Experiments and Quasi-Experiments

Outline

1. Potential Outcomes, Causal Effects, and Idealized Experiments
2. Threats to Validity of Experiments
3. Application: The Tennessee STAR Experiment
4. Quasi-Experiments: Differences-in-Differences, IV Estimation, Regression Discontinuity Design, and Other Approaches.
5. Threats to Validity of Quasi-Experiments
6. Heterogeneous Causal Effects

Why study experiments?

- Ideal randomized controlled experiments provide a conceptual benchmark for assessing observational studies.
- Actual experiments are rare ($$$) but influential.
- Experiments can overcome the threats to internal validity of observational studies, however they have their own threats to internal and external validity.
- Thinking about experiments helps us to understand quasi-experiments, or "natural experiments," in "natural" variation induces "as if" random assignment.

Terminology: experiments and quasi-experiments

- An ***experiment*** is designed and implemented consciously by human researchers. An experiment randomly assigns subjects to treatment and control groups (think of clinical drug trials)
- A ***quasi-experiment*** or ***natural experiment*** has a source of randomization that is "as if" randomly assigned, but this variation was not the result of an explicit randomized treatment and control design.
- ***Program evaluation*** is the field of statistics aimed at evaluating the effect of a program or policy, for example, an ad campaign to cut smoking, or a job training program.

Different Types of Experiments: Three Examples

- Clinical drug trial: does a proposed drug lower cholesterol?
  - Y = cholesterol level
  - X = treatment or control group (or dose of drug)
- Job training program (Job Training Partnership Act)
  - Y = has a job, or not (or Y = wage income)
  - X = went through experimental program, or not
- Class size effect (Tennessee class size experiment)
  - Y = test score (Stanford Achievement Test)
  - X = class size treatment group (regular, regular + aide, small)

Potential Outcomes, Causal Effects, and Idealized Experiments (SW Section 13.1)

A treatment has a causal effect for a given individual: give the individual the treatment and something happens, which is (possibly) different than what happens if you don't get the treatment.

- A ***potential outcome*** is the outcome for an individual under a potential treatment or potential non-treatment.
- For an individual, the causal effect is the difference in potential outcomes if you do or don't get the treatment.
- An individual's causal effect cannot be observed because you can give the subject the treatment, or not – but not both!

From potential outcomes to regression: the math (1 of 2)

Consider subject i drawn at random from a population and let:

X_i = 1 if subject i treated, 0 if not (binary treatment)
$Y_i(0)$ = potential outcome for subject i if untreated
$Y_i(1)$ = potential outcome for subject i if treated

We observe (Y_i, X_i), where Y_i is the observed outcome:

$$Y_i = Y_i(1)X_i + Y_i(0)(1 - X_i) \quad \text{(think about it)}$$
$$= Y_i(0) + [Y_i(1) - Y_i(0)]X_i \quad \text{(algebra)}$$
$$= E[Y_i(0)] + [Y_i(1) - Y_i(0)]X_i + [Y_i(0) - E(Y_i(0))] \quad \text{(more algebra)}$$

where the expectation is over the population distribution.

From potential outcomes to regression: the math (2 of 2)

Thus

$$Y_i = E[Y_i(0)] + [Y_i(1) - Y_i(0)]X_i + [Y_i(0) - E(Y_i(0))]$$
$$= \beta_0 + \beta_{1i} X_i + u_i$$

where

$\beta_0 = E[Y_i(0)]$

$\beta_{1i} = Y_i(1) - Y_i(0)$ = individual i's causal effect

$u_i = Y_i(0) - E(Y_i(0))$, so $Eu_i = 0$.

The regression model with heterogeneous treatment effects – each person has his or her own treatment effect, is:

$$Y_i = \beta_0 + \beta_{1i} X_i + u_i$$

where β_{1i} is individual i's causal effect ("treatment effect").

Average Treatment Effect

Heterogeneous treatment effect regression model:

$$Y_i = \beta_0 + \beta_{1i} X_i + u_i$$

where β_{1i} is individual i's causal effect ("treatment effect").

- In general, different people have different treatment effects. For people drawn from a population, the ***average treatment effect*** is the population mean value of the individual treatment effects:

  Average Treatment Effect = ATE = $E(\beta_{1i})$

- In many applications, the object of interest is the ATE (the average effect in the population of interest).
- For now we suppose there is no heterogeneity in treatment effects, so that all individuals have the same treatment effect β_1.
  - We return to heterogeneous treatment effects in Section 13.6

Estimating the treatment effect in an ideal randomized controlled experiment (1 of 2)

An ideal randomized controlled experiment randomly assigns subjects to treatment and control groups.

- Let X be the treatment variable and Y the outcome variable of interest. If X is randomly assigned (for example by computer) then X is independent of all individual characteristics.
- Let the (homogeneous) treatment effect be β_1:

$$Y_i = \beta_0 + \beta_1 X_i + u_i.$$

  If X_i is randomly assigned, then X_i is independent of u_i, so $E(u_i|X_i) = 0$, so OLS yields an unbiased estimator of β_1.
- The causal effect is the population value of β_1 in an ideal randomized controlled experiment

Estimating the treatment effect in an ideal randomized controlled experiment (2 of 2)

$$Y_i = \beta_0 + \beta_1 X_i + u_i$$

- When the treatment is binary, $\hat{\beta}_1$ is just the difference in mean outcome (Y) in the treatment vs. control group $(\bar{Y}^{treated} - \bar{Y}^{control})$.
- This difference in means is sometimes called the ***differences estimator***.

Additional regressors

Let X_i = treatment variable and W_i = control variable(s).

