Two cases have been made for the application of social experimentation.

One is the classical argument for experimental design.

Inducing variation in regressors increases precision of estimates and the power of tests.

The other case focuses on solving endogeneity and self-selection problems.

Randomization is an instrumental variable.

The two cases are compatible, but entail different emphases.
Both cases can be motivated within a linear regression model for outcome $Y$ with treatment indicator $D$ and covariates $X$:

$$Y = X\alpha + D\beta + U,$$  \hspace{1cm} (1.1)

where $U$ is an unobservable.

“$D$” is treatment indicator, e.g., schooling, medical intervention.

$\beta$ may be the same for all observations (conditional on $X$) as in the common coefficient setup, or it may be a variable coefficient of the type used extensively in the recent literature.

$D$ (and the $X$) may be statistically dependent on $U$.

Entertain the possibility that when $\beta$ is random it is dependent on $D$, as in the generalized Roy model.

A lot of recent evidence suggests that this is an important consideration.
Both cases for social experimentation seek to secure identification of some parameters of (1.1) or parameters that can be generated from (1.1).

Analysts advocating the first case for experimentation typically assume a common coefficient model for $\alpha$ and $\beta$.

They address the problem that variation in $(X, D)$ may be insufficient to identify or precisely estimate $(\alpha, \beta)$.

Manipulating $(X, D)$ through randomization, or more generally, through controlled variation, can secure identification.

It is typically assumed that $(X, D)$ is independent of $U$ or at least mean independent.
Examples in economics of experimentation designed to increase the variation in the regressors are studies by Conlisk (1973), Conlisk and Watts (1969), and Aigner (1979a,b, 1985).

The papers by Conlisk show how experimental manipulation can solve a multicollinearity problem.

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 (1979b) shows how variation in the prices paid for electricity across the day can identify price effects that cannot be identified in regimes with uniform prices across all hours of the day.
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.
The second case for social experiments and the one that now receives the most attention in applied work in economics and in this paper focuses on the dependence between \((X, D)\) and \(U\) that invalidates least squares as an estimator of the causal effect of \(X\) and \(D\) on \(Y\).

This is the problem of least squares bias raised by Haavelmo (1943).

In the second case, experimental variation in \((X, D)\) is sought to make it “exogenous” or “external” to \(U\).

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.

It implicitly assumes that no economic model is needed to interpret the evidence.
Randomization is an instrument.

It shares all of the assets and liabilities of IV (Heckman and Vytlacil, 2007).

In particular, randomization applied to a correlated random coefficient (or a model of essential heterogeneity) raises issues about the multiplicity of parameters identified by different randomizations as arise in connection with the multiplicity of parameters identified by different instruments (Heckman, Urzua, and Vytlacil, 2006).
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

- The argument justifying randomization as an instrument assumes that randomization (or more generally the treatment assignment rule) does not alter subjective or objective potential outcomes.

- Also maintain absence of general equilibrium effects throughout.
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.

Agents choose and experimenters can only manipulate choice sets.
Let $\xi = 1$ if an agent is eligible to participate in the program; $\xi = 0$ otherwise.

$\tilde{\xi} = \{0, 1\}$ is the set of possible values of $\xi$.

Let $D$ indicate participation under ordinary conditions.

In the absence of randomization, $D$ is an indicator of whether the agent actually participates in the program.

Let actual participation be $A$.

By construction, under invariance conditions,

$$A = D\xi. \quad (2.1)$$

This assumes that eligibility is strictly enforced.
Consider two types of randomization of eligibility.

**Randomization of Type 1.**

A random mechanism (possibly conditional on \((X, Z)\)) is used to determine \(\xi\). The probability of eligibility is \(\Pr(\xi = 1 \mid X, Z)\).
This randomization affects the eligibility of the agent for the program but because agents still self-select, there is no assurance that eligible agents will participate in the program.

