_1} F_{X|Z} \times \frac{\nabla_{z_2} F_{Y|XZ}}{\nabla_{z_2} F_{X|Z}}
\end{aligned}$$

Non-additive latent variable model: excess heterogeneity

- With **continuous** Y

$$\nabla_x F_{X|Z}(x|z) = \frac{1}{\nabla_v Q_{X|Z}(v|z)} \qquad \nabla_{z_i} F_{X|Z}(x|z) = -\frac{\nabla_{z_i} Q_{X|Z}(v|z)}{\nabla_v Q_{X|Z}(v|z)}$$

where $v \equiv F_{X|Z}(x|z)$

- There is identification of $\nabla_x \theta / \nabla_{z_1} \theta$ if $\nabla_{z_2} F_{X|Z} \neq 0$ because

$$\begin{aligned}
\boldsymbol{\nabla}_\theta \mathbf{s} \times \boldsymbol{\nabla}_x \boldsymbol{\theta} &= \nabla_x F_{Y|XZ} - \nabla_x F_{X|Z} \times \frac{\nabla_{z_2} F_{Y|XZ}}{\nabla_{z_2} F_{X|Z}} \\
\boldsymbol{\nabla}_\theta \mathbf{s} \times \boldsymbol{\nabla}_{z_1} \boldsymbol{\theta} &= \nabla_{z_1} F_{Y|XZ} - \nabla_{z_1} F_{X|Z} \times \frac{\nabla_{z_2} F_{Y|XZ}}{\nabla_{z_2} F_{X|Z}}
\end{aligned}$$

- In terms of quantile derivatives:

$$\begin{aligned}
\boldsymbol{\nabla}_\theta \mathbf{s} \times \boldsymbol{\nabla}_x \boldsymbol{\theta} &= -\frac{1}{\nabla_\tau Q_{Y|XZ}} \left(\nabla_x Q_{Y|XZ} + \frac{\nabla_{z_2} Q_{Y|XZ}}{\nabla_{z_2} Q_{X|Z}} \right) \\
\boldsymbol{\nabla}_\theta \mathbf{s} \times \boldsymbol{\nabla}_{z_1} \boldsymbol{\theta} &= -\frac{1}{\nabla_\tau Q_{Y|XZ}} \left(\nabla_{z_1} Q_{Y|XZ} - \nabla_{z_1} Q_{X|Z} \frac{\nabla_{z_2} Q_{Y|XZ}}{\nabla_{z_2} Q_{X|Z}} \right)
\end{aligned}$$

Multiple discrete choice with endogeneity

- 3 choices, random utilities

$$\begin{aligned} Y_1^* &= h_1(\theta_1(X,Z),U_1) \\ Y_2^* &= h_2(\theta_2(X,Z),U_2) \\ Y_3^* &= h_3(\theta_3(X,Z),U_3) \end{aligned}$$

$$Y_1 = 1[Y_1^* > \max(Y_2^*, Y_3^*)]$$

$$X = g(Z,V)$$

g strictly increasing in continuous V. $(U_1, U_2, U_3, V) \perp\!\!\!\perp Z$

- Probabilities

$$P[Y_1 = 1 | X = x, Z = x] = s_1(\theta_1(x,z), \theta_2(x,z), \theta_3(x,z), g^{-1}(z,x))$$

and with restrictions (exclusion) there can be identification of e.g.

$$\frac{\nabla_X \theta_i}{\nabla_{Z_a} \theta_i}$$

Discrete endogenous variable X

- Consider the case with Y **continuous**: h strictly increasing in U.

$$\text{A:} \quad \begin{array}{c} Y = h(X, U) \\ X = g(Z, V) \end{array} \qquad (U, V) \perp\!\!\!\perp Z$$

- Chesher (2005) also deals with discrete Y case.
- h strictly increasing in U so there is an inverse function $h^{-1}(x, u)$ such that

$$y = h(x, h^{-1}(x, y)) \qquad u = h^{-1}(x, h(x, u))$$

- $(U, V) \perp\!\!\!\perp Z$. DF of U given V: $F_{U|V}(u|v)$ and quantile function of U given V:

$$c(\tau, v) \equiv Q_{U|V}(\tau|v)$$

  are now restricted strictly **monotonic** in v.
- X discrete - some definitions...

