Randomized Evaluations

from *Econometric Evaluation of Social Programs Part II*

James J. Heckman
University of Chicago
American Bar Foundation
University College Dublin, Ireland

Edward J. Vytlacil
Yale University

Econ 312, Spring 2021

- Two cases have been made application of social experimentation.
- (1) is a classical argument for experimental design.
- Inducing variation in regressors increases precision of estimates and the power of tests.
- (2) focuses on solving endogeneity and self-selection problems.
- The two cases are mutually compatible, but involve different emphases.

- Both cases can be motivated within a linear regression model for outcome Y with treatment indicator D and covariates X:

$$Y = X\alpha + D\beta + U, \tag{1}$$

  where U is an unobservable, $E(U) = 0$.
- D (and the X) may be statistically dependent on U; $D \not\perp\!\!\!\perp \beta$.

- Both cases for social experimentation seek to secure identification of some parameters of (1)
- Advocates for the first case for experimentation typically assume a common coefficient model for α and β.
- Variation in (X, D) may be insufficient to identify or precisely estimate (α, β).
- Manipulating (X, D) through randomization, or more generally, through controlled variation, can secure identification.
- Typically assumed that (X, D) is independent of U or at least mean independent but not essential.
- Traditional case analyzed in experimental design in statistics.

- In analyzing the effects of taxes on labor supply, it is necessary to isolate the effect of wages (the substitution effect) from the effect of pure asset income (the income effect) on labor supply.
- In observational data, empirical measures of wages and asset income are highly intercorrelated.
- In addition, asset income is often poorly measured.

- By experimentally assigning these variables as in the negative income tax experiments, it is possible to identify both income and substitution effects in labor supply equations (see Cain and Watts, 1973).
- Aigner (1979): time of day pricing

- **Random assignment is not essential to this approach**.
- Any regressor assignment rule based on variables Q that are stochastically independent of U will suffice, although the efficiency of the estimates will depend on the choice of Q and care must be taken to avoid inducing multicollinearity by the choice of an assignment rule (if predictors Q are highly correlated with X and D).

- Second case focuses on dependence between (X, D) and U that invalidates least squares as an estimator of the causal effect of X and D on Y.
- In the second case, experimental variation in (X, D) is sought to make it "exogenous" or "external" to U.
- This is misleading
- As in matching experiments **balance bias** between treatment and control

- A popular argument in favor of experiments is that they produce simple, transparent estimates of the effects of the programs being evaluated in the presence of such biases.

- A quotation from Banerjee (2006) is apt:

> The beauty of randomized evaluations is that the results are what they are: we compare the outcome in the treatment *group* with the outcome in the control group, see whether they are different, and if so by how much. Interpreting quasi-experiments sometimes requires statistical legerdemain, which makes them less attractive . . .

- This argument assumes that interesting evaluation questions can be answered by the marginal distributions produced from experiments.
- IV, matching experiments without further assumptions only identify $F(Y_1|X)$, $F(Y_0|X)$ **not** $F(Y_1 - Y_0|X)$ unless
  a. $Y_1 - Y_0$ is a constant (ignores heterogeneity) or
  b. Persons have identical quantiles for $(Y_1 - Y_0)$
- "Quantile treatment effects"
- **Question: What configurations of Generalized Roy model have identical quantiles for Y_1 and Y_0 for all persons?**

- Randomization is an instrument.
- As such, it shares all of the assets and liabilities of IV already discussed.

- Only if the randomization (instrument) corresponds **exactly** to the policy that is sought to be evaluated will the IV (randomization) identify the parameters of economic interest.

Randomization as an Instrumental Variable

- Potential outcomes are (Y_0, Y_1) and cost of participation is C.
- Assume perfect certainty in the absence of randomization.
- Under self-selection, the treatment choice is governed by

$$D = \mathbf{1}(Y_1 - Y_0 - C \geq 0).$$

- Assume additive separability between the observables (X, W) and the unobservables (U_0, U_1, U_C) for convenience:

$$Y_1 = \mu_1(X) + U_1, \qquad Y_0 = \mu_0(X) + U_0,$$
$$C = \mu_C(W) + U_C, \qquad V = U_1 - U_0 - U_C,$$
$$\mu_I(X, W) = \mu_1(X) - \mu_0(X) - \mu_C(W), \qquad Z = (X, W).$$

  Only some components of X and/or W may be randomized.

- Randomization can be performed unconditionally or conditional on strata, Q, where the strata may or may not include components of (X, W) that are not randomized.
- Specifically, it can be performed conditional on X, just as in our analysis of IV.
- Parameters can be defined conditional on X.

- The intervention is assumed to change the arguments of functions without shifting the functions themselves.
- Thus for the intervention of randomization, the functions are assumed to be structural in the sense of Hurwicz (1962).
- The distributions of (U_0, U_1, U_C) conditional on X, and hence the distribution of V conditional on X, are also invariant.

- Assuming under full compliance, that the manipulated Z are the same as the Z facing the agent.

(R-1)

The Z assigned agent ω conditional on X are the Z realized and acted on by the agent conditional on X.

- The most commonly used randomizations restrict eligibility either in advance of agent decisions about participation in a program or after agent decisions are made, but before actual participation begins.
- Unlike statistical discussions of randomization, we build agent choice front and center into our analysis.
- Agents choose and experimenters can only manipulate choice sets.

Choice Problem

- Let $\xi = 1$ if an agent is eligible to participate in the program; $\xi = 0$ otherwise.
- $\tilde{\xi} \in \{0, 1\}$ ξ.
- D: indicates participation under ordinary conditions.
- In the absences of randomization, D is an indicator of whether the agent actually participates in the program.
- Actual participation: A.
- By construction, under invariance,

$$A = D\xi. \tag{2}$$

- This assumes that eligibility is strictly enforced.

