ALEX VON HAFFTEN

ECON 710 EXAM MAY 6, 2021

(1) For β, we can find the MOM estimator:

$$E[Ze]=0 \Rightarrow E[Z(Y-X'\beta)]=0$$

Replace w/ sample analog

$$\frac{1}{n}\sum_{i=1}^{n}[Z_i(Y_i - X_i'\hat{\beta})]=0$$

$$\Rightarrow \hat{\beta} = \left(\frac{1}{n}\sum_{i=1}^{n} Z_i X_i'\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} Z_i Y_i\right)$$

Thus, we can estimate $\hat{\beta}$ using IV estimator.

From $\hat{\beta}$, we can derive residuals: $\hat{e}_i = Y_i - X_i'\hat{\beta}$.

For γ, the functional form is known and it is nonlinear in γ. Thus, we can use NLLS:

$$\Rightarrow \hat{\gamma} = \arg\min_{\gamma} \frac{1}{n}\sum_{i=1}^{n}\left[\hat{e}_i^2 - \exp(\gamma' Z_i)\right]^2$$

Thus, $$\begin{pmatrix}\hat{\beta} \\ \hat{\gamma}\end{pmatrix} = \begin{pmatrix}\left(\frac{1}{n}\sum_i Z_i X_i'\right)^{-1}\left(\frac{1}{n}\sum_i Z_i Y_i\right) \\ \arg\min_{\gamma} \frac{1}{n}\sum_{i=1}^{n}\left[\hat{e}_i^2 - \exp(\gamma' Z_i)\right]^2\end{pmatrix}$$

where $\hat{e}_i = Y_i - X_i'\hat{\beta}$.

(2) Define $\ddot{y}_{it} = y_{it} - \bar{y}_t - \bar{y}_i + \bar{y}$ as the two-way transformed version of y_{it}.

$$\Rightarrow \quad \ddot{y}_{it} = \theta D_{it} + \varepsilon_{it}$$

Since $E[\varepsilon_{it} | z_{i1}, z_{i2}] = 0$, we can use IV to estimate θ.

$$\hat{\theta} = \left(\frac{1}{n} \sum_i z_{it}' D_{it} \right)^{-1} \left(\frac{1}{n} \sum z_{it}' \ddot{y}_{it} \right)$$

Here, I assume that z_{i1} and z_{i2} are the ith values of z for each period. If there are two different instruments we can use 2SLS to estimate $\hat{\theta}$. [1st stage, estimate $\hat{D}_{it} := \hat{\gamma}_0 + \hat{\gamma}_1' z_{i1} + \hat{\gamma}_2' z_{i2}$ w/ OLS. 2nd stage, estimate $\hat{\theta}$ w/ OLS: $\ddot{y}_{it} = \hat{\theta} \hat{D}_{it} + \hat{\varepsilon}_{it}$].

Per email w/ Bruce, I assume that we observe z_{i1} & z_{i2}.

(3) (a) We can estimate $m(x)$ w/ a polynomial series regression

$$m(x) \approx m_K(x) = \beta_0 + \beta_1 x + \dots + \beta_K x^K$$

We can estimate $(\beta_0, \beta_1, \dots, \beta_K)$ w/ OLS, [Since $E[e|X]=0$].

$$\Rightarrow \hat{m}_K(x) = \hat{\beta}_0 + \hat{\beta}_1 x + \dots + \hat{\beta}_K x^K$$

Thus, we can estimate $D(x)$ w/

$$\hat{D}_K(x) = \frac{\partial}{\partial x} \hat{m}_K(x)$$

$$= \hat{\beta}_1 + 2\hat{\beta}_2 x + \dots + K\hat{\beta}_K x^{K-1}$$

We can choose K using cross-validation

(b). Suppose $\theta = a(m) = D(x)$. Then $a(m_K) = a_K' \beta_K$ where $a_K = (0, 1, 2, \dots, K-1)'$. Then under some regularity conditions.

$$\frac{\sqrt{n}\,(\hat{\theta} - \theta + a(r_K))}{V_K^{1/2}} \xrightarrow{d} N(0, 1)$$

