

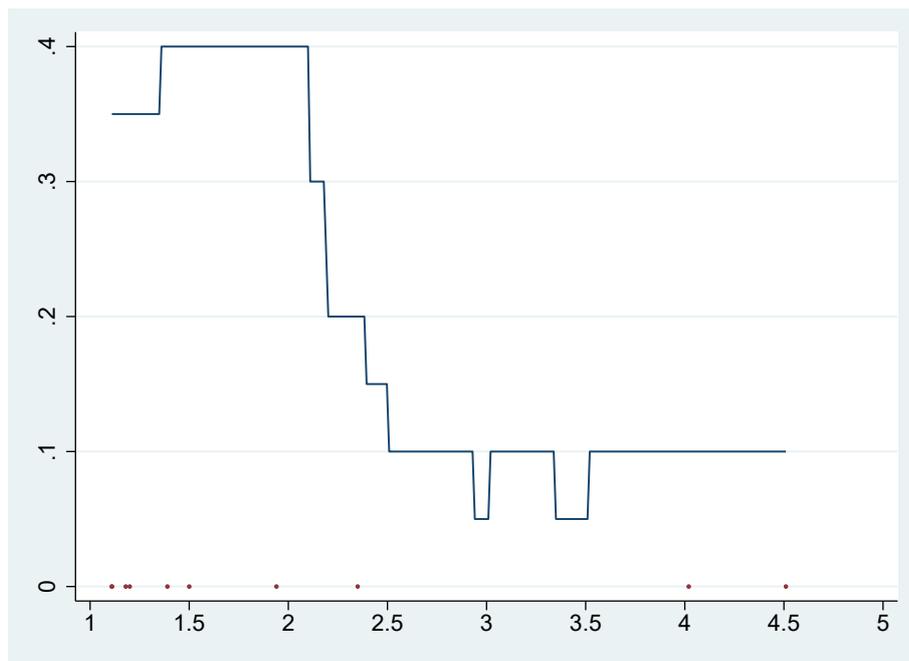
Exam 5106/9106B – Fall 2021 Postponed

IMPORTANT: Always motivate your answers. Show knowledge and understanding of the concepts taught in the course. Your answers should be as short as possible and as long as necessary. Total marks of the examination is 140. Each sub-question carries 10 marks.

1. Suppose y_i $i = 1, 2, \dots, 10$ are i.i.d draw from a unknown continuous distribution where the CDF and PDF are $F(\cdot)$ and $f(\cdot)$. You are supposed to estimate the distribution functions based on the attached Stata output. Suppose that y_i follows the following distribution

$$f(y; \alpha) = \begin{cases} \alpha y^{-(1+\alpha)}, & y \geq 1 \\ 0, & \text{otherwise} \end{cases}$$

- (a) Write down the likelihood function and derive the score.
- (b) Derive the ML estimator of α from the score. Compute it, and derive the standard error of your estimate.
- (c) Test for $H_0 : \alpha = 1$ against $H_1 : \alpha \neq 1$ at the 5% significance level using the LR, Wald and LM test.
- (d) Explain how would you use bootstrapping to obtain a standard error estimate for $\hat{\alpha}_{MLE}$. Based on the attached Stata output, what is the bootstrapped standard error of $\hat{\alpha}_{MLE}$?
- (e) Use the Stata output to generate the bootstrap bias corrected estimate of α .
- (f) Now suppose that you don't know the distribution family and want to use nonparametric methods to estimate the density. The following figure shows a kernel density estimate based on the same 10 observations. The red dots represent the observations.



One of the following five kernels is used, and the bandwidth takes value in the set $\{0.5, 1, 2\}$. Determine which kernel and bandwidth are used.

Kernel	$K(z)$
Uniform	$\frac{1}{2} \times 1[z < 1]$
Triangular	$(1 - z) \times 1[z < 1]$
Epanechnikov	$\frac{3}{4}(1 - z^2) \times 1[z < 1]$
Biweight	$\frac{15}{16}(1 - z^2)^2 \times 1[z < 1]$
Gaussian	$(2\pi)^{-1/2} \exp(-z^2/2)$

Let Z be a random variable that is distributed according to pdf $f(\cdot)$. In the following, you are asked to find an estimate for $p = \text{prob}(Z \leq 3)$.

