Problem Set 9
Econ 312, Spring 2019
James J. Heckman
Due June 4th, 2019

1. [25 points] Compare the assumptions used to identify ex ante and ex post ATE, TOT, MTE, and LATE based on (a) matching, (b) IV, and (c) selection approaches. Use the cross-section model where observations are independently distributed:

$$
\begin{aligned}
Y &= DY_1 + (1-D)Y_0 \\
Y_1 &= \mu_1(X) + U_1 \quad E(U_1) = 0 \\
Y_0 &= \mu_0(X) + U_0 \quad E(U_0) = 0 \\
D &= 1(\mu_D(Z) \geq V) \quad D \in \{0,1\} \\
(U_0&, U_1, V) \perp\!\!\!\perp X, Z
\end{aligned}
$$

   Z and X may contain some elements in common. Specifically, compare the information available from (Y, Z, X) and that (Y, Z, D, X), i.e.,

   (a) Information from IV assumptions using $E(Y|Z, X)$.

   (b) Selection model information using $E(Y|Z, X, D)$.

   (c) Matching information using Y, X, D, Z.

Use the relationship:

$$
\begin{aligned}
E(Y|Z,X) = & E(Y|X,Z,D=1)Pr(D=1|X,Z) \\
& + E(Y|X,Z,D=0)Pr(D=0|X,Z)
\end{aligned}
$$

to compare the estimators. Compute:

$$\frac{\partial E(Y|Z,X)}{\partial P(Z,X)}$$

for this model and relate to ATE and MTE. How do the derivatives of the control function compare to MTE?

Compare the information sets of the agent and the econometrician under each approach.

2. [25 points] Consider a dynamic model of life cycle consumption $(C_t)_{t=1}^T$ over horizon T, for agents facing an exogenous, but uncertain income flow $\{Y_t\}_{t=1}^T$. Assets evolve according to

$$A_{t+1} = (1+r)A_t + Y_t - C_t \qquad \text{(I)}$$

and initial endowment A_0 is specified as exogenous. r is a known constant exogenous interest rate.

Agents maximize

$$\sum_{t=0}^{T} \beta^t U(C_t)$$

subject to (I). $\beta(1+r) = 1$.

Assume $U(C_t)$ is quadratic:

$$U(C_t) = \delta_0 + \gamma_1 C_t + \gamma_2 C_t^2$$
$$\gamma_2 < 0, \gamma_1 > 0.$$

$$Y_t = \underbrace{P_t}_{\text{permanent}} + \underbrace{\tau_t}_{\text{transitory}} \qquad \text{(A)}$$
$$P_t = \rho P_{t-1} + \lambda_t \quad E(\lambda_t) = 0$$
$$\tau_t = \omega_t + \theta\omega_{t-1}$$
$$E(\omega_j) = 0, \forall j$$

the process is stationary, the ω_j are mutually independent, P_0 is fixed and exogenous. Agents at t know $\varepsilon_{t-j} j \leq t$, but not future ε_t. Suppose $T < \infty$.

(a) Characterize the response of consumption C_t of an agent to a unit shock in ω_t to a unit shock in λ_t. How do they change with age? With increases in ρ? With t? With increases in $\sigma^2_{\omega_t}$?

(b) Show how to identify the model using panel data on individuals $(C_t, Y_t)_{t=1}^T$. Assume shocks are independent across agents.

(c) Suppose we add to equation (A), measurement error M_t, M_t iid? $E(M_t) = 0, Var(M_t = \sigma^2_M)$. What parameters can be identified using only income? Using income and consumption?

(d) How does access to data on C_t aid in identifying agent information sets?

(e) Using the posted panel data sets on income and consumption, estimate the parameters of (A) and $U(C_t)$.

3. [15 points] Read the posted handout, "RIP to HIP," based on Hyrshko.

   (a) Define HIP and RIP. What information is assumed to be known to the agent in each model? How would you test this information assumption from income data alone?

   (b) How can you distinguish these two earnings processes using only income data?

   (c) How does access to consumption data help?

   (d) Why does any of this matter for determining agent welfare?

   (e) Using the posted data set, estimate the models in Hyrshko's paper. (*Hint:* The data are not Hyrshko's.)

4. [25 points] Consider the following model of earnings dynamics based on

Heckman and Robb (*Journal of Econometrics*, 1985): Earnings evolve as

$$
\begin{aligned}
y_{it} =& \alpha_0 + \alpha_1 X_t + \alpha_2 D_k + U_{it} \\
U_{it} =& \rho_t P_t + \lambda_t \qquad t = 1, \dots, T, 1 < k < T \\
U_{it} =& \underbrace{P_t}_{\text{permanent}} + \underbrace{\tau_t}_{\text{transitory}} \qquad t = 1, T \\
P_{it} =& \rho P_{i,t-1} + \lambda_{i,t} + \eta_i \\
& B(L)\tau_{i,t} = A(L)\varepsilon_{i,t} \\
& E(\lambda_t) = 0 \quad , \quad \lambda_t \perp\!\!\!\perp \eta_i, (\varepsilon_{i,t})_{t=1}^{T} \\
& E(\varepsilon_{i,t}) = 0 \\
& E(\eta_i) = 0 \quad B(L) = 1 - \rho L \\
& A(L) = 1 + \theta L
\end{aligned}
$$

Income maximizing agents are able to choose to participate in a program at times $1 < k < \infty$.

If they participate, they get (*ex post*) a boost of $\alpha > 0$ per period in their earnings after period k. Assume $T \to \infty$ (for simplicity) so $\frac{\alpha}{r}$ is the *ex post* gain in terms of present value of income maximization. r is the non-stochastic interest rate.

Participation in the program costs $Y_{i,k}$ (forgone earnings) and some psychic costs $C = Z\delta + V$.

(a) Justify the decision role for participation in the program:

$$D_k = 1(E_k(\frac{\alpha}{r} - C - Y_{i,k}) > 0)$$

where E_k means the information at k. Characterize the possible agent information sets under two conditions:

i. α is a constant

ii. α_i varies among individuals, assuming either $Y_{i,k}$ is known or is not known at k. But C is known.

(b) For each information set, characterize $Cov(D_k, Y_t)$

$1 < t < \infty$.

(c) For each information set, discuss identification of α (or $E(\alpha)$) and the MTE both the fixed coefficient and for random coefficient version using

i. IV

ii. Matching

iii. Two-step selection models

iv. MLE

Assume (only for convenience) normally distributed errors.

(d) Suppose that you have two independent samples taken at the same time. They have identical distributions of regressors and errors. One sample has no access to the program ever. One sample has access to the program only at time period k. How does access to this data improve your ability to identify α or $E(\alpha)$? Consider two cases:

   i. Both samples are exposed to a common time trend $\mu(t) = f exp(-gt)$.

   ii. $g_{\text{treated}} > g_{\text{control}}$

   iii. Compare the identification different achieved from difference in difference estimators with that of the estimators discussed in (c).

(e) For the posted data sets, estimate $\alpha, E(\alpha), g, f$ and other parameters of the model using the four methods in part (c) and the difference in difference estimator.

5. [10 points] What is Marschak's Maxim? Discuss in the context of the two-equation model in the simultaneous causality handout restriction. (*Hint:* ISI, 2008, *Econometric Causality.*)