ECON0108 2022-23 Part 1

Week 4 Slides for Lecture 4

Andrew Chesher

CeMMAP & UCL

Conditional independence restrictions: topics

1. Average structural functions.
2. Treatment effects models.

Conditional independence - Average Structural Function

- Consider the model with Y and X endogenous

$$Y = m(X, U) \qquad U \perp\!\!\!\perp X | Z$$

and the **Average Structural Function (ASF)**:

$$\mu(x) \equiv E_U[m(x, U)] = \int m(x, u) f_U(u) du$$

- An example: a binary outcome model with endogeneity and maybe **random coefficients**

$$Y = m(X, U) \equiv \begin{cases} 1 & , \quad (\beta_0 + U_0) + (\beta_1 + U_1)X > 0 \\ 0 & , \quad (\beta_0 + U_0) + (\beta_1 + U_1)X \leq 0 \end{cases}$$

  - Here the **ASF** is the probability $Y = 1$ when $X = x$ and U has its *marginal* distribution - called a **counterfactual probability**.

Conditional independence - Average Structural Function

- Consider the model
$$Y = m(X, U) \qquad U \perp\!\!\!\perp X | Z$$
and the **Average Structural Function (ASF)**:
$$\mu(x) \equiv E_U[m(x, U)] = \int m(x, u) f_U(u) du$$

Conditional independence - Average Structural Function

- Consider the model

$$Y = m(X, U) \qquad U \perp\!\!\!\perp X | Z$$

and the **Average Structural Function (ASF)**:

$$\mu(x) \equiv E_U[m(x, U)] = \int m(x, u) f_U(u) du$$

- $E[Y|X = x]$ may *not* equal the **ASF**.

$$E[Y|X = x] = \int m(x, u)\, f_{U|X}(u|x)\, du \neq \mu(x) \text{ unless } U \perp\!\!\!\perp X$$

Conditional independence - Average Structural Function

- Consider the model

$$Y = m(X, U) \qquad U \perp\!\!\!\perp X|Z$$

and the **Average Structural Function (ASF)**:

$$\mu(x) \equiv E_U[m(x, U)] = \int m(x, u) f_U(u) du$$

- $E[Y|X = x]$ may *not* equal the **ASF**.

$$E[Y|X = x] = \int m(x, u)\, f_{U|X}(u|x)\, du \neq \mu(x) \text{ unless } U \perp\!\!\!\perp X$$

- Since $U \perp\!\!\!\perp X|Z$:

$$\begin{aligned} E[Y|X = x, Z = z] &= \int m(x, u)\, f_{U|XZ}(u|x, z)\, du \\ &= \int m(x, u)\, f_{U|Z}(u|z)\, du \end{aligned}$$

and take expectation with respect to Z.

Conditional independence - Average Structural Function

- Taking expectation with respect to Z:

$$
\begin{aligned}
E_Z[E[Y|X=x,Z=z]] &= \int \left(\int m(x,u) f_{U|Z}(u|z)du \right) f_Z(z)dz \\
&= \int m(x,u) \left(\int f_{U|Z}(u|z) f_Z(z)dz \right) du \\
&= \int m(x,u)\, f_U(u)\, du = \mu(x)
\end{aligned}
$$

Conditional independence - Average Structural Function

- Taking expectation with respect to Z:

$$\begin{aligned} E_Z[E[Y|X=x,Z=z]] &= \int \left(\int m(x,u) f_{U|Z}(u|z) du \right) f_Z(z) dz \\ &= \int m(x,u) \left(\int f_{U|Z}(u|z) f_Z(z) dz \right) du \\ &= \int m(x,u)\, f_U(u)\, du = \mu(x) \end{aligned}$$

- So under the restrictions of the model

$$Y = m(X,U) \qquad U \perp\!\!\!\perp X|Z$$

the Average Structural Function, $\mu(x)$, is

$$\mu(x) = E_Z[E[Y|X=x,Z]].$$

Conditional independence: treatment effects 1

- Outcomes: U_0 if not treated and U_1 if treated. A treatment indicator D equals 1 if treated and 0 if not treated. We observe values of D and Y

$$Y = DU_1 + (1 - D)U_0$$

There may be interest in the average treatment effect - ATE:

$$\text{ATE: } \mu \equiv E[U_1 - U_0]$$

or the average effect of treatment on the treated - ATT:

$$\text{ATT} : \mu_1 \equiv E[U_1 - U_0 | D = 1]$$

Conditional independence: treatment effects 2

- We observe values of D and Y

$$Y = DU_1 + (1 - D)U_0$$

But (U_0, U_1) may not be independent of D and then ATE, ATT are not (point) identified.

