Quantitative Methods in Economics
Causality and treatment effects

Maximilian Kasy

Harvard University, fall 2016

1) Causality, Potential Outcomes, and the Estimation of Treatment Effects in Randomized Studies

(cf. “Mostly Harmless Econometrics,” chapter 2)

Purpose, Scope, and Examples

The goal of **program evaluation** is to assess the causal effect of public policy interventions. Examples include effects of:

- Job training programs on earnings and employment
- Class size on test scores
- Minimum wage on employment
- Military service on earnings and employment
- Tax-deferred saving programs on savings accumulation

In addition, we may be interested in the effect of variables that do not represent public policy interventions. Examples:

- Interest rate on credit card usage
- Incentive scheme on employer productivity
- Immigration on wages

Causality with Potential Outcomes

Treatment

D_i: Indicator of treatment intake for *unit i*

$$D_i = \begin{cases} 1 & \text{if unit } i \text{ received the treatment} \\ 0 & \text{otherwise.} \end{cases}$$

Causality with Potential Outcomes

Treatment

D_i: Indicator of treatment intake for *unit i*

$$D_i = \begin{cases} 1 & \text{if unit } i \text{ received the treatment} \\ 0 & \text{otherwise.} \end{cases}$$

Outcome

Y_i: Observed outcome variable of interest for unit i

Causality with Potential Outcomes

Treatment

D_i: Indicator of treatment intake for *unit i*

$$D_i = \begin{cases} 1 & \text{if unit } i \text{ received the treatment} \\ 0 & \text{otherwise.} \end{cases}$$

Outcome

Y_i: Observed outcome variable of interest for unit i

Potential Outcomes

Y_{0i} and Y_{1i}: Potential outcomes for unit i

$$\begin{aligned} Y_{1i} &: \quad \text{Potential outcome for unit } i \text{ with treatment} \\ Y_{0i} &: \quad \text{Potential outcome for unit } i \text{ without treatment} \end{aligned}$$

Causality with Potential Outcomes

Treatment Effect

The treatment effect or causal effect of the treatment on the outcome for unit i is the difference between its two potential outcomes:

$$Y_{1i} - Y_{0i}$$

Causality with Potential Outcomes

Treatment Effect

The treatment effect or causal effect of the treatment on the outcome for unit i is the difference between its two potential outcomes:

$$Y_{1i} - Y_{0i}$$

Observed Outcomes

Observed outcomes are realized as

$$Y_i = Y_{1i} D_i + Y_{0i}(1 - D_i) \text{ or } Y_i = \begin{cases} Y_{1i} & \text{if } D_i = 1 \\ Y_{0i} & \text{if } D_i = 0 \end{cases}$$

Causality with Potential Outcomes

Treatment Effect

The treatment effect or causal effect of the treatment on the outcome for unit i is the difference between its two potential outcomes:

$$Y_{1i} - Y_{0i}$$

Observed Outcomes

Observed outcomes are realized as

$$Y_i = Y_{1i} D_i + Y_{0i}(1 - D_i) \text{ or } Y_i = \begin{cases} Y_{1i} & \text{if } D_i = 1 \\ Y_{0i} & \text{if } D_i = 0 \end{cases}$$

Fundamental Problem of Causal Inference

Cannot observe both potential outcomes (Y_{1i}, Y_{0i})

Identification Problem for Causal Inference

Problem

Causal inference is difficult because it involves missing data. How can we find $Y_{1i} - Y_{0i}$*?*

- A large amount of homogeneity would solve this problem:
  - (Y_{1i}, Y_{0i}) constant across individuals
  - (Y_{1i}, Y_{0i}) constant across time
- However, often there is a large degree of heterogeneity in the individual responses to participation in public programs or to exposure to other treatment of interest

Stable Unit Treatment Value Assumption (SUTVA)

Assumption

Observed outcomes are realized as

$$Y_i = Y_{1i} D_i + Y_{0i}(1 - D_i)$$

- Implies that potential outcomes for unit i are unaffected by the treatment of unit j
- Rules out interference across units
- Examples:
  - Effect of fertilizer on plot yield
  - Effect of flu vaccine on hospitalization
- This assumption may be problematic, so we should choose the units of analysis to minimize interference across units.

Quantities of Interest (Estimands)

ATE

Average treatment effect is:

$$\alpha_{ATE} = E[Y_1 - Y_0]$$

ATET

Average treatment effect on the treated is:

$$\alpha_{ATET} = E[Y_1 - Y_0 | D = 1]$$

Average Treatment Effect (ATE)

Imagine a population with 4 units:

i	Y_{1i}	Y_{0i}	Y_i	D_i	$Y_{1i} - Y_{0i}$
1	3	?	3	1	?
2	1	?	1	1	?
3	?	0	0	0	?
4	?	1	1	0	?

