Econ 241 C.

Let $D = 1$ be an indicator for whether subject is treated or not. Let $Y(1)$ and $Y(0)$ be the potential outcomes for being treated and untreated, respectively, and let $Y = DY(1) + (1 - D)Y(0)$ be the realized outcome. The observed data $W = (X, D, Y)$ consists of the pre-treatment covariates X, the treatment indicator D and the realized outcome Y. The Average Treatment Effect parameter is defined as

$$\theta = \mathrm{E}[Y(1) - Y(0)]. \tag{1}$$

The propensity score shows the conditional probability of being treated:

$$\pi(X) = \Pr(D = 1 \mid X).$$

The regression function is

$$\mu(d, x) = \mathrm{E}[Y \mid D = d, X = x], \quad d \in \{1, 0\}.$$

I will use the following conditional ignorability assumption

Assumption 1 *The potential outcomes are independent of treatment assignment conditional on X:*

$$(Y(1), Y(0)) \perp D \mid X.$$

1. *(50p). Treatment effects.*

   (a) (50/3p) Show that $\mathrm{E}Y(1)$ and $\mathrm{E}Y(0)$ can be identified as

   $$\mathrm{E}[Y(1)] = \mathrm{E}\mu(1, X), \quad \mathrm{E}[Y(0)] = \mathrm{E}\mu(0, X).$$

   (b) (50/3p) Alternatively, show that $\mathrm{E}Y(1)$ and $\mathrm{E}Y(0)$ can be identified as

   $$\mathrm{E}[Y(1)] = \mathrm{E}\frac{D}{\pi(X)}Y, \quad \mathrm{E}[Y(0)] = \mathrm{E}\frac{1-D}{1-\pi(X)}Y.$$

   (c) (50/3p) State an orthogonal/doubly robust moment equation for the ATE parameter θ_0.

(d) (Optional, not for grade.) Suppose the following perfect balance condition holds:

$$\mu_1(X) = 1/2 \text{ a.s. } .$$

Show that the ATE parameter θ_0 in (1) can be identified as the parametric component β_0 in the *Partially Linear Model*

$$Y = D'\beta_0 + g_0(X) + U, \quad \mathrm{E}[U \mid D, X] = 0.$$

In other words,

$$\theta_0 = \beta_0.$$

(Optional. Clarify the interpretation of β_0 if Assumption 1 holds but $\mu_1(X)$ is not constant.)

2. *(50p). Policy learning.* A policy $G \subseteq \mathcal{X}$ prescribes treatment as follows: a subject is treated if and only if the covariate value $X \in G$. The average welfare of the policy G is

$$W(G) = \mathrm{E}[Y(1)1\{X \in G\} + Y(0)1\{X \notin G\}].$$

(a) (10p) Assuming conditional ignorability, show that the average welfare $W(G)$ of the policy G is identified as

$$W(G) = \mathrm{E}\left(\frac{D}{\pi_1(X)}1\{X \in G\} + \frac{1-D}{1-\pi_1(X)}1\{X \in G^c\}\right). \tag{2}$$

(b) (15p) Show that the largest possible welfare

$$W(G^*) = \max_{G \subseteq \mathcal{X}} W(G)$$

is attained by the policy

$$G^* = \{X : \quad \tau_0(X) \geq 0\},$$

where $\tau_0(X)$ is the conditional ATE:

$$\tau_0(X) = \mathrm{E}[Y \mid D = 1, X] - \mathrm{E}[Y \mid D = 0, X].$$

(c) (5p) Given an i.i.d sample $(W_i)_{i=1}^n$,

i. Propose a sample estimate of $W(G)$, where G is a given policy. Hint: replace $\mathrm{E}[\cdot]$ in (2) by a sample average $N^{-1}\sum_{i=1}^{n}[\cdot]$.

ii. State the Empirical Welfare Maximization procedure

(d) (20p). Open the DML-CATE notebook The question aims to understand the effect of 401K eligibility on Net Financial Assets as a function of baseline income. To do:

- Explain what the notebook does and summarize your key empirical findings.
- Estimate the welfare $W(\widehat{G})$ using the plug-in estimate of the conditional ATE $\widehat{G} := \{x : \widehat{\tau}(x) \geq 0\}$.