- In applied work **conditional independence** is a popular restriction:

$$(U_1, U_0) \perp\!\!\!\perp D | Z$$

- Under this condition

$$E[Y|D = 1, Z = z] = E[U_1|D = 1, Z = z] = E[U_1|Z = z]$$

$$E[Y|D = 0, Z = z] = E[U_0|D = 0, Z = z] = E[U_0|Z = z]$$

subtract, take expectation with respect to Z

$$E_Z[E[Y|D = 1, Z = z] - E[Y|D = 0, Z = z]] = \mu.$$

Conditional independence restrictions - conditional on what?

- Consider the Gaussian case

$$X = [X_1, X_2, \ldots, X_K] \sim N_K(\mu, \Sigma)$$

where $\Sigma = ||\sigma_{ij}||$. We ask - when is $Cov(X_1, X_2|X_k) = 0$?

  - When X is Gaussian $X_1 \perp\!\!\!\perp X_2|X_k$ implies $X_1 \perp\!\!\!\perp X_2|X_k$.

- Partitioning X thus: $[X_A, X_B]$ with

$$\Sigma = \begin{bmatrix} \Sigma_{AA} & \Sigma_{AB} \\ \Sigma_{BA} & \Sigma_{BB} \end{bmatrix}$$

where $\Sigma_{BA} = \Sigma_{AB}^T$ there is

$$Var[X_A|X_B = x_B] = \Sigma_{AA} - \Sigma_{AB}\Sigma_{BB}^{-1}\Sigma_{BA}$$

Conditional independence restrictions - conditional on what?

There is

$$Var[X_A|X_B = x_B] = \Sigma_{AA} - \Sigma_{AB}\Sigma_{BB}^{-1}\Sigma_{BA}$$

so the conditional variance of (X_1, X_2) given X_k is

$$\begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix} - \frac{1}{\sigma_{kk}} \begin{bmatrix} \sigma_{1k}^2 & \sigma_{1k}\sigma_{2k} \\ \sigma_{1k}\sigma_{2k} & \sigma_{2k}^2 \end{bmatrix}$$

- There is $X_1 \perp\!\!\!\perp X_2 | X_k$ if and only if

$$\sigma_{12} - \frac{\sigma_{1k}\sigma_{2k}}{\sigma_{kk}} = 0.$$

Conditional independence restrictions - conditional on what?

There is

$$Var[X_A|X_B = x_B] = \Sigma_{AA} - \Sigma_{AB}\Sigma_{BB}^{-1}\Sigma_{BA}$$

so the conditional variance of (X_1, X_2) given X_k is

$$\begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix} - \frac{1}{\sigma_{kk}} \begin{bmatrix} \sigma_{1k}^2 & \sigma_{1k}\sigma_{2k} \\ \sigma_{1k}\sigma_{2k} & \sigma_{2k}^2 \end{bmatrix}$$

- There is $X_1 \perp\!\!\!\perp X_2 | X_k$ if and only if

$$\sigma_{12} - \frac{\sigma_{1k}\sigma_{2k}}{\sigma_{kk}} = 0.$$

- If $X_1 \perp\!\!\!\perp X_2 | X_k$ then in general $X_1 \perp\!\!\!\perp X_2 | \left(X_k, X_j\right)$ does NOT hold.

Conditional independence restrictions - conditional on what?

There is

$$Var[X_A|X_B = x_B] = \Sigma_{AA} - \Sigma_{AB}\Sigma_{BB}^{-1}\Sigma_{BA}$$

so the conditional variance of (X_1, X_2) given X_k is

$$\begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix} - \frac{1}{\sigma_{kk}} \begin{bmatrix} \sigma_{1k}^2 & \sigma_{1k}\sigma_{2k} \\ \sigma_{1k}\sigma_{2k} & \sigma_{2k}^2 \end{bmatrix}$$

- There is $X_1 \perp\!\!\!\perp X_2|X_k$ if and only if

$$\sigma_{12} - \frac{\sigma_{1k}\sigma_{2k}}{\sigma_{kk}} = 0.$$

- If $X_1 \perp\!\!\!\perp X_2|\left(X_k, X_j\right)$ then in general NEITHER $X_1 \perp\!\!\!\perp X_2|X_k$ NOR $X_1 \perp\!\!\!\perp X_2|X_j$ holds.

Conditional independence restrictions - conditional on what?

