Name: ______________________________

ECON 21020: Econometrics

The University of Chicago, Spring 2022

Instructor: Thomas Wiemann

Final Exam

Date: May 25, 2022

1. **Write your name on the exam.**
2. The exam is closed book and closed notes.
3. No calculators are allowed.
4. There are a total of 100 possible points (+ 5 bonus points).
5. Answer as many questions as you can. **You do not need to answer questions in order.** Try to answer later parts of a question even if you struggle with earlier parts.
6. Please write your answers in the space provided on the exam paper. There is additional scratch paper at the end of the exam.
7. Label your final answers clearly where appropriate.
8. Do not take this exam with you.
9. Any students caught cheating will fail the course.
10. **Good luck!**

Problem 1   20 Points

The following are "True or false?"-questions. If the statement is true, provide a brief proof (≈ 3 lines). If the statement is false, provide a counter example. There are no points awarded for answers without a proof or counter example.

a)   5 Points

True or false? Let X be a random variable. If $F(x)$ is the cumulative density function (CDF) of X, then $F(x) \in [0, 1], \forall x \in \mathbb{R}$.

b)   5 Points

True or false? Let Y and X be random variables. If $E[Y|X] = E[Y]$, then $Y \perp\!\!\!\perp X$.

c) 5 Points

True or false? Let X be a random variable. If $X_1, \dots, X_n \overset{iid}{\sim} X$ and $\lambda \in \mathbb{R}$, then

$$\frac{1}{n+\lambda} \sum_{i=1}^{n} X_i \xrightarrow{p} E[X]. \tag{1}$$

d) 5 Points

True or false? Every consistent estimator is unbiased.

Problem 2 22 Points

Let Y, X, and Z be random variables. Consider another random variable X^* defined by

$$X^* = X + Z, \tag{2}$$

where $Var(X^*) > 0$, $Var(X) > 0$, and $Var(Z) > 0$.

Suppose the econometrician is interested in the BLP($Y|X$)-coefficient β_X but only observes a sample $(Y_1, X_1^*), \ldots, (Y_n, X_n^*) \overset{iid}{\sim} (Y, X^*)$.

a) 4 Points

Let β_{X^*} denote the BLP($Y|X^*$)-coefficient. Show that

$$\beta_{X^*} = \frac{Var(X)\beta_X + Cov(X,Z)\beta_X + Cov(\varepsilon, Z)}{Var(X) + Var(Z) + 2Cov(X,Z)}, \tag{3}$$

where $\varepsilon \equiv Y - \text{BLP}(Y|X) = Y - (\alpha_X + X\beta_X)$.

(Hint: Recall that whenever $Var(X^*) > 0$, it holds that $\beta_{X^*} = Cov(Y, X^*)/Var(X^*)$.)

b) 3 Points

Suppose for the remainder of the question that $Z \perp\!\!\!\perp (Y, X)$.

Show that

$$Cov(\varepsilon, Z) = 0. \tag{4}$$

c) 4 Points

Use part a) and b) to show that

$$\beta_{X^*} = \frac{Var(X)}{Var(X) + Var(Z)}\beta_X. \tag{5}$$

Conclude that this implies $|\beta_{X^*}| \leq |\beta_X|$.

d) 4 Points

Suppose for the remainder of the question that $Var(Z) = \sigma_Z^2$ is known.

Construct sample analogue estimators for

$$\frac{Var(X) + Var(Z)}{Var(X)}, \tag{6}$$

and β_{X^*}.

e) 3 Points

Use part c) and d) to construct a sample analogue estimator for β_X.

f) 4 Points

Prove that the estimator constructed in part e) is consistent.

Problem 3 26 Points

Let (Y, W, X, U) be a random vector with joint distribution characterized by

$$Y = g(W, U), \tag{7}$$

where W is binary. Suppose further that both selection on observables (SO) – i.e., $W \perp\!\!\!\perp U|X$ – and common support (CS) hold. Define the *propensity score* by

$$p(X) \equiv P(W = 1|X), \tag{8}$$

where $\operatorname{supp} p(X) \subset (0, 1)$.

