Semi and Nonparametric Econometrics

Part 2: IV methods in nonparametric/nonlinear models

Xavier D'Haultfœuille

ENSAE and Paris Saclay Master

Outline

Outline

The importance of IV

- Assessing the causal effect of a factor X on an outcome Y is one of the most difficult task in social sciences (economics but also epidemiology, sociology...).
- The main problem is that X is seldom affected randomly, but rather chosen, at least partly. This choice may then be related to unobserved factors that also affect Y: the endogeneity problem.
- Other sources of endogeneity: measurement error on X, simultaneity.
- One of the most common way to tackle this issue is to use instrumental variables (IV), namely variables affecting X but not directly Y.
- Intuition behind: the variations of X induced by Z are exogenous and can thus be used to identify the causal effect of X.

The importance of nonlinear and nonparametric models

- Why focusing on nonlinear models? Very pervasive:
  - Discrete choice models: models based on the maximization of a random utility are nonlinear (logit, probit, multinomial logit...).
  - Other limited dependent variable models are also nonlinear: censored models or integer valued variables.
  - Using quantile restrictions make the model nonlinear as well.
- Why nonparametric? Theory may predict shape restrictions (monotonicity, convexity/concavity...) but rarely the functional form of the dependence between X and Y.
- Important to understand if identification stems from the functional form restrictions or the instrumental variable itself.

A search for "universal solution"

- The linear model, where the situation is simple, provides insights on general solutions to handle IV estimation in more complex cases.
- In the linear case, three equivalent ways can be used to define β_0, the slope parameter of X.
- Two of them will extend to nonlinear/nonparametric models. However, they are not equivalent anymore, neither in terms of identification nor for estimation.
- We consider hereafter nonparametric models. In general, semiparametric identification / estimation can be easily treated as particular cases.

Outline

Introduction

- Consider the IV linear model:

$$Y = X'\beta_0 + \varepsilon, \quad E(Z\varepsilon) = 0.$$

- In this model, there are three equivalent ways to define β_0:
  1. through a projection;
  2. through an estimating equation;
  3. through a control variable approach.
- The conditions for identification are the same in the three cases. The corresponding estimator are also the same for 1 and 3, but not necessarily for 2.

Projection

- In the first, we project linearly X on Z:

$$X = \Gamma_0 Z + \nu, \text{ with } E(Z\nu) = 0$$

- Then, instead of regressing Y on X, we regress Y on $\widehat{X} = \Gamma_0 Z$:

$$\begin{aligned} Y &= X'\beta_0 + \varepsilon \\ &= \widehat{X}'\beta_0 + \nu'\beta_0 + \varepsilon. \end{aligned}$$

- In this regression, $\widehat{X}$ is exogenous because $E(Z\nu) = 0$ and $E(Z\varepsilon) = 0$.
- Identification is ensured as soon as the regressors $\widehat{X}$ are linearly independent, or, equivalently $E(ZX')$ being full rank.
- This idea directly translates into the 2SLS estimator.

Estimating equation

- The second way is to write:

$$E(ZY) = E(ZX')\beta_0,$$

  and solve the linear equation to find β_0.
- Identification directly follows from the standard condition of $E(ZX')$ being full rank.
- An estimator following from this strategy is the GMM, since $E[Z(Y - X'\beta_0)] = 0$.
- Note that the GMM estimator is equal to the 2SLS when $\dim(X) = \dim(Z)$, but they may not coincide in the overidentified case where $\dim(X) < \dim(Z)$.

The control variable approach

- The third way to identify β_0 is to project ε on $\nu(= X - \Gamma_0 Z)$:

$$\varepsilon = \nu'\delta_0 + \zeta, \text{ with } E(\nu\zeta) = 0.$$

- Then we regress Y on (X, ν):

$$Y = X'\beta_0 + \nu\delta_0 + \zeta.$$

- Regressors are exogenous because $E(\nu\zeta) = 0$ and

$$E(X\zeta) = E((\Gamma_0 Z + \nu)\zeta) = \Gamma_0 E(Z(\varepsilon - \nu\delta_0)) = 0.$$

- Intuition behind: by exogeneity of Z, ν contains all the endogeneity of X. Once we control for ν in the regression, X is exogenous.
- Identification is ensured as soon as X and ν are not linearly dependent, which once more is equivalent to $E(ZX')$ being full rank.
- The corresponding estimator is, as in the first case, the 2SLS estimator.