There is

$$Var[X_A|X_B = x_B] = \Sigma_{AA} - \Sigma_{AB}\Sigma_{BB}^{-1}\Sigma_{BA}$$

so the conditional variance of (X_1, X_2) given X_k is

$$\begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix} - \frac{1}{\sigma_{kk}} \begin{bmatrix} \sigma_{1k}^2 & \sigma_{1k}\sigma_{2k} \\ \sigma_{1k}\sigma_{2k} & \sigma_{2k}^2 \end{bmatrix}$$

- There is $X_1 \perp\!\!\!\perp X_2 | X_k$ if and only if

$$\sigma_{12} - \frac{\sigma_{1k}\sigma_{2k}}{\sigma_{kk}} = 0.$$

- Measurement error: ε: if $X_1 \perp\!\!\!\perp X_2 | X_k$ then in general $X_1 \perp\!\!\!\perp X_2 | (X_k + \varepsilon)$ does NOT hold.

Complete or incomplete models?

- Specifying complete models require much knowledge of the process being studied.
  - Misspecification of just part of a model can lead to misleading inference about parameters of correctly specified elements.

Complete or incomplete models?

- Specifying complete models require much knowledge of the process being studied.
  - Misspecification of just part of a model can lead to misleading inference about parameters of correctly specified elements.

- So we turn to consider *incomplete* models.

- **Definition** (Koopmans, Cowles Monograph 10, 1950): an incomplete model is a model that admits structures which can deliver *non-unique* values of endogenous outcomes given values of observed and unobserved exogenous variables.

Incomplete models

- Example: *single equation models* with more than one endogenous variable are incomplete.

$$Y_1 = \alpha + \beta Y_2 + \gamma Z + U$$

  - Any $Y = (Y_1, Y_2)$ satisfying $Y_1 - \beta Y_2 = \alpha + \gamma Z + U$ satisfies this equation. There is a *set* of such values.

Incomplete models

- Models in which structural restrictions involve inequalities are usually incomplete.

- Models admitting structures that deliver multiple equilibria are incomplete, for example the simultaneous firm entry models of Jovanovich (Ecta, 1989), Bresnahan and Reiss (ReSt 1990), Tamer (ReSt, 2003), Ciliberto and Tamer (Ecta, 2009).

- The constructs and methods developed so far allow analysis of cases in which models require unobservables to be *single-valued* functions of observable variables.

Single equation models

- Consider models with endogenous Y (scalar) and X and exogenous Z

$$Y = g(X, Z, U)$$

  - U is an unobserved latent variable, $U \perp\!\!\!\perp Z$, or some other independence restriction.
  - the model is "incomplete", "single equation", "limited information".

- Types of model we consider:
  - Linear and nonlinear parametric models, continuous and discrete X.
  - Nonparametric models.
  - Models with "set-valued" residuals, i.e. U is not a single-valued function of Y and X.

IV and parametric nonlinear models

- Consider the linear in parameters model, endogenous Y and X

$$Y = \alpha + \beta X + \gamma X^2 + U \qquad U \perp\!\!\!\perp Z$$

with $E[U] = 0$, implies $E[g(Z)U] = 0$ for $g(Z)$ with bounded support.

- Multiplying by Z, Z^2 gives

$$\begin{aligned} E[Y] &= \alpha + \beta E[X] + \gamma E[X^2] \\ E[YZ] &= \alpha E[Z] + \beta E[XZ] + \gamma E[X^2 Z] \\ E[YZ^2] &= \alpha E[Z^2] + \beta E[XZ^2] + \gamma E[X^2 Z^2] \end{aligned}$$

and α, β, γ are solutions to 3 simultaneous equations. Point identification under a rank condition. Requires at least 3 points in the support of Z.

- Overidentification: could use any functions of Z, not just Z and Z^2. How to choose? Smallest variance analogue estimator uses $E[X|Z]$ and $E[X^2|Z]$.

IV and parametric nonlinear models

- Consider the model with a conditional mean independence restriction

$$Y = \alpha + \beta X + \gamma X^2 + U \qquad \forall z \quad E[U|Z = z] = c$$

- With just 3 values of Z there can be point identification of parameter values

$$\begin{aligned} E[Y|z_1] &= \alpha + \beta E[X|z_1] + \gamma E[X^2|z_1] + c \\ E[Y|z_2] &= \alpha + \beta E[X|z_2] + \gamma E[X^2|z_2] + c \\ E[Y|z_3] &= \alpha + \beta E[X|z_3] + \gamma E[X^2|z_3] + c \end{aligned}$$

- Require rich support for Z as polynomial (spline etc.) becomes more flexible.