- There is a distinction between desired participation by the agent (D) and actual participation (A).
- This distinction is conceptually distinct from the *ex-ante*, *ex-post* distinction.
- At all stages of the application and enrollment process, agents may be perfectly informed about their value of ξ and desire to participate (D), but may not be allowed to participate.
- On the other hand, the agent may be surprised by ξ after applying to the program.
- In this case, there is revelation of information and there is a distinction between *ex ante* expectations and *ex post* realizations.
- We cover both cases.

We consider two types of randomization of eligibility.

Randomization of Type 1. *A random mechanism (possibly conditional on* (X, Z)*) is used to determine* ξ*. The probability of eligibility is* $\Pr(\xi = 1 \mid X, Z)$.

- For this type of randomization, in the context of the generalized Roy model, it is assumed that

- (e-1a) $\xi \perp\!\!\!\perp (U_0, U_1, U_C) \mid X, Z$ **(Randomization of Eligibility)**

and

- (e-1b) $\Pr(A = 1 \mid X, Z, \xi)$ depends on ξ.

- This condition does **not** impose exogeneity on X, Z.
- Thus Z can fail as an instrument but ξ remains a valid instrument.

- (e-1a)′ $\xi \perp\!\!\!\perp \left(Y_0, Y_1, \{A(z,e)\}_{(z,e)\in\mathcal{Z}\times\tilde{\xi}} \right) \mid X, Z$

and

- (e-1b)′ $\Pr(A = 1 \mid X, Z, \xi)$ depends on ξ.

- A second type of randomization conditions on individuals manifesting a desire to participate through their decision to apply to the program.
- This type of randomization is widely used.

Randomization of Type 2: *Eligibility may be a function of D (conditionally on some or all components of X, Z, Q or unconditionally). It is common to deny entry into programs among people who applied and were accepted into the program ($D = 1$) so the probability of eligibility is* $\Pr(\xi = 1 \mid X, Z, Q, D = 1)$

- For this randomization of eligibility, it is assumed that:
- (e-2a) $\xi \perp\!\!\!\perp (U_0, U_1) \mid X, Z, Q, D = 1$

and

- (e-2b) $\Pr(A = 1 \mid X, Z, D = 1, \xi = 1) = 1;$
  $\Pr(A = 1 \mid X, Z, D = 1, \xi = 0) = 0.$

- Agent failure to comply with the eligibility rules or protocols of experiments can lead to violations of (e-1) and/or (e-2).

- Both randomizations are instruments.

- A variety of conditioning variables is permitted by these definitions.
- Thus, (e-1) and (e-2) allow for the possibility that the conventional instruments Z fail (IV-1) and (IV-2), but nonetheless the randomization generates a valid instrument ξ.
- The simplest randomizations do not condition on any variables.

What Does Randomization Identify?

- Under invariance assumption and under one set of randomization assumptions just presented, IV is an instrument that identifies some treatment effect for an ongoing program.
- Which treatment effect?
- Following our discussion of IV with essential heterogeneity, different randomizations (or instruments) identify different parameters unless there is a common coefficient model ($Y_1 - Y_0 = \beta(X)$ is the same for everyone given X) or unless there is no dependence between the treatment effect $(Y_1 - Y_0)$ and the indicator D of the agents' desire to participate in the treatment.
- **In these two special cases, all mean treatment parameters are the same.**

- Using IV, we can identify the marginal distributions $F_0(y_0 \mid X)$ and $F_1(y_1 \mid X)$.
- In a model with essential heterogeneity, the instruments generated by randomization can identify parameters that are far from the parameters of economic interest.

- Under autonomy and (e-1), or equivalently (e-1)′, the first type of eligibility randomization identifies $\Pr(D = 1 \mid X, Z)$ (the choice probability) and hence relative subjective evaluations, and the marginal outcome distributions $F_0(y_0 \mid X, D = 0)$ and $F_1(y_1 \mid X, D = 1)$ for the eligible population ($\xi = 1$).
- Agents made eligible for the program self-select as usual.
- Noncompliance is a sort of information — rarely used

- For **those deemed ineligible** $(\xi = 0)$, under our assumptions, we identify the distribution of Y_0, which can be partitioned into components for those who would have participated in the program had it not been for the randomization and components for those who would not have participated if offered the opportunity to do so:

$$F_0(y_0 \mid X) = F_0(y_0 \mid X, D = 0)\Pr(D = 0 \mid X) + F_0(y_0 \mid X, D = 1)\Pr(D = 1 \mid X).$$

- Since we know $F_0(y_0 \mid X, D = 0)$ and $\Pr(D = 1 \mid X)$ from the eligible population, we can identify $F_0(y_0 \mid X, D = 1)$.
- This is the new piece of information produced by the randomization compared to what can be obtained from standard observational data.

- In particular, we can identify the parameter TT, $E(Y_1 - Y_0 \mid X, D = 1)$, but without further assumptions, we cannot identify the other treatment parameters
- ATE $(= E(Y_1 - Y_0 \mid X))$
- or the joint distributions $F(y_0, y_1 \mid X)$
- or $F(y_0, y_1 \mid X, D = 1)$.
- To show that ξ is a valid instrument for A, form the Wald estimand,

$$\text{IV}_{\text{(e-1)}} = \frac{E(Y \mid \xi = 1, Z, X) - E(Y \mid \xi = 0, Z, X)}{\Pr(A = 1 \mid \xi = 1, Z, X) - \Pr(A = 1 \mid \xi = 0, Z, X)}. \quad (3)$$

Under invariance assumption autonomy, $\Pr(D = 1 \mid Z, X)$ is the same in the presence or absence of randomization.