This condition does not impose exogeneity on $X, Z$. 
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 $\Pr(\xi = 1 \mid X, Z, Q, D = 1)$. This assumes invariance.
What Does Randomization Identify?

\[ Y = DY_1 + (1 - D) Y_0 \]
\[ Y_1 = \mu_1(X) + U_1 \]
\[ Y_0 = \mu_0(X) + U_0 \]

- Under an invariance assumption and under one of the sets of randomization assumptions just presented, IV is an instrument that identifies some treatment effect for an ongoing program.
- The question is: which treatment effect?
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 $D$ correlated with $\beta$, the instruments generated by randomization can identify parameters that are far from the parameters of economic interest.
Thus, unless we condition on the other instruments, the IV defined by randomization can be negative even if all of the underlying treatment effects or LATEs and MTEs generating choice behavior are positive.

The weighted average of the MTE generated by the instrument may be far from the policy relevant treatment effect.
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.

For those deemed ineligible \((\xi = 0)\), under our assumptions, we would 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 $\xi$ is a valid instrument for $A$, form the Wald estimand,

$$\text{IV}_{(e-1)} = \frac{E(Y \mid \xi = 1, Z, X) - E(Y \mid \xi = 0, Z, X)}{\Pr(A = 1 \mid \xi = 1, Z, X) - \Pr(A = 1 \mid \xi = 0, Z, X)} = E(Y_1 - Y_0 \mid D = 1, Z, X).$$ \hspace{1cm} (3.1)

Assuming full compliance so that agents randomized to ineligibility do not show up in the program,

$$\Pr(A = 1 \mid \xi = 0, Z, X) = 0.$$
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 \(D\).

The distributions \(F_0(y_0 \mid X, D = 1)\) and \(F_1(y_1 \mid X, D = 1)\) can be identified from observational data.

Thus we can identify the outcome distributions for \(Y_0\) and for \(Y_1\) separately, conditional on \(D = 1, X, Z\), but without additional assumptions we cannot identify the joint distribution of outcomes or the other treatment parameters.
The second type of eligibility randomization proceeds conditionally on $D = 1$.

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; Moffitt, 1992).

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

\[
IV_{(e-2)} = \frac{E(Y \mid D = 1, \xi = 1, X, Z) - E(Y \mid D = 1, \xi = 0, X, Z)}{Pr(A = 1 \mid D = 1, \xi = 1, X, Z) - Pr(A = 1 \mid D = 1, \xi = 0, X, Z)}.
\]

\[
= E(Y_1 - Y_0 \mid D = 1, X, Z).
\]
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 analyzed to answer the full range of policy questions addressed there.
Many randomizations alter the environment they are studying and inject what may be unwelcome sources of uncertainty into agent decision making.
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.

- 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.
Randomization may still be a valid instrument for the altered program.

Although the program studied may be changed, randomization can produce “internally valid” treatment effects for the altered program.

Thus randomization can answer policy questions for a program changed by randomization, but not for the program as it would operate in the absence of randomization.
Randomization may alter risk-averse agent decision behavior but has no effects on potential outcomes.
In this case, the parameter \( \text{ATE}(X) = E (Y_1 - Y_0 | X) \) is the same in the ongoing program as in the population generated by the randomized trial.

However, treatment parameters conditional on choices such as \( \text{TT}(X) = E (Y_1 - Y_0 | X, D = 1) \), \( \text{TUT}(X) = E (Y_1 - Y_0 | X, D = 0) \) are not, in general, invariant.

If the subjective valuations are altered, so are the parameters based on choices produced by the subjective valuations.
Randomization alters the composition of participants in the conditioning set that defines the treatment parameter.

This analysis applies with full force to LATE.

In general, randomization alters LATE.
In general, 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.

We now show another advantage of the economic approach in an analysis of noncompliance and its implications for interpreting experimental evidence.
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 (3.1), 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 Thorton, 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.
The econometric approach builds in the possibility of self-selection as an integral part of model specification.

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 $\text{ATE}(X) = E(Y_1 - Y_0 | X)$. 
Experiments are conceived as tools for direct allocation of treatments.

For that reason, the experimental literature elevates ATE to pre-eminence as the parameter of interest because it is thought that this parameter can be produced by experiments.
In social experiments, it is rare that the experimenter can force anyone to do anything.

As the old adage goes, “you can lead a horse to water but you cannot make it drink.”

Agent choice behavior intervenes.

Thus it is no accident that if they are not compromised, the two randomizations most commonly implemented directly identify parameters conditional on choices.
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.
A fully dynamic analysis, along the lines of the models developed in Abbring and Heckman (2007), analyzes each of these decisions, accounting for the updating of agent and program administrators’ information.

