A Running Example Based on:

Alternative Methods For Evaluating the Impact of Interventions: An Overview

Excerpt from the *Journal of Econometrics*, 1985

James J. Heckman & Richard Robb Jr.

Econ 312, Spring 2023

For the rest of the course, we will work with a model of program participation for the impact of job training on earnings.

1. Earnings functions

- Assume that individuals experience only one opportunity to participate in training.
- This opportunity occurs in period k.
- Training takes a single period for participants to complete.
- During training, participants earn no labor income.

- Earnings of individual i in period t by Y_{it}.
- Earnings depend on a vector of observed characteristics, X_{it}.
- Post-program earnings $(t > k)$ also depend on a dummy variable, d_i, which equals one if the ith individual participates and is zero if he does not.
- U_{it} represent the error term in the earnings equation and assume that $E[U_{it}] = 0$.

- For a linear specification, latent earnings is

$$Y_{it}^* = X_{it}\beta + U_{it},$$

β is a vector of parameters.

- $E(U_{it}|X_{it}) = 0$ all X_{it}.

- β is the coefficient of X in the conditional expectation of Y^* given X.
- Observed earnings Y_{it}:

$$\begin{aligned} & X_{it}\beta + U_{it} \text{ if } t < k \\ & (1 - d_i)(X_{it}\beta + U_{it}) \text{ if } t = k \\ & X_{it}\beta + U_{it} + \alpha_i d_i \text{ if } t > k \end{aligned} \tag{1}$$

- $d_i = 1$ if the person takes training and $d_i = 0$ otherwise:
- α is one definition of the causal or structural effect of training on earnings.
- Observed earnings *are the sum of latent earnings* Y_{it}^* *and the structural shift term* $d_i\alpha$ *that is a consequence of training.* Y_{it} is thus the sum of two random variables when $t > k$.

- Problem of selection bias arises because d_i may be correlated with U_{it}.
- Consequence of selection decisions by agents. Thus, selection bias is present if

$$E(U_{it}d_i) \neq 0.$$

- If $E(U_{it} d_i) \neq 0$,

$$E(Y_{it} \mid X_{it}, d_i) \neq X_{it}\beta + d_i\alpha.$$

- We will also develop the analysis when α varies among individuals α_i

1.2 Enrollment rules

- The decision to participate in training may be determined by a prospective trainee, by a program administrator or both.
- Whatever the specific content of the rule, it can be described in terms of an index function framework.
- Let IN_i be an index of benefits to the appropriate decision-makers from taking training.
- It is a function of observed (Z_i) and unobserved by the economist (V_i) variables.
- Thus
$$IN_i = Z_i\gamma + V_i. \tag{2}$$
- What is observed by the agent is a different matter.
- Z may include X_i

- In terms of this function,

$$d_i = \begin{cases} 1 & \text{iff } IN_i > 0 \\ 0 & \text{otherwise.} \end{cases}$$

The distribution function of V_i: $F(v_i) = \Pr(V_i < v_i)$.

- V_i: independently and identically distributed across persons.
- Let $p = E[d_i] = \Pr[d_i = 1]$ and assume $1 > p > 0$.

- If that V_i is distributed independently of Z_i, $\Pr(d_i = 1 \mid Z_i) = F(-Z_i\gamma)$
- Sometimes called the "propensity score" in statistics (see, e.g., Rosenbaum and Rubin, 1983).

- The condition for the existence of selection bias

$$E(U_{it}d_i) \neq 0$$

may occur because of stochastic dependence between U_{it} and the unobservable V_i in equation (2) (selection on the unobservables) or because of stochastic dependence between U_{it} and Z_i in equation (2) (selection on observables).

Simple Behavioral Model

- Natural starting point: model of trainee self-selection based on a comparison of the expected value of earnings with and without training.
- For simplicity, assume that training programs accept all applicants (no distinction between ATE and ITT).