- Consider the case with Y **continuous**: h strictly increasing in U.

$$\text{A:} \quad \begin{array}{l} Y = h(X, U) \\ X = g(Z, V) \end{array} \qquad (U, V) \perp\!\!\!\perp Z$$

- X discrete

$$X \in \{x_1, x_2, \ldots .x_M\}$$

- $g(Z, V)$ restricted **weakly** increasing in $V \sim Unif(0, 1)$

$$g(z, v) = \begin{cases} x_1 & , & 0 < & v & \leq p^1(z) \\ x_2 & , & p^1(z) < & v & \leq p^2(z) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_M & , & p^{M-1}(z) < & v & \leq 1 \end{cases}$$

- Cumulative probabilities

$$P[X \leq x_m | Z = z] = p^m(z)$$

- Point probabilities

$$P[X = x_m | Z = z] = p_m(z) \equiv p^m(z) - p^{m-1}(z)$$

Discrete endogenous variable

$$Y = h(X, U) \qquad X = g(Z, V)$$

- Distribution function of Y given X and Z.

$$P[Y \leq y | X = x_m, Z = z] = \frac{1}{p_m(z)} P[Y \leq y \wedge X = x_m | Z = z]$$

$$P[\,(Y \leq y) \wedge (X = x_m)\,|Z = z] =$$

$$P[\,\left(U \leq h^{-1}(x_m, y)\right) \wedge \left(p^{m-1}(z) \leq V \leq p^m(z)\right)|Z = z]$$

- Given by

$$P[(Y \leq y) \wedge (X = x_m)\,|Z = z] = \int_{p^{m-1}(z)}^{p^m(z)} F_{U|V}(h^{-1}(x_m, y)|v)dv$$

- So

$$F_{Y|XZ}(y|x_m, z) = \frac{1}{p_m(z)} \int_{p^{m-1}(z)}^{p^m(z)} F_{U|V}(h^{-1}(x_m, y)|v)dv$$

Bounds on distribution functions

$$Y = h(X, U) \qquad X = g(Z, V)$$

- Distribution function of Y given X and Z.

$$F_{Y|XZ}(y|x_m, z) = \frac{1}{p_m(z)} \int_{p^{m-1}(z)}^{p^m(z)} F_{U|V}(h^{-1}(x_m, y)|v)dv$$

- Restrict $F_{U|V}(u|v)$ monotonic in v. When *increasing*:

$$F_{U|V}(h^{-1}(x_m, y)|p^{m-1}(z)) \leq F_{Y|XZ}(y|x_m, z) \leq F_{U|V}(h^{-1}(x_m, y)|p^m(z))$$

- Consider solutions for y with each term set equal to τ.

$$F_{U|V}(h^{-1}(x_m, y)|p^{m-1}(z)) = \tau$$

implies

$$h^{-1}(x_m, y) = Q_{U|V}(\tau|p^{m-1}(z))$$

- Invert again:

$$y = h(x_m, Q_{U|V}(\tau|p^{m-1}(z)))$$

Bounds on quantile functions 1

$$Y = h(X, U) \qquad X = g(Z, V)$$

- Restrict $F_{U|V}(u|v)$ monotonic in v. When *increasing*:

$$F_{U|V}(h^{-1}(x_m, y)|p^{m-1}(z)) \leq F_{Y|XZ}(y|x_m, z) \leq F_{U|V}(h^{-1}(x_m, y)|p^m(z))$$

- Consider solutions for y with each term set equal to τ.

$$F_{U|V}(h^{-1}(x_m, y)|p^{m-1}(z)) = \tau \implies y = h(x_m, Q_{U|V}(\tau|p^{m-1}(z)))$$

- Similarly:

$$F_{U|V}(h^{-1}(x_m, y)|p^m(z)) = \tau \implies y = h(x_m, Q_{U|V}(\tau|p^m(z)))$$

- Finally:

$$F_{Y|XZ}(y|x_m, z) = \tau \implies y = Q_{Y|XZ}(\tau|x_m, z)$$

Bounds on quantile functions 2

$$Y = h(X, U) \qquad X = g(Z, V)$$

- Restrict $F_{U|V}(u|v)$ monotonic in v. When *increasing*:

$$F_{U|V}(h^{-1}(x_m, y)|p^{m-1}(z)) \leq F_{Y|XZ}(y|x_m, z) \leq F_{U|V}(h^{-1}(x_m, y)|p^m(z))$$

- On inverting the bounds are reversed.

$$h(x_m, Q_{U|V}(\tau|p^m(z))) \leq Q_{Y|XZ}(\tau|x_m, z) \leq h(x_m, Q_{U|V}(\tau|p^{m-1}(z)))$$

- We have bounds on $Q_{Y|XZ}(\tau|x_m, z)$ for $F_{U|V}(u|v) \uparrow$ in v.
- BUT - we want to bound h not Q!