This paper briefly discusses some new issues that arise in a more dynamic formulation of the dropout problem.

In this section, we analyze the effects of dropouts on inferences from social experiments and assume no attrition.

Our analysis of this case is of interest both in its own right and as a demonstration of the power of our approach.
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.

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, $\sum_{j=0}^{J} D_j = 1$.

• The sequential updating model developed below in Abbring and Heckman (2007) can be used to formalize these decisions and their associated outcomes.

• We can also use the simple multinomial choice model developed and analyzed in appendix B of Heckman and Vytlacil (2007a).
Let \( \{D_j(z)\}_{z \in Z} \) be the set of potential treatment choices for choice \( j \) associated with setting \( Z = z \).

For each \( Z = z \), \( \sum_{j=0}^{J} D_j(z) = 1 \).

Using the methods exposited in Abbring and Heckman (2007), we could update the information sets at each stage.

We keep this updating implicit.

Different components of \( Z \) may determine different choice indicators.

Array the collections of choice indicators evaluated at each \( Z = z \) into a vector

\[
D(z) = \left( \{D_1(z)\}_{z \in Z}, \ldots, \{D_J(z)\}_{z \in Z} \right).
\]
The potential outcomes associated with each of the \( J + 1 \) states are

\[ Y_j = \mu_j (X, U_j), \quad j = 0, \ldots, J. \]

\( Y_0 \) 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 $D_j = 0$. 
• Appendix B of Heckman and Vytlacil (2007a) gives conditions under which the distributions of the $Y_j$ and the subjective evaluations $R_j$, $j = 0, \ldots, J$, that generate choices $D_j$ are identified.

• Using the tools for multiple outcome models developed in Heckman and Vytlacil (2007b), we can use IV and our extensions of IV to identify the treatment parameters discussed there.
In this section, we consider what randomizations at various stages identify.

We assume that the randomizations do not disturb the program.

Thus we invoke Assumption (PI-3).

Recall that we also assume absence of general equilibrium effects (PI-4).

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 $\xi_j$ are ordered in a natural way: $\xi_j = 1$ only if $\xi_\ell = 1$, $\ell = 0, \ldots, j - 1$.

Array the $\xi_j$ into a vector $\xi$ 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

\[
Y = \sum_{j=0}^{J} D_j \left( \sum_{k=0}^{j} \left( \prod_{\ell=0}^{k-1} \xi_\ell \right) (1 - \xi_k) Y_k \right).
\]
To understand this formula, consider a program with three stages \((J = 3)\) after the initial participation stage.

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 express $A_k$ as

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

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 $\xi$ are satisfied under the following sequential randomization assumptions.

- They parallel the sequential randomization conditions developed in the dynamic models analyzed in Abbring and Heckman (2007):

\begin{align}
&\xi_i \perp \left( \tilde{Y}, \{A(z,e)\}_{(z,e)\in Z \times \tilde{\xi}} \mid X, Z, D_\ell = 1 \text{ for } \ell < i, \prod_{\ell=0}^{i-1} \xi_\ell = 1 \right), \\
&\text{for } i = 1, \ldots, J,
\end{align}

and

\begin{align}
&\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) \text{ depends on } \\
&\xi_i, \text{ for } i = 1, \ldots, J.
\end{align}
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 $\xi_0$, 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 (PI-3) and (PI-4) hold.
For those declared eligible,

\[ E(Y | \xi_0 = 1) = \sum_{j=0}^{J} E(Y_j | D_j = 1) \Pr(D_j = 1). \quad (6.1) \]

For those declared ineligible,

\[ E(Y | \xi_0 = 0) = \sum_{j=0}^{J} E(Y_0 | D_j = 1) \Pr(D_j = 1), \quad (6.2) \]

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 (6.1).
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 (6.2) for \( J > 1 \).

When \( J = 1 \), we can identify all of the components of (6.2).

The individual components of (6.2) 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 $\xi_0$ alone, setting $\xi_1 = \cdots = \xi_J = 1$ for anyone for whom $\xi_0 = 1$.