The nonparametric IV model: discrete X

- Discrete, $X \in \{x_1, x_2, \dots, x_M\}$

$$Y = h(X) + U \qquad \forall z \quad E[U|Z = z] = 0$$

$$\theta_m \equiv h(x_m) \qquad D_m(X) \equiv 1[X = x_m]$$

- Then - Das (2005)

$$Y = \sum_{m=1}^{M} \theta_m D_m(X) + U \qquad E[Y|z] = \sum_{m=1}^{M} \theta_m P[X = x_m|z]$$

The nonparametric IV model: discrete X

- Discrete, $X \in \{x_1, x_2, \ldots, x_M\}$

$$Y = h(X) + U \qquad \forall z \quad E[U|Z = z] = 0$$

$$\theta_m \equiv h(x_m) \qquad D_m(X) \equiv 1[X = x_m]$$

- Then - Das (2005)

$$Y = \sum_{m=1}^{M} \theta_m D_m(X) + U \qquad E[Y|z] = \sum_{m=1}^{M} \theta_m P[X = x_m|z]$$

- If $Z \in \{z_1, \ldots, z_K\}$ identifiability of h depends on the rank of the $K \times M$ matrix of probabilities;

$$\begin{bmatrix} P[X = x_1|z_1] & P[X = x_2|z_1] & \cdots & P[X = x_M|z_1] \\ \vdots & \vdots & & \vdots \\ P[X = x_1|z_K] & P[X = x_2|z_K] & \cdots & P[X = x_M|z_K] \end{bmatrix}$$

ill conditioned matrix when M large; smoothness; dimension of X and Z.

The nonparametric IV model: continuous X

- A nonparametric model with continuous X (Newey & Powell (2003)) and additive U

$$Y = h(X) + U \qquad E[U|Z = z] = 0 \qquad z \in \mathcal{R}_Z$$

$$E[Y|Z = z] = E[h(X)|Z = z] \equiv \int h(x) f_{X|Z}(x|z) dx. \tag{*}$$

The nonparametric IV model: continuous X

- A nonparametric model with continuous X (Newey & Powell (2003)) and additive U

$$Y = h(X) + U \qquad E[U|Z = z] = 0 \qquad z \in \mathcal{R}_Z$$

$$E[Y|Z = z] = E[h(X)|Z = z] \equiv \int h(x) f_{X|Z}(x|z) dx. \qquad (*)$$

- Suppose $h(\cdot)$ and $h^*(\cdot)$ both solve (*), that is, for all z:

$$\begin{aligned} E[Y|Z &= z] = \int h(x) f_{X|Z}(x|z) dx \\ E[Y|Z &= z] = \int h^*(x) f_{X|Z}(x|z) dx \end{aligned}$$

then, on subtracting, for all z

$$0 = \int (h(x) - h^*(x)) f_{X|Z}(x|z) dx$$

The nonparametric IV model: continuous X

- A nonparametric model with continuous X (Newey & Powell (2003)) and additive U

$$Y = h(X) + U \qquad E[U|Z = z] = 0 \qquad z \in \mathcal{R}_Z$$

$$E[Y|Z = z] = E[h(X)|Z = z] \equiv \int h(x) f_{X|Z}(x|z) dx. \qquad (*)$$

- Suppose $h(\cdot)$ and $h^*(\cdot)$ both solve (*), that is, for all z:

$$\begin{aligned} E[Y|Z &= z] = \int h(x) f_{X|Z}(x|z) dx \\ E[Y|Z &= z] = \int h^*(x) f_{X|Z}(x|z) dx \end{aligned}$$

then, on subtracting, for all z

$$0 = \int (h(x) - h^*(x)) f_{X|Z}(x|z) dx$$

- So, the model (*) identifies $h(\cdot)$ if a "completeness condition" holds:

$$\forall z \in \mathcal{R}_Z : \quad \int a(x) f_{X|Z}(x|z) dx = 0 \implies \forall x : \quad a(x) = 0$$

The nonparametric IV model: continuous X

- The IV model with continuous X

$$Y = h(X) + U \qquad E[U|Z = z] = 0 \qquad z \in \mathcal{R}_Z$$

and the completeness condition

$$\forall z \in \mathcal{R}_Z : \quad \int a(x) f_{X|Z}(x|z) dx = 0 \implies \forall x : \quad a(x) = 0$$

identifies the function $h(\cdot)$.

The nonparametric IV model: continuous X

- The IV model with continuous X

$$Y = h(X) + U \qquad E[U|Z = z] = 0 \qquad z \in \mathcal{R}_Z$$

and the completeness condition

$$\forall z \in \mathcal{R}_Z : \quad \int a(x) f_{X|Z}(x|z) dx = 0 \implies \forall x : \quad a(x) = 0$$

identifies the function $h(\cdot)$.

