Econ 312 Part B, Spring 2021

Problem Set 4

James J. Heckman

Due May 25th, 2021

This draft, May 17, 2021

Problem 1. There are many ways to estimate well defined economic parameters and their progeny-causal effects. Consider a model

$$Y = \alpha + \beta D + \gamma X + U, \quad E(U) = 0. \text{ All variables are scalar.}$$

$D = 1$ if a person receives an intervention; $= 0$ otherwise. Give conditions for identifying β (a fixed coefficient) by:

(a) Matching on X; How does matching differ from OLS?

(b) Using Z as an instrument ($\text{COV}(Z, D) \neq 0$ $\text{COV}(Z, U) = 0$)

(c) Random assignment of D (with full compliance)

(d) Selection methods (assume a general non-normal model)

(e) Proxy variable methods

where $U = \phi V + \varepsilon$, $V \perp\!\!\!\perp \varepsilon$, and $\varepsilon \perp\!\!\!\perp (D, X)$ when M_1, M_2, M_3 are measures available to you:

$$M_1 = \gamma_{10} + \gamma_{11} V + \eta_1,$$

$$M_2 = \gamma_{20} + \gamma_{22} V + \eta_2,$$

$$M_3 = \gamma_{30} + \gamma_{33} V + \eta_3.$$

and $\{n_i\}_{i=1}^{3}$ are mean zero, mutually independent, random variables independent of (X, V).

(f) In problem (e), give conditions under which the $\gamma_{ij}, i, \in \{1, 2, 3\}$ are identified along with the variances in η_1, η_2, η_3.

Be explicit how in each case there is exogenous variation that produces identification and state what it is.

Problem 2. (Building on (e) from Problem 1) Suppose you have panel data on

$$Y_{it} = \alpha_t + \beta D_{it} + \gamma X_{it} + U_{it},\ t = 1, \ldots, T$$

$$\text{where } U_{it} = \phi V_i + \varepsilon_{it},\ V \perp\!\!\!\perp \{\varepsilon_{it}\}_{t=1}^{T},\ \{\varepsilon_t\}_{t=1}^{T} \perp\!\!\!\perp \{X_t\},$$

$$E(\varepsilon_{it}) = 0\ \forall t = 1, \ldots T, \text{ and}$$

$$\begin{cases} D_{it} = 0 & t \leq t^* \\ D_{it} = 1 & t > t^* \end{cases} \quad \text{where the intervention occurs at } 1 < t^* < T.$$

How would you identify β in the following cases?

(a) Using panel data on individuals $i = 1, \ldots, I$

(b) Suppose instead you have repeated cross sections of persons sampled from the same population but in general samples have different individuals

(c) Suppose instead of (b) you have only aggregate time series data before and after t^* and you do not know the treatment status of any person, but the aggregate proportion treated? Consider two cases: (i) $\alpha_t = \alpha\ \forall t$ and (ii) α_t freely varies.

(d) In what way, if any, does this model differ from the "difference in differences" model you studied in part 1 of this course?

Problem 3. State and prove Yitzhaki's theorem (See the appendix of Heckman et al. (2006) for OLS weights and LATE (IV) weights and explore their connection. Relate those weights to Theil's theorem characterizing OLS as a weighted average of bivariate regression coefficients. (Take scalar X)

Problem 4. Compare and contrast the roles of the probability of selection ($\Pr(D = 1|Z)$) in LATE, matching, and selection models.

Problem 5. Answer (1) when β is heterogeneous, β_i, and $\beta_i \not\perp\!\!\!\perp D_i$.

Problem 6. Answer all of the questions posed in the 4 posted panel data model for the final week.

Problem 7. Consider the panel problem set on earnings posted at [**Will insert here**]. Compute the parameters of the following earnings dynamic models using the NLSY $X_{it} \perp\!\!\!\perp U'_{it}$ at all $U_{it} \perp\!\!\!\perp U'_i t''$ all i, i' for $t \neq t'$.

(a) $\ln Y_{it} = \alpha' X_{it} + \beta Y_{i,t-1} + U_{it}, t = 1, T, \ldots, i = 1, \ldots, I$

$U_{it} = F_i + \varepsilon_{it}$ ε_{it} iid over i, t.

(b) $\beta = 0, U_{it} = \rho U_{i,t-1} + \varepsilon_{it}, \varepsilon_{it}$ iid.

(c) $\beta = 1, U_{it} = \rho U_{i,t-1} + \varepsilon_{it}, \varepsilon_{it}$ iid.

(d) Same as (b) but $\rho = 1$.

(e) All of the above but ε_{it} is $MA1$.

Which model best fits the data? Trace out the implied dynamics of an exogenous unit change in $Y_{i,t-1}$.

References

Heckman, J. J., S. Urzúa, and E. J. Vytlacil (2006). Understanding instrumental variables in models with essential heterogeneity. *Review of Economics and Statistics 88*(3), 389–432.