- If we assume full compliance so that agents randomized to ineligibility do not show up in the program,

$$\Pr\left(A=1 \mid \xi=0, Z, X\right)=0,$$

and

$$\begin{aligned} E\left(Y \mid \xi=0, Z, X\right) &= E\left(Y_0 \mid Z, X\right) \\ &= E\left(Y_0 \mid D=1, X, Z\right) \Pr\left(D=1 \mid X, Z\right) \\ &\quad + E\left(Y_0 \mid D=0, X, Z\right) \Pr\left(D=0 \mid X, Z\right). \end{aligned}$$

- Going back to IV, Z also satisfies the requirement (IV-1) that it is an instrument, then $E\left(Y_0 \mid Z, X\right)=E\left(Y_0 \mid X\right)$.
- However, under (e-1) or (e-1)$'$ we do not have to assume that Z is a valid instrument.

- Using (e-1) and assumption autonomy, the first term in the numerator of (3) can be written as

$$\begin{aligned} E(Y \mid \xi = 1, Z, X) &= E(Y_1 \mid D = 1, Z, X) \Pr(D = 1 \mid Z, X) \\ &\quad + E(Y_0 \mid D = 0, Z, X) \Pr(D = 0 \mid Z, X). \end{aligned}$$

- Substituting this expression into the numerator of equation (3) and collecting terms, $\text{IV}_{\text{(e-1)}}$ identifies the parameter treatment on the treated:

$$\text{IV}_{\text{(e-1)}} = E(Y_1 - Y_0 \mid D = 1, Z, X).$$

- It does not identify the other mean treatment effects, such as LATE or the average treatment effect ATE, unless the common coefficient model governs the data or $(Y_1 - Y_0)$ is mean independent of D.

- Note $F(y \mid X) = E(\mathbf{1}(Y \leq y) \mid X)$,
- $\text{IV}_{\text{(e-1)}}$ also identifies $F_0(y_0 \mid X, D = 1)$,
- since we can compute conditional means of $\mathbf{1}(Y \leq y)$ for all y.
- The distribution $F_1(y_1 \mid X, D = 1)$ can be identified from observational data.
- Thus we can identify the outcome distributions for Y_0 and for Y_1 separately, conditional on $D = 1, X, Z$
- Without additional assumptions we cannot identify the joint distribution of outcomes or the other treatment parameters.
- This is a fundamental problem (the evaluation problem: we observe $Y = DY_1 + (1 - D)Y_0$ but not both Y_0, Y_1)

- The second type of eligibility randomization proceeds conditionally on $D = 1$.
- Accordingly, data generated from such experiments do not identify choice probabilities ($\Pr(D = 1 \mid X, Z)$) and hence do not identify the subjective evaluations of agents (Heckman, 1992).
- Under autonomy and (e-2) (or equivalent conditions (e-2)$'$) randomization identifies $F_0(y_0 \mid D = 1, X, Z)$ from the data on the randomized-out participants.
- This conditional distribution cannot be constructed from ordinary observational data unless additional assumptions are invoked.
- From the data for the eligible ($\xi = 1$) population, we identify $F_1(y_1 \mid D = 1, X, Z)$.

- The Wald estimator for mean outcomes in this case is

$$\mathrm{IV}_{\text{(e-2)}} = \frac{E\left(Y \mid D=1, \xi=1, X, Z\right) - E\left(Y \mid D=1, \xi=0, X, Z\right)}{\Pr\left(A=1 \mid D=1, \xi=1, X, Z\right) - \Pr\left(A=1 \mid D=1, \xi=0, X, Z\right)}.$$

- Under (e-2)/(e-2)′,

$$\begin{aligned}
\Pr\left(A=1 \mid D=1, \xi=1, X, Z\right) &= 1, \\
\Pr\left(A=1 \mid D=1, \xi=0, X, Z\right) &= 0, \\
E\left(Y \mid A=0, D=1, \xi=0, X, Z\right) &= E\left(Y_0 \mid D=1, X, Z\right), \text{ and} \\
E\left(Y \mid A=1, D=1, \xi=1, X, Z\right) &= E\left(Y_1 \mid D=1, X, Z\right).
\end{aligned}$$

- Thus,

$$\mathrm{IV}_{\text{(e-2)}} = E\left(Y_1 - Y_0 \mid D=1, X, Z\right).$$

- In the general model with essential heterogeneity, randomized trials with full compliance that do not disturb the activity being evaluated **answer a limited set of questions**,
- Do not in general identify the policy relevant treatment effect (PRTE).
- Randomizations have to be carefully chosen to make sure that they answer interesting economic questions.
- Their analysis has to be supplemented with the methods previously studies to answer the full range of policy questions addressed there.

Randomization Bias

- If randomization alters the program being evaluated, the outcomes of a randomized trial may bear little resemblance to the outcomes generated by an ongoing version of the program that has not been subject to randomization.
- In this case, variance is violated.
- Such violations are termed "Hawthorne effects" and are called "Randomization Bias" in the economics literature.
- The process of randomization may affect objective outcomes, subjective outcomes or both.

- Although the program studied may be changed, randomization can produce "internally valid" treatment effects for the altered program.
- Thus randomization can answer policy question (P-1) for a program changed by randomization, but not for the program as it would operate in the absence of randomization.

- Randomization bias?
- This case might arise when randomization alters risk-averse agent decision behavior but has no effects on potential outcomes.
- Thus the $R(s, \omega)$ (subjective values) are affected, but not the $Y(s, \omega)$.