- Completeness restriction is on $f_{X|Z}(x|z)$ and the model is silent about the way in which X is determined.
- In the discrete endogenous explanatory variable case equivalent to restricting matrix of probabilities $P[X = x_m|z_k]$ to have rank M.
- Requires rich support for Z.
- Smoothness or parametric restrictions helpful - Chetverikov and Wilhelm (Ecta 2017).

The IV model nonadditive U

- Chernozhukov and Hansen (Ecta 2005), Chernozhukov, Imbens, Newey (JEcts 2007):

$$Y = h(X, U) \qquad U \perp\!\!\!\perp Z \qquad z \in \mathcal{R}_Z$$

h strictly increasing in *scalar* continuous $U \sim Unif(0, 1)$.

The IV model nonadditive U

- Chernozhukov and Hansen (Ecta 2005), Chernozhukov, Imbens, Newey (JEcts 2007):

$$Y = h(X, U) \qquad U \perp\!\!\!\perp Z \qquad z \in \mathcal{R}_Z$$

h strictly increasing in *scalar* continuous $U \sim Unif(0,1)$.

- h strictly increasing implies for each x

$$\{u : u \leq \tau\} = \{u : h(x,u) \leq h(x,\tau)\}$$

so

$$\mathbb{P}[U \leq \tau | X = x, Z = z] = \mathbb{P}[\underbrace{h(X,U)}_{Y} \leq h(X,\tau) | X = x, Z = z]$$

The IV model nonadditive U

- Chernozhukov and Hansen (Ecta 2005), Chernozhukov, Imbens, Newey (JEcts 2007):

$$Y = h(X, U) \qquad U \perp\!\!\!\perp Z \qquad z \in \mathcal{R}_Z$$

h strictly increasing in *scalar* continuous $U \sim \mathit{Unif}(0, 1)$.

- h strictly increasing implies for each x

$$\{u : u \leq \tau\} = \{u : h(x, u) \leq h(x, \tau)\}$$

so

$$\mathbb{P}[U \leq \tau | X = x, Z = z] = \mathbb{P}[\underbrace{h(X, U)}_{Y} \leq h(X, \tau) | X = x, Z = z]$$

- Take expectations over X given Z:

$$\mathbb{P}[U \leq \tau | Z = z] = \mathbb{P}[Y \leq h(X, \tau) | Z = z]$$

So for all $\tau \in (0, 1)$, $z \in \mathcal{R}_Z$

$$\tau = \Pr[Y \leq h(X, \tau) | Z = z]$$

The IV model nonadditive U

- We have shown: if h is strictly increasing in scalar U.

$$Y = h(X, U) \qquad U \perp\!\!\!\perp Z \qquad z \in \mathcal{R}_Z$$

then for all $\tau \in (0, 1)$, $z \in \mathcal{R}_Z$

$$\Pr[Y \leq h(X, \tau) | Z = z] = \tau$$

The IV model nonadditive U

- We have shown: if h is strictly increasing in scalar U.

$$Y = h(X, U) \qquad U \perp\!\!\!\perp Z \qquad z \in \mathcal{R}_Z$$

then for all $\tau \in (0, 1)$, $z \in \mathcal{R}_Z$

$$\Pr[Y \leq h(X, \tau)|Z = z] = \tau$$

- There is **point** identification of h if restrictions on h and $F_{UX|Z}$ guarantee a unique solution $h^*(\cdot, \cdot) = h(\cdot, \cdot)$ to:

$$\forall \tau, z : \quad \int 1[h(x, u) \leq h^*(x, \tau)] f_{UX|Z}(u, x|z) du dx = \tau$$

The IV model nonadditive U

- We have shown: if h is strictly increasing in scalar U.

$$Y = h(X, U) \qquad U \perp\!\!\!\perp Z \qquad z \in \mathcal{R}_Z$$

then for all $\tau \in (0, 1)$, $z \in \mathcal{R}_Z$

$$\Pr[Y \leq h(X, \tau) | Z = z] = \tau$$

- There is **point** identification of h if restrictions on h and $F_{UX|Z}$ guarantee a unique solution $h^*(\cdot, \cdot) = h(\cdot, \cdot)$ to:

$$\forall \tau, z : \quad \int 1[h(x, u) \leq h^*(x, \tau)] f_{UX|Z}(u, x|z) du dx = \tau$$

- A unique solution is guaranteed if this completeness condition holds:

$$\forall z : \int a(x, u) f_{UX|Z}(u, x|z) du dx = 0 \implies \forall x, u : a(x, u) = 0$$