What is $\alpha_{ATE} = E[Y_1 - Y_0]$?

Average Treatment Effect (ATE)

Imagine a population with 4 units:

i	Y_{1i}	Y_{0i}	Y_i	D_i	$Y_{1i} - Y_{0i}$
1	3	0	3	1	3
2	1	1	1	1	0
3	1	0	0	0	1
4	1	1	1	0	0

What is $\alpha_{ATE} = E[Y_1 - Y_0]$?

Average Treatment Effect (ATE)

Imagine a population with 4 units:

i	Y_{1i}	Y_{0i}	Y_i	D_i	$Y_{1i} - Y_{0i}$
1	3	0	3	1	3
2	1	1	1	1	0
3	1	0	0	0	1
4	1	1	1	0	0
$E[Y_1]$	1.5				
$E[Y_0]$		0.5			
$E[Y_1 - Y_0]$					1

$$\alpha_{ATE} = E[Y_1 - Y_0] = 3 \cdot (1/4) + 0 \cdot (1/4) + 1 \cdot (1/4) + 0 \cdot (1/4) = 1$$

Average Treatment Effect on the Treated (ATET)

Imagine a population with 4 units:

i	Y_{1i}	Y_{0i}	Y_i	D_i	$Y_{1i} - Y_{0i}$
1	3	0	3	1	3
2	1	1	1	1	0
3	1	0	0	0	1
4	1	1	1	0	0

What is $\alpha_{ATET} = E[Y_1 - Y_0 | D = 1]$?

Average Treatment Effect on the Treated (ATET)

Imagine a population with 4 units:

i	Y_{1i}	Y_{0i}	Y_i	D_i	$Y_{1i} - Y_{0i}$
1	3	0	3	1	3
2	1	1	1	1	0
3	1	0	0	0	1
4	1	1	1	0	0
$E[Y_1 \mid D=1]$	2				
$E[Y_0 \mid D=1]$		0.5			
$E[Y_1 - Y_0 \mid D=1]$					1.5

$$\alpha_{ATET} = E[Y_1 - Y_0 \mid D=1] = 3 \cdot (1/2) + 0 \cdot (1/2) = 1.5$$

Selection Bias

Problem

Comparisons of earnings for the treated and the untreated do not usually give the right answer:

$$\begin{aligned} E[Y|D=1]-E[Y|D=0] &= E[Y_1|D=1]-E[Y_0|D=0] \\ &= \underbrace{E[Y_1-Y_0|D=1]}_{ATET}+\underbrace{\{E[Y_0|D=1]-E[Y_0|D=0]\}}_{BIAS} \end{aligned}$$

- Bias term is not likely to be zero in most applications
- Selection into treatment often depends on potential outcomes

Selection Bias

Problem

Comparisons of earnings for the treated and the untreated do not usually give the right answer:

$$E[Y|D=1]-E[Y|D=0]=E[Y_1|D=1]-E[Y_0|D=0]$$
$$=\underbrace{E[Y_1-Y_0|D=1]}_{ATET}+\underbrace{\{E[Y_0|D=1]-E[Y_0|D=0]\}}_{BIAS}$$

Example: Job training program for disadvantaged

- participants are self-selected from a subpopulation of individuals in difficult labor situations
- post-training period earnings would be lower for participants than for nonparticipants in the absence of the program ($E[Y_0|D=1]-E[Y_0|D=0]<0$)

Training Program for the Disadvantaged in the U.S.

Data from the National Supported Work Demonstration (NSW)

TABLE 1.—MEAN EARNINGS PRIOR, DURING, AND SUBSEQUENT TO TRAINING FOR 1964 MDTA CLASSROOM TRAINEES AND A COMPARISON GROUP

	White Males		Black Males		White Females		Black Females	
	Trainees	Comparison Group	Trainees	Comparison Group	Trainees	Comparison Group	Trainees	Comparison Group
1959	$1,443	$2,588	$ 904	$1,438	$ 635	$ 987	$ 384	$ 616
1960	1,533	2,699	976	1,521	687	1,076	440	693
1961	1,572	2,782	1,017	1,573	719	1,163	471	737
1962	1,843	2,963	1,211	1,742	813	1,308	566	843
1963	1,810	3,108	1,182	1,896	748	1,433	531	937
1964	1,551	3,275	1,273	2,121	838	1,580	688	1,060
1965	2,923	3,458	2,327	2,338	1,747	1,698	1,441	1,198
1966	3,750	4,351	2,983	2,919	2,024	1,990	1,794	1,461
1967	3,964	4,430	3,048	3,097	2,244	2,144	1,977	1,678
1968	4,401	4,955	3,409	3,487	2,398	2,339	2,160	1,920
1969	$4,717	$5,033	$3,714	$3,681	$2,646	$2,444	$2,457	$2,133
Number of Observations	7,326	40,921	2,133	6,472	2,730	28,142	1,356	5,192

Assignment Mechanism

Assignment Mechanism

Assignment mechanism is the procedure that determines which units are selected for treatment intake. Examples include:

- random assignment
- selection on observables
- selection on unobservables

Typically, treatment effects models attain identification by restricting the assignment mechanism in some way.