- In this case, the parameter $\text{ATE}(X) = E(Y_1 - Y_0 \mid X)$ is the same in the ongoing program as in the population generated by the randomized trial.
- However, treatment parameters conditional on choices such as $\text{TT}(X) = E(Y_1 - Y_0 \mid X, D = 1)$, $\text{TUT}(X) = E(Y_1 - Y_0 \mid X, D = 0)$ are not, in general, invariant.
- If the subjective valuations are altered, so are the parameters based on choices produced by the subjective valuations.
- Different random variables generate the conditioning sets in the randomized and nonrandomized regimes and, in general, they will have a different dependence structure with the outcomes $Y(s, \omega)$.

- This arises because randomization alters the composition of participants in the conditioning set that defines the treatment parameter, just like two different sets of instruments do in LATE. In general, this alters LATE.

Parameters Defined on Choices Altered are Shifted

- Treatment parameters defined conditional on choices are not invariant to the choice of randomization.
- This insight shows the gain in clarity in interpreting what experiments identify from adopting a choice-theoretic, econometric approach to the evaluation of social programs, as opposed to the conventional approach adopted by statisticians.

Compliance

- The statistical treatment effect literature extensively analyzes the problem of noncompliance.
- Persons assigned to a treatment may not accept it.
- In the notation of equation (3), let $\xi = 1$ if a person is assigned to treatment, $\xi = 0$ otherwise.
- Compliance is said to be perfect when $\xi = 1 \Rightarrow A = 1$ and $\xi = 0 \Rightarrow A = 0$.
- In the presence of self selection by agents, these conditions do not, in general, hold.
- People assigned to treatment may not comply ($\xi = 1$ but $D = 0$).

- This is also called the "**dropout**" problem (Mallar, Kerachsky, and Thornton, 1980; Bloom, 1984).
- In its formulation of this problem, the literature assumes that outcomes are measured for each participant but that outcomes realized are not always those intended by the randomizers.
- In addition, people denied treatment may find substitutes for the treatment outside of the program.
- This is the problem of **substitution bias**.
- Since self-selection is an integral part of choice models, noncompliance, as the term is used by the statisticians, is a feature of most social experiments.

- Noncompliance is a source of information about subjective evaluations of programs.

- Noncompliance is a problem if the goal of the social experiment is to estimate $\text{ATE}(X) = E(Y_1 - Y_0 \mid X)$ without using the econometric methods previously discussed.

The Dynamics of Dropout and Program Participation

- Actual programs are more dynamic in character than the stylized program just analyzed.
- Multiple actors are involved, such as the agents being studied and the groups administering the programs.
- People apply, are accepted, enroll, and complete the program.

- Analyze the effects of dropouts on inferences from social experiments and assume no attrition.

- Consider a stylized multiple stage program.
- In stage "0 ", the agent (possibly in conjunction with program officials) decides to participate or not to participate in the program.
- This is an enrollment phase prior to treatment.
- Let $D_0 = 1$ denote that the agent does not choose to participate.
- $D_0 = 0$ denotes that the agent participates and receives some treatment among J possible program levels beyond the no treatment state (this may be confusing because here alone $D = 1$ means you stop).
- The outcome associated with state "0" is Y_0.
- This assumes that acts of inquiry about a program or registration in it have no effect on outcomes.

- One could disaggregate stage "0" into recruitment, application, and acceptance stages, but for expositional simplicity we collapse these into one stage.
- If the J possible treatment stages are ordered, say, by the intensity of treatment, then "1" is the least amount of treatment and "J" is the greatest amount.
- A more general model would allow people to transit to stage j but not complete it.
- The J distinct stages can be interpreted quite generally.
- If a person no longer participates in the program after stage j, $j = 1, \ldots, J$, we set indicator $D_j = 1$.
- The person is assumed to receive stage j treatment.

- $D_J = 1$ corresponds to completion of the program in all J stages of its treatment phase.
- Note that, by construction, $\sum_{j=0}^{J} D_j = 1$.

- Let $\{D_j(z)\}_{z \in \mathcal{Z}}$ be the set of potential treatment choices for choice j associated with setting $Z = z$.
- For each $Z = z$, $\sum_{j=0}^{J} D_j(z) = 1$.
- Different components of Z may determine different choice indicators.

- Array the collections of choice indicators evaluated at each $Z = z$ into a vector

$$D(z) = \left(\{D_1(z)\}_{z \in \mathcal{Z}}, \ldots, \{D_J(z)\}_{z \in \mathcal{Z}}\right).$$

The potential outcomes associated with each of the $J + 1$ states are

$$Y_j = \mu_j(X, U_j), \qquad j = 0, \ldots, J.$$

Y_0 is the no treatment state, and the Y_j, $j \geq 1$, correspond to outcomes associated with dropping out at various stages of the program.

- In the absence of randomization, the observed Y is

$$Y = \sum_{j=0}^{J} D_j Y_j,$$

  the Roy-Quandt switching regime model.
- Let $\tilde{Y} = (Y_0, \ldots, Y_J)$ denote the vector of potential outcomes associated with all phases of the program.
- Through selection, the Y_j for persons with $D_j = 1$ may be different from the Y_j for persons with $D_j = 0$.

- Consider what randomizations at various stages identify.
- We assume that the randomizations do not disturb the program.
- **Autonomy**.
- Recall that we also assume absence of general equilibrium effects.
- Let $\xi_j = 1$ denote whether the person is eligible to move beyond stage j.
- $\xi_j = 0$ means the person is randomized out of the program after completing stage j.

- A randomization at stage j with $\xi_j = 1$ means the person is allowed to continue on to stage $j+1$, although the agent may still choose not to.
- We set $\xi_J \equiv 1$ to simplify the notation.
- The ξ_j are ordered in a natural way: $\xi_j = 1$ only if $\xi_\ell = 1$, $\ell = 0, \ldots, j-1$.
- Array the ξ_j into a vector ξ and denote its support by $\tilde{\xi}$.