- (g) Compute the estimate of the MLE estimator of p and its standard error.
 - (h) Propose a non-parametric estimator of p , and compute its estimate and standard error.
2. Using data from a universal child care program aimed at 3-6 years olds, a group of researchers want to better understand the effectiveness of universal child care programs. Their main outcome variable is a measure of overall school readiness y_i at age 7. They also observe individual child care attendance D_i during age 3 to age 6.
- (a) Explain whether you can use the following regression to estimate the effect of child care attendance on school readiness:

$$y_i = a + \gamma D_i + \varepsilon_i \quad (1)$$

What is the key assumption for $\hat{\gamma}$ to have a causal interpretation?

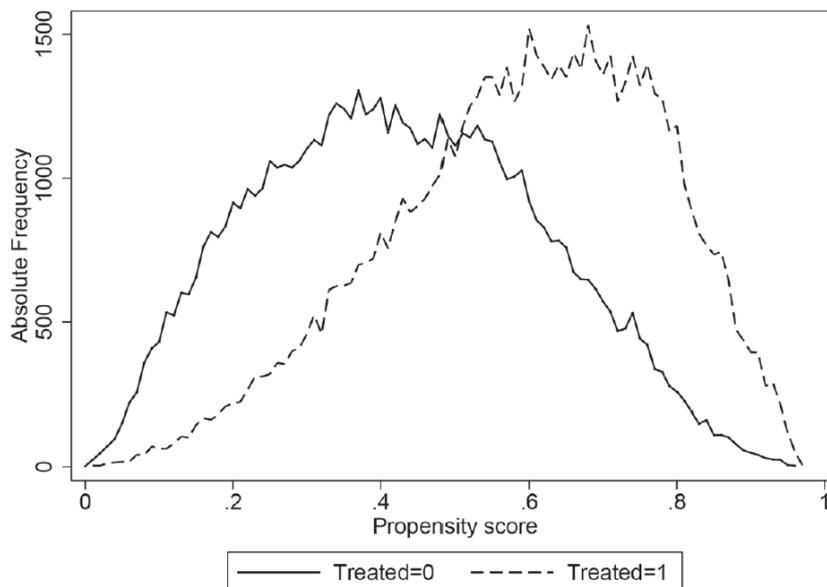
- (b) They suggest using a measure of local child care supply, Z_i , which is the local child care coverage rate at age 3 as the instrument for D_i . Consider first the case where we use a binary version of the instrument, namely $\tilde{Z}_i = Z_i > 0.5$. In IV analysis, it is often useful to classify people into different groups, such as always takers, etc. Define these groups and classify the children who participate in the study into these groups. Name the assumptions underpinning instrumental variables as a strategy for identifying the effect of an endogenous treatment on the outcome for compliers. Suppose that these assumptions hold, based on the Stata output (d : treatment status, z : binary instrument), calculate the share of children in each group.
- (c) In the following, we use the continuous instrument Z_i . Assume that the child care attendance follows the following latent index model:

$$D_i^* = \beta_0 + \beta Z_i - V_i$$

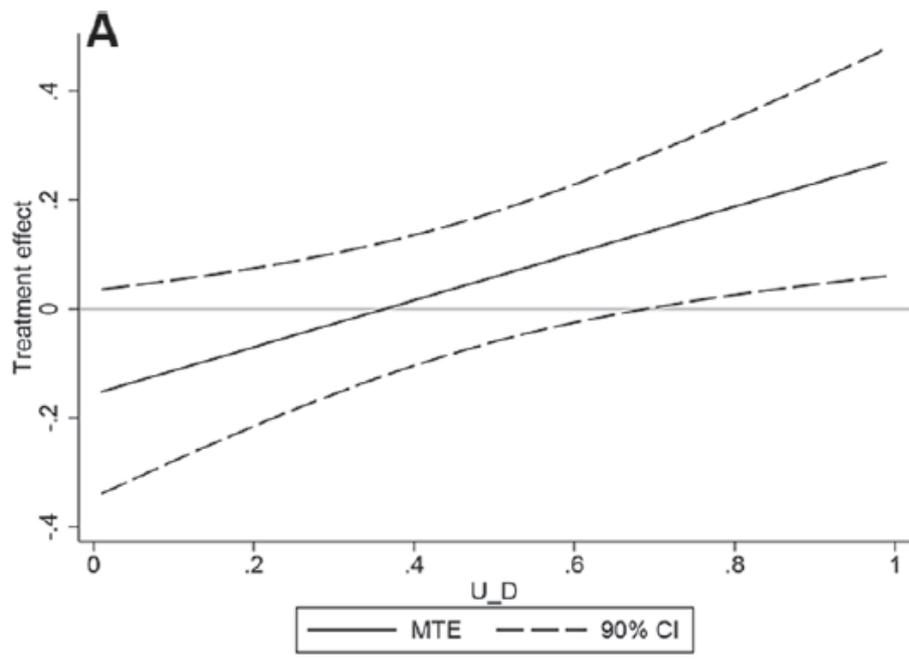
$$D_i = 1, \text{ if } D_i^* > 0 \quad D_i = 0, \text{ otherwise.}$$