Bounds on h 1

$$Y = h(X, U) \qquad X = g(Z, V)$$

- Bounds on $Q_{Y|XZ}(\tau|x_m, z)$ for $F_{U|V}(u|v) \uparrow$ in v.

$$h(x_m, Q_{U|V}(\tau|p^m(z))) \leq Q_{Y|XZ}(\tau|x_m, z) \leq h(x_m, Q_{U|V}(\tau|p^{m-1}(z)))$$

- Consider a value $\bar{v}$ and suppose there exist values of Z, z_{m-1} and z_m such that

$$p^m(z_m) \leq \bar{v} \leq p^{m-1}(z_{m-1})$$

- Cannot be satisfied if $X \in \{x_1, x_2\}$ is **binary**.

$$m = 1: \qquad p^m(z_m) \leq \bar{v} \leq p^{m-1}(z_{m-1}) = 0$$

$$m = 2: \qquad 1 = p^m(z_m) \leq \bar{v} \leq p^{m-1}(z_{m-1})$$

- Note:

$$F_{X|Z}(x_m|z_m) = p^m(z_m) \leq \bar{v} \leq p^{m-1}(z_{m-1}) = F_{X|Z}(x_{m-1}|z_{m-1})$$

implies

$$Q_{X|Z}(\bar{v}|z_{m-1}) \leq x_{m-1} < x_m \leq Q_{X|Z}(\bar{v}|z_m)$$

Lower bound on h

$$Y = h(X, U) \qquad X = g(Z, V)$$

- Bounds on $Q_{Y|XZ}(\tau|x_m, z)$ for $F_{U|V}(u|v) \uparrow$ in v.

$$h(x_m, Q_{U|V}(\tau|p^m(z))) \leq Q_{Y|XZ}(\tau|x_m, z) \leq h(x_m, Q_{U|V}(\tau|p^{m-1}(z)))$$

- Consider a value $\bar{v}$ and suppose there exist values of Z, z_{m-1} and z_m such that

$$p^m(z_m) \leq \bar{v} \leq p^{m-1}(z_{m-1})$$

- Set $z = z_{m-1}$. $F_{U|V}(u|v) \uparrow$ in v implies $Q_{U|V}(\tau|v) \downarrow$ in v. So:

$$Q_{Y|XZ}(\tau|x_m, z_{m-1}) \leq h(x_m, Q_{U|V}(\tau|p^{m-1}(z_{m-1}))) \leq h(x_m, Q_{U|V}(\tau|\bar{v}))$$

Upper bound on h

$$Y = h(X, U) \qquad X = g(Z, V)$$

- Bounds on $Q_{Y|XZ}(\tau|x_m, z)$ for $F_{U|V}(u|v) \uparrow$ in v.

$$h(x_m, Q_{U|V}(\tau|p^m(z))) \leq Q_{Y|XZ}(\tau|x_m, z) \leq h(x_m, Q_{U|V}(\tau|p^{m-1}(z)))$$

- Consider a value $\bar{v}$ and suppose there exist values of Z, z_{m-1} and z_m such that

$$p^m(z_m) \leq \bar{v} \leq p^{m-1}(z_{m-1})$$

- Set $z = z_m$. $F_{U|V}(u|v) \uparrow$ in v implies $Q_{U|V}(\tau|v) \downarrow$ in v. So:

$$h(x_m, Q_{U|V}(\tau|\bar{v})) \leq h(x_m, Q_{U|V}(\tau|p^m(z_m))) \leq Q_{Y|XZ}(\tau|x_m, z_m)$$

Interval for h 1

$$Y = h(X, U) \qquad X = g(Z, V)$$

- Bounds on $Q_{Y|XZ}(\tau|x_m, z)$ for $F_{U|V}(u|v) \uparrow$ in v.

$$h(x_m, Q_{U|V}(\tau|p^m(z))) \leq Q_{Y|XZ}(\tau|x_m, z) \leq h(x_m, Q_{U|V}(\tau|p^{m-1}(z)))$$

- Consider a value $\bar{v}$ and suppose there exist values of Z, z_{m-1} and z_m such that

$$p^m(z_m) \leq \bar{v} \leq p^{m-1}(z_{m-1})$$

- Bring the bounds together:

$$Q_{Y|XZ}(\tau|x_m, z_{m-1}) \leq h(x_m, Q_{U|V}(\tau|p^{m-1}(z_{m-1}))) \leq h(x_m, Q_{U|V}(\tau|\bar{v}))$$

$$h(x_m, Q_{U|V}(\tau|\bar{v})) \leq h(x_m, Q_{U|V}(\tau|p^m(z_m))) \leq Q_{Y|XZ}(\tau|x_m, z_m)$$

Interval for h 2

$$Y = h(X, U) \qquad X = g(Z, V)$$

- There exist values of Z, z_{m-1} and z_m such that

$$p^m(z_m) \leq \bar{v} \leq p^{m-1}(z_{m-1})$$

- Recall

$$c(\tau, \bar{v}) \equiv Q_{U|V}(\tau|\bar{v})$$

- We have bounds on $h(x_m, c(\tau, \bar{v}))$ for $F_{U|V}(u|v)$ **increasing** in v.

$$Q_{Y|XZ}(\tau|x_m, z_{m-1}) \leq h(x_m, c(\tau, \bar{v})) \leq Q_{Y|XZ}(\tau|x_m, z_m)$$

- For $F_{U|V}(u|v)$ **decreasing** in v the bounds are reversed.

$$Q_{Y|XZ}(\tau|x_m, z_m) \leq h(x_m, c(\tau, \bar{v})) \leq Q_{Y|XZ}(\tau|x_m, z_{m-1})$$

- Data will reveal the direction of dependence!

