Are Job-Training Programs Effective?

Donald B. Rubin

Harvard University

Presentation based on joint work with Fabrizia Mealli, Paolo Frumento, and Barbara Pacini.

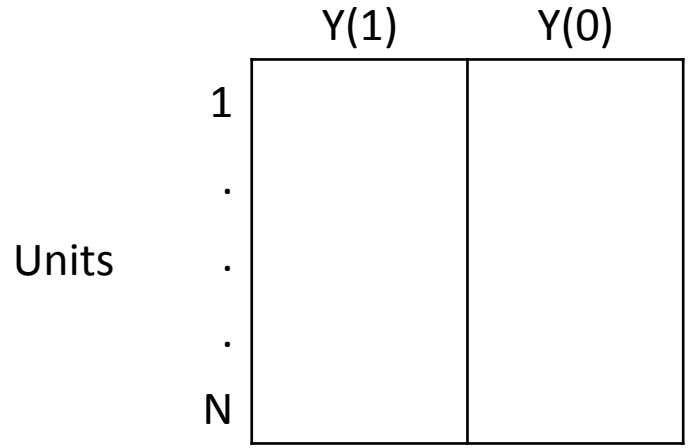
The National Job Corps Study

- A randomized study to evaluate the effects of a training program on employment and wages
  - Randomization assures fair comparison, in expectation, between treatment groups
  - Sampled youths (n=15,386) were assigned randomly to a job training program group or a control group
  - Only those assigned to the job training program group were able to enroll in Job Corps
- Post-treatment complications
  - Noncompliance (only 73% attended the offered training)
  - Truncation of wages for the unemployed
  - Missing outcomes due to nonresponse

Potential Outcomes Approach to Causal Inference – Simplest Setting

T = 1 $\Rightarrow$ active treatment (e.g., job training)

T = 0 $\Rightarrow$ control treatment (e.g., no training)



Y(1) = outcomes if exposed to active treatment

Y(0) = outcomes if exposed to control treatment

Ave[$Y_i(1) - Y_i(0)$] = Average causal effect of active versus control treatment

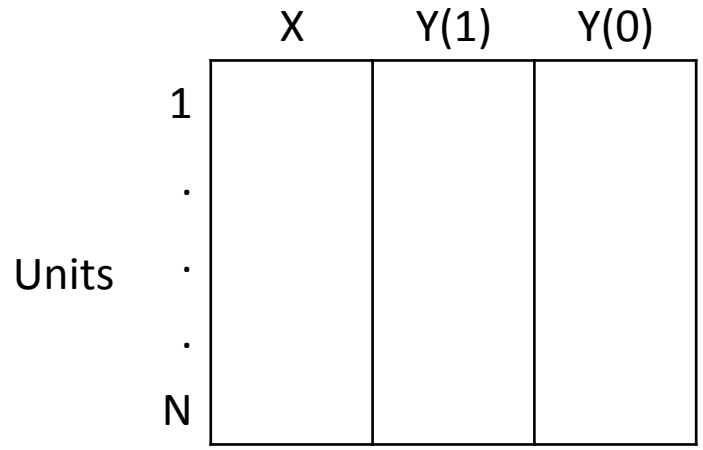
Potential Outcomes Approach to Causal Inference – Simplest Setting

- Fundamental problem of causal inference
  - For each *i*, only $Y_i(1)$ or $Y_i(0)$ can be observed

Units	Y(1)	Y(0)	T
1	✓	?	1
.	✓	?	1
.	✓	?	1
.	?	✓	0
.	?	✓	0
N	?	✓	0

- Random assignment of active versus control $\Rightarrow$ representative sample of $Y_i(1)$ will be compared to representative sample of $Y_i(0)$

Potential Outcomes Approach to Causal Inference – Simplest Setting with Covariates



- Same as before, except includes pretreatment covariates, e.g., age, sex, background education

$$\underset{i:\text{Female}}{\text{Ave}}\,[Y_i(1) - Y_i(0)] = \text{Average causal effect of active vs control treatment for females}$$

- Randomization still works for females

Potential Outcomes Approach to Causal Inference – Simple Noncompliance with Active Treatment

		D(1)	D(0)	Y(1)	Y(0)	
Units	1	1	0			compliers
	.	1	0			
	.	1	0			
	.	0	0			noncompliers
	N	0	0			

D(1) = treatment taken when assigned active treatment
D(1) = 1 ⇒ active taken, D(1) = 0 ⇒ control taken

D(0) = treatment taken when assigned control treatment
simple setting, always control ⇒ D(0) = 0

$$\underset{i:D_i(1)=1}{\text{Ave}}\,[Y_i(1) - Y_i(0)] = \text{Average causal effect for true compliers}$$