- Because of agent self-selection, a person who does not choose to participate at stage j cannot be forced to do so.
- For a person who would choose k ($D_k = 1$) in a nonexperimental environment, Y_k is observed if $\prod_{\ell=0}^{k-1} \xi_\ell = 1$.
- Otherwise, if $\xi_{k-1} = 0$ but, say, $\prod_{\ell=0}^{k'-1} \xi_\ell = 1$ and $\prod_{\ell=0}^{k'} \xi_\ell = 0$ for $k' < k$, we observe $Y_{k'}$ for the agent.
- From an experiment with randomization administered at different stages, we observe

$$Y = \sum_{j=0}^{J} D_j \overbrace{\left(\sum_{k=0}^{j} \underbrace{\left(\prod_{\ell=0}^{k-1} \xi_\ell \right)}_{\text{person stops at } k} (1 - \xi_k)\, Y_k \right)}^{\text{outcome for a dropout at stage } j}.$$

To understand this formula, consider a program with three

- For a person who would like to complete the program ($D_3 = 1$), but is stopped by randomization after stage 2, we observe Y_2 instead of Y_3.
- If the person is randomized out after stage 1, we observe Y_1 instead of Y_3.
- Let A_k be the indicator that we observe the agent with a stage k outcome.
- This can happen if a person would have chosen to stop at stage k ($D_k = 1$) and survives randomization through k ($\prod_{\ell=0}^{k-1} \xi_\ell = 1$), or if a person would have chosen to stop at a stage later than k but was thwarted from doing so by the randomization and settles for the best attainable state given the constraint imposed by the randomization.

- We can define A_k as

$$A_k = D_k \prod_{\ell=0}^{k-1} \xi_\ell + \sum_{j \geq k} D_j \left(\prod_{\ell=0}^{k-1} \xi_\ell \right) (1 - \xi_k), \ k = 1, \ldots, J.$$

  If a person who chooses $D_k = 1$ survives all stages of randomization through $k-1$ and hence is allowed to transit to k, we observe Y_k for that person.

- For persons who would choose $D_j = 1$, $j > k$, but get randomized out at k, i.e., $\left(\prod_{\ell=0}^{k-1} \xi_\ell \right) (1 - \xi_k) = 1$, we also observe Y_k.

- We now state the conditions under which sequential randomizations are instrumental variables for estimation stage-wise treatment effects A_j.
- Let $A_i(z, e_i)$ be the value of A_i when $Z = z$ and $\xi_i = e_i$.
- Array the A_i, $i = 1, \ldots, J$, into a vector

$$A(z, e) = (A_1(z, e_1), A_2(z, e_2), \ldots, A_J(z, e_J)).$$

- A variety of randomization mechanisms are possible in which randomization depends on the information known to the randomizer at each stage of the program.

- IV conditions for ξ are satisfied under the following sequential randomization assumptions.
- (e-3a): $\xi_i \perp\!\!\!\perp \left(\tilde{Y}, \{A(z,e)\}_{(z,e)\in\mathcal{Z}\times\tilde{\xi}} \mid X, Z, D_\ell = 1 \text{ for } \ell < i, \prod_{\ell=0}^{i-1}\xi_\ell = 1\right)$, for $i = 1, \ldots, J$,

and

- (e-3b): $\Pr\left(A_i = 1 \mid X,\ Z,\ D_\ell = 1 \text{ for } \ell < i,\ \xi_i,\ \prod_{\ell=0}^{i-1}\xi_\ell = 1\right)$ depends on ξ_i, for $i = 1, \ldots, J$.
- These conditions also appear in a lot of discrete dynamic choice models, i.e., Rust

- These expressions assume that the components of $\tilde{Y} = (Y_0, \ldots, Y_J)$ that are realized are known to the randomizer after the dropout decision is made, and thus cannot enter the conditioning set for the sequential randomizations.

- To fix ideas, consider a randomization of eligibility ξ_0, setting $\xi_1 = \cdots = \xi_J = 1$.
- This is a randomization that makes people eligible for participation at all stages of the program.
- We investigate what this randomization identifies, assuming invariance conditions hold.

- For those declared eligible,

$$E\left(Y \mid \xi_0 = 1\right) = \sum_{j=0}^{J} E\left(Y_j \mid D_j = 1\right) \Pr\left(D_j = 1\right). \qquad (4)$$

For those declared ineligible,

$$\underbrace{E\left(Y \mid \xi_0 = 0\right)}_{\text{this we know}} = \sum_{j=0}^{J} \underbrace{E\left(Y_0 \mid D_j = 1\right)}_{?} \overbrace{\Pr\left(D_j = 1\right)}^{\text{this we know}}, \qquad (5)$$

since agents cannot participate in any stage of the program and are all in the state "0" with outcome Y_0.

- From observed choice behavior, we can identify each of the components of (4).

- We observe $\Pr(D_j = 1)$ from observed choices of treatment, and we observe $E(Y_j \mid D_j = 1)$ from observed outcomes for each treatment choice.
- Except for the choice probabilities $(\Pr(D_j = 1), j = 0, \ldots, J)$ and $E(Y_0 \mid D_0 = 1)$, we cannot identify individual components of (5) for $J > 1$.
- When $J = 1$, we can identify all of the components of (5).
- The individual components of (5) cannot, without further assumptions, be identified by the experiment, although the sum can be.
- Comparing the treatment group with the control group, we obtain the "intention to treat" parameter with respect to the randomization of ξ_0 alone, setting $\xi_1 = \cdots = \xi_J = 1$ for anyone for whom $\xi_0 = 1$.