Set identification of h

$$Y = h(X, U) \qquad X = g(Z, V)$$

- There exist values of Z, z_{m-1} and z_m such that
$$p^m(z_m) \leq \bar{v} \leq p^{m-1}(z_{m-1}) \tag{*}$$
- Bounds on $h(x_m, Q_{U|V}(\tau|\bar{v}))$ for $F_{U|V}(u|v) \uparrow$ or $\downarrow$ in v.
$$\min\left\{ \begin{array}{c} Q_{Y|XZ}(\tau|x_m, z_{m-1}) \\ Q_{Y|XZ}(\tau|x_m, z_m) \end{array} \right\} \leq h(x_m, c(\tau, \bar{v})) \leq \max\left\{ \begin{array}{c} Q_{Y|XZ}(\tau|x_m, z_{m-1}) \\ Q_{Y|XZ}(\tau|x_m, z_m) \end{array} \right\}$$
- The restriction (*) implies
$$Q_{X|Z}(\bar{v}|z_{m-1}) \leq x_{m-1} < x_m \leq Q_{X|Z}(\bar{v}|z_m)$$
- May have $\bar{v}$ equal to $p^m(z_m)$ and $p^{m-1}(z_{m-1})$:
$$\bar{v} = p^m(z_m) \implies x_m = Q_{X|Z}(\bar{v}|z_m)$$
$$\bar{v} = p^{m-1}(z_{m-1}) \implies x_{m-1} = Q_{X|Z}(\bar{v}|z_{m-1})$$
- Tightest bounds are got choosing
$$z_m = \sup\{z : p^m(z) \leq \bar{v}\} \qquad z_{m-1} = \inf\{z : p^{m-1}(z) \geq \bar{v}\}$$

Set identification of h

$$Y = h(X, U) \qquad X = g(Z, V)$$

- There exist values of Z, z_{m-1} and z_m such that

$$F_{X|Z}(x_m|z_m) \equiv p^m(z_m) \leq \bar{v} \leq p^{m-1}(z_{m-1}) \equiv F_{X|Z}(x_{m-1}|z_{m-1})$$

- Bounds on $h(x_m, Q_{U|V}(\tau|\bar{v}))$ for $F_{U|V}(u|v) \uparrow$ or $\downarrow$ in v.

$$\min \left\{ \begin{array}{c} Q_{Y|XZ}(\tau|x_m, z_{m-1}) \\ Q_{Y|XZ}(\tau|x_m, z_m) \end{array} \right\} \leq h(x_m, c(\tau, \bar{v})) \leq \max \left\{ \begin{array}{c} Q_{Y|XZ}(\tau|x_m, z_{m-1}) \\ Q_{Y|XZ}(\tau|x_m, z_m) \end{array} \right\}$$

- Over-partial-identifcation - bounds got using any selection of instrumental values should have non-empty intersections.
- Can bound effect of a covariate if $Y = h(X, Z, U)$.
- Can localize restrictions.
- Continuous endogenous variables - point identification - can have $z_{m-1} = z_m$ as $x_{m-1} \rightarrow x_m$.
- Exogenous X when $F_{U|V} = F_U$ - then $Q_{Y|XZ} = Q_{Y|X}$.

Set identification of partial differences

$$Y = h(X, U) \qquad X = g(Z, V)$$

- Consider identification of
$$h(x_m, \bar{u}) - h(x_n, \bar{u})$$
at some value $\bar{u}$ that can be known and has structural intepretation. Partial difference; ceteris paribus effect of a shift from x_m to x_n.
- Suppose there are values of Z, $\{z_{m-1}, z_m\}$ and $\{z_{n-1}, z_n\}$such that
$$p^m(z_m) \leq \bar{v} \leq p^{m-1}(z_{m-1})$$
$$p^n(z_n) \leq \bar{v} \leq p^{n-1}(z_{n-1})$$
- Then with $c(\tau, \bar{v}) \equiv Q_{U|V}(\tau|\bar{v})$

$$\min \left\{ \begin{array}{c} Q_{Y|XZ}(\tau|x_m, z_{m-1}) \\ Q_{Y|XZ}(\tau|x_m, z_m) \end{array} \right\} \leq h(x_m, c(\tau, \bar{v})) \leq \max \left\{ \begin{array}{c} Q_{Y|XZ}(\tau|x_m, z_{m-1}) \\ Q_{Y|XZ}(\tau|x_m, z_m) \end{array} \right\}$$

$$-\max \left\{ \begin{array}{c} Q_{Y|XZ}(\tau|x_n, z_{n-1}) \\ Q_{Y|XZ}(\tau|x_n, z_n) \end{array} \right\} \leq -h(x_n, c(\tau, \bar{v})) \leq -\min \left\{ \begin{array}{c} Q_{Y|XZ}(\tau|x_n, z_{n-1}) \\ Q_{Y|XZ}(\tau|x_n, z_n) \end{array} \right\}$$

- Set identification of $h(x_m, \bar{u}) - h(x_n, \bar{u})$ at $\bar{u} = c(\tau, \bar{v})$ on adding these intervals.