Key Ideas

- Causality is defined by potential outcomes, not by realized (observed) outcomes
- Observed association is neither necessary nor sufficient for causation
- Estimation of causal effects of a treatment (usually) starts with studying the assignment mechanism

Selection Bias

Recall the selection problem when comparing the mean outcomes for the treated and the untreated:

$$\underbrace{E[Y|D=1]-E[Y|D=0]}_{\text{Difference in Means}} = E[Y_1|D=1]-E[Y_0|D=0]$$
$$= \underbrace{E[Y_1-Y_0|D=1]}_{\text{ATET}} + \underbrace{\{E[Y_0|D=1]-E[Y_0|D=0]\}}_{\text{BIAS}}$$

- Random assignment of units to the treatment forces the selection bias to be zero
- The treatment and control group will tend to be similar along all characteristics (including Y_0)

Identification in Randomized Experiments

Randomization implies:

$$(Y_1, Y_0) \text{ independent of } D, \quad \text{or} \quad (Y_1, Y_0) \perp\!\!\!\perp D.$$

We have that $E[Y_0|D=1] = E[Y_0|D=0]$ and therefore

$$\alpha_{ATET} = E[Y_1 - Y_0|D=1] = E[Y|D=1] - E[Y|D=0]$$

Also, we have that

$$\alpha_{ATE} = E[Y_1 - Y_0] = E[Y_1 - Y_0|D=1] = E[Y|D=1] - E[Y|D=0]$$

As a result,

$$\underbrace{E[Y|D=1] - E[Y|D=0]}_{\text{Difference in Means}} = \alpha_{ATE} = \alpha_{ATET}$$

Identification in Randomized Experiments

The identification result extends beyond average treatment effects. Given random assignment $(Y_1, Y_0) \perp\!\!\!\perp D$:

$$\begin{aligned} F_{Y_0}(y) &= \Pr(Y_0 \leq y) = \Pr(Y_0 \leq y | D = 0) \\ &= \Pr(Y \leq y | D = 0) \end{aligned}$$

Similarly,

$$F_{Y_1}(y) = \Pr(Y \leq y | D = 1).$$

So effect of the treatment at any quantile, $Q_\theta(Y_1) - Q_\theta(Y_0)$ is identified.

- Randomization identifies the entire marginal distributions of Y_0 and Y_1
- Does not identify the quantiles of the effect: $Q_\theta(Y_1 - Y_0)$ (the difference of quantiles is not the quantile of the difference)

Estimation in Randomized Experiments

Consider a randomized trial with N individuals. Suppose that the estimand of interest is ATE:

$$\alpha_{ATE} = E[Y_1 - Y_0] = E[Y|D=1] - E[Y|D=0].$$

Using the analogy principle, we construct an estimator:

$$\widehat{\alpha} = \bar{Y}_1 - \bar{Y}_0,$$

where

$$\bar{Y}_1 = \frac{\sum Y_i \cdot D_i}{\sum D_i} = \frac{1}{N_1} \sum_{D_i=1} Y_i\,;$$

$$\bar{Y}_0 = \frac{\sum Y_i \cdot (1 - D_i)}{\sum (1 - D_i)} = \frac{1}{N_0} \sum_{D_i=0} Y_i$$

with $N_1 = \sum_i D_i$ and $N_0 = N - N_1$.
$\widehat{\alpha}$ is an unbiased and consistent estimator of α_{ATE}.

Testing in Large Samples: Two Sample t-Test

Notice that:

$$\frac{\widehat{\alpha} - \alpha_{ATE}}{\sqrt{\dfrac{\widehat{\sigma}_1^2}{N_1} + \dfrac{\widehat{\sigma}_0^2}{N_0}}} \xrightarrow{d} N(0,1),$$

where

$$\widehat{\sigma}_1^2 = \frac{1}{N_1 - 1} \sum_{D_i = 1} (Y_i - \bar{Y}_1)^2,$$

and $\widehat{\sigma}_0^2$ is analogously defined. In particular, let

$$t = \frac{\widehat{\alpha}}{\sqrt{\dfrac{\widehat{\sigma}_1^2}{N_1} + \dfrac{\widehat{\sigma}_0^2}{N_0}}}.$$

We reject the null hypothesis H_0: $\alpha_{ATE} = 0$ against the alternative H_1: $\alpha_{ATE} \neq 0$ at the 5% significance level if $|t| > 1.96$.