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

(6.3)
For $J > 1$, this simple experimental estimator does not identify the effect of full participation in the program for those who participate ($E (Y_J - Y_0 | D_J = 1)$) unless additional assumptions are invoked, such as the assumption that partial participation has the same mean effect as full participation for persons who drop out at the early stages, i.e.,

$$E (Y_j - Y_0 | D_j = 1) = E (Y_J - Y_0 | D_j = 1)$$

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 \( E(Y_j - Y_0 \mid D_j = 1) = 0 \) for all \( j < J \).

Under those conditions, if we divide the mean difference by \( \Pr(D_J = 1) \), we obtain the “Bloom” estimator (Mallar, Kerachsky, and Thorton, 1980; Bloom, 1984)

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

assuming \( \Pr(D_J = 1) \neq 0 \).

This is an IV estimator using \( \xi_0 \) as the instrument for \( A_J \).

In general, the mean difference between the treated and the controlled identifies only the composite term shown in (6.3).

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

$$E \left[ Y \, \middle| \, \prod_{\ell=0}^{J} \xi_{\ell} = 1 \right] - E \left[ Y \, \middle| \left( \prod_{\ell=0}^{k-1} \xi_{\ell} \right) (1 - \xi_k) = 1 \right]$$

$$= \sum_{j=k}^{J} E \left( Y_j - Y_k \, \middle| \, D_j = 1 \right) \Pr (D_j = 1) .$$
Hence, since we know \( E(Y_k \mid D_k = 1) \) and \( \Pr(D_k = 1) \) from observational data, we can identify the combination of parameters

\[
\sum_{j=k+1}^{J} E(Y_k \mid D_j = 1) \Pr(D_j = 1),
\]

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 $J$

\[ ([\prod_{\ell=0}^{J-2} \xi_{\ell}] (1 - \xi_{J-1}) = 1) \text{ identifies } E(Y_J - Y_{J-1} \mid D_J = 1): \]

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

Thus,

\[
E (Y \mid \xi_0 = 1, \ldots, \xi_J = 1) - E (Y \mid \xi_0 = 1, \ldots, \xi_{J-1} = 1, \xi_J = 0) \\
= E (Y_J - Y_{J-1} \mid D_J = 1) \Pr(D_J = 1).
\]

Since $\Pr(D_J = 1)$ is observed from choice data, as is $E(Y_J \mid D_J = 1)$, we can identify $E(Y_{J-1} \mid D_J = 1)$ from the experiment.
In the general case under Assumptions (PI-3) and (PI-4), a randomization that prevents agents from moving beyond stage $\ell$ ($\xi_0 = 1, \ldots, \xi_{\ell-1} = 1, \xi_\ell = 0$) directly identifies

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

$$= \sum_{j=0}^{\ell} E(Y_j \mid D_j = 1) \Pr(D_j = 1)$$

all components known from observational data

$$+ \sum_{j=\ell+1}^{J} E(Y_\ell \mid D_j = 1) \Pr(D_j = 1).$$

sum and probability weights known, but not individual $E(Y_\ell \mid D_j = 1)$. 
All of the components of the first set of terms on the right-hand side are known from observational data.

The probabilities 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 $\ell$ 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 $\xi_\ell$ as an instrument for $A_\ell$. 
Keeping the conditioning on $X, Z$ implicit, we obtain

$$\text{IV}_{\xi_\ell} = \frac{E [Y \mid \xi_\ell = 0] - E [Y \mid \xi_\ell = 1]}{\Pr (A_\ell = 1 \mid \xi_\ell = 0) - \Pr (A_\ell = 1 \mid \xi_\ell = 1)}$$

$$= \sum_{j=\ell+1} J E [Y_\ell - Y_j \mid D_j = 1] \Pr (D_j = 1) \sum_{j=\ell+1} J \Pr (D_j = 1),$$

$\ell = 0, \ldots, J - 1$.

By the preceding analysis, we know the numerator term but not the individual components.

We know the denominator from choices measured in observational data and invariance Assumption (PI-3).

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 $E(Y_j | D_\ell = 1), j = 0, \ldots, J$ and $\ell = 0, \ldots, J$.

The conditional mean outcomes obtained from observational data are on the diagonal of the table $(E(Y_j | D_j = 1), j = 0, \ldots, J)$.

Because of choices of agents, experiments do not directly identify the elements in the table that are above the diagonal.