Estimating bounds on h

$$Y = h(X, U) \qquad X = g(Z, V)$$

- There exist values of Z, z_{m-1} and z_m such that

$$p^m(z_m) \leq \bar{v} \leq p^{m-1}(z_{m-1})$$

- Estimate $F_{X|Z}$:

$$\hat{z}_m = \sup\{z : \hat{F}_{X|Z}(x_m|z) \leq \bar{v}\} \qquad \hat{z}_{m-1} = \inf\{z : \hat{F}_{X|Z}(x_m|z) \geq \bar{v}\}$$

- Estimate $Q_{Y|XZ}$. Estimated bounds:

$$\min\left\{ \begin{array}{c} \hat{Q}_{Y|XZ}(\tau|x_m, \hat{z}_{m-1}) \\ \hat{Q}_{Y|XZ}(\tau|x_m, \hat{z}_m) \end{array} \right\} \leq h(x_m, c(\tau, \bar{v})) \leq \max\left\{ \begin{array}{c} \hat{Q}_{Y|XZ}(\tau|x_m, \hat{z}_{m-1}) \\ \hat{Q}_{Y|XZ}(\tau|x_m, \hat{z}_m) \end{array} \right\}$$

THE
QUARTERLY JOURNAL
OF ECONOMICS

Vol. CVI November 1991 Issue 4

DOES COMPULSORY SCHOOL ATTENDANCE AFFECT SCHOOLING AND EARNINGS?*

JOSHUA D. ANGRIST AND ALAN B. KRUEGER

We establish that season of birth is related to educational attainment because of school start age policy and compulsory school attendance laws. Individuals born in the beginning of the year start school at an older age, and can therefore drop out after completing less schooling than individuals born near the end of the year. Roughly 25 percent of potential dropouts remain in school because of compulsory schooling laws. We estimate the impact of compulsory schooling on earnings by using quarter of birth as an instrument for education. The instrumental variables estimate of the return to education is close to the ordinary least squares estimate, suggesting that there is little bias in conventional estimates.

Every developed country in the world has a compulsory schooling requirement, yet little is known about the effect these laws have on educational attainment and earnings.[1] This paper exploits an unusual natural experiment to estimate the impact of compulsory schooling laws in the United States. The experiment stems from the fact that children born in different months of the year start school at different ages, while compulsory schooling laws generally require students to remain in school until their sixteenth or seventeenth birthday. In effect, the interaction of school-entry requirements and compulsory schooling laws compel students born

*We thank Michael Boozer and Lisa Krueger for outstanding research assistance. Financial support was provided by the Princeton Industrial Relations Section, an NBER Olin Fellowship in Economics, and the National Science Foundation (SES-9012149). We are also grateful to Lawrence Katz, John Pencavel, an anonymous referee, and many seminar participants for helpful comments. The data and computer programs used in the preparation of this paper are available on request.

1. See OECD [1983] for a comparison of compulsory schooling laws in different countries.

© 1991 by the President and Fellows of Harvard College and the Massachusetts Institute of Technology.

Weak instruments: Angrist-Krueger QJE (1991)

1930-39 cohort: W: log wage, S: years of schooling, B: quarter of birth

$$\begin{aligned} W &= h(S, U) \\ S &= g(B, V) \end{aligned}$$

- There is no value m of years of schooling and no $\bar{v} \in (0, 1)$ such that:

$$p^m(z_m) \leq \bar{v} \leq p^{m-1}(z_{m-1})$$

- The AK model is not nonparametrically identifying.

Table 1: Estimated distribution function of years of schooling for each quarter of birth

Years of schooling	QOB 1	QOB 2	QOB 3	QOB 4
1	.002	.003	.002	.002
2	.005	.005	.004	.004
3	.008	.009	.007	.007
4	.013	.013	.011	.010
5	.020	.019	.016	.016
6	.032	.032	.028	.027
7	.052	.050	.046	.045
8	.104	.101	.094	.090
9	.146	.142	.133	.128
10	.197	.194	.182	.177
11	.238	.236	.223	.218
12	.612	.604	.598	.595
13	.666	.658	.653	.652
14	.742	.734	.734	.732
15	.771	.763	.764	.763
16	.877	.873	.874	.872
17	.913	.909	.911	.909
18	.947	.945	.947	.946
19	.966	.964	.966	.966
20	1.00	1.00	1.00	1.00

Inspecting Table 1 it can be seen that for every value of m

$$\max_i(\hat{P}[Y_2 \leq y_2^{m-1}|X_i = 1]) < \min_i(\hat{P}[Y_2 \leq y_2^{m-1}|X_i = 1])$$

and it follows that there is *no* value of τ_2 and m for which there are configurations of the four binary indicators, x^m and x^{m-1} such that:

$$\hat{p}^m(x^m) \leq \tau_2 \leq \hat{p}^{m-1}(x^{m-1}).$$

This strongly suggests that the AK quarter of birth instruments are too weak to nonparametrically identify the returns to schooling.