Testing in Small Samples: Fisher's Exact Test

- Test of differences in means with large N:

$$H_0 : E[Y_1] = E[Y_0], \quad H_1 : E[Y_1] \neq E[Y_0]$$

- Fisher's Exact Test with small N:

$$H_0 : Y_1 = Y_0, \quad H_1 : Y_1 \neq Y_0 \qquad \text{(sharp null)}$$

- Let Ω be the set of all possible randomization realizations.
- We only observe the outcomes, Y_i, for one realization of the experiment. We calculate $\hat{\alpha} = \bar{Y}_1 - \bar{Y}_0$.
- Under the sharp null hypothesis we can calculate the value that the difference of means would have taken under any other realization, $\hat{\alpha}(\omega)$, for $\omega \in \Omega$.

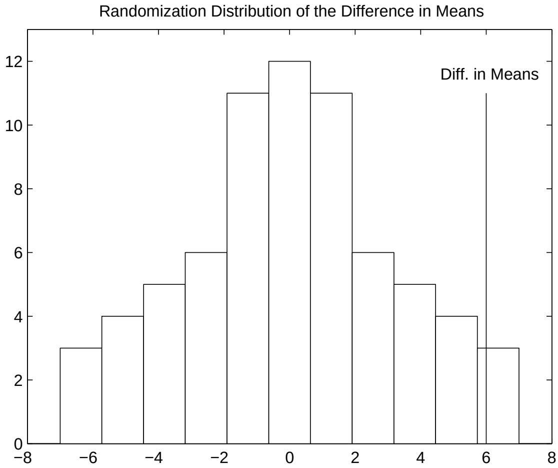
Testing in Small Samples: Fisher's Exact Test

Suppose that we assign 4 individuals out of 8 to the treatment:

Y_i	12	4	6	10	6	0	1	1	
D_i	1	1	1	1	0	0	0	0	$\hat{\alpha} = 6$
									$\hat{\alpha}(\omega)$
$\omega = 1$	1	1	1	1	0	0	0	0	6
$\omega = 2$	1	1	1	0	1	0	0	0	4
$\omega = 3$	1	1	1	0	0	1	0	0	1
$\omega = 4$	1	1	1	0	0	0	1	0	1.5
				...					
$\omega = 70$	0	0	0	0	1	1	1	1	-6

- The randomization distribution of $\widehat{\alpha}$ (under the sharp null hypothesis) is $\Pr(\widehat{\alpha} \leq z) = \frac{1}{70} \sum_{\omega \in \Omega} 1\{\widehat{\alpha}(\omega) \leq z\}$
- Now, find $\bar{z} = \inf\{z : P(|\widehat{\alpha}| > z) \leq 0.05\}$
- Reject the null hypothesis, H_0: $Y_{1i} - Y_{0i} = 0$ for all i, against the alternative hypothesis, H_1: $Y_{1i} - Y_{0i} \neq 0$ for some i, at the 5% significance level if $|\widehat{\alpha}| > \bar{z}$

Testing in Small Samples: Fisher's Exact Test



$$\Pr(|\hat{\alpha}(\omega)| \geq 6) = 0.0857$$

Covariate Balance

- Randomization balances observed but also unobserved characteristics between treatment and control group
- Can check random assignment using so called “balance tests” (e.g., t-tests) to see if distributions of the observed covariates, X, are the same in the treatment and control groups
- X are pre-treatment variables that are measured prior to treatment assignment (i.e., at “baseline”)

Experimental Design: Relative Sample Sizes for Fixed N

Suppose that you have N experimental subjects and you have to decide how many will be in the treatment group and how many in the control group. We know that:

$$\bar{Y}_1 - \bar{Y}_0 \sim \left(\mu_1 - \mu_0, \frac{\sigma_1^2}{N_1} + \frac{\sigma_0^2}{N_0} \right).$$

We want to choose N_1 and N_0, subject to $N_1 + N_0 = N$, to minimize the variance of the estimator of the average treatment effect.
The variance of $\bar{Y}_1 - \bar{Y}_0$ is:

$$\text{var}(\bar{Y}_1 - \bar{Y}_0) = \frac{\sigma_1^2}{pN} + \frac{\sigma_0^2}{(1-p)N}$$

where $p = N_1/N$ is the proportion of treated in the sample.