Throughout, (Y, W, X) are observables and U are unobservables.

a) 3 Points

Give a brief interpretation of $P(W = 1|X)$.

b) 5 Points

Show that

$$E[g(1, U)] = E\left[\frac{YW}{p(X)}\right]. \tag{9}$$

(Hint: Recall that $Y = Wg(1, U) + (1 - W)g(0, U)$.)

c) 5 Points

Show that

$$E[g(0,U)] = E\left[\frac{Y(1-W)}{1-p(X)}\right]. \tag{10}$$

d) 5 Points

Use part b) and c) to argue that

$$\text{ATE} \equiv E[g(1,U) - g(0,U)] \tag{11}$$

is point-identified.

e) 4 Points

Suppose that the econometrician observes a sample $(Y_1, W_1, X_1), \ldots, (Y_n, W_n, X_n) \overset{iid}{\sim} (Y, W, X)$. Further, suppose that the propensity score $p(\cdot)$ is a known function.

Construct a sample analogue estimator of the ATE.

f) 4 Points

Briefly explain why your estimator from part e) may not be broadly applicable in practice even if W is binary and SO and CS hold.

Problem 4 32 Points

Suppose an econometrician is interested in studying preferences for leisure and work among Americans. For this purpose, she considers leveraging random variation in lottery winnings to assess whether substantial increases in (unearned) income decrease labor supply.[1]

To think carefully about the effect of a lottery winnings on employment decisions, consider the random vector (Y, W, U) with joint distribution characterized by

$$Y = g(W, U), \tag{12}$$

and $g : \operatorname{supp} W \times \operatorname{supp} U \to \operatorname{supp} Y$ denotes a labor supply function. Here, Y denotes a person's labor earnings, W is the USD amount won in the lottery, and U are all determinants of Y other than W.

a) 2 Points

Give an example for an unobserved determinant U of Y.

b) 3 Points

Interpret the potential outcomes $g(w, U)$ for $w \geq 0$.

[1] This question is inspired by Imbens et al. (2001).

c) 3 Points

Interpret $E[g(0, U)]$ and $E[g(1000000, U)]$.

d) 3 Points

Interpret $E[Y|W = 0]$ and $E[Y|W = 1000000]$.

e) 2 Points

Do your interpretations in part c) and d) differ? Explain briefly.

f) 4 Points

Interpret the assumption $W \perp\!\!\!\perp U$. Does it appear plausible here? Explain briefly.

g) 5 Points

For the remainder of the exercise, consider the random vector (Y, W, U, X), where (Y, W, U) as before and X denotes the number person's average number of lottery tickets bought per week. For simplicity, suppose that $\operatorname{supp} X = \{0, 1, \dots, 100\}$.

Interpret the assumption $W \perp\!\!\!\perp U|X$. Does it appear plausible here? Explain briefly.

h) 5 Points

For the remainder of the exercise, suppose that $W \perp\!\!\!\perp U|X$ holds.

Define the conditional average treatment effect of winning 1 million USD:

$$\text{CATE}(x) = E[g(1000000, U) - g(0, U)|X = x]. \tag{13}$$

Is the CATE(x) point-identified for all $x \in \{0, 1, \dots, 100\}$? Explain briefly.

i) 5 Points

Define the average treatment effect of winning 1 million USD:

$$\text{ATE} = E[g(1000000, U) - g(0, U)]. \quad (14)$$

Is the ATE point-identified? Explain briefly.

Problem 5 Bonus

This is an optional exercise. Correct solutions are awarded 5 bonus points.

Name the three distinct tasks arising in the analysis of causal questions according to Heckman and Vytlacil (2007).

References

Heckman, J. J. and Vytlacil, E. J. (2007). Econometric evaluation of social programs, part I: Causal models, structural models and econometric policy evaluation. In Heckman, J. J. and Leamer, E., editors, *Handbook of Econometrics*, volume 6, chapter 70, pages 4779–4874. Elsevier, Amsterdam.

Imbens, G. W., Rubin, D. B., and Sacerdote, B. I. (2001). Estimating the effect of unearned income on labor earnings, savings, and consumption: Evidence from a survey of lottery players. *American Economic Review*, 91(4):778–794.

(scratch work)

(scratch work)

(scratch work)