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 W_i + u_i$$

Two reasons to include W in a regression analysis of the effect of a randomly assigned treatment:

1. If X_i is randomly assigned then X_i is uncorrelated with W_i so omitting W_i doesn't result in omitted variable bias. But including W_i reduces the error variance and can result in smaller standard errors.
2. If the probability of assignment depends on W_i, so that X_i is randomly assigned *given* W_i, then omitting W_i can lead to OV bias, but including it eliminates that OV bias. This situation is called…

Randomization based on covariates (1 of 2)

Example: men ($W_i = 0$) and women ($W_i = 1$) are randomly assigned to a course on table manners (X_i), but women are assigned with a higher probability than men. Suppose women have better table manners than men prior to the course. Then even if the course has no effect, the treatment group will have better post-course table manners than the control group because the treatment group has a higher fraction of women than the control group.

That is, the OLS estimator of β_1 in the regression,

$$Y_i = \beta_0 + \beta_1 X_i + u_i$$

has omitted variable bias, which is eliminated by the regression,

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 W_i + u_i$$

Randomization based on covariates (2 of 2)

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 W_i + u_i$$

- In this example, X_i is randomly assigned, given W_i, so $E(u_i|X_i, W_i) = E(u_i|W_i)$.
  - In words, among women, treatment is randomly assigned, so among women, the error term is independent of X_i so, among women, its mean doesn't depend on X_i. Same is true among men.
- Thus if randomization is based on covariates, conditional mean independence holds, so that once W_i is included in the regression the OLS estimator is unbiased.

Estimating causal effects that depend on observables

The causal effect in the previous example might depend on observables, perhaps $\beta_{1,men} > \beta_{1,women}$ (men could benefit more from the table manners course than women).

- We already know how to estimate different coefficients for different groups – use interactions.
- In the table manners example, we would simply estimate the interactions model,

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 X_i \times W_i + \beta_3 W_i + u_i$$

- Because interactions were covered in Ch. 8, for simplicity in Ch. 13 we ignore differences in causal effects that depend on observable W's – not because they aren't important, but because they are easy to handle.
  - We return to differences in β_{1i}'s (unobserved heterogeneity – in contrast to heterogeneity that depends on observable variables like sex) in Section 13.6.

Threats to Validity of Experiments (SW Section 13.2) (1 of 4)

Threats to Internal Validity

1. ***Failure to randomize*** (or imperfect randomization)
   - for example, openings in job treatment program are filled on first-come, first-serve basis; latecomers are controls
   - result is correlation between X and u

Threats to Validity of Experiments (SW Section 13.2) (2 of 4)

2. Failure to follow treatment protocol (or "***partial compliance***")

- some controls get the treatment
- some of those who should be treated aren't
- If you observe whether the subject actually receives treatment (X), if you know whether the individual was initially assigned to a treatment group (Z), and if initial assignment was random, then you can estimate the causal effect using initial assignment as an instrument for actual treatment.

Threats to Validity of Experiments (SW Section 13.2) (3 of 4)

3. Attrition (some subjects drop out)

- Suppose the controls who get jobs move out of town; then corr(X, u) $\neq$ 0
- This is a reincarnation of sample selection bias from Ch. 9 (the sample is selected in a way related to the outcome variable).

Threats to Validity of Experiments (SW Section 13.2) (4 of 4)

4. Experimental effects

- experimenter bias (conscious or subconscious): treatment X is associated with "extra effort" or "extra care," so corr(X, u) $\neq$ 0
- subject behavior might be affected by being in an experiment, so corr(X, u) $\neq$ 0 (Hawthorne effect)

Just as in regression analysis with observational data, threats to the internal validity of regression with experimental data implies that corr(X, u) $\neq$ 0 so OLS (the differences estimator) is biased.

- In an experiment, this can sometimes be mitigated by using a "double blind" protocol in which neither the experimenter or the subject knows who is in the treatment or control groups.

Threats to External Validity

1. Nonrepresentative sample
2. Nonrepresentative "treatment" (that is, program or policy)
3. General equilibrium effects (effect of a program can depend on its scale; admissions counseling)

Experimental Estimates of the Effect of Class Size Reductions (SW Section 13.3)

Project STAR (Student-Teacher Achievement Ratio)

- 4-year study, $12 million
- Upon entering the school system, a student was randomly assigned to one of three groups:
  - regular class (22 – 25 students)
  - regular class + aide
  - small class (13 – 17 students)
- regular class students re-randomized after first year to regular or regular + aide
- Y = Stanford Achievement Test scores

Deviations from experimental design

- Partial compliance:
  - 10% of students switched treatment groups because of "incompatibility" and "behavior problems" – how much of this was because of parental pressure?
  - Newcomers: incomplete receipt of treatment for those who move into district after grade 1
- Attrition
  - students move out of district
  - students leave for private/religious schools
  - This is only a problem if their departure is related to Y_i; for example if high-achieving kids leave because they are assigned to a large class, then large classes will spuriously appear to do relatively worse ($\text{corr}(u_i, X_i) > 0$)

Regression analysis

- The "differences" regression model:

$$Y_i = \beta_0 + \beta_1 SmallClass_i + \beta_2 RegAide_i + u_i$$

where

$SmallClass_i$ = 1 if in a small class

$RegAide_i$ = 1 if in regular class with aide

- Additional regressors (W's)
  - teacher experience
  - free lunch eligibility
  - gender, race

Differences estimates (no W's) (1 of 2)

TABLE 13.1 Project STAR: Differences Estimates of Effect on Standardized Test Scores of Class Size Treatment Group

	Grade			
Regressor	K	1	2	3
Small class	13.90 (4.23) [5.48, 22.32]	29.78 (4.79) [20.24, 39.32]	19.39 (5.12) [9.18, 29.61]	15.59 (4.21) [7.21, 23.97]
Regular-sized class with aide	0.31 (3.77) [−7.19, 7.82]	11.96 (4.87) [2.27, 21.65]	3.48 (4.91) [−6.31, 13.27]	−0.29 (4.04) [−8.35, 7.77]
Intercept	918.04 (4.82)	1039.39 (5.82)	1157.81 (5.29)	1228.51 (4.66)
Number of observations	5786	6379	6049	5967

The regressions were estimated using the Project STAR public access data set described in Appendix 13.1. The dependent variable is the student's combined score on the math and reading portions of the Stanford Achievement Test. Standard errors, clustered at the school level, appear in parentheses, and 95% confidence intervals appear in brackets.