4.2. "Continuous" years of schooling. It is interesting to consider what the identifying power of the quarter of birth instruments would have been had years of schooling been recorded *continuously.* This also provides the opportunity to illustrate some of the results set out earlier in the paper.

Additive errors

$$Y = h(X) + U \qquad X = g(Z, V)$$

$$E[U|V, Z] = E[U|V] \qquad V \perp\!\!\!\perp Z$$

- If Z has rich support there is point identification as in Lecture 1, Das (2005).
- If Z has poor support relative to X there is a similar set identification result if $E[U|V = v]$ is monotonic in v and there are instrumental values such that

$$p^m(z_m) \leq \bar{v} \leq p^{m-1}(z_{m-1})$$

- Define

$$E[U|V = v, Z = z] = c(v)$$

then - Chesher (2004)

$$\min \left\{ \begin{array}{c} E_{Y|XZ}(Y|x_m, z_{m-1}) \\ E_{Y|XZ}(Y|x_m, z_m) \end{array} \right\} \leq h(x_m) + c(\bar{v}) \leq \max \left\{ \begin{array}{c} E_{Y|XZ}(Y|x_m, z_{m-1}) \\ E_{Y|XZ}(Y|x_m, z_m) \end{array} \right\}$$

IV model for a binary outcome

- Consider the threshold crossing model for binary Y with observed X and latent U

$$Y = h(X,U) \equiv \begin{cases} 0 \quad , & 0 < U \leq p(X) \\ 1 \quad , & p(X) < U \leq 1 \end{cases} \qquad Z \perp\!\!\!\perp U \sim Unif(0,1)$$

- First consider binary $X \in \{x_1, x_2\}$. Define $\gamma \equiv (\gamma_1, \gamma_2)$.

$$\gamma_1 \equiv p(x_1) \qquad \gamma_2 \equiv p(x_2)$$

- If $Z = X$ - that is - $X \perp\!\!\!\perp U$ then the model identifies γ.

$$\Pr[Y = 0 | X = x_i] = \gamma_i .$$

- When U and X are not independently distributed

$$\Pr[Y = 0 | X = x_i] = \Pr[U \leq \gamma_i | X = x_i] \neq \gamma_i$$

- We ask: does the IV model identify γ? We will show that it **set** identifies γ.
- There are other attacks on the binary Y binary X problem seeking point identification of average effects of X on Y, e.g. the triangular model of Vytlacil and Yildiz (2007).

IV model for a binary outcome

- Consider the threshold crossing model for binary Y with observed X and latent U

$$Y = \begin{cases} 0 \quad , \quad 0 < U \leq p(X) \\ 1 \quad , \quad p(X) < U \leq 1 \end{cases} \qquad Z \perp\!\!\!\perp U \sim Unif(0,1)$$

$$X \in \{x_1, x_2\} \qquad \gamma_1 \equiv p(x_1) \qquad \gamma_2 \equiv p(x_2)$$

- Define $\theta \equiv (\alpha_1, \alpha_2, \delta_1, \delta_2)$ **which depend on** z:

$$\alpha_1 \equiv \Pr[Y = 0|x_1, z] \qquad \alpha_2 \equiv \Pr[Y = 0|x_2, z]$$

$$\delta_1 \equiv \Pr[X = x_1|z] \qquad \delta_2 \equiv \Pr[X = x_2|z]$$

$$\begin{array}{ll} \beta_{11} \equiv P[U \leq \gamma_1|x_1, z] & \beta_{21} \equiv P[U \leq \gamma_2|x_1, z] \\ \beta_{12} \equiv P[U \leq \gamma_1|x_2, z] & \beta_{22} \equiv P[U \leq \gamma_2|x_2, z] \end{array}$$

- Implication of $Z \perp\!\!\!\perp U \sim Unif(0,1)$

$$\begin{aligned} \delta_1\beta_{11} + \delta_2\beta_{12} &= P[U \leq \gamma_1|Z = z] = \gamma_1 \\ \delta_1\beta_{21} + \delta_2\beta_{22} &= P[U \leq \gamma_2|Z = z] = \gamma_2 \end{aligned}$$

IV model for a binary outcome

- Define $\theta \equiv (\alpha_1, \alpha_2, \delta_1, \delta_2)$ **which depend on** z:

$$\alpha_1 \equiv \Pr[Y = 0 | x_1, z] \qquad \alpha_2 \equiv \Pr[Y = 0 | x_2, z]$$

$$\delta_1 \equiv \Pr[X = x_1 | z] \qquad \delta_2 \equiv \Pr[X = x_2 | z]$$

$$\begin{array}{ll} \beta_{11} \equiv P[U \leq \gamma_1 | x_1, z] & \beta_{21} \equiv P[U \leq \gamma_2 | x_1, z] \\ \beta_{12} \equiv P[U \leq \gamma_1 | x_2, z] & \beta_{22} \equiv P[U \leq \gamma_2 | x_2, z] \end{array}$$