The IV model nonadditive U

- A unique solution is guaranteed if completeness condition holds:

$$\forall z : \int a(x, u) f_{UX|Z}(u, x|z) du dx = 0 \implies \forall x, u : a(x, u) = 0$$

The IV model nonadditive U

- A unique solution is guaranteed if completeness condition holds:

$$\forall z : \int a(x,u) f_{UX|Z}(u,x|z)dudx = 0 \implies \forall x,u : a(x,u) = 0$$

- To see why, suppose $h(\cdot,\cdot)$ and $h^*(\cdot,\cdot)$ are solutions

$$\forall \tau, z \quad : \quad \int 1[h(x,u) \leq h(x,\tau)] f_{UX|Z}(u,x|z)dudx = \tau$$
$$\forall \tau, z \quad : \quad \int 1[h(x,u) \leq h^*(x,\tau)] f_{UX|Z}(u,x|z)dudx = \tau$$

$$\implies \forall \tau, z : \quad \int a(z,u) f_{UX|Z}(u,x|z)dudx = 0$$

where

$$a(x,u) = 1[h(x,u) \leq h(x,\tau)] - 1[h(x,u) \leq h^*(x,\tau)]$$

The IV model nonadditive U

- If h strictly increasing in scalar U.

$$Y = h(X, U) \qquad U \perp\!\!\!\perp Z \qquad z \in \mathcal{R}_Z$$

then for all $\tau \in (0,1)$, $z \in \mathcal{R}_Z$

$$\Pr[Y \leq h(X, \tau) | Z = z] = \tau$$

  - point identification of h requires rich support for Z.
  - local independence: could just have $Q_{U|Z}(\tau|z) = \tau$ for $\tau \in \mathcal{T}$, then we may be able to identify $h(x, \tau)$ for $\tau \in \mathcal{T}$.
  - estimation - for any random variable S

$$P[S \leq Q_{S|Z}(\tau|z) | Z = z] = \tau$$

so, find $h(\cdot, \cdot)$ such that the quantile regression of $Y - h(X, \tau)$ on z does not depend on z.

Other types of incomplete models

- We have considered single equation models for an outcome Y_1 (scalar) with potentially endogenous explanatory variables Y_2 in general form:

$$Y_1 = s(Y_2, U)$$

U is an unobserved latent variable, $Y_2 \perp\!\!\!\perp U$ may not hold, but maybe $U \perp\!\!\!\perp Z$.

- This is a special case of a class of models in which values of endogenous variables $Y = (Y_1, \ldots, Y_M)$ are solution(s) to

$$h(Y, Z, U) = 0$$

and U is a vector of unobservable variables and Z is a vector of observable exogenous variables.

Example: Kline and Tamer (2016) (KT16)

- Data on 7882 markets - air routes with two airline types, Low Cost Carriers(LCC) and Other Airlines (OA).
- Binary Y_{LCC} and Y_{OA} indicate the presence of respectively LCC and OA operating on an air route in the USA. There are exogenous variables listed in vector $Z \in \mathcal{R}_Z$, structural equations

$$\begin{aligned} Y_{LCC} &= 1\left[Z\beta_{LCC} + Y_{OA}\,\Delta_{LCC} + U_{LCC} > 0\right] \\ Y_{OA} &= 1\left[Z\beta_{OA} + Y_{LCC}\,\Delta_{OA} + U_{OA} > 0\right] \end{aligned}$$

- This type of model is studied in many papers including Heckman (1978), Bresnahan and Reiss (1990,1991), and Tamer (2003).
- In KT16 and most other applications of this model U is restricted to be normally distributed independent of Z.

GIV structural functions

- For example in the Kline Tamer model

$$h((y_{LCC}, y_{OA}), z, (u_{LCC}, u_{OA})) =$$

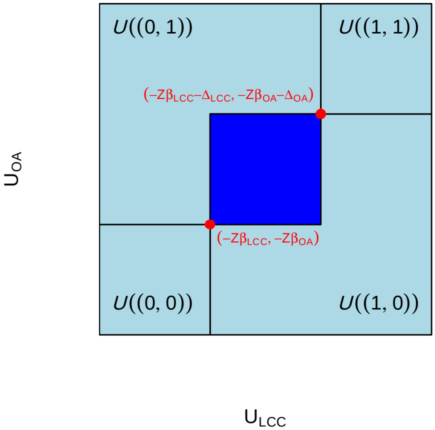
$$y_{LCC} \cdot \max(-(z\beta_{LCC} + y_{OA}\Delta_{LCC} + u_{LCC}), 0) +$$

$$(1 - y_{LCC}) \cdot \max(z\beta_{LCC} + y_{OA}\Delta_{LCC} + u_{LCC}, 0) +$$

$$y_{OA} \cdot \max(-(z\beta_{OA} + y_{LCC}\Delta_{OA} + u_{OA}), 0) +$$

$$(1 - y_{OA}) \cdot \max(z\beta_{OA} + y_{LCC}\Delta_{OA} + u_{OA}, 0)$$

$\Delta_{\mathrm{LCC}} < 0$ and $\Delta_{\mathrm{OA}} < 0$



Example: Mazzeo (2002) (M02)