Differences estimates (no W's) (2 of 2)

TABLE 13.2 Project STAR: Differences Estimates with Additional Regressors for Kindergarten

Regressor	(1)	(2)	(3)	(4)
Small class	13.90 (4.23) [5.48, 22.32]	14.00 (4.25) [5.55, 22.46]	15.93 (4.08) [7.81, 24.06]	15.89 (3.95) [8.03, 23.74]
Regular-sized class with aide	0.31 (3.77) [−7.19, 7.82]	−0.60 (3.84) [−8.25, 7.05]	1.22 (3.64) [−6.04, 8.47]	1.79 (3.60) [−5.38, 8.95]
Teacher's years of experience		1.47 (0.44) [0.60, 2.34]	0.74 (0.35) [0.04, 1.45]	0.66 (0.36) [−0.05, 1.37]
Boy				−12.09 (1.54)
Free lunch eligible				−34.70 (2.47)
Black				−25.43 (4.52)
Race other than black or white				−8.50 (12.64)
School indicator variables?	no	no	yes	yes
$\overline{R}^2$	0.01	0.02	0.22	0.28
Number of observations	5786	5766	5766	5748

The regressions were estimated using the Project STAR public access data set described in Appendix 13.1. The dependent variable is the student's combined test score on the math and reading portions of the Stanford Achievement Test. All regressions include an intercept (not reported). The number of observations differs in the different regressions because of some missing data. Standard errors, clustered at the school level, appear in parentheses, and 95% confidence intervals appear in brackets.

How big are these estimated effects? (1 of 2)

- Put on same basis by dividing by std. dev. of Y
- Units are now standard deviations of test scores

TABLE 13.3 Estimated Class Size Effects in Units of Standard Deviations of the Test Score Across Students

	Grade			
Treatment Group	K	1	2	3
Small class	0.19 (0.06)	0.33 (0.05)	0.23 (0.06)	0.21 (0.06)
Regular-sized class with aide	0.00 (0.05)	0.13 (0.05)	0.04 (0.06)	0.00 (0.06)
Sample standard deviation of test scores (s_Y)	73.75	91.25	84.08	73.27

The estimates and standard errors in the first two rows are the estimated effects in Table 13.1, divided by the sample standard deviation of the Stanford Achievement Test for that grade (the final row in this table), computed using data on the students in the experiment. Standard errors, clustered at the school level, appear in parentheses.

How big are these estimated effects? (2 of 2)

How do these estimates compare to those from the California & Mass. observational studies? (Ch. 4 – 9)

TABLE 13.4 Estimated Effects of Reducing the Student–Teacher Ratio by 7.5 Based on the STAR Data and the California and Massachusetts Observational Data

Study	$\hat{\beta}_1$	Change in Student–Teacher Ratio	Standard Deviation of Test Scores Across Students	Estimated Effect	95% Confidence Interval
STAR (grade K)	–13.90** (2.45)	Small class vs. regular class	73.8	0.19** (0.03)	(0.13, 0.25)
California	– 0.73** (0.26)	–7.5	38.0	0.14** (0.05)	(0.04, 0.24)
Massachusetts	– 0.64* (0.27)	–7.5	39.0	0.12* (0.05)	(0.02, 0.22)

Note: The estimated coefficient $\hat{\beta}_1$ for the STAR study is taken from column (1) of Table 13.2. The estimated coefficients for the California and Massachusetts studies are taken from the first column of Table 9.3. The estimated effect is the effect of being in a small class versus a regular class (for STAR) or the effect of reducing the student–teacher ratio by 7.5 (for the California and Massachusetts studies). The 95% confidence interval for the reduction in the student–teacher ratio is this estimated effect ± 1.96 standard errors. Standard errors are given in parentheses under estimated effects. The estimated effects are statistically significantly different from zero at the *5% significance level or **1% significance level using a two-sided test.

A conditional mean independence example from STAR: What is the effect on *Y* of *X* = *Teacher's years of experience*?

<u>More on the design of Project STAR</u>

Teachers were randomly assigned to small/regular/reg+aide classrooms *within their normal school* – teachers didn't change schools as part of the experiment.

Because teacher experience differed systematically across schools (more experienced teachers in more affluent school districts), a regression of test scores on teacher experience would have omitted variable bias and the estimated effect on test scores of teacher experience would be biased up (overstated).

However, the design implies conditional mean independence (1 of 2)

- W = full set of school binary indicators
- Given W (school), X is randomly assigned (teachers are randomly assigned to classrooms and students)
- W is plausibly correlated with u (nonzero school fixed effects: some schools are richer than others)
- Thus $E(u|X) \neq 0$ but $E(u|X,W) = E(u|W)$ (conditional mean independence)
- The key is that teacher randomization is "stratified" by school: X is randomly assigned given W.
- The coefficient on the school identity (W) is not a causal effect (think about it)

However, the design implies conditional mean independence (2 of 2)

TABLE 13.2 Project STAR: Differences Estimates with Additional Regressors for Kindergarten

Regressor	(1)	(2)	(3)	(4)
Small class	13.90 (4.23) [5.48, 22.32]	14.00 (4.25) [5.55, 22.46]	15.93 (4.08) [7.81, 24.06]	15.89 (3.95) [8.03, 23.74]
Regular-sized class with aide	0.31 (3.77) [−7.19, 7.82]	−0.60 (3.84) [−8.25, 7.05]	1.22 (3.64) [−6.04, 8.47]	1.79 (3.60) [−5.38, 8.95]
Teacher's years of experience		1.47 (0.44) [0.60, 2.34]	0.74 (0.35) [0.04, 1.45]	0.66 (0.36) [−0.05, 1.37]
Boy				−12.09 (1.54)
Free lunch eligible				−34.70 (2.47)
Black				−25.43 (4.52)
Race other than black or white				−8.50 (12.64)
School indicator variables?	no	no	yes	yes
$\overline{R}^2$	0.01	0.02	0.22	0.28
Number of observations	5786	5766	5766	5748

The regressions were estimated using the Project STAR public access data set described in Appendix 13.1. The dependent variable is the student's combined test score on the math and reading portions of the Stanford Achievement Test. All regressions include an intercept (not reported). The number of observations differs in the different regressions because of some missing data. Standard errors, clustered at the school level, appear in parentheses, and 95% confidence intervals appear in brackets.