- Implication of $Z \perp\!\!\!\perp U \sim Unif(0,1)$ and $\beta_{11} = \alpha_1$, $\beta_{22} = \alpha_2$

$$\delta_1 \alpha_1 + \delta_2 \beta_{12} = \gamma_1 \qquad \delta_1 \beta_{21} + \delta_2 \alpha_2 = \gamma_2$$

- Since $\beta_{12} \in [0,1]$ and $\beta_{21} \in [0,1]$

$$\delta_1 \alpha_1 \leq \gamma_1 \leq \delta_1 \alpha_1 + \delta_2 \qquad \delta_2 \alpha_2 \leq \gamma_2 \leq \delta_2 \alpha_2 + \delta_1$$

- There is a pair of bounds for each z.

Identifying the sign

- Binary Y with observed X and latent U

$$Y = \begin{cases} 0 \quad , & 0 < \quad U \quad \leq p(X) \\ 1 \quad , & p(X) < \quad U \quad \leq 1 \end{cases} \qquad Z \perp\!\!\!\perp U \sim Unif(0,1)$$

$$X \in \{x_1, x_2\} \qquad \gamma_1 \equiv p(x_1) \qquad \gamma_2 \equiv p(x_2)$$

$$\alpha_1 \equiv \Pr[Y = 0 | x_1, z] \qquad \alpha_2 \equiv \Pr[Y = 0 | x_2, z]$$

$$\delta_1 \equiv \Pr[X = x_1 | z] \qquad \delta_2 \equiv \Pr[X = x_2 | z]$$

- Set identification of γ_1 and γ_2:

$$\delta_1 \alpha_1 \leq \gamma_1 \leq \delta_1 \alpha_1 + \delta_2 \qquad \delta_2 \alpha_2 \leq \gamma_2 \leq \delta_2 \alpha_2 + \delta_1$$

- These intervals **always intersect** because:

$$\delta_1 \alpha_1 + \delta_2 < \delta_2 \alpha_2 \Rightarrow \delta_1 \alpha_1 < \delta_2 (\alpha_2 - 1) < 0$$

$$\delta_2 \alpha_2 + \delta_1 < \delta_1 \alpha_1 \Rightarrow \delta_2 \alpha_2 < \delta_1 (\alpha_1 - 1) < 0$$

Monotonicity restriction

- Recall

$$\alpha_1 = \beta_{11} \equiv P[U \leq \gamma_1 | x_1, z] \qquad \beta_{21} \equiv P[U \leq \gamma_2 | x_1, z]$$
$$\beta_{12} \equiv P[U \leq \gamma_1 | x_2, z] \qquad \beta_{22} \equiv P[U \leq \gamma_2 | x_2, z] = \alpha_2$$

- A **monotonicity restriction** tightens set identification, e.g. $\gamma_1 \leq \gamma_2$ implies

$$\beta_{21} \in [\alpha_1, 1] \qquad \beta_{12} \in [0, \alpha_2]$$

  and

$$\delta_1 \alpha_1 + \delta_2 \beta_{12} = \gamma_1 \qquad \delta_1 \beta_{21} + \delta_2 \alpha_2 = \gamma_2$$

  leading to

$$\delta_1 \alpha_1 \leq \gamma_1 \leq \delta_1 \alpha_1 + \delta_2 \alpha_2 \leq \gamma_2 \leq \delta_1 + \delta_2 \alpha_2$$

- The restriction $\gamma_2 \leq \gamma_1$ implies that

$$\beta_{21} \in [0, \alpha_1] \qquad \beta_{12} \in [\alpha_2, 1]$$

$$\delta_2 \alpha_2 \leq \gamma_2 \leq \delta_1 \alpha_1 + \delta_2 \alpha_2 \leq \gamma_1 \leq \delta_1 \alpha_1 + \delta_2$$

Monotonicity restriction

- $\gamma_1 \leq \gamma_2$ implies

$$\delta_1\alpha_1 \leq \gamma_1 \leq \delta_1\alpha_1 + \delta_2\alpha_2 \leq \gamma_2 \leq \delta_1 + \delta_2\alpha_2$$

- $\gamma_2 \leq \gamma_1$ implies

$$\delta_2\alpha_2 \leq \gamma_2 \leq \delta_1\alpha_1 + \delta_2\alpha_2 \leq \gamma_1 \leq \delta_1\alpha_1 + \delta_2$$