- Two types (T) of motel operator: low (L) and high (H) quality.
- M02 data records the number of motels $y = (y_L, y_H)$ and exogenous z at 492 freeway exits.
- At a location with y_L **low** and y_H **high** quality motels the profit on opening an **additional** motel of type $T \in \{H, L\}$ is

$$\pi_T(z, y_L, y_H, u_T) = g_T(z, y_L, y_H) + u_T$$

  with z exogenous and $u = (u_L, u_H)$ not observed by the econometrician.
- If there are nonnegative profits at all locations then at a location with y_L low and y_H high quality motels

$$\pi_L(z, y_L, y_H, u_L) < 0 \qquad \pi_H(z, y_L, y_H, u_H) < 0$$

$$\pi_L(z, y_L - 1, y_H, u_L) > 0 \qquad \pi_H(z, y_L, y_H - 1, u_H) > 0$$

  - Simply positive profit conditions - minimal restrictions a la Haile and Tamer 2003). Also arise as conditions under which (y_L, y_H) is a simultaneous move pure strategy Nash equilibrium.

Chesher and Rosen (2017) - CR17

- Here is a suitable structural function for the Mazzeo model

$$
\begin{aligned}
h(y,z,u) \equiv\ & \max(g_L(z,y_L,y_H)+u_L,0)+\max(g_H(z,y_L,y_H)+u_H,0) \\
& -\min(g_L(z,y_L-1,y_H)+u_L,0)-\min(g_H(z,y_L,y_H-1)+u_H,0)
\end{aligned}
$$

- This has

$$h(y,z,u)=0$$

if and only if:

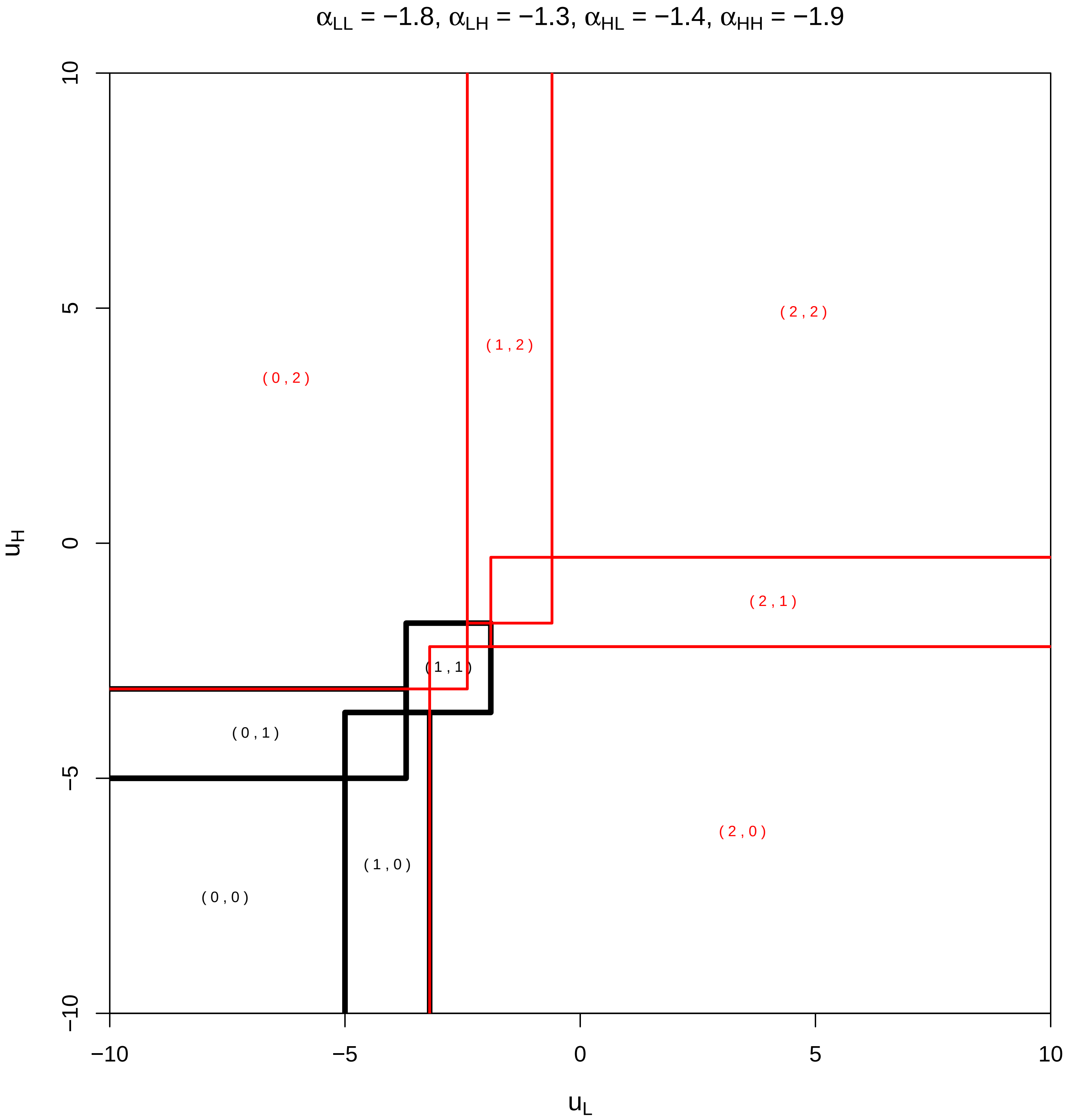
$$
\begin{array}{rr}
g_L(z,y_L,y_H)+u_L<0 & \qquad g_H(z,y_L,y_H)+u_H<0 \\
g_L(z,y_L-1,y_H)+u_L>0 & \qquad g_H(z,y_L,y_H-1)+u_H>0
\end{array}
$$