Example: effect of teacher experience

- Without school fixed effects (2), the estimated effect of an additional year of experience is 1.47 (*SE* = .17)
- "Controlling for the school" in (3), the estimated effect of an additional year of experience is .74 (*SE* = .17)
- *Does the difference between (2) and (3) make sense?*
- The OLS estimator of the coefficient on years of experience is biased up without school effects; with school effects, OLS is an unbiased estimator of the causal effect of *X*.

Another example of a well-done randomized controlled experiment: Program Evaluation of Teach for America

Full report is available (follow links) at:

https://www.mathematica-mpr.com/our-publications-and-findings/projects/teach-for-america

Summary: The Tennessee Class Size Experiment

Remaining threats to internal validity

- partial compliance/incomplete treatment
  - can use TSLS with Z = initial assignment
  - Turns out, TSLS and OLS estimates are similar (Krueger (1999)), so this bias seems not to be large

Main findings:

- The effects are small quantitatively (same size as gender difference)
- Effect is sustained but not cumulative or increasing (biggest effect at the youngest grades)

Quasi-Experiments (SW Section 13.4)

A ***quasi-experiment*** or ***natural experiment*** has a source of randomization that is "as if" randomly assigned, but this variation was not the result of an explicit randomized treatment and control design.

Two types of quasi-experiments

a) Treatment (X) is "as if" randomly assigned (perhaps conditional on some control variables W)
   - *Example*: Effect of marginal tax rates on labor supply
   - X = marginal tax rate (the tax rate changes in one state, not another; state is "as if" randomly assigned)

b) A variable (Z) which influences receipt of treatment (X) is "as if" randomly assigned, so we can run IV and use Z as an instrument for X.
   - *Example*: Effect on survival of cardiac catheterization
     - X = cardiac catheterization;
     - Z = differential distance to CC hospital

Econometric methods (1 of 3)

(a) Treatment (X) is "as if" randomly assigned: OLS

The differences-in-differences estimator uses two pre- and post-treatment measurements of Y, and estimates the treatment effect as the difference between the pre- and post-treatment values of Y for the treatment and control groups.

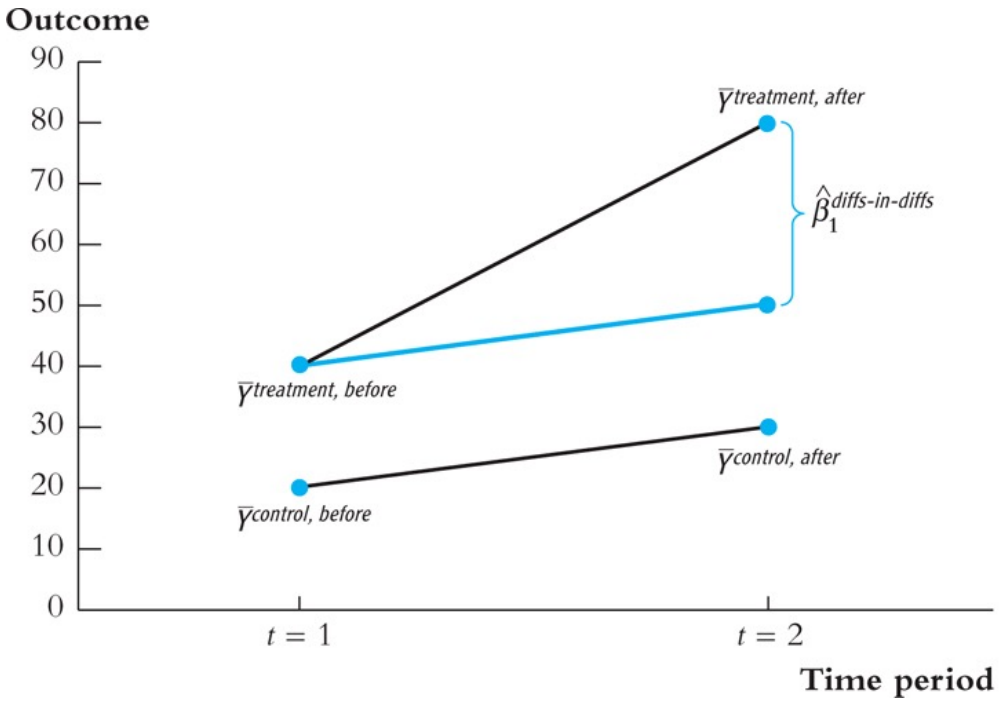
Let:

$$Y_i^{before} = \text{value of } Y \text{ for subject } i \text{ before the expt}$$

$$Y_i^{after} = \text{value of } Y \text{ for subject } i \text{ after the expt}$$

Econometric methods (2 of 3)

$$\hat{\beta}_1^{diffs-in-diffs} = (\bar{Y}^{treat,after} - \bar{Y}^{treat,before}) - (\bar{Y}^{control,after} - \bar{Y}^{control,before})$$



Econometric methods (3 of 3)

"Differences" regression formulation:

$$\Delta Y_i = \beta_0 + \beta_1 X_i + u_i$$

where

$$\Delta Y_i = Y_i^{after} - Y_i^{before}$$

$$X_i = 1 \text{ if treated, } = 0 \text{ otherwise}$$

$$\hat{\beta}_1 \text{ is the diffs-in-diffs estimator}$$

The differences-in-differences estimator allows for systematic differences in pre-treatment characteristics, which can happen in a quasi-experiment because treatment is not randomly assigned.