Experimental Design: Relative Sample Sizes for Fixed N

Find the value p^* that minimizes $\text{var}(\bar{Y}_1 - \bar{Y}_0)$:

$$-\frac{\sigma_1^2}{p^{*2}N} + \frac{\sigma_0^2}{(1-p^*)^2 N} = 0.$$

Therefore:

$$\frac{1-p^*}{p^*} = \frac{\sigma_0}{\sigma_1},$$

and

$$p^* = \frac{\sigma_1}{\sigma_1 + \sigma_0} = \frac{1}{1 + \sigma_0/\sigma_1}.$$

A "rule of thumb" for the case $\sigma_1 \approx \sigma_0$ is $p* = 0.5$

For practical reasons it is sometimes better to choose unequal sample sizes (even if $\sigma_1 \approx \sigma_0$)

Experimental Design: Power Calculations to Choose *N*

- Recall that for a statistical test:
  - Type I error: Rejecting the null if the null is true.
  - Type II error: Not rejecting the null if the null is false.
- Size of a test is the probability of type I error, usually 0.05.
- Power of a test is one minus the probability of type II error, i.e. the probability of rejecting the null if the null is false.
- Statistical power increases with the sample size.
- But when is a sample "large enough"?
- We want to find N such that we will be able to detect an average treatment effect of size α or larger with high probability.

Experimental Design: Power Calculations to Choose N

Assume a particular value, α, for $\mu_1 - \mu_0$.
Let $\widehat{\alpha} = \bar{Y}_1 - \bar{Y}_0$ and

$$\text{s.e.}(\widehat{\alpha}) = \sqrt{\frac{\sigma_1^2}{N_1} + \frac{\sigma_0^2}{N_0}}.$$

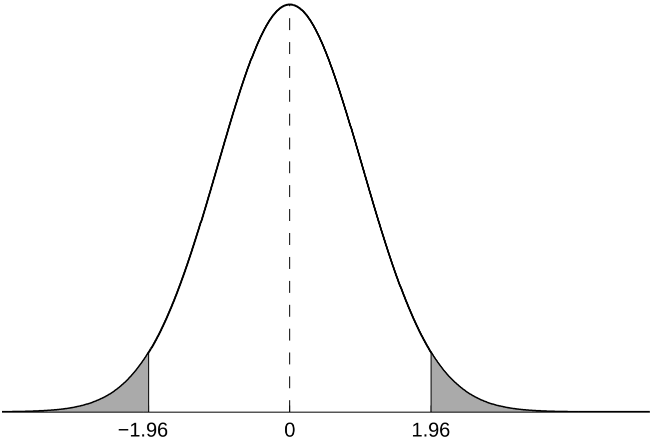
For a large enough sample, we can approximate:

$$\frac{\widehat{\alpha} - \alpha}{\text{s.e.}(\widehat{\alpha})} \sim N(0,1).$$

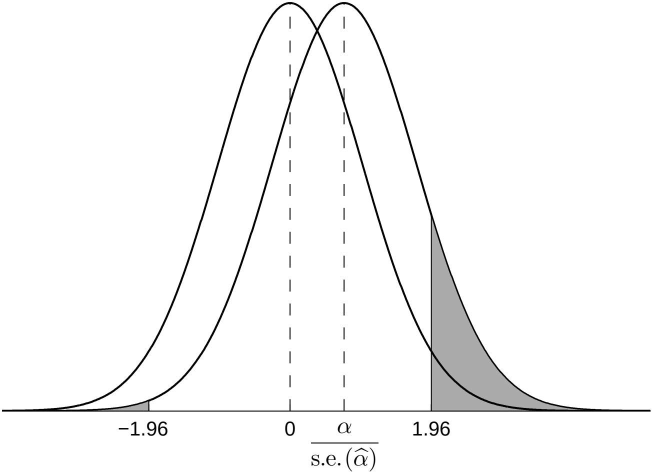
Therefore, the t-statistic for a test of significance is:

$$t = \frac{\widehat{\alpha}}{\text{s.e.}(\widehat{\alpha})} \sim N\left(\frac{\alpha}{\text{s.e.}(\widehat{\alpha})}, 1\right).$$

Probability of Rejection if $\mu_1 - \mu_0 = 0$



Probability of Rejection if $\mu_1 - \mu_0 = \alpha$



Experimental Design: Power Calculations to Choose N

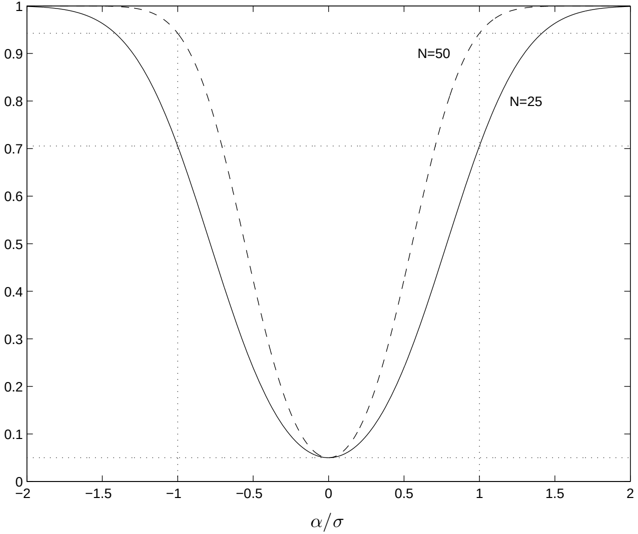
The probability of rejecting the null $\mu_1 - \mu_0 = 0$ is:

$$\begin{aligned}
\Pr(|t| > 1.96) &= \Pr(t < -1.96) + \Pr(t > 1.96) \\
&= \Pr\left(t - \frac{\alpha}{\text{s.e.}(\widehat{\alpha})} < -1.96 - \frac{\alpha}{\text{s.e.}(\widehat{\alpha})}\right) \\
&+ \Pr\left(t - \frac{\alpha}{\text{s.e.}(\widehat{\alpha})} > 1.96 - \frac{\alpha}{\text{s.e.}(\widehat{\alpha})}\right) \\
&= \Phi\left(-1.96 - \frac{\alpha}{\text{s.e.}(\widehat{\alpha})}\right) + \left(1 - \Phi\left(1.96 - \frac{\alpha}{\text{s.e.}(\widehat{\alpha})}\right)\right)
\end{aligned}$$

Suppose that $p = 1/2$ and $\sigma_1^2 = \sigma_0^2 = \sigma^2$. Then,

$$\begin{aligned}
\text{s.e.}(\widehat{\alpha}) &= \sqrt{\frac{\sigma^2}{N/2} + \frac{\sigma^2}{N/2}} \\
&= \frac{2\sigma}{\sqrt{N}}.
\end{aligned}$$

Power Functions with $p = 1/2$ and $\sigma_1^2 = \sigma_0^2$



General formula for the power function ($p \neq 1/2$, $\sigma_0^2 \neq \sigma_1^2$)

$$\begin{aligned}
&\Pr(\text{reject } \mu_1 - \mu_0 = 0 | \mu_1 - \mu_0 = \alpha) \\
&\quad = \Phi\left(-1.96 - \alpha \Big/ \sqrt{\frac{\sigma_1^2}{pN} + \frac{\sigma_0^2}{(1-p)N}}\right) \\
&\qquad + \left(1 - \Phi\left(1.96 - \alpha \Big/ \sqrt{\frac{\sigma_1^2}{pN} + \frac{\sigma_0^2}{(1-p)N}}\right)\right).
\end{aligned}$$

To choose N we need to specify:

1. α: minimum detectable magnitude of treatment effect
2. Power value (usually 0.80 or higher)
3. σ_1^2 and σ_0^2 (usually $\sigma_1^2 = \sigma_0^2$) (e.g., using previous measures)
4. p: proportion of observations in the treatment group If $\sigma_1 = \sigma_0$, then the power is maximized by $p = 0.5$

Threats to the Validity of Randomized Experiments

- Internal validity: can we estimate treatment effect for our particular sample?
  - Fails when there are differences between treated and controls (other than the treatment itself) that affect the outcome and that we cannot control for
- External validity: can we extrapolate our estimates to other populations?
  - Fails when the treatment effect is different outside the evaluation environment

Most Common Threats to Internal Validity

- Failure of randomization
- Non-compliance with experimental protocol
- Attrition

Most Common Threats to External Validity

- Non-representative sample
- Non-representative program
  - The treatment differs in actual implementations
  - Scale effects
  - Actual implementations are not randomized (nor full scale)

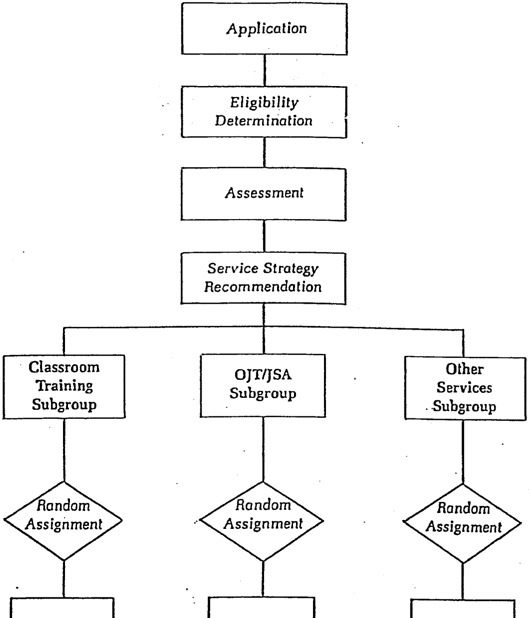
Example: Job Training Partnership Act (JTPA)

- Largest randomized training evaluation ever undertaken in the U.S.; started in 1983 at 649 sites throughout the country
- Sample: Disadvantaged persons in the labor market (previously unemployed or low earnings)
- D: Assignment to one of three general service strategies
  - classroom training in occupational skills
  - on-the-job training and/or job search assistance
  - other services (eg. probationary employment)
- Y: earnings 30 months following assignment
- X: Characteristics measured before assignment (age, gender, previous earnings, race, etc.)