Outline

Why this fails in general

- Consider a general model

$$Y = \varphi(X, \varepsilon),$$

  where Z is exogenous, namely satisfy some restrictions w.r.t. ε: mean independence $E(\varepsilon|Z) = 0$, quantile independence $q_{\varepsilon|Z}(\tau) = 0$, full independence $Z \perp\!\!\!\perp \varepsilon$... We denote this restriction by $r(f_{Z,\varepsilon}) = 0$.
- Then $Y = \varphi(\widehat{X} + \nu, \varepsilon)$ but in general there exists no ζ such that

$$\varphi(\widehat{X} + \nu, \varepsilon) = \varphi(\widehat{X}, \zeta), \text{ with } r(f_{\widehat{X},\zeta}) = 0.$$

- This works in the linear model where $\varphi(X, \varepsilon) = X'\beta_0 + \varepsilon$ and $E(Z\varepsilon) = 0$ but not in general when φ or r are nonlinear.

A first example: quadratic model with a mean restriction

- Suppose that $\varphi(X, \varepsilon) = \alpha_0 + X\beta_0 + X^2\gamma_0 + \varepsilon$ and the regression of X on Z is heteroskedastic:

$$X = \widehat{X}(1 + \widetilde{\nu}), \quad \text{with } \widetilde{\nu} \perp\!\!\!\perp \widehat{X}.$$

- Then:

$$Y = \alpha_0 + \widehat{X}\beta_0 + \widehat{X}^2\gamma_0 + \left[\widehat{X}^2\widetilde{\nu}^2\gamma_0 + \left(\varepsilon + \widehat{X}(\beta_0 + 2\widehat{X}\gamma_0)\widetilde{\nu}\right)\right]$$

- The first term into the brackets is correlated with $\widehat{X}$ and induces a bias in the regression.

Another example: linear model with a quantile restriction

- Consider the model
$$Y = X'\beta_\tau + \varepsilon_\tau, \quad \text{with } q_\tau(\varepsilon_\tau|Z) = 0. \qquad (1)$$
- This is the same idea as linear IV models, except that we replace $E(\varepsilon|Z) = 0$ by a quantile restriction.
- In such models, some people have proposed (i) to regress X on Z and (ii) to run a quantile regression of Y on the projection $\widehat{X}$.
- However, this is valid only under the very weird condition that
$$q_\tau(\varepsilon_\tau + (X - \widehat{X} - q_\tau(X - \widehat{X}))\beta_\tau|Z) = 0,$$
- This does not hold in general, even when $q_\tau(\varepsilon_\tau|Z) = 0$ and $q_\tau(X - \widehat{X}|Z) = q_\tau(X - \widehat{X})$ because in general,
$$q_\tau(U + V) \neq q_\tau(U) + q_\tau(V).$$
- Thus, this method leads in general to an inconsistent estimator.

Outline

A first generalization to nonparametric additive models

- Instead of $Y = X'\beta_0 + \varepsilon$, consider the nonparametric additive model (see Newey and Powell, 2003, and Darolles et al., 2005)

$$Y = \varphi(X) + \varepsilon, \text{ with } E(\varepsilon|Z) = 0.$$

- Then one can identify $\varphi(.)$ through the estimating equation:

$$E(Y|Z) = E(\varphi(X)|Z)$$

or, equivalently, the conditional moment condition

$$E(Y - \varphi(X)|Z) = 0.$$

A first generalization to nonparametric additive models

- The identifying condition is

$$E(g(X)|Z) = 0 \Rightarrow g(X) = 0. \tag{2}$$

- This is known as the *completeness condition* (because of the link with complete statistics).
- Condition (2) is far less intuitive than in the linear case. Suppose for instance that $X = Z + U$:
  - Then if $U \sim \mathcal{N}(0, \sigma^2)$, the completeness condition holds;
  - But if $U \sim \mathcal{U}[-1/2, 1/2]$, it fails to hold because there are periodic functions for which

$$\int_{-1/2}^{1/2} g(z+u)du = 0 \quad \forall z.$$

- Not much is known about this condition: see Newey and Powell (2003) and D'Haultfœuille (2011) for sufficient conditions.