- What is identified from the experiment?

$$E\left(Y \mid \xi_0 = 1\right) - E\left(Y \mid \xi_0 = 0\right) = \sum_{j=1}^{J} E\left(Y_j - Y_0 \mid D_j = 1\right) \Pr\left(D_j = 1\right). \tag{6}$$

- For $J > 1$, this simple experimental estimator does not identify the effect of full participation in the program for those who participate ($E\left(Y_J - Y_0 \mid D_J = 1\right)$) unless additional assumptions are invoked
- E.g., the assumption that partial participation has the same mean effect as full participation for persons who drop out at the early stages, i.e., $E\left(Y_j - Y_0 \mid D_j = 1\right)$ $= E\left(Y_J - Y_0 \mid D_j = 1\right)$ for all j.
- This assumption might be appropriate if just getting into the program is all that matters — a pure signalling effect.

- A second set of conditions for identification of this parameter is that $E\left(Y_j - Y_0 \mid D_j = 1\right) = 0$ for all $j < J$.
- Under those conditions, if we divide the mean difference by $\Pr\left(D_J = 1\right)$, we obtain the "Bloom" estimator (Mallar, Kerachsky, and Thornton, 1980, Bloom, 1984)

$$\text{IV}_{\text{Bloom}} = \frac{E\left(Y \mid \xi_0 = 1\right) - E\left(Y \mid \xi_0 = 0\right)}{\Pr\left(D_J = 1\right)},$$

  assuming $\Pr\left(D_J = 1\right) \neq 0$.
- This is an IV estimator using ξ_0 as the instrument for A_J.
- In general, the mean difference between the treated and the controlled identifies only the composite term shown in (6).
- In this case, the simple randomization estimator identifies a not-so-simple or easily interpreted parameter.

- More generally, if we randomize persons out after completing stage k ($[\prod_{\ell=0}^{k-1} \xi_\ell](1-\xi_k)=1$) and for another group establish full eligibility at all stages ($\prod_{\ell=0}^{J} \xi_\ell = 1$), we obtain

$$E\left[Y \middle| \prod_{\ell=0}^{J} \xi_\ell = 1\right] - E\left[Y \middle| \left(\prod_{\ell=0}^{k-1} \xi_\ell\right)(1-\xi_k) = 1\right]$$
$$= \sum_{j=k}^{J} E\left(Y_j - Y_k \mid D_j = 1\right) \Pr\left(D_j = 1\right),$$

- **Problem: Prove this**

- Hence, since we know $E\left(Y_k \mid D_k = 1\right)$ and $\Pr\left(D_k = 1\right)$ from observational data, we can identify the combination of parameters

$$\sum_{j=k+1}^{J} E\left(Y_k \mid D_j = 1\right) \Pr\left(D_j = 1\right), \tag{7}$$

for each randomization that stops persons from advancing beyond level k, $k = 0, \ldots, J-1$.

- Observe that a randomization of eligibility that prevents people from going to stage $J-1$ but not to stage J ($[\prod_{\ell=0}^{J-2}\xi_\ell](1-\xi_{J-1})=1$) identifies $E(Y_J - Y_{J-1} \mid D_J = 1)$:

$$\begin{aligned} & E\left(Y \mid \xi_0 = 1, \ldots, \xi_{J-2} = 1, \xi_{J-1} = 0\right) \\ = & \left[\sum_{j=0}^{J-1} E\left(Y_j \mid D_j = 1\right) \Pr\left(D_j = 1\right)\right] + E\left(Y_{J-1} \mid D_J = 1\right) \Pr\left(D_J = 1\right). \end{aligned}$$

- Thus,

$$\begin{aligned} & E\left(Y \mid \xi_0 = 1, \ldots, \xi_J = 1\right) - E\left(Y \mid \xi_0 = 1, \ldots, \xi_{J-1} = 1, \xi_J = 0\right) \\ = & E\left(Y_J - Y_{J-1} \mid D_J = 1\right) \Pr\left(D_J = 1\right). \end{aligned}$$

Since $\Pr\left(D_J = 1\right)$ is observed from choice data, as is $E\left(Y_J \mid D_J = 1\right)$, we can identify $E\left(Y_{J-1} \mid D_J = 1\right)$ from the experiment.

- In the general case under autonomy, a randomization that prevents agents from moving beyond stage ℓ $(\xi_0 = 1, \ldots, \xi_{\ell-1} = 1, \xi_\ell = 0)$ directly identifies

$$E\left(Y \mid \xi_0 = 1, \ldots, \xi_{\ell-1} = 1, \xi_\ell = 0\right) =$$

$$= \underbrace{\sum_{j=0}^{\ell} E\left(Y_j \mid D_j = 1\right) \Pr\left(D_j = 1\right)}_{\text{all components known from observational data}}$$

$$+ \underbrace{\sum_{j=\ell+1}^{J} E\left(Y_\ell \mid D_j = 1\right) \Pr\left(D_j = 1\right)}_{\text{sum and probability weights known, but not individual } E(Y_\ell | D_j=1)}.$$

- All of the components of the first set of terms on the right-hand side are known from observational data.

- The probability terms in the second set of terms are known, but the individual conditional expectations $E(Y_\ell \mid D_j = 1)$, $j = \ell + 1, \ldots, J$, are not known without further assumptions.
- Randomization at stage ℓ is an IV.
- To show this, decompose the observed outcome Y into components associated with each value of A_j, the indicator associated with observing a stage j outcome:

$$Y = \sum_{j=0}^{J} A_j Y_j.$$

- We can interpret ξ_ℓ as an instrument for A_ℓ.