Exhibit 5 Impacts on Total 30-Month Earnings: Assignees and Enrollees, by Target Group

	Mean earnings		Impact per assignee		
	Treatment group (1)	Control group (2)	In dollars (3)	As a percent of (2)	Impact per enrollee in dollars
Adult women	$ 13,417	$ 12,241	$ 1,176***	9.6%	$ 1,837***
Adult men	19,474	18,496	978*	5.3	1,599*
Female youths	10,241	10,106	135	1.3	210
Male youth non-arrestees	15,786	16,375	-589	-3.6	-868
Male youth arrestees					
Using survey data	14,633	18,842	-4,209**	-22.3	-6,804**
Using scaled UI data	14,148	14,152	-4	0.0	-6

Sources: Estimates based on First and Second Follow-up Survey responses and earnings data from state unemployment insurance (UI) agencies.
Sample sizes: adult women, 6,102; adult men, 5,102; female youths, 2,657; male youth non-arrestees, 1,704; male youth arrestees, 416.
* Statistically significant at the .10 level, ** at the .05 level, *** at the .01 level (two-tailed test).


Application
Eligibility Determination
Assessment
Service Strategy Recommendation
Classroom Training Subgroup
OJT/JSA Subgroup
Other Services Subgroup
Random Assignment
Random Assignment
Random Assignment

MEANS AND STANDARD DEVIATIONS

	Entire Sample	Assignment: Treatment	Assignment: Control	Difference (t-stat.)
A. Men				
Number of observations	5,102	3,399	1,703	
Treatment				
Training	.42 [.49]	.62 [.48]	.01 [.11]	.61 (70.34)
Outcome variable				
30 month earnings	19,147 [19,540]	19,520 [19,912]	18,404 [18,760]	1,116 (1.96)
Baseline Characteristics				
Age	32.91 [9.46]	32.85 [9.46]	33.04 [9.45]	-.19 (-.67)
High school or GED	.69 [.45]	.69 [.45]	.69 [.45]	-.00 (-.12)
Married	.35 [.47]	.36 [.47]	.34 [.46]	.02 (1.64)
Black	.25 [.44]	.25 [.44]	.25 [.44]	.00 (.04)
Hispanic	.10 [.30]	.10 [.30]	.09 [.29]	.01 (.70)
Worked less than 13 weeks in past year	.40 [.47]	.40 [.47]	.40 [.47]	.00 (.56)

B. Women

B. Women				
Number of observations	6,102	4,088	2,014	
Treatment				
Training	.45 [.50]	.66 [.47]	.02 [.13]	.64 (80.24)
Outcome variable				
30 month earnings	13,029 [13,415]	13,439 [13,614]	12,197 [12,964]	1,242 (3.46)
Baseline Characteristics				
Age	33.33 [9.78]	33.33 [9.77]	33.35 [9.81]	-.02 (-.09)
High school or GED	.72 [.43]	.73 [.43]	.70 [.44]	.03 (2.01)
Married	.22 [.40]	.22 [.40]	.21 [.39]	.01 (1.55)
Black	.26 [.44]	.27 [.44]	.26 [.44]	.01 (.95)
Hispanic	.12 [.32]	.12 [.32]	.12 [.33]	-.00 (-.89)
Worked less than 13 weeks in past year	.52 [.47]	.52 [.47]	.52 [.47]	-.00 (-.08)
AFDC	.31 [.46]	.30 [.46]	.31 [.46]	-.01 (-1.03)