Differences-in-differences with control variables

$$\Delta Y_{it} = \beta_0 + \beta_1 X_{it} + \beta_2 W_{1it} + \cdots + \beta_{1+r} W_{rit} + u_{it},$$

X_{it} = 1 if the treatment is received, = 0 otherwise

Why include control variables?

For the usual reason: If the treatment (X) is "as if" randomly assigned, given W, then u is conditionally mean independent of X: $E(u|X, W) = E(u|W)$ and including W results in the OLS estimator of β_1 being unbiased.

Differences-in-differences with multiple time periods

The drunk driving law analysis of Ch. 10 can be thought of as a quasi-experiment panel data design: if (given the control variables) the beer tax is as if randomly assigned, then the causal effect of the beer tax (the elasticity) can be estimated by panel data regression.

- The tools of Ch. 10 apply. Ignoring W's, the differences-in-differences estimator obtains from including individual fixed effects and time effects:

$$Y_{it} = \alpha_i + \delta_t + \beta_1 X_{it} + u_{it}.$$

- If $T = 2$ (2 periods) and the treatment is in the second period, then $\Delta X_{it} = X_{it}$ (*why?*) and the fixed effects/time effects regression becomes $\Delta Y_{it} = \beta_0 + \beta_1 X_{it} + \Delta u_{it}$

IV estimation

- If a variable (Z) that influences treatment (X) is "as if" randomly assigned, conditional on W, then Z can be used as an instrumental variable for X in an IV regression that includes the control variables W.
  - We encountered this in Ch. 12 (IV regression).
  - The concept of "as-if" randomization has proven to be a fruitful way to think of instrumental variables.
- One example is the location of a heart attack in the cardiac catheterization study (the location being "as if" randomly assigned)

Regression Discontinuity Estimators

If treatment occurs when some continuous variable W crosses a threshold w_0, then you can estimate the treatment effect by comparing individuals with W just below the threshold (treated) to these with W just above the threshold (untreated). If the direct effect on Y of W is continuous, the effect of treatment should show up as a jump in the outcome. The magnitude of this jump estimates the treatment effect.

- In ***sharp*** regression discontinuity design, everyone above (or below) the threshold w_0 gets treatment.
- In ***fuzzy*** regression discontinuity design, crossing the threshold w_0 influences the probability of treatment, but that probability is between 0 and 1.

Sharp Regression Discontinuity

Everyone with $W < w_0$ gets treated while everyone with $W \geqslant w_0$ does not, so

$$X_i = 1 \text{ if } W_i < w_0 \text{ and } X_i = 0 \text{ otherwise.}$$

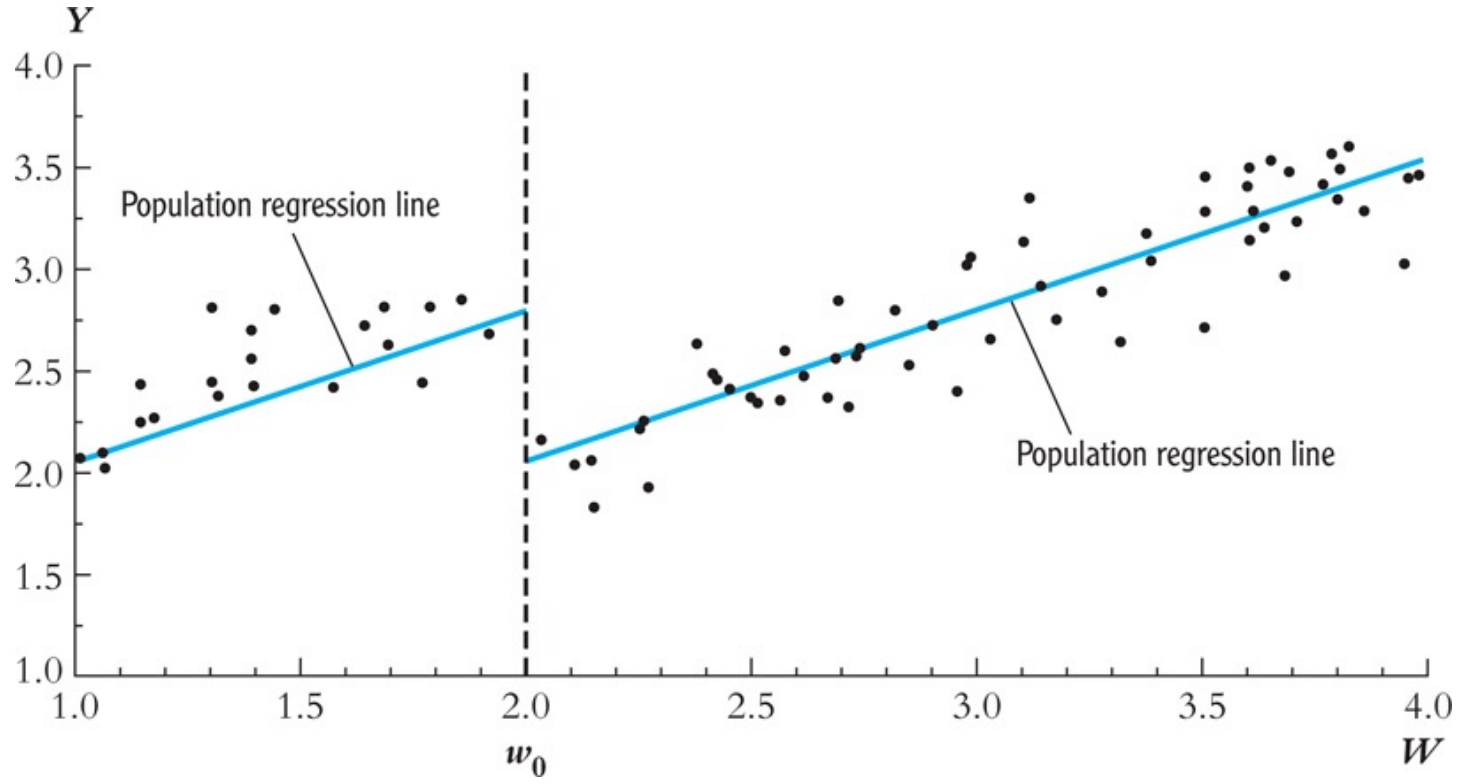
The treatment effect, β_1, can be estimated by OLS:

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 W_i + u_i$$

If crossing the threshold affects Y_i only through the treatment, then $E(u_i|X_i, W_i) = E(u_i|W_i)$ so $\hat{\beta}_1$ is unbiased.

Sharp regression discontinuity design in a picture:

Treatment occurs for everyone with $W < w_0$, and the treatment effect is the jump or "discontinuity."



Fuzzy Regression Discontinuity

Let

X_i = binary treatment variable

$Z_i = 1$ if $W < w_0$ and $Z_i = 0$ otherwise.

If crossing the threshold has no direct effect on Y_i, so only affects Y_i by influencing the probability of treatment, then $E(u_i|Z_i, W_i) = 0$. Thus Z_i is an exogenous instrument for X_i.

- Example:
  - Matsudaira, Jordan D. (2008). "Mandatory Summer School and Student Achievement." *Journal of Econometrics* 142: 829–850. This paper studies the effect of mandatory summer school by comparing subsequent performance of students who fell just below, and just above, the grade cutoff at which summer school was required.

Potential Problems with Quasi-Experiments (SW Section 13.5) (1 of 2)

The **threats to the internal validity** of a quasi-experiment are the same as for a true experiment, with one addition.

1. ***Failure to randomize*** (imperfect randomization)
   a. Is the "as if" randomization really random, so that X (or Z) is uncorrelated with u?
2. ***Failure to follow treatment protocol***
3. ***Attrition*** (n.a.)
4. ***Experimental effects*** (n.a.)
5. ***Instrument invalidity*** (relevance + exogeneity) (Maybe healthier patients *do* live closer to CC hospitals –they might have better access to care in general)

Potential Problems with Quasi-Experiments (SW Section 13.5) (2 of 2)

The **threats to the external validity** of a quasi-experiment are the same as for an observational study.

1. Nonrepresentative sample
2. Nonrepresentative "treatment" (that is, program or policy)

Example: Cardiac catheterization

- The CC study has better external validity than controlled clinical trials because the CC study uses observational data based on real-world implementation of cardiac catheterization.
- However they used data from the early '90s – do the findings apply to CC usage today?

Experimental and Quasi-Experimental Estimates in Heterogeneous Populations (SW Section 13.6) (1 of 5)

By a "heterogeneous population" we mean a population in which the treatment effect differs from one person to the next.

- In the potential outcome terminology, each individual's treatment effect is $\beta_{1i} = Y_i(1) - Y_i(0)$, where $Y_i(1)$ is individual i's potential outcome if treated and $Y_i(0)$ is i's potential outcome if untreated.
- In general, β_{1i} differs across people in unobservable ways:
  - Effect of job training program probably depends on motivation
  - Effect of a cholesterol-lowering drug could depend on unobserved health factors

Experimental and Quasi-Experimental Estimates in Heterogeneous Populations (SW Section 13.6) (2 of 5)

$$\beta_{1i} = Y_i(1) - Y_i(0) = \text{person } i\text{'s treatment effect}$$

If this variation depends on observed variables, then this is a job for interaction variables! We know how to do this already.

What if the source of variation is unobserved? This raises two questions:

1. What do we *want* to estimate?
2. What do our usual tools (OLS, IV) deliver when there is population heterogeneity? Do OLS and IV give us what we want, or something different?

Experimental and Quasi-Experimental Estimates in Heterogeneous Populations (SW Section 13.6) (3 of 5)

1. What do we *want* to estimate?
   - Most commonly, we want to estimate the **average treatment effect** in the population:

     $$\beta_1 = E\beta_{1i} = E[Y_i(1) - Y_i(0)]$$

     This is the mean outcome we would get if everyone in the population were treated (all high-cholesterol patients took the drug; all qualified unemployed took the job training program).
   - Another possibility is the **effect of treatment on the treated**, that is, the treatment effect for those two get treated: $E[\beta_{1i} | X_i = 1]$.
   - We will focus on the average treatment effect.

Experimental and Quasi-Experimental Estimates in Heterogeneous Populations (SW Section 13.6) (4 of 5)

2. What do our usual tools (OLS, IV) deliver when there is population heterogeneity?

(a) OLS with heterogeneity and random assignment

Recall that, starting with the definitions of potential outcomes, we derived the heterogeneous effects regression model,

$$\begin{aligned} Y_i &= \beta_0 + \beta_{1i} X_i + u_i && \text{(from above)} \\ &= \beta_0 + \beta_1 X_i + (\beta_{1i} - \beta_1) X_i + u_i && \text{(algebra)} \\ &= \beta_0 + \beta_1 X_i + v_i && \text{(redefinition)} \end{aligned}$$

where

$v_i = (\beta_{1i} - \beta_1) X_i + u_i$

$\beta_1 = E\beta_{1i}$ = average treatment effect.