A first generalization to nonparametric additive models

- Note that this model is not well suited when Y is limited, and $Y = g(\mu(X) + \varepsilon)$. On the other hand, X can be limited.
- As for estimation, this is a rather difficult problem since we have to solve an infinite dimensional inverse problem.
- A simple solution is to rely on "sieve estimation", namely replace the nonparametric model by a parametric one, but of growing dimension.
- For instance, we could approximate φ by a polynomial of degree $k_n \to \infty$ at an appropriate speed.
- Then, to estimate φ, we would simply solve the empirical counterpart of the moment conditions

$$E\left[Z^k\left(Y - \sum_{j=0}^{k_n} \lambda_j X^j\right)\right] = 0 \text{ with } 1 \leq k \leq K(\geq k_n).$$

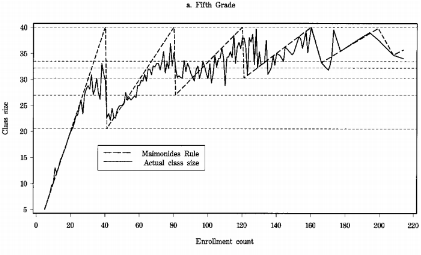
Nonparametric additive models: example

- Example: revisiting Angrist & Lavy (1999)'s paper on the effect of class size on students' achievement.
- Idea of A & L: use an exogenous rule on class openings to build an IV. In Israel, such a rule, established by Mainmonides in the 12th century, states that classroom size cannot not exceed 40.
- This implies that if there are 80 pupils of a given cohort in a school, there may be only 2 classrooms (each of size 40), but there should be at least 3, of average size 27, with 81 such pupils.
- Let S denote the cohort size. A& L use ($[x]$ =integer part of x):

$$Z = S/([(S-1)/40]+1),$$

the expected average classroom size, as an instrument for the class size X.

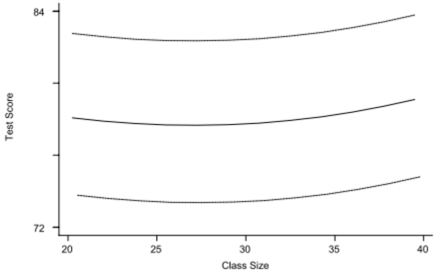
Nonparametric additive models: example



Link between predicted and average class size (taken from A& L)

- Using 2SLS, A& L show that increasing class size by ten students decreases students' average achievement by 4 to 6 points, for average scores around 60-70.
- It is unlikely, however, that class size has a linear effect on test scores. Horowitz (2011) revisit their paper, by supposing instead that $Y = \varphi(X) + \varepsilon$.

Nonparametric additive models: example



Estimated effect between class size and students' achievement, under a nonparametric model (taken from Horowitz, 2011)

- Conclusion: when using nonparametric methods, we do not find a significant effect anymore.
- Important policy implications!

A second generalization to nonadditive models

- Consider a nonadditive model (see Chernozhukov and Hansen, 2005):
$$Y = \varphi(X, \varepsilon), \quad \text{with } q_\tau(\varepsilon|Z) = q_\tau(\varepsilon). \tag{3}$$
and $\varphi(x, .)$ is strictly increasing. The condition $q_\tau(\varepsilon|Z) = q_\tau(\varepsilon)$ is a "quantile independence" restriction similar to the mean independence condition $E(\varepsilon|Z) = 0$.
- We can suppose without loss of generality (provided that ε is continuous) that $\varepsilon \sim \mathcal{U}[0, 1]$.
- Then

$$\begin{aligned} \tau &= P(\varepsilon \leq \tau) = P(\varepsilon \leq \tau|Z) \\ &= P(\varphi(X, \varepsilon) \leq \varphi(X, \tau)|Z) = P(Y \leq \varphi(X, \tau)|Z). \end{aligned}$$

A second generalization to nonadditive models

- Thus, $\varphi(., \tau)$ solves the conditional moment conditions:

$$E\left(\mathbb{1}\{Y \leq \varphi(X, \tau)\} - \tau | Z\right) = 0. \tag{4}$$

- Identification of $\varphi(., \tau)$ based on (4) is even more complicated to establish than in the additive case $Y = \varphi(X) + \varepsilon$. It is known to hold only in very particular cases.
- Estimation is also more difficult than previously because $g \mapsto E\left(\mathbb{1}\{Y \leq g(X, \tau)\} - \tau | Z\right)$ is not linear. Several solutions proposed recently: Chernozhukov, Imbens and Newey (2007), Horowitz and Lee (2007) and Chen and Pouzo (2012).
- Though model (3) generalizes the previous additive model, it still cannot handle limited Y. For a binary threshold model for instance, $Y = \mathbb{1}\{X'\beta_0 + \varepsilon \geq 0\}$ so that $\varphi(x, \varepsilon) = \mathbb{1}\{x'\beta_0 + \varepsilon \geq 0\}$ is not strictly increasing in ε.