- Randomization still works for compliers

Potential Outcomes Approach to Causal Inference – Simple Noncompliance with Active Treatment: Observed Data

Units	D(1)	D(0)	Y(1)	Y(0)	T	
1	1	0	✓	?	1	complier status observed
.	1	0	✓	?	1	
.	?	0	?	✓	0	complier status missing
.	?	0	?	✓	0	
.	0	0	✓	?	1	noncomplier status observed
N	?	0	?	✓	0	noncomplier status missing

- **Compliers**
  - For individuals assigned treatment (T=1), D(1)=1 & D(0)=0
  - For individuals assigned control (T=0), D(1)=? because true compliance under treatment is unknown & D(0)=0
- **Noncompliers**
  - For individuals assigned treatment (T=1), D(1)=0 & D(0)=0
  - For individuals assigned control (T=0), D(1)=? because true compliance under treatment is unknown & D(0)=0
- **Randomization still works for compliers**

Key Idea: Principal Stratification
(Frangakis and Rubin, 2002)

- Stratify on values of post-treatment intermediate outcome
- Convert $D_i(1)$, $D_i(0)$ into stratification variable
  - True complier "c" if $D_i(1)=1$
  - Noncomplier "n" if $D_i(1)=0$
- Idea works more generally

Intermediate Outcome - Employment

- Employed (yes, no) at a given time post-treatment is an important outcome, but is also needed to define principal strata for "final" outcomes, Y, describing attributes of possible employment, such as wages, retirement plan benefits, etc., which are not well-defined if unemployed
- Principal strata are defined by employment status
  - EE = employed whether assigned to training or not
  - EU = employed if trained, unemployed if not trained
  - UE = unemployed if trained, employed if not trained
  - UU = unemployed whether assigned to training or not
- Causal effects of training on Y only well-defined for EE
- UE empty? Reservation wage issue

Causal Effects of Training within Principal Strata

- Principal strata are defined by compliance with assignment to job training and by employment status
  - c&EE, c&EU, c&UE, c&UU
  - n&EE, n&EU, n&UE, n&UU
- By assumption (exclusion restriction on employment), we rule out n&EU and n&UE
  - If assignment does not affect entry into training, assignment cannot affect employment status
  - Also assume exclusion for attributes of employment, Y
- Causal effects of T on Y are only well-defined for c&EE and n&EE principal strata (no effect on Y in n&EE by exclusion restriction)

Not Done Yet with Needed Principal Strata

- Indicators for response to survey items asking about employment status and wages, etc.
  - R(1) and R(0), each indicating respondent or not
- Do not make exclusion restriction here
  - e.g., males could have $R_i(1) \Longrightarrow$ respond if assigned training, but $R_i(0) \Longrightarrow$ not respond if assigned control
- But do assume missing at random (MAR)
  - A nuisance, not of scientific interest

Causal Effects

- Assignment to be trained on being job-trained
  Pr(c) = proportion compliers
- Assignment to be trained on being employed
  Pr(c&EU) – Pr(c&UE)
- Assignment to be trained on being employed for compliers
  [Pr(c&EU) – Pr(c&UE)]/Pr(c)
- Relative sizes of principal strata
  – c&EE, c&EU, c&UE, c&UU, n&EE, n&UU
- Distributions of X within principal strata

Causal Effects on Wages

- For the always employed

  Ave[$Y_i(1) - Y_i(0)|$c&EE or n&EE]

- For the always employed compliers

  Ave[$Y_i(1) - Y_i(0)|$c&EE]

- By exclusion, for the always employed noncompliers
  - Ave[$Y_i(1) - Y_i(0)|$n&EE] = 0

Method of Analysis

- Direct likelihood at each of three post-treatment points in time
- Search for parsimonious model to help guide policy
- Needs scientific judgement

Estimated Means of Covariates within Principal Strata Week 52

Principal Stratum	c&EE	c&EU	c&UE	c&UU	n&EE	n&UU
Percent in Stratum	24	4	5	42	12	13
Female	0.4	0.3	0.3	0.4	0.4	0.5
Age at baseline	19.0	19.1	19.3	18.4	19.5	18.9
White	0.3	0.4	0.3	0.2	0.3	0.2
With a Partner	0.1	0.0	0.0	0.1	0.1	0.1
Has children	0.2	0.1	0.1	0.2	0.2	0.3
Education	0.4	0.4	0.5	0.3	0.6	0.4
Ever arrested	0.2	0.3	0.3	0.3	0.3	0.3
Mother's education	0.2	0.2	0.3	0.2	0.2	0.2
Father's education	0.3	0.3	0.3	0.3	0.3	0.3
Household income > $6000	0.6	0.6	0.5	0.5	0.6	0.5
Person income > $6000	0.1	0.1	0.0	0.1	0.2	0.1
Have job	0.3	0.3	0.3	0.2	0.3	0.2
Had job, previous year	0.8	0.8	0.7	0.6	0.8	0.6
Months in Job, previous year	5.0	5.1	4.9	2.9	5.6	3.1
Earnings, previous year	0.2	0.3	0.1	-0.2	0.4	-0.1