- Keeping the conditioning on X, Z implicit, we obtain

$$\begin{aligned} \mathrm{IV}_{\xi_\ell} &= \frac{E\left[Y \mid \xi_\ell = 0\right] - E\left[Y \mid \xi_\ell = 1\right]}{\Pr\left(A_\ell = 1 \mid \xi_\ell = 0\right) - \Pr\left(A_\ell = 1 \mid \xi_\ell = 1\right)} \\ &= \frac{\sum_{j=\ell+1}^{J} E\left[Y_\ell - Y_j \mid D_j = 1\right] \Pr\left(D_j = 1\right)}{\sum_{j=\ell+1}^{J} \Pr\left(D_j = 1\right)}, \ \ell = 0, \ldots, J-1. \end{aligned}$$

- By the preceding analysis, we know the numerator term as a sum but not the individual components.
- We know the denominator from choices measured in observational data and invariance assumption autonomy.
- The IV formalism is less helpful in the general case.

- Table 1 summarizes the parameters or combinations of parameters that can be identified from randomizations performed at different stages.
- It presents the array of factual and counterfactual conditional mean outcomes $E\left(Y_j \mid D_\ell = 1\right)$, $j = 0, \ldots, J$ and $\ell = 0, \ldots, J$.
- The conditional mean outcomes obtained from observational data are on the diagonal of the table $(E(Y_j \mid D_j = 1), j = 0, \ldots, J)$.
- Because of choices of agents, experiments do not directly identify the elements in the table that are above the diagonal.
- Under autonomy, experiments described at the base of the table identify the *combinations* of the parameters below the diagonal.

- Recall that if $\xi_\ell = 0$, the agent cannot advance beyond stage ℓ.
- If we randomly deny eligibility to move to J ($\xi_{J-1} = 0$), we point identify $E(Y_{J-1} \mid D_J = 1)$, as well as the parameters that can be obtained from observational data.
- In general, we can only identify the combinations of parameters shown at the base of the table.
- Terms below the diagonal are not known in general

Evidence on Randomization Bias

- Violations of autonomy assumption in the general case with essential heterogeneity affect the interpretation of the outputs of social experiments.

- In the 1980s, the U.S. Department of Labor financed a large-scale experimental evaluation of the ongoing, large-scale manpower training program authorized under the Job Training Partnership Act (JTPA).
- A study by Doolittle and Traeger (1990) gives some indirect information from which it is possible to determine whether randomization bias was present in an ongoing program.
- Job training in the United States is organized through geographically decentralized centers.

- These centers receive incentive payments for placing unemployed persons and persons on welfare in “high-paying” jobs.
- The participation of centers in the experiment was not compulsory.
- Funds were set aside to compensate job centers for the administrative costs of participating in the experiment.
- The funds set aside range from 5 percent to 10 percent of the total operating costs of the centers.

- In attempting to enroll geographically dispersed sites, MDRC experienced a training center refusal rate in excess of 90 percent.
- By randomizing, the centers had to widen the available pool of persons deemed eligible, and there was great concern about the effects of this widening on applicant quality—precisely the behavior ruled out by autonomy.

Table 1: Percentage of Local JTPA Agencies Citing Specific Concerns About Participating in the Experiment

Concern	Percentage of Training Centers Citing the Concern
1. Ethical and public relations implications of:	
a. Random assignment in social programs	61.8
b. Denial of services to controls	54.4
2. Potential negative effect of creation of a control group on achievement of client recruitment goals	47.8
3. Potential negative impact on performance standards	25.4
4. Implementation of the study when service providers do intake	21.1
5. Objections of service providers to the study	17.5
6. Potential staff administrative burden	16.2
7. Possible lack of support by elected officials	15.8
8. Legality of random assignment and possible grievances	14.5
9. Procedures for providing controls with referrals to other services	14.0
10. Special recruitment problems for out-of-school youth	10.5
Sample size	228

- In attempting to entice centers to participate, MDRC had to reduce the randomized rejection probability from $\frac{1}{2}$ to as low as $\frac{1}{6}$ for certain centers.
- The resulting reduction in the size of the control group impairs the power of statistical tests designed to test the null hypothesis of no program effect.
- Compensation for participation was expanded sevenfold in order to get any centers to participate in the experiment.

- The MDRC analysts conclude:

> Implementing a complex random assignment research design in an ongoing program providing a variety of services does inevitably change its operation in some ways. The most likely difference arising from a random assignment field study of program impacts is a change in the mix of clients served. Expanded recruitment efforts, needed to generate the control group, draw in additional applicants who are not identical to the people previously served. A second likely change is that the treatment categories may somewhat restrict program staff's flexibility to change service recommendations

(Doolittle and Traeger, 1990, p. 121).

These authors go on to note that

> some [training centers] because of severe recruitment problems or up-front services cannot implement the type of random assignment model needed to answer the various impact questions without major changes in procedures

(Doolittle and Traeger, 1990, p. 123).

- During the JTPA experiment conducted at Corpus Christi, Texas, center administrators successfully petitioned the government of Texas for a waiver of its performance standards on the ground that the experiment disrupted center operations.
- Self-selection likely guarantees that participant sites are the least likely sites to suffer disruption.

- Such selective participation in the experiment calls into question the validity of experimental estimates as a statement about the JTPA system as a whole, as it clearly poses a threat to external validity — problem (P-2) as defined in Part I.
- Torp et al. (1993) report similar problems in a randomized evaluation of a job training program in Norway.

- Kramer and Shapiro (1984) note that subjects in drug trials were less likely to participate in randomized trials than in nonexperimental studies.

- In a study of treatment of adults with cirrhosis, no effect of the treatment was found for participants in a randomized trial.
- But the death rates for those randomized out of the treatment were substantially lower than among those individuals who refused to participate in the experiment, despite the fact that both groups were administered the same alternative treatment.
- Part of any convincing identification strategy by randomization requires that the agent document the absence of randomization bias.