Semiparametric example: quantile IV models

- Suppose also that $\varphi(X,\varepsilon) = X'\beta_\varepsilon$, with $u \mapsto x'\beta_u$ strictly increasing.
- This is the same as the linear quantile IV model (1). We then get
$$E\left[\mathbb{1}\{Y \leq X'\beta_\tau\} - \tau | Z\right] = 0,$$
which implies that for any g,
$$E\left[(\mathbb{1}\{Y \leq X'\beta_\tau\} - \tau)g(Z)\right] = 0. \quad (5)$$
- Identification of β_τ holds if there exists g such that (5) has a unique solution. As with linear IV, this requires X and Z to be related, but the conditions are more difficult to write formally because of the nonlinearity of the equations.
- The first idea to estimate β_τ would be to do some GMM, using $K \geq$dim(X) real functions $g_1, ..., g_K$.
- Problem: the moment conditions (5) are discontinuous in β_τ. They are therefore difficult to solve numerically.

Semiparametric example: quantile IV models

- Computationally convenient method proposed by Chernozhukov and Hansen (2006, 2008). Let $X = (X_0, X_1)$ where X_0 is endogenous while X_1 is exogenous, let $\beta_\tau = (\alpha_0, \beta_0)$ be the corresponding parameters. Then:

$$Y - X_0'\alpha_0 = X_1'\beta_0 + Z_0'0 + \varepsilon_\tau, \quad q_\tau(\varepsilon_\tau | X_1, Z_0) = 0.$$

- In other words,

$$(\beta_0, 0) = \arg\min_{(\beta,\gamma)} E\left[\rho_\tau(Y - X_0'\alpha_0 - X_1'\beta - Z_0'\gamma)\right] \qquad (6)$$

- For a given value of α (not necessarily α_0), it is easy to obtain the parameters of the quantile regression of $Y - X_0'\alpha$ on (X_1, Z_0). Let $\beta(\alpha)$ and $\gamma(\alpha)$ be the corresponding parameters.
- Then the idea is to choose α such that $\gamma(\alpha)$ is “small”.

Semiparametric example: quantile IV models

- In practice, define a grid on α, $\{\alpha_1, ..., \alpha_J\}$. Then, for $j = 1$ to J:
  - Compute the quantile regression of $Y - X_0'\alpha_j$ on (X_1, Z_0). Let $(\widehat{\beta}(\alpha_j), \widehat{\gamma}(\alpha_j))$ be the corresponding estimators.
  - Compute the Wald statistic corresponding to the test of $\gamma(\alpha_j) = 0$:

$$W_n(\alpha_j) = n\widehat{\gamma}(\alpha_j)' \widehat{V}_{as}^{-1}(\widehat{\gamma}(\alpha_j))\widehat{\gamma}(\alpha_j).$$

- Then define the estimator of α_0 as

$$\widehat{\alpha} = \arg\min_{j=1...J} W_n(\alpha_j)$$

  and $\widehat{\beta} = \widehat{\beta}(\widehat{\alpha})$.

- See Chernozhukov and Hansen (2006) for the asymptotics and inference, and Christian Hansen's webpage for the Matlab code.
- N.B.: the method is especially convenient when $\dim(\alpha)$ is low (1 or 2), otherwise it may be time consuming.

Quantile IV: an example

- There has been much debate on the efficiency of subsidized training programs (classroom training, on-the-job training, job search assistance...) on earnings.
- The usual problem for evaluating its causal effect is endogeneity (why here?).
- Abadie et al. (2002) use a large random experiment conducted in the US on the Job Training Partnership Act (JTPA).
- In this experiment, 11,202 people were assigned randomly in a "treatment" or "control". However, among people of the treatment group, only 60% actually receive training. Thus, receiving training is probably endogenous.
- On the other hand, the experiment provides us with a valid instrument.