Exhibit 2.4 DERIVING 30-MONTH EARNINGS SAMPLE FROM FULL EXPERIMENTAL SAMPLE

	All target groups	Adult women	Adult men	Female youths	Male youth non-arrestees	Male youth arrestees
Full experimental sample	20,601	8,058	6,853	3,132	2,041	517
Sample after exogenous deletions for:						
Extra treatment group members[a]	20,123	7,936	6,724	3,015	1,949	499
Late cohorts[b]	19,019	7,497	6,303	2,864	1,871	484
Persons in non-UI sites randomly excluded from Second Follow-up survey[c]	16,347	6,191	5,223	2,712	1,755	466
Male youth arrestees in non-UI sites[d]	16,304	6,191	5,223	2,712	1,755	423
Sample after deletions for missing data:						
30-month earnings sample	15,981	6,102	5,102	2,657	1,704	416
Potentially nonrandom attrition rate	2.0%	1.4%	2.3%	2.0%	2.9%	1.7%

a. A total of 473 treatment group members in 5 sites were randomly excluded to ensure a 2/1 treatment/control group ratio in all sites. Also, the 5 sample members under 22 years of age from Oakland, Calif., were deleted because youths were excluded from the experimental design in Oakland.

b. Deleted were all treatment and control group members randomly assigned after December 1988 in Jackson, Miss.; after April 1989 in Butte, Mont., Jersey City, N.J., and Marion, Ohio; and after June 1989 in Omaha, Neb.

c. The "non-UI" sites (where UI earnings data were not available) are Butte, Jersey City, Marion, and Oakland.

d. The remaining sample at this stage has the statistical properties of a randomized experiment.

Exhibit 3.3 SELECTED ECONOMIC CONDITIONS AT 16 STUDY SITES

Site	Mean unemployment rate, 1987–89 (1)	Mean earnings, 1987 (2)	Percentage employed in manufacturing, mining, or agriculture, 1988 (3)	Annual growth in retail and wholesale earnings, 1989 (4)
Fort Wayne, Ind.	4.7%	$18,700	33.3%	−0.1%
Coosa Valley, Ga.	6.5	16,000	42.8	2.1
Corpus Christi, Tex.	10.2	18,700	16.8	−15.5
Jackson, Miss.	6.1	17,600	12.8	−2.4
Providence, R.I.	3.8	17,900	28.0	9.7
Springfield, Mo.	5.5	15,800	19.4	−1.8
Jersey City, N.J.	7.3	21,400	20.9	9.9
Marion, Ohio	7.0	18,600	37.7	1.7
Oakland, Calif.	6.8	23,000	14.6	3.0
Omaha, Neb.	4.3	18,400	11.8	1.8
Larimer County, Colo.	6.5	17,800	21.2	−3.1
Heartland, Fla.	8.5	15,700	23.8	−0.3
Northwest Minnesota	8.0	14,100	23.0	2.4
Butte, Mont.	6.8	16,900	9.6	−5.7
Decatur, Ill.	9.2	21,100	27.1	−1.1
Cedar Rapids, Iowa	3.6	17,900	21.9	−0.5
16-site average	6.6	18,100	22.8	0.0
National average, all SDAs	6.6	18,167	23.4	1.5

Source: Unweighted annual averages calculated from JTPA Annual Status Report computer files produced by U.S. Department of Labor.
Note: Missing data for certain measures precluded using same year across columns.

Exhibit 3.6 SELECTED CHARACTERISTICS OF JTPA TITLE II PROGRAMS AT 16 STUDY SITES, PROGRAM YEARS 1987–89

Site	Mean number of adult and youth terminees[a] (1)	Mean number of weeks enrolled: Adults (2)	Mean number of weeks enrolled: Youths[a] (3)	Mean federal program cost per adult terminee (4)
Fort Wayne, Ind.	1,195	16	31	$1,561
Coosa Valley, Ga.	1,063	12	15	2,481
Corpus Christi, Tex.	1,049	34	33	2,570
Jackson, Miss.	1,227	8	15	1,897
Providence, R.I.	503	7	5	2,841
Springfield, Mo.	938	17	17	1,898
Jersey City, N.J.	853	16	14	3,637
Marion, Ohio	714	27	26	2,199
Oakland, Calif.	1,396	16	17	2,539
Omaha, Neb.	1,111	11	12	2,404
Larimer County, Colo.	354	32	26	1,937
Heartland, Fla.	1,793	15	24	1,782
Northwest Minnesota	430	29	28	2,371
Butte, Mont.	576	21	19	2,665
Decatur, Ill.	525	29	25	3,039
Cedar Rapids, Iowa	658	31	23	2,212
16-site average	899	20	21	2,377
National average, all SDAs	1,177	20	22	2,241

Source: Unweighted annual averages calculated from JTPA Annual Status Report computer files produced by U.S. Department of Labor.

a. Includes adults and both out-of-school and in-school youths ages 14 to 21. Experimental sample does not include in-school youths or youths under age 16.

A Final Word about Policy Outcome

After the results of the National JTPA study were released, in 1994, funding for JTPA training for the youth was drastically cut:

SPENDING ON JTPA PROGRAMS

Year	Youth Training Grants	Adult Training Grants
1993	677	1015
1994	609	988
1995	127	996
1996	127	850
1997	127	895