Under Assumptions (PI-3) and (PI-4), experiments described at the base of the table identify the combinations of the parameters below the diagonal.
Table 1: Params. and Combinations of Params. That Can be Identified by Different Randomizations

<table>
<thead>
<tr>
<th>Choice Prob. (known)</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pr($D_0 = 1$)</td>
<td>$D_0$</td>
</tr>
<tr>
<td>Pr($D_1 = 1$)</td>
<td>$D_1$</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Pr($D_j = 1$)</td>
<td>$D_j$</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Pr($D_{J-1} = 1$)</td>
<td>$D_{J-1}$</td>
</tr>
<tr>
<td>Pr($D_J = 1$)</td>
<td>$D_J$</td>
</tr>
</tbody>
</table>

Randomization:

$\xi_0 = 0$ \quad \ldots \quad \xi_j = 0$ \quad \ldots \quad \xi_J = 0

New Identified Combinations of Parameters:

$\sum_{\ell=1}^{J} \{ E(Y_0 | D_\ell = 1) \} \times Pr(D_\ell = 1)$ \quad \ldots \quad $\sum_{\ell=j+1}^{J} \{ E(Y_j | D_\ell = 1) \} \times Pr(D_\ell = 1)$
Recall that if $\xi_{\ell} = 0$, the agent cannot advance beyond stage $\ell$.

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.

Following Balke and Pearl (1997), Manski (1989, 1990, 1996, 2003), and Robins (1989), we can use the identified combinations from different randomizations to bound the admissible values of counterfactuals below the diagonal of Table 1.
Heckman, Smith, and Taber (1998) present a test for a strengthened version of the identifying assumptions made by Bloom.

They perform a sensitivity analysis to analyze departures from the assumption that dropouts have the same outcomes as nonparticipants.

Hotz, Mullin, and Sanders (1997) apply the Manski bounds in carefully executed empirical examples and show the difficulties involved in using the Bloom estimator in experiments with multiple outcomes.

We next turn to some evidence on the importance of randomization bias.
Evidence on Randomization Bias

- Manifestations of a more general problem termed “Hawthorne effects” that arise from observing any population (see Campbell and Stanley, 1963; Cook and Campbell, 1979).

- How important is this theoretical possibility in practice?

- Surprisingly, very little is known about the answer to this question for the social experiments conducted in economics.

- This is so because randomized social experimentation has usually only been implemented on “pilot projects” or “demonstration projects” designed to evaluate new programs never previously estimated.
Disruption by randomization cannot be confirmed or denied using data from these experiments.

In one ongoing program evaluated by randomization by the Manpower Demonstration Research Corporation (MDRC), participation was compulsory for the target population (Doolittle and Traeger, 1990).

Hence randomization did not affect applicant pools or assessments of applicant eligibility by program administrators.
There is some information on the importance of randomization, although it is indirect.

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.

The reasons for refusal to participate are given in Tables 2A – 2C.
<table>
<thead>
<tr>
<th>Concern</th>
<th>Percentage of Training Centers Citing the Concern</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Ethical and public relations implications of:</td>
<td></td>
</tr>
<tr>
<td>a. Random assignment in social programs</td>
<td>61.8</td>
</tr>
<tr>
<td>b. Denial of services to controls</td>
<td>54.4</td>
</tr>
<tr>
<td>2. Potential negative effect of creation of a control group on achievement of client recruitment goals</td>
<td>47.8</td>
</tr>
<tr>
<td>3. Potential negative impact on performance standards</td>
<td>25.4</td>
</tr>
</tbody>
</table>
**Table 2B:** Percentage of Local JTPA Agencies Citing Specific Concerns About Participating in the Experiment

<table>
<thead>
<tr>
<th>Concern</th>
<th>Percentage of Training Centers Citing the Concern</th>
</tr>
</thead>
<tbody>
<tr>
<td>4. Implementation of the study when service providers do intake</td>
<td>21.1</td>
</tr>
<tr>
<td>5. Objections of service providers to the study</td>
<td>17.5</td>
</tr>
<tr>
<td>6. Potential staff administrative burden</td>
<td>16.2</td>
</tr>
<tr>
<td>7. Possible lack of support by elected officials</td>
<td>15.8</td>
</tr>
</tbody>
</table>
Table 2C: Percentage of Local JTPA Agencies Citing Specific Concerns About Participating in the Experiment