Quantile IV: an example

Here are the results obtained by Abadie et al. (2002):

Impact of training on 30-month earnings (in percentage of earnings)

	Men		Women	
Method	Without IV	IV	Without IV	IV
Linear reg.	21.2	8.6	18.5	14.6
$q_{0.15}$	135.6	5.2	60.8	35.5
$q_{0.25}$	75.2	12.0	44.4	23.1
$q_{0.50}$	34.5	9.6	32.3	18.4
$q_{0.75}$	17.2	10.7	14.5	10.1
$q_{0.85}$	13.4	9.0	8.1	7.4

Outline

A first generalization to additive model

- Consider the model (see Newey, Powell and Vella, 1999):

$$\begin{cases} Y &= \varphi(X) + \varepsilon \\ X &= \psi(Z) + \nu \end{cases} \qquad Z \perp\!\!\!\perp (\varepsilon, \nu) \text{ (also, } E(\varepsilon) = 0)$$

- The exogeneity condition on Z implies that $E(\varepsilon|X,\nu) = E(\varepsilon|\nu)$. Then:

$$\begin{aligned} E(Y|X,\nu) &= \varphi(X) + E(\varepsilon|\nu, X) \\ &= \varphi(X) + E(\varepsilon|\nu). \end{aligned}$$

- We can identify φ by 1) regressing nonparametrically X on Z to obtain ν and 2) regressing nonparametrically Y on X and ν.

A first generalization to additive model

- Identification is secured if

$$h_1(X) + h_2(\nu) = 0 \Rightarrow h_1(X) = -h_2(\nu) = \text{constant}.$$

- This is a mild restriction that holds if Z is continuous but even with a discrete (at least ternary) Z.
- Thus this approach requires far less than the completeness condition in terms of the dependence between X and Z.
- On the other hand, it is more restrictive than the estimating equation approach on the instrument: $Z \perp\!\!\!\perp (\varepsilon, \nu)$ is stronger than $E(\varepsilon|Z) = 0$.
- Also, this approach rules out limited X or limited Y.

A first generalization to additive model

- To estimate φ, we first estimate ν_i. We regress nonparametrically X on Z and let $\widehat{\nu}_i = X_i - \widehat{\psi}(Z_i)$.
- We then have to recover φ in the nonparametric additive model $E(Y|X,\nu) = \varphi(X) + g(\nu)$, with $E[g(\nu)] = 0$.
- A first solution is *marginal integration*, which is based on the following equality:

  $$\int E(Y|X,\nu = u)dF_\nu(u) = \varphi(X) + E(g(\nu)) = \varphi(X).$$

  Then: 1) estimate by a kernel estimator $\widehat{E}(Y|X,\widehat{\nu})$ and 2) define $\widehat{\varphi}(.)$ by:

  $$\widehat{\varphi}(x) = \frac{1}{n}\sum_{i=1}^{n} \widehat{E}(Y|X = x, \widehat{\nu} = \widehat{\nu}_i)$$

- Another solution is to rely on sieves: regress Y on (separate) functions of X and functions of $\widehat{\nu}$.

A second generalization to nonadditive model

- Consider the nonadditive model (see Imbens and Newey, 2009):

$$\begin{cases} Y = \varphi(X, \varepsilon) \\ X = \psi(Z, \nu) \end{cases} \quad Z \perp\!\!\!\perp (\nu, \varepsilon)$$

- Suppose also that $\psi(Z, .)$ is strictly increasing.
- In this model it is difficult to recover φ directly. However, we can recover other quantities of interest such as:
    - the average structural function $E(\varphi(x, \varepsilon))$ or quantile structural function $\tau \mapsto q_\tau(\varphi(x, \varepsilon))$, i.e. averages or quantiles if everybody had $X = x$;
    - Average effects if X moved to $\ell(X)$: $E[\varphi(\ell(X), \varepsilon) - Y]$;
    - Average marginal effects $\Delta = E\left[\frac{\partial \varphi}{\partial x}(X, \varepsilon)\right]$.
- We focus hereafter on Δ. The identification of structural functions requires far more restrictive support conditions.

A second generalization to nonadditive model

- First, we suppose without loss of generality that ν is uniform. Let $\psi_2^{-1}(z,.)$ denote the inverse of $\psi(z,.)$. Then

$$\begin{aligned} F_{X|Z}(x|Z) &= P(X \leq x|Z) = P(\psi(Z,\nu) \leq x|Z) \\ &= P(\nu \leq \psi_2^{-1}(Z,x)|Z) = \psi_2^{-1}(Z,x). \end{aligned}$$

  Thus, $\nu = \psi_2^{-1}(Z,X) = F_{X|Z}(X|Z)$ is identified (as a generalized residual).