- Example: $(\alpha_1, \alpha_2, \delta_1, \delta_2) = (0.9, 0.1, 0.5, 0.5)$:

$$\gamma_1 \leq \gamma_2 : \qquad 0.45 \leq \gamma_1 \leq 0.5 \leq \gamma_2 \leq 0.55$$

$$\gamma_2 \leq \gamma_1 : \qquad 0.05 \leq \gamma_2 \leq 0.5 \leq \gamma_1 \leq 0.95$$

$$0.45 \leq \gamma_1 \leq 0.95 \qquad 0.05 \leq \gamma_2 \leq 0.55$$

- $\gamma_1 = \gamma_2$ yields point identification

$$\gamma_1 = \gamma_2 = \delta_1\alpha_1 + \delta_2\alpha_2$$

Parametric restrictions

- Parametric restrictions are easily introduced, e.g. suppose that for some value of (π_0, π_1):

$$p(x) = \frac{1}{1 + \exp(\pi_0 + \pi_1 x)}$$

- In this case

$$\begin{aligned} \pi_0 + \pi_1 x_1 &= \log(\gamma_1^{-1} - 1) \\ \pi_0 + \pi_1 x_2 &= \log(\gamma_2^{-1} - 1) \end{aligned}$$

  and an identified set of values of γ_1 and γ_2, e.g.

$$\delta_1 \alpha_1 \leq \gamma_1 \leq \delta_1 \alpha_1 + \delta_2 \qquad \delta_2 \alpha_2 \leq \gamma_2 \leq \delta_2 \alpha_2 + \delta_1$$

  leads directly to an identified set of values for π_0 and π_1.
- This is more interesting when X has more points of support!

Many instrumental values

- For **each** value z there are bounds

$$\begin{aligned}\delta_1(z)\alpha_1(z) &\le \gamma_1 \le \delta_1(z)\alpha_1(z)+\delta_2(z)\\ \delta_2(z)\alpha_2(z) &\le \gamma_2 \le \delta_2(z)\alpha_2(z)+\delta_1(z)\end{aligned}$$

- With values $z \in \Omega$ available:

$$\begin{aligned}\max_{z\in\Omega}\left(\delta_1(z)\alpha_1(z)\right) &\le \gamma_1 \le \min_{z\in\Omega}\left(\delta_1(z)\alpha_1(z)+\delta_2(z)\right)\\ \max_{z\in\Omega}\left(\delta_2(z)\alpha_2(z)\right) &\le \gamma_2 \le \min_{z\in\Omega}\left(\delta_2(z)\alpha_2(z)+\delta_1(z)\right)\end{aligned}$$

- Now there is the possibility of identifying the sign of $\gamma_1 - \gamma_2$.

Strong instruments

- With values $z \in \Omega$ available:

$$\begin{aligned}
\max_{z\in\Omega}\left(\delta_1(z)\alpha_1(z)\right) &\leq \gamma_1 \leq \min_{z\in\Omega}\left(\delta_1(z)\alpha_1(z)+\delta_2(z)\right)\\
\max_{z\in\Omega}\left(\delta_2(z)\alpha_2(z)\right) &\leq \gamma_2 \leq \min_{z\in\Omega}\left(\delta_2(z)\alpha_2(z)+\delta_1(z)\right)
\end{aligned}$$

- If there is a value z such that $\Pr[X = x_i | Z = z]$ for $i = 1$ (2) is close to 1 then there is near point identification of γ_1 (γ_2).
- For example with:

$$(\alpha_1, \alpha_2, \delta_1, \delta_2) = (0.9, 0.1, 0.95, 0.05)$$

$$\gamma_1 \leq \gamma_2 : \qquad 0.855 \leq \gamma_1 \leq 0.860 \leq \gamma_2 \leq 0.955$$

$$\gamma_2 \leq \gamma_1 : \qquad 0.005 \leq \gamma_2 \leq 0.860 \leq \gamma_1 \leq 0.905$$

$$0.855 \leq \gamma_1 \leq 0.905 \qquad 0.005 \leq \gamma_2 \leq 0.955$$

A general approach

- The threshold crossing model

$$Y = \begin{cases} 0 & , \quad 0 < U \leq p(X) \\ 1 & , \quad p(X) < U \leq 1 \end{cases} \qquad Z \perp\!\!\!\perp U \sim Unif(0,1)$$

  is a special case of this model.

$$Y = h(X,U) \qquad h \text{ weakly increasing in } U \qquad Z \perp\!\!\!\perp U \sim Unif(0,1)$$

- Chesher (2007b) shows that under the restrictions of this model for all $u \in (0,1)$ and $z \in \Omega$:

$$\Pr[Y \leq h(X,u)|Z=z] \geq u \qquad \Pr[Y < h(X,u)|Z=z] < u$$

- All functions h that satisfy these inequalities for a particular $F_{YX|Z}$ and $z \in \Omega$, are elements of structures observationally equivalent to any admissible structure which generates $F_{YX|Z}$.
- Generically this model delivers set not point identification even when there are parametric restrictions.
- Topic of Wednesday's seminar.