Evidence on Dropping Out and Substitution Bias

- Dropouts are a feature of all social programs.
- Most social programs have good substitutes, so that the estimated effect of a program as typically estimated has to be defined relative to the full range of substitute activities in which non-participants engage.
- Experiments exacerbate this problem by creating a pool of persons who attempt to take training who then flock to substitute programs when they are placed in an experimental control group.

- Table 2 (reproduced from Heckman, Hohmann, Smith, and Khoo, 2000) demonstrates the practical importance of both dropout and substitution bias in experimental evaluations.

Table 2: Fraction of Experimental Treatment and Control Groups Receiving Services in Experimental Evaluations of Employment and Training Programs

Study	Authors/time period	Target group(s)	Fraction of treatments receiving services	Fraction of controls receiving services
1. NSW	Hollister, et al. (1984) (9 months after RA)	Long-term AFDC women	0.95	0.11
		Ex-addicts	NA	0.03
		17-20 year old high school dropouts	NA	0.04
2. SWIM	Friedlander and Hamilton (1993) (Time period not reported)	AFDC women: applicants and recipients		
		a. Job search assistance	0.54	0.01
		b. Work experience	0.21	0.01
		c. Classroom training/OJT	0.39	0.21
		d. Any activity	0.69	0.30
		AFDC-U unemployed fathers		
		a. Job search assistance	0.60	0.01
		b. Work experience	0.21	0.01
		c. Classroom training/OJT	0.34	0.22
		d. Any activity	0.70	0.23
3. JOBSTART	Cave, et al. (1993) (12 months after RA)	Youth high school dropouts		
		Classroom training/OJT	0.90	0.26
4. Project Independence	Kemple, et al. (1995) (24 months after RA)	AFDC women: applicants and recipients		
		a. Job search assistance	0.43	0.19
		b. Classroom training/OJT	0.42	0.31
		c. Any activity	0.64	0.40

Table 15 [Continued]

Study	Authors/time period	Target group(s)	Fraction of treatments receiving services	Fraction of controls receiving services
5. New Chance	Quint, et al. (1994)	Teenage single mothers		
	(18 months after RA)	Any education services	0.82	0.48
		Any training services	0.26	0.15
		Any education or training	0.87	0.55
6. National JTPA Study	Heckman and Smith (1998)	Self-reported from survey data		
	(18 months after RA)	Adult males	0.38	0.24
		Adult females	0.51	0.33
		Male youth	0.50	0.32
		Female youth	0.81	0.42
		Combined Administrative Survey Data		
		Adult males	0.74	0.25
		Adult females	0.78	0.34
		Male youth	0.81	0.34
		Female youth	0.81	0.42

- There are counterpart findings in the application of randomized clinical trials.
- For example, Palca (1989) notes that AIDS patients denied potentially life-saving drugs took steps to undo random assignment.
- Patients had the pills they were taking tested to see if they were getting a placebo or an unsatisfactory treatment, and were likely to drop out of the experiment in either case or to seek more effective medication, or both.
- In the MDRC experiment, in some sites qualified trainees found alternative avenues for securing exactly the same training presented by the same subcontractors by using other methods of financial support.

Aigner, D. J. (1979, September). Sample design for electricity pricing experiments: Anticipated precision for a time-of-day pricing experiment. *Journal of Econometrics 11*(1), 195–205.

Banerjee, A. V. (2006, July/August). Making aid work: How to fight global poverty—effectively. *Boston Review 31*(4), 7–9.

Bloom, H. S. (1984, April). Accounting for no-shows in experimental evaluation designs. *Evaluation Review 8*(2), 225–246.

Cain, G. G. and H. W. Watts (1973). *Income Maintenance and Labor Supply: Econometric Studies*. New York: Academic Press.

Doolittle, F. C. and L. Traeger (1990). *Implementing the National JTPA Study*. New York: Manpower Demonstration Research Corporation.

Heckman, J. J. (1992). Randomization and social policy evaluation. In C. F. Manski and I. Garfinkel (Eds.), *Evaluating Welfare and Training Programs*, Chapter 5, pp. 201–230. Cambridge, MA: Harvard University Press.

Heckman, J. J., N. Hohmann, J. Smith, and M. Khoo (2000, May). Substitution and dropout bias in social experiments: A study of an influential social experiment. *Quarterly Journal of Economics 115*(2), 651–694.

Hurwicz, L. (1962). On the structural form of interdependent systems. In E. Nagel, P. Suppes, and A. Tarski (Eds.), *Logic, Methodology and Philosophy of Science*, pp. 232–239. Stanford University Press.

Kramer, M. S. and S. H. Shapiro (1984, November). Scientific challenges in the application of randomized trials. *JAMA: the Journal of the American Medical Association 252*(19), 2739–2745.

Mallar, C., S. Kerachsky, and C. Thornton (1980). The short-term economic impact of the Job Corps program. In E. W. Stromsdorfer and G. Farkas (Eds.), *Evaluation Studies Review*

Annual, Volume 5, pp. 332–359. Beverly Hills, CA: Sage Publications.

Palca, J. (1989, October 6). AIDS drug trials enter new age. *Science, New Series 246*(4926), 19–21.

Torp, H., O. Raaum, E. Hernæs, and H. Goldstein (1993). The first Norwegian experiment. In G. Burtless (Ed.), *Measuring Labour Market Measures: Evaluating the Effects of Active Labour Market Policy Initiatives*, Copenhagen. Proceedings from the Danish Presidency Conference "Effects and Measuring of Effects of Labour Market Policy Initiatives" Kolding, May 1993: Ministry of Labour.