- Second, because as previously, $X \perp\!\!\!\perp \varepsilon|\nu$,

$$E(Y|X,\nu) = \int \varphi(X,e)dF_{\varepsilon|\nu}(e).$$

- Therefore, under regularity conditions,

$$\frac{\partial E(Y|X,\nu)}{\partial x} = \int \frac{\partial \varphi}{\partial x}(X,e)dF_{\varepsilon|\nu}(e|\nu) = E\left[\frac{\partial \varphi}{\partial x}(X,\varepsilon)|X,\nu\right].$$

A second generalization to nonadditive model

- Hence,
$$E\left[\frac{\partial E(Y|X,\nu)}{\partial x}\right] = E\left[E\left[\frac{\partial \varphi}{\partial x}(X,\varepsilon)|X,\nu\right]\right] = \Delta.$$
- As a result, Δ is identified under rather mild restrictions.
- Main one: conditional on ν, X should have a continuous distribution $\Rightarrow Z$ should be continuous as well.
- We can estimate Δ by a three step procedure:
  - estimate ν. For that we run a nonparametric regression of $\mathbb{1}\{X \leq x\}$ on Z, for several x. We obtain an estimator $\widehat{F}_{X|Z}$ of $F_{X|Z}$ and then let $\widehat{\nu}_i = \widehat{F}_{X|Z}(X_i|Z_i)$.
  - run a nonparametric regression of Y on X and $\widehat{\nu}$ and takes its derivative wrt x to get $\partial \widehat{E}(Y|X,\nu)/\partial x$.
  - Define
$$\widehat{\Delta} = \frac{1}{n}\sum_{i=1}^{n} \frac{\partial \widehat{E}(Y|X,\nu)}{\partial x}(X_i,\widehat{\nu}_i).$$

Application to the "IV probit" model

- Following Rivers and Vuong (1989), consider the case $\varphi(x, \varepsilon) = \mathbb{1}\{x'\beta_0 + \varepsilon \geq 0\}$, with $\varepsilon \sim \mathcal{N}(0, 1)$.
- Suppose that $X = (X_0, X_1)$ where $X_1 \in \mathbb{R}^k$ is exogenous but $X_0 \in \mathbb{R}$ is endogenous (and $\beta_0 = (\beta_{01}, \beta_{02})$).
- Suppose also that we have an instrument Z affecting X_0 but not Y directly. Specifically, assume that

$$X_0 = X_1\gamma_1 + Z\gamma_2 + \nu,$$

with $(X_1, Z) \perp\!\!\!\perp (\varepsilon, \nu)$, $\varepsilon = \rho\nu + \eta$ and $\eta|\nu \sim \mathcal{N}(0, \sigma^2)$. The last two conditions hold if (ε, ν) is gaussian, but they are weaker.

Application to the "IV probit" model

- Under these restrictions what we saw above simplifies drastically:
  - ν can be estimated by the residual of a linear regression of X_0 on (X_1, Z).
  - We have
    $$Y = \mathbb{1}\{X_0\beta_{00} + X_1\beta_{01} + \rho\nu + \eta \geq 0\},$$
    where, under the conditions above, $\eta \perp\!\!\!\perp (X_0, X_1, \nu)$ and $\eta \sim \mathcal{N}(0, \sigma^2)$. Thus, we can estimate $(\beta_{00}/\sigma, \beta_{01}/\sigma, \rho/\sigma)$ by a simple probit of Y on $(X_1, X_2, \widehat{\nu})$.
  - Remark that
    $$\frac{\partial \widehat{E}(Y|X,\nu)}{\partial x_0} = \frac{\beta_{00}}{\sigma}\varphi\left(\frac{X_0\beta_{00} + X_1\beta_{01} + \rho\nu}{\sigma}\right).$$
  - Thus, we can get $\widehat{\Delta}$ using the usual formula of average marginal effects in probit models.

Conclusion

- The nonseparable model is convenient because it imposes no restriction on φ. Thus, we can handle limited Y.
- On the other hand, $\psi(Z,.)$ is strictly increasing, which imposes X to be continuous...
- To sum up, we have solutions for either Y limited and X continuous (with control variables) or Y continuous and X limited (with estimating equations), but not when both Y and X are limited.
- To date, there is no "universal" solution for this problem. Particular solutions do exist, however:
  - Fully parametric models such as biprobit models;
  - Approaches based on "special regressors", see e.g. Lewbel (2000);
  - Use of control variable or estimating equations, but providing partial identification only (Chesher, 2010, Shaikh and Vytlacil, 2010...).