README for replication code for Numerical Illustration of IV Bounds in Section 3.2 of "Estimating Endogenous Effects on Ordinal Outcomes" by Chesher, Rosen, and Siddique.

This file accompanies the R file MTO_numerical_example_code.R used to compute the bounds on counterfactual probabilities and marginal effects presented in Table 1 of the paper.

Notation

The code computes lower and upper bounds on counterfactual probabilities and conditional marginal effects from an ordered outcome IV model. This is the model in the paper, although the code uses alternative notation. The model employed, in the notation of the code, is:

$$Y_1 = \left\{ \begin{array}{ccc} 0 & , & \alpha Y_2 + U \leq \gamma_1 \\ 1 & , & \gamma_1 \leq \alpha Y_2 + U \leq \gamma_2 \\ 2 & , & \gamma_2 \leq \alpha Y_2 + U \end{array} \right\}, \qquad U \perp\!\!\!\perp Z, \qquad U \sim N(0,1).$$

The data generating process from which population probabilities are derived, in the notation used in the code, is

$$Y_1 = \left\{ \begin{array}{ccc} 0 & , & Zb + aY_2 + U_1 \leq c_1 \\ 1 & , & c_1 \leq Zb + aY_2 + U_1 \leq c_2 \\ 2 & , & c_2 \leq Zb + aY_2 + U_1 \end{array} \right\}, \qquad Y_2 = d_0 + Zd_1 + U_2,$$

$$U \equiv (U_1, U_2) \perp\!\!\!\perp Z, \qquad U \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & s_{12} \\ s_{12} & s_{22} \end{bmatrix} \right),$$

with $\mathrm{Supp}(Z) = \{-1, 1\}$ and population parameters:

$$c_1 = -0.5, \quad c_2 = 0.8, \quad b = 0, \quad a = 1, \quad d_0 = 0, \quad d_1 = 0.5, \quad s_{12} = 0, \quad s_{22} \in \{1.00, 0.01\}.$$

Thus, underlying population parameters are denoted by Roman letters and unknown model parameters for the ordered outcome IV model are denoted using Greek letters. Here, translating to the notation used in the paper, Z denotes the full set of exogenous variables (X, Z), Y_2 is the endogenous variable W, U_1 stands in place for U, and U_2 for V. In addition there is the following translation from notation for the model and DGP used in the code to the notation of the paper: $b \sim \gamma_0$, $a \sim \beta_0$, $d_0 \sim \delta_x$, $d_1 \sim \delta_z$, $s_{12} \sim R$, $s_{22} \sim \sigma_v$, $c_1 \sim c_{0,1}$, $c_2 \sim c_{0,2}$, $\gamma_1 \sim c_1$, $\gamma_2 \sim c_2$, $\alpha \sim \beta$. For notation from the paper, subscripts of 0 denote true population DGP values as opposed to parameters. So, for example, β_0 denotes the population DGP parameter value (a in the notation in the code) and β (α in the notation in the code) denotes a conjectured value of this parameter in the model that the researcher is using for estimation and inference.

Replication for Table 1

Up through line 446 the code defines several functions used for the computations reported in the table. Execution of $\mathsf{gettruth}(\mathsf{spar}, \mathsf{y2})$ for $y_2 = 0.5, 0, -0.5$ on lines 465, 479, and 492 produces the DGP population counterfactual probabilities and marginal effects reported the third and sixth columns in the table. The function ochoice_projection with appropriate arguments is then used in sequence to compute the 72 lower and upper bounds on counterfactual probabilities and marginal effects reported in columns 4,5,7,8. The magnitude of the bound in each case corresponds to the objective attribute of the return value. For lower bounds, the bound is the objective attribute, while for upper bounds it is -1 times the objective attribute. This is because a minimization routine is used to find the extremal values of the region. For example, after executing line 460 holdasy_eps1l$objective provides the lower bound of 0.03 (after rounding to two decimal places) on the counterfactual probability corresponding to $y_2 = 0.5$ and $s_{22} = 1$. After executing line 467 holdasy_eps1u$objective provides -1 times the upper bound of 0.64 (again after rounding to two decimal places) on the same counterfactual probability.