where V_i follows a standard normal distribution. The potential outcome can be specified as $y_i(1) = a + \gamma + \varepsilon_{i1}$ and $y_i(0) = a + \varepsilon_{i0}$ with $E(\varepsilon_{ij}|D_i) = 0$. Explain what the random variables ε_{ij} and V_i represent? What does it imply if ε_{ij} and V_i is correlated?

- (d) Suppose that Z_i is a valid instrument for D_i , define the marginal treatment effect (MTE), and outline how you will estimate it.
- (e) The following figure shows the distribution of the predicted propensity score by treatment status. What does it imply for the MTE estimation?



- (f) The following figure shows the estimated MTE function with 90% confidence interval. What do the axes represent? Does the figure tell us anything about the selection mechanism? Based on the MTE function, arrange ATE, ATT, and ATNT in the order of increasing magnitude.



Stata Output

```

##### Problem 1.
. describe
-----+-----
Variable      Storage   Display   Value     Variable
  name        type     format    label     label
-----+-----
x              float    %9.0g     x         x
lnx            float    %9.0g     natural   natural log of x
-----+-----

. list
-----+-----
      |      x      lnx |
-----+-----
1.    |  1.2    .1823216 |
2.    |  1.5    .4054651 |
3.    |  4.02   1.391282 |
4.    |  1.94   .662688  |
5.    |  1.11   .10436   |
-----+-----
6.    |  1.11   .10436   |
7.    |  4.51   1.506297 |
8.    |  1.18   .1655144 |
9.    |  2.35   .8544153 |
10.   |  1.39   .3293037 |
-----+-----

. su
-----+-----
Variable |      Obs      Mean      Std. dev.      Min      Max
-----+-----
x        |         10      2.031      1.247918      1.11      4.51
lnx      |         10      .5706007      .5242123      .10436      1.506297

. * program to compute MLE estimates of alpha
. * it returns a scalar alpha_hat

. program ml_est, rclass
  version 17
  ....
  code skipped
  ....
  return scalar alpha_hat
. end

. bootstrap r(alpha_hat), reps(10) saving(mybs, replace): ml_est

. use mybs, clear
(bootstrap: ml_est)
. list
-----+-----
      |      _bs_1 |
-----+-----
1.    |  2.216439 |
2.    |  1.536116 |
3.    |  1.769239 |
4.    |  1.499552 |
5.    |  2.741992 |
-----+-----
6.    |  1.577018 |
7.    |  3.28103  |
8.    |  2.306023 |
9.    |  2.749356 |
10.   |  1.597841 |
-----+-----

. su
-----+-----
Variable |      Obs      Mean      Std. dev.      Min      Max
-----+-----
_bs_1    |         10      2.127461      .6316372      1.499552      3.28103

Critical Values of Chi-square
df      .50      .25      .10      .05      .025      .01      .001
1       0.45      1.32      2.71      3.84      5.02      6.63      10.83
2       1.39      2.77      4.61      5.99      7.38      9.21      13.82
3       2.37      4.11      6.25      7.81      9.35      11.34      16.27
4       3.36      5.39      7.78      9.49      11.14      13.28      18.47
5       4.35      6.63      9.24      11.07      12.83      15.09      20.52
6       5.35      7.84      10.64      12.59      14.45      16.81      22.46
7       6.35      9.04      12.02      14.07      16.01      18.48      24.32
8       7.34      10.22      13.36      15.51      17.53      20.09      26.12
9       8.34      11.39      14.68      16.92      19.02      21.67      27.88
10      9.34      12.55      15.99      18.31      20.48      23.21      29.59

##### Problem 2:
. tab z d
-----+-----
      |      d      |
      |      0      |      1      |      Total
-----+-----+-----
0     |      2600   |      2800   |      5400
1     |      1800   |      5600   |      7400
-----+-----+-----
Total |      4400   |      8400   |     12800

```