Restrictions in Mazzeo (2002)

- M02 restricts $g_T(z, y_L, y_H)$ to be a linear index, U and Z independent, $U \sim N_2(0, \Sigma)$, so $\mathcal{G}_{U|Z}$ becomes G_U.
  - here in illustrative calculations

$$\begin{aligned} g_L(z, y_L, y_H) &= z'\beta_L + \alpha_{LL} y_L + \alpha_{LH} y_H \\ g_H(z, y_L, y_H) &= z'\beta_H + \alpha_{HL} y_L + \alpha_{HH} y_H \end{aligned}$$

- In M02 there is censoring: $y_T = 3$ denotes 3 *or more* motels of type $T \in \{L, H\}$. We further censor, $y_T = 2$ signifies 2 or more motels of type T.



$\alpha_{LL} = -1.8, \alpha_{LH} = -1.3, \alpha_{HL} = -1.4, \alpha_{HH} = -1.9$
(0 , 2)
(1 , 2)
(2 , 2)
(2 , 1)
(1 , 1)
(0 , 1)
(2 , 0)
(1 , 0)
(0 , 0)
u_H
u_L
−10
−5
0
5
10

Incomplete models for discrete outcomes

- In the last lecture we study incomplete models in which values of *discrete* endogenous variables Y are solution(s) to

$$h(Y, Z, U) = 0$$

  where U is a vector of unobservable variables and Z is a vector of observable exogenous variables.

- When Y is discrete, unobserved U *cannot* be written as a single-valued function of (Y, Z).

Frame Title

- In models restricting U to be a *function* of (Y, Z), developing identifying correspondences is often straightforward.

- Consider models with structural functions $h(y, z, u)$ such that there exists a function $g(y, z)$ such that

$$\forall y, z \quad h(y, z, g(y, z)) = 0$$

  that is $u = g(y, z)$.

  - Models that impose the restriction $E[UZ] = 0$ identify the set of functions

$$\{g : E[g(Y, Z)Z] = 0\}$$

  - Models that impose the restriction $U \perp\!\!\!\perp Z$ identify the set of functions

$$\left\{g : g(Y, Z) \perp\!\!\!\perp Z\right\}$$

- But when U is a *set-valued* function of Y and Z this simple approach is not available.

Incomplete models for discrete outcomes

- In the last lecture we characterize identified sets delivered by incomplete models in which U is a set-valued function of observed Y and Z.
  - Also applies when models are complete and/or U is a function of Y and Z.
- We will study data used in Angrist and Evans (AER 1998) to estimate the Local Average Treatment Effect (LATE) of an additional child on labour force participation of married females.
- The LATE is point identified using the "same-sex" and "twins" instruments. We develop sharp bounds on the Average Treatment Effect (ATE).