Estimated Means of Covariates within Principal Strata Week 130

Principal Stratum	c&EE	c&EU	c&UE	c&UU	n&EE	n&UU
Percent in Stratum	31	7	5	32	13	13
Female	0.4	0.3	0.2	0.5	0.4	0.5
Age at baseline	18.9	18.9	18.9	18.4	19.5	18.9
White	0.3	0.4	0.4	0.2	0.3	0.2
With a Partner	0.1	0.1	0.1	0.0	0.1	0.1
Has children	0.2	0.2	0.2	0.2	0.2	0.3
Education	0.4	0.4	0.4	0.3	0.5	0.4
Ever arrested	0.2	0.3	0.3	0.3	0.3	0.3
Mother's education	0.2	0.2	0.3	0.2	0.2	0.2
Father's education	0.3	0.3	0.2	0.3	0.3	0.3
Household income > $6000	0.6	0.6	0.6	0.5	0.6	0.5
Person income > $6000	0.1	0.1	0.1	0.1	0.1	0.1
Have job	0.3	0.3	0.3	0.2	0.3	0.2
Had job, previous year	0.7	0.8	0.7	0.6	0.8	0.6
Months in Job, previous year	4.3	5.2	4.3	2.9	5.0	3.5
Earnings, previous year	0.1	0.3	0.2	-0.2	0.3	-0.0

Estimated Means of Covariates within Principal Strata Week 208

Principal Stratum	c&EE	c&EU	c&UE	c&UU	n&EE	n&UU
Percent in Stratum	39	6	5	26	15	10
Female	0.4	0.3	0.2	0.5	0.4	0.5
Age at baseline	18.9	18.7	18.6	18.4	19.4	18.9
White	0.3	0.4	0.4	0.2	0.3	0.2
With a Partner	0.1	0.0	0.1	0.0	0.1	0.1
Has children	0.2	0.1	0.2	0.2	0.2	0.2
Education	0.4	0.4	0.4	0.3	0.5	0.4
Ever arrested	0.2	0.3	0.3	0.3	0.3	0.3
Mother's education	0.2	0.3	0.2	0.2	0.2	0.2
Father's education	0.3	0.3	0.2	0.3	0.3	0.3
Household income > $6000	0.6	0.7	0.7	0.5	0.6	0.5
Person income > $6000	0.1	0.1	0.1	0.0	0.1	0.1
Have job	0.3	0.3	0.2	0.1	0.3	0.2
Had job, previous year	0.7	0.7	0.7	0.5	0.7	0.6
Months in Job, previous year	4.5	4.4	4.1	2.6	4.9	3.5
Earnings, previous year	0.1	0.2	0.1	-0.3	0.2	-0.0

Percent within Principal Strata by Time Period

Principal Stratum	c&EE	c&EU	c&UE	c&UU	n&EE	n&UU
Week 52	24	4	5	42	12	13
Week 130	31	7	5	32	13	13
Week 208	39	6	5	26	15	10

- For compliers, % EE increases in time, and % UU decreases
- For noncompliers, EE remains fairly stable
- Causal effect of training slightly increases in time, i.e., the difference between proportions in c&EU and c&UE appears to increase in time
  - Economists' "lock-in effect" during the period of training

Estimated Average Hourly Wages for Those Employed in Dollars within Principal Strata by Time Period

Principal Stratum	c&EE(1)	c&EE(0)	c&EU(1)	c&UE(0)	n&EE
Week 52	5.8	5.6	7.2	6.6	6.6
Week 130	6.7	6.5	9.3	7.5	8.0
Week 208	7.7	7.5	9.3	9.6	9.1

- Estimated causal effect on wages for always employed compliers is approximately 0.2 for all time periods
- Always employed compliers, whether trained or not, have the lower hourly wages than the sometimes employed (c&EU or c&UE) or n&EE
- Wages tend to increase in time

Final Conclusions for This Job Training Program

- In long run, for compliers, minor positive effect on employment status
- For always employed compliers, minor positive effect on wages at all time periods
- Background characteristics of individuals differ across principal strata
- Suggests need for more targeted programs
- Even if evaluation is based on randomized experiment, difficult to analyze correctly