If $E(v_i|X_i) = 0$, then $\hat{\beta}_1$ is an unbiased estimator of the average treatment effect, β_1

Experimental and Quasi-Experimental Estimates in Heterogeneous Populations (SW Section 13.6) (5 of 5)

$Y_i = \beta_0 + \beta_1 X_i + v_i$, where $v_i = (\beta_{1i} - \beta_1)X_i + u_i$

Check:

$$E(v_i|X_i) = E[(\beta_{1i} - \beta_1)X_i + u_i|X_i]$$
$$= E[(\beta_{1i} - \beta_1)X_i|X_i] + E(u_i|X_i)$$
$$= E[(\beta_{1i} - \beta_1)|X_i]\, X_i + E(u_i|X_i) = 0$$

because $E[(\beta_{1i} - \beta_1)|X_i] = 0$ and $E(u_i|X_i) = 0$, which both follow from X_i being randomly assigned and therefore independent of all individual characteristics.

Thus, if X is randomly assigned, $E(\hat{\beta}_1) = \beta_1$, so OLS estimates the ATE!

(b) IV with heterogeneity and random assignment (1 of 3)

Suppose the treatment effect is heterogeneous *and* the effect of the instrument on X is heterogeneous:

$$Y_i = \beta_0 + \beta_{1i} X_i + u_i \qquad \text{(equation of interest)}$$

$$X_i = \pi_0 + \pi_{1i} Z_i + v_i \qquad \text{(first stage of TSLS)}$$

In general, TSLS estimates the causal effect for those whose value of X (probability of treatment) is most influenced by the instrument – which is called the ***Local Average Treatment Effect (LATE)***

(b) IV with heterogeneity and random assignment (2 of 3)

$$Y_i = \beta_0 + \beta_{1i} X_i + u_i \qquad \text{(equation of interest)}$$

$$X_i = \pi_0 + \pi_{1i} Z_i + v_i \qquad \text{(first stage of TSLS)}$$

Intuition:

- If for some people $\pi_{1i} = 0$, then their predicted value of X_i wouldn't depend on Z, so the IV estimator would ignore them.
- The IV estimator puts most of the weight on individuals for whom Z has a large influence on X.
- TSLS measures the treatment effect for those whose probability of treatment is most influenced by X.

(b) IV with heterogeneity and random assignment

(3 of 3)

The math…

$$Y_i = \beta_0 + \beta_{1i} X_i + u_i \qquad \text{(equation of interest)}$$

$$X_i = \pi_0 + \pi_{1i} Z_i + v_i \qquad \text{(first stage of TSLS)}$$

To simplify things, suppose:

- Z_i is randomly assigned so is independent of $(\beta_{1i}, \pi_{1i}, u_i, v_i)$
  - Thus $E(u_i|Z_i) = 0$ and $E(v_i|Z_i) = 0$
- $E(\pi_{1i}) \neq 0$ (so that the instrument is relevant on average)

$$\text{Then } \hat{\beta}_1^{TSLS} \xrightarrow{p} \frac{E(\beta_{1i}\pi_{1i})}{E(\pi_{1i})} \text{ (derived in SW App. 13.2)}$$

- TSLS estimates the causal effect for those individuals for whom Z is most influential (those with large π_{1i}).

The Local Average Treatment Effect (LATE): (1 of 8)

$$\hat{\beta}_1^{TSLS} \xrightarrow{p} \frac{E(\beta_{1i}\pi_{1i})}{E(\pi_{1i})}$$

- TSLS estimates the causal effect for those individuals for whom Z is most influential (those with large π_{1i}).

- $\frac{E(\beta_{1i}\pi_{1i})}{E(\pi_{1i})}$, is called the ***local average treatment effect*** (LATE) – it is the average treatment effect for those in a "local" region who are most heavily influenced by Z.

- In general, LATE is neither the average treatment effect nor the effect of treatment on the treated

The Local Average Treatment Effect (LATE): (2 of 8)

$$\hat{\beta}_1^{TSLS} \xrightarrow{p} \frac{E(\beta_{1i}\pi_{1i})}{E(\pi_{1i})} = \text{LATE}$$

Recall the covariance fact,

$$E(\beta_{1i}\pi_{1i}) = E(\beta_{1i})\,E(\pi_{1i}) + \text{cov}(\beta_{1i}, \pi_{1i})$$

so $$\text{LATE} = \frac{E(\beta_{1i}\pi_{1i})}{E(\pi_{1i})} = \frac{E(\beta_{1i})E(\pi_{1i}) + \text{cov}(\beta_{1i}, \pi_{1i})}{E(\pi_{1i})}$$

$$= E(\beta_{1i}) + \frac{\text{cov}(\beta_{1i}, \pi_{1i})}{E(\pi_{1i})}$$

or

$$\boxed{\text{LATE} = \text{ATE} + \frac{\text{cov}(\beta_{1i}, \pi_{1i})}{E(\pi_{1i})}}$$

The Local Average Treatment Effect (LATE): (3 of 8)

$$\text{LATE} = \text{ATE} + \frac{\text{cov}(\beta_{1i}, \pi_{1i})}{E(\pi_{1i})}$$

- If the treatment effect is large for individuals for whom the effect of the instrument is also large, then $\text{cov}(\beta_{1i}, \pi_{1i}) > 0$ and LATE > ATE (if $E(\pi_{1i}) > 0$).
- In the binary case, LATE is the treatment effect for those whose probability of receipt of treatment is most heavily influenced by Z.
  - If you always (or never) get treated, you don't show up in limit of the IV estimator (in LATE).
  - If you never take the treatment, or always do, regardless of Z, it is like you aren't in the experiment in the first place.

The Local Average Treatment Effect (LATE): (4 of 8)

When there are heterogeneous causal effects, what TSLS estimates depends on the choice of instruments!

- With different instruments, TSLS estimates different weighted averages!
- Suppose you have two **valid** instruments, Z_1 and Z_2.
  - In general these instruments will be influential for different members of the population.
  - Using Z_1, TSLS will estimate the treatment effect for those people whose probability of treatment (X) is most influenced by Z_1
  - The LATE for Z_1 might differ from the LATE for Z_2
  - If so, the J-statistic will tend to reject even if both Z_1 and Z_2 are exogenous! (*Why?*)

When does TSLS estimate the average causal effect?