Where $V_K = a_K' Q_K^{-1} \Omega_K Q_K^{-1} a_K$

$$Q_K = E[x_{Ki} x_{Ki}']$$

$$\Omega_K = E[x_{Ki} x_{Ki}' e_i^2]$$

$$r_K(x) = m(x) - m_K(x).$$

(3) (b) Thus,

$$\hat{V}_K = a_K' \hat{Q}_K^{-1} \hat{\Omega}_K Q_K^{-1} a_K$$

where $\hat{Q}_K = \frac{1}{n} \sum_{i=1}^{n} x_{Ki} x_{Ki}'$

$$\hat{\Omega}_K = \frac{1}{n} \sum_{i=1}^{n} x_{Ki} x_{Ki}'$$

$$X_{Ki} = (x, \ldots, x^K)$$

The standard error is

$$\hat{s}(x) = \sqrt{\hat{R}(x)' \hat{V}_K \hat{R}(x)}$$

where $\hat{R}(x) = \begin{pmatrix} \frac{\partial}{\partial \beta_1} \hat{D}(x) \\ \vdots \\ \frac{\partial}{\partial \beta_K} \hat{D}(x) \end{pmatrix}$

(c) the pointwise confidence interval is

$$[\hat{D}(x) \pm 1.96\, \hat{s}(x)].$$

(4) (a) As $h \to \infty$, the bandwidth that you calculate the expected value of y increases.

Thus, $E[Y_1 | X = c]$ because $E[Y_1 | X \geq c]$ and $E[Y_0 | X = c]$ because $E[Y_0 | X < c]$. The estimated treatment effect is the difference between the outcome variable for observations above the cutoff and below:

$$\hat{\theta} = \frac{1}{\mathbb{1}\{X_i \geq c\}'\mathbb{1}\{X_i \geq c\}} \sum_{i=1}^{n} Y_i \mathbb{1}\{X_i \geq c\} - \frac{1}{\mathbb{1}\{X_i < c\}'\mathbb{1}\{X_i < c\}} \sum_{i=1}^{n} Y_i \mathbb{1}\{X_i < c\}$$

4 (b) If the bandwidth goes to zero, then the estimated treatment effect would just be the difference between the treated and untreated observations at the cut off:

$$\hat{\Theta} = Y_i - Y_j$$

5 (a) Yes, unlike an ordered probit/logit the thresholds are known to the researcher. Thus, we can identify the error variance.

(5) (b) Notice that we can rewrite the model w/ $W_i^* := \log(Y_i^*)$:

$$\Rightarrow W^* = X'\beta + e$$
$$e \mid X \sim N(0, \sigma^2)$$

$$W = \begin{cases} 1 & \text{if} \quad W^* < \log(10{,}000) \\ 2 & \text{if} \quad \log(10{,}000) \leq W^* < \log(20{,}000) \\ 3 & \text{if} \quad \log(20{,}000) \leq W^* < \log(50{,}000) \\ 4 & \text{if} \quad \log(50{,}000) \leq W^* < \log(100{,}000) \\ 5 & \text{if} \quad \log(100{,}000) \leq W^* \end{cases}$$

The response probabilities are

$$\begin{aligned} P_j(x) &= P[W = j \mid X = x] \\ &= P[\alpha_{j-1} < W^* < \alpha_j \mid X = x] \\ &= P[\alpha_{j-1} - x'\beta < e < \alpha_j - x'\beta \mid X = x] \\ &= \Phi\left(\frac{\alpha_j - x'\beta}{\sigma}\right) - \Phi\left(\frac{\alpha_{j-1} - x'\beta}{\sigma}\right) \end{aligned}$$

Where α_κ represent the thresholds

Thus, the log-likelihood function is

$$l_n(\beta, \sigma) = \sum_{j=1}^{J} \sum_{i=1}^{n} \mathbb{1}\{W_i = j\} \log P_j(x \mid \beta, \sigma)$$

We can substitute in Y_i & the thresholds.

$$
\begin{aligned}
l_n(\beta, \sigma) = \sum_{i=1}^{n} & \mathbb{1}\{Y_i = 1\} \log \Phi\left(\frac{\log(10{,}000) - x'\beta}{\sigma}\right) \\
& + \mathbb{1}\{Y_i = 2\} \log\left[\Phi\left(\frac{\log(20{,}000) - x'\beta}{\sigma}\right) - \Phi\left(\frac{\log(10{,}000) - x'\beta}{\sigma}\right)\right] \\
& + \mathbb{1}\{Y_i = 3\} \log\left[\Phi\left(\frac{\log(50{,}000) - x'\beta}{\sigma}\right) - \Phi\left(\frac{\log(20{,}000) - x'\beta}{\sigma}\right)\right] \\
& + \mathbb{1}\{Y_i = 4\} \log\left[\Phi\left(\frac{\log(100{,}000) - x'\beta}{\sigma}\right) - \Phi\left(\frac{\log(50{,}000) - x'\beta}{\sigma}\right)\right] \\
& + \mathbb{1}\{Y_i = 5\} \log\left[1 - \Phi\left(\frac{\log(100{,}000) - x'\beta}{\sigma}\right)\right]
\end{aligned}
$$