<table>
<thead>
<tr>
<th>Concern</th>
<th>Percentage of Training Centers Citing the Concern</th>
</tr>
</thead>
<tbody>
<tr>
<td>8. Legality of random assignment and possible grievances</td>
<td>14.5</td>
</tr>
<tr>
<td>9. Procedures for providing controls with referrals to other services</td>
<td>14.0</td>
</tr>
<tr>
<td>10. Special recruitment problems for out-of-school youth</td>
<td>10.5</td>
</tr>
<tr>
<td>Sample size</td>
<td>228</td>
</tr>
</tbody>
</table>
Notes: Concerns noted by fewer than 5 percent of the training centers are not listed. Percentages add up to more than 100.0 because training centers could raise more than one concern.

Source: Based on responses of 228 local JTPA agencies contacted about possible participation in the National JTPA Study.

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:

Doolittle and Traeger (1990, p. 121)

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.
These authors go on to note that

Doolittle and Traeger (1990, p. 123)

...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.
This indirect evidence is hardly decisive even about the JTPA experiment, much less all experiments.

Training centers may offer these arguments only as a means of avoiding administrative scrutiny, and there may be no “real” effect of randomization.

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.

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

They discuss one study of drugs administered to children afflicted with a disease.

The study had two components.

The nonexperimental phase of the study had a 4 percent refusal rate, while 34 percent of a subsample of the same parents refused to participate in a randomized subtrial, although the treatments were equally nonthreatening.
These authors cite further evidence suggesting that refusal to participate in randomization schemes is selective.

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.

We next consider some evidence on the importance of dropping out and noncompliance with experimental protocols.
Evidence on Dropping Out and Substitution Bias

- Dropouts are a feature of all social programs.
- Randomization may raise dropout rates, but the evidence for such effects is weak.
- In addition, 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.
Tables 3A and 3B (reproduced from Heckman, Hohmann, Smith, and Khoo, 2000) demonstrate the practical importance of both dropout and substitution bias in experimental evaluations.

It reports the rates of treatment group dropout and control group substitution from a variety of social experiments.

It reveals that the fraction of treatment group members receiving program services is often less than 0.7, and sometimes less than 0.5.

Furthermore, the observed characteristics of the treatment group members who drop out often differ from those who remain and receive the program services.
### Table 3A: Fraction of Experimental Treatment and Control Groups Receiving Services in Experimental Evaluations of Employment and Training Programs

<table>
<thead>
<tr>
<th>Study</th>
<th>Authors/time period</th>
<th>Target group(s)</th>
<th>Fraction of treatments receiving services</th>
<th>Fraction of controls receiving services</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. NSW</td>
<td>Hollister, et al. (1984) (9 months after RA)</td>
<td>Long-term AFDC women</td>
<td>0.95</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ex-addicts</td>
<td>NA</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td></td>
<td>17-20 year old high school dropouts</td>
<td>NA</td>
<td>0.04</td>
</tr>
<tr>
<td>2. SWIM</td>
<td>Friedlander and Hamilton (1993) (Time period not reported)</td>
<td>AFDC women: applicants and recipients</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>a. Job search assistance</td>
<td>0.54</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>b. Work experience</td>
<td>0.21</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>c. Classroom training/OJT</td>
<td>0.39</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td></td>
<td>d. Any activity</td>
<td>0.69</td>
<td>0.30</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AFDC-U unemployed fathers</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>a. Job search assistance</td>
<td>0.60</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>b. Work experience</td>
<td>0.21</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>c. Classroom training/OJT</td>
<td>0.34</td>
<td>0.22</td>
</tr>
<tr>
<td></td>
<td></td>
<td>d. Any activity</td>
<td>0.70</td>
<td>0.23</td>
</tr>
<tr>
<td>3. JOBSTART</td>
<td>Cave, et al. (1993) (12 months after RA)</td>
<td>Youth high school dropouts</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Classroom training/OJT</td>
<td>0.90</td>
<td>0.26</td>
</tr>
</tbody>
</table>
Table 3B: Fraction of Experimental Treatment and Control Groups Receiving Services in Experimental Evaluations of Employment and Training Programs