- Prospective trainees to discount earnings streams by a common discount factor $1/(1 + r)$.
- From (1) training raises trainee earnings by α per period.
- While in training, individual i receives a subsidy S_i which may be negative (so there may be direct costs of program participation).
- Trainees forego income in training period k.
- To simplify the expressions, assume that people live forever.

- As of period k, the present value of earnings for a person who does not receive training is

$$PV_i(0) = E_{k-1}\left(\sum_{j=0}^{\infty}\left(\frac{1}{1+r}\right)^j Y_{i,k+j}\right).$$

- E_{k-1} means that the expectation is taken with respect to information available to the prospective trainee in period $k-1$.
- The expected present value of earnings for a trainee is

$$PV_i(1) = E_{k-1}\left(S_i + \sum_{j=1}^{\infty}\left(\frac{1}{1+r}\right)^j Y_{i,k+j} + \sum_{j=1}^{\infty}\frac{\alpha}{(1+r)^j}\right).$$

- The risk-neutral wealth-maximizing decision rule is to enroll in the program if $PV_i(I) > PV_i(0)$ or, letting IN_i denote the index function in decision rule (2),

$$IN_i = PV_i(1) - PV_i(0) = E_{k-1}[S_i - Y_{ik} + \alpha/r], \qquad (3)$$

- Decision to train is characterized by

$$d_i = \begin{cases} 1 & \text{iff } E_{k-1}[S_i - Y_{ik} + \alpha/r] > 0 \\ 0 & \text{otherwise.} \end{cases} \qquad (4)$$

- Let W_i be the determinant of the subsidy that the econometrician observes (with associated coefficient ϕ) and let τ_i be the part which he does not observe:

$$S_i = W_i\phi + \tau_i.$$

- A special case of this model arises when agents possess perfect foresight so that $E_{k-1}[S_i] = S_i$, $E_{k-l}[Y_{ik}] = Y_{ik}$ and $E_{k-1}[\alpha/r] = \alpha/r$.

- Collecting terms,

$$d_i = \begin{cases} 1 & \text{iff } S_i - Y_{ik} + \alpha/r = W_i\phi + \alpha/r - X_{ik}\beta + \tau_i - U_{ik} > 0 \\ 0 & \text{otherwise.} \end{cases} \quad (5)$$

- Then $(\tau_i - U_{ik}) = V_i$ in (2) and (W_i, X_{ik}) corresponds to Z_i in (2).
- Assuming that (W_i, X_{ik}) is distributed independently of V_i makes (5) a standard discrete choice model.

THE UNIVERSITY OF CHICAGO

- Suppose decision rule (5) determines enrollment.
- If the costs of program participation are independent of U_{it} for all t (so both W_i and τ_i are independent of U_{it}), then $E[U_{it} d_i] = 0$ only if the unobservables in period t are (mean) independent of the unobservables in period k or

$$E[U_{it} \mid U_{ik}] = 0 \text{ for } t > k.$$

- Whether or not U_{it} and d_i are uncorrelated hinges on the serial dependence properties of U_{it}.

- If U_{it} is a moving average of order m so

$$U_{it} = \sum_{j=1}^{m} a_j \varepsilon_{i,t-j},$$

where the $\varepsilon_{i,t-j}$ are iid, then for $t-k>m$, $E[U_{it}d_i]=0$.
- On the other hand, if U_{it} follows a first-order autoregressive scheme, then $E[U_{it} \mid U_{ik}] \neq 0$ for all finite t and k.

- The enrollment decision rules give context to the selection bias problem.
- The estimators differ greatly in their dependence on particular features of these rules.
- Some estimators do not require that these decision rules be specified at all, while other estimators require a great deal of *a priori* specification of these rules.
- Given the inevitable controversy that surrounds specification of enrollment rules, there is always likely to be a preference by analysts for estimators that require little prior knowledge about the decision rule.
- But this often throws away valuable information and ignores the subjective evaluation implicit in $d_i = 1$.