$Y_i = \beta_0 + \beta_{1i} X_i + u_i$ (equation of interest)

$X_i = \pi_0 + \pi_{1i} Z_i + v_i$ (first stage of TSLS)

$$\hat{\beta}_1^{TSLS} \xrightarrow{p} \text{LATE} = \text{ATE} + \frac{\text{cov}(\beta_{1i}, \pi_{1i})}{E(\pi_{1i})}$$

- LATE = ATE (a.k.a. average causal effect), that is $\hat{\beta}_1^{TSLS} \xrightarrow{p} E(\beta_{1i})$, if at least one of these three conditions holds:
  a) If $\beta_{1i} = \beta_1$ (no heterogeneity in equation of interest); **or**
  b) If $\pi_{1i} = \pi_1$ (no heterogeneity in first stage equation); **or**
  c) If β_{1i} and π_{1i} vary but are independently distributed.

- Otherwise, $\hat{\beta}_1^{TSLS}$ does *not* estimate $E(\beta_{1i})$

- Whether this is important depends on the application…

The Local Average Treatment Effect (LATE): (5 of 8)

Example #1: Cigarette elasticity

Y_i = Cigarette consumption in state i

X_i = Price per pack in state i

Z_i = cigarette tax in state i

β_{1i} = price elasticity in state i

π_{1i} = effect of cigarette tax on price in state i

LATE = the average causal effect if:

a) If $\beta_{1i} = \beta_1$ (no heterogeneity in equation of interest); **or**

b) If $\pi_{1i} = \pi_1$ (no heterogeneity in first stage equation); **or**

c) If β_{1i} and π_{1i} vary but are independently distributed.

How close is LATE to the average causal effect in this example?

The Local Average Treatment Effect (LATE): (6 of 8)

Example #2: Cardiac catheterization

Y_i = survival time (days) for AMI patients

X_i = received cardiac catheterization (or not)

Z_i = differential distance to CC hospital

Equation of interest:

$$SurvivalDays_i = \beta_0 + \beta_{1i} CardCath_i + u_i$$

First stage (*linear probability model*):

$$CardCath_i = \pi_0 + \pi_{1i} Distance_i + v_i$$

- For whom does distance have the great effect on the probability of treatment?
- For those patients, what is *their* causal effect β_{1i}?

The Local Average Treatment Effect (LATE): (7 of 8)

Equation of interest:

$$SurvivalDays_i = \beta_{0i} + \beta_{1i} CardCath_i + u_i$$

First stage (*linear probability model*):

$$CardCath_i = \pi_{0i} + \pi_{1i} Distance_i + v_i$$

LATE = the average causal effect if:

a) If $\beta_{1i} = \beta_1$ (no heterogeneity in equation of interest); **or**

b) If $\pi_{1i} = \pi_1$ (no heterogeneity in first stage equation); **or**

c) If β_{1i} and π_{1i} vary but are independently distributed.

How close is LATE to the average causal effect in this example?

The Local Average Treatment Effect (LATE): (8 of 8)

- TSLS estimates the causal effect for those whose value of X_i is most heavily influenced by Z_i
- TSLS estimates the causal effect for those for whom distance most influences the probability of treatment: What is *their* causal effect?
  - If in the expert judgment of the EMT, CC wouldn't have substantial benefits, relative to the cost of making a longer trip, then they should just go to the closest hospital. These are the patients who are most heavily influenced by relative distance, so it is their LATE that is being estimated.
- This is a plausible explanation of why the TSLS estimate is smaller than the clinical trial OLS estimate.

Heterogeneous Causal Effects: Summary

- Heterogeneous causal effects means that the causal (or treatment) effect varies across individuals.
- The average treatment effect is the average value in the population, $E(\beta_{1i})$.
- When these differences depend on observable variables, heterogeneous causal effects can be estimated using interactions (nothing new here).
- When differences in β_{1i} are unobservable, then the behavior of OLS and IV can change.
  - If X_i is randomly assigned, then OLS is consistent.
  - If Z_i is (as-if) randomly assigned, then IV estimates the LATE, which depends on the instrument

Summary: Experiments and Quasi-Experiments (SW Section 13.7) (1 of 4)

Ideal experiments and potential outcomes

- The average treatment effect is the population mean of the individual treatment effect, which is the difference in potential outcomes when treated and not treated.
- The treatment effect estimated in an ideal randomized controlled experiment is unbiased for the average treatment effect.

Summary: Experiments and Quasi-Experiments (SW Section 13.7) (2 of 4)

Actual experiments

- Actual experiments have threats to internal validity
- Depending on the threat, these threats to internal validity can be addressed by:
  - panel data regression (differences-in-differences)
  - multiple regression (including control variables), and
  - IV (using initial assignment as an instrument, possibly with control variables)
- External validity also can be an important threat to the validity of experiments

Summary: Experiments and Quasi-Experiments (SW Section 13.7) (3 of 4)

Quasi-experiments

- Quasi-experiments have an "as-if" randomly assigned source of variation.
- This as-if random variation can generate:
  - X_i which plausibly satisfies $E(u_i|X_i) = 0$ (so estimation proceeds using OLS); or
  - instrumental variable(s) which plausibly satisfy $E(u_i|Z_i) = 0$ (so estimation proceeds using TSLS)
- Quasi-experiments also have threats to internal validity

Summary: Experiments and Quasi-Experiments (SW Section 13.7) (4 of 4)

Heterogeneous treatment effects

- Heterogeneous treatment effects refers to variation in individual causal effects that are unrelated to observables
  - Variation in treatment effects that depends on observable variables can be handled using interactions (e.g., one treatment effect for men, another for women)
- If X_i is randomly assigned, then OLS estimates the average causal effect (ATE).
- If $E(u_i|Z_i) = 0$, then TSLS estimates the local average treatment effect (LATE), which is the average effect of treatment for those most influenced by Z_i.