<table>
<thead>
<tr>
<th>Study</th>
<th>Authors/time period</th>
<th>Target group(s)</th>
<th>Fraction of treatments receiving services</th>
<th>Fraction of controls receiving services</th>
</tr>
</thead>
<tbody>
<tr>
<td>4. Project Independence</td>
<td>Kemple, et al. (1995)</td>
<td>AFDC women: applicants and recipients</td>
<td>0.43</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td>(24 months after RA)</td>
<td>a. Job search assistance</td>
<td>0.43</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>b. Classroom training/OJT</td>
<td>0.42</td>
<td>0.31</td>
</tr>
<tr>
<td></td>
<td></td>
<td>c. Any activity</td>
<td>0.64</td>
<td>0.40</td>
</tr>
<tr>
<td>5. New Chance</td>
<td>Quint, et al. (1994)</td>
<td>Teenage single mothers</td>
<td>0.82</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td>(18 months after RA)</td>
<td>Any education services</td>
<td>0.26</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Any training services</td>
<td>0.87</td>
<td>0.55</td>
</tr>
<tr>
<td>6. National JTPA Study</td>
<td>Heckman and Smith (1998)</td>
<td>Self-reported from survey data</td>
<td>0.38</td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td>(18 months after RA)</td>
<td>Adult males</td>
<td>0.51</td>
<td>0.33</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Adult females</td>
<td>0.50</td>
<td>0.32</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Male youth</td>
<td>0.81</td>
<td>0.42</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Female youth</td>
<td>0.81</td>
<td>0.42</td>
</tr>
<tr>
<td>Combined Administrative Survey Data</td>
<td></td>
<td></td>
<td>0.74</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Adult males</td>
<td>0.78</td>
<td>0.34</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Adult females</td>
<td>0.81</td>
<td>0.34</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Male youth</td>
<td>0.81</td>
<td>0.34</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Female youth</td>
<td>0.81</td>
<td>0.42</td>
</tr>
</tbody>
</table>
Notes: RA = random assignment. H.S. = high school. AFDC = Aid to Families with Dependent Children. OJI = On the Job Training.

Service receipt includes any employment and training services. The services received by the controls in the NSW study are CETA and WIN jobs. For the Long Term AFDC Women, this measure also includes regular public sector employment during the period.

Sources for data: Maynard and Brown (1980), p. 169, Table A14; Masters and Maynard (1981), p. 148, Table A.15; Friedlander and Hamilton (1993), p. 22, Table 3.1; Cave, et al. (1993), p. 95, Table 4.1; Quint, et al. (1994), p. 110, Table 4.9; and Kemple, et al. (1995), p. 58, Table 3.5; Heckman and Smith (1998) and calculations by the authors.

Source: Heckman, LaLonde and Smith (1999) and Heckman et al. (2000).
• With regard to substitution bias, Tables 3A and 3B show that as many as 40% of the controls in some experiments received substitute services elsewhere.

• In a simple one treatment experiment with full compliance ($\xi = 1 \Rightarrow A = 1$ and $\xi = 0 \Rightarrow A = 0$), all individuals assigned to the treatment group receive the treatment and there is no control group substitution, so that the difference between the fraction of treatments and controls that receive the treatment equals 1.0.

• In practice, this difference is often well below 1.0.

• Randomization reduced and delayed receipt of training in the experimental control group but by no means eliminated it.

• Many of the treatment group members received no treatment.
The extent of both substitution and dropout depends on the characteristics of the treatment being evaluated and the local program environment.

In the NSW study, where the treatment was relatively unique and of high enough quality to be clearly perceived as valuable by participants, dropout and substitution rates were low enough to approximate the ideal case.

In contrast, for the NJS and for other programs that provide low cost services widely available from other sources, substitution and dropout rates are high.

In the NJS, the substitution problem is accentuated by the fact that the program relied on outside vendors to provide most of its training.
Many of these vendors, such as community colleges, provided the same training to the general public, often with subsidies from other government programs such as Pell Grants.

In addition, in order to help in recruiting sites to participate in the NJS, evaluators allowed them to provide control group members with a list of alternative training providers in the community.

Of the 16 sites in the NJS, 14 took advantage of this opportunity to alert control group members to substitute training opportunities.
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.

- Heckman, LaLonde, and Smith (1999) discuss a variety of other problems that often plague social experiments.