Experiments and Quasi-Experiments
(SW Chapter 13)

Why study experiments?

- Ideal randomized controlled experiments provide a benchmark for assessing observational studies.
- Actual experiments are rare ($$$) but influential.
- Experiments can solve the threats to internal validity of observational studies, but they have their own threats to internal validity.
- Thinking about experiments helps us to understand quasi-experiments, or "natural experiments," in which there some variation is "as if" randomly assigned.

Terminology: experiments and quasi-experiments

- An ***experiment*** is designed and implemented consciously by human researchers. An experiment entails conscious use of a treatment and control group with random assignment (e.g. clinical trials of a drug)
- A ***quasi-experiment*** or ***natural experiment*** has a source of randomization that is "as if" randomly assigned, but this variation was not part of a conscious 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.

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)

Our treatment of experiments: brief outline

- Why (precisely) do ideal randomized controlled experiments provide estimates of causal effects?
- What are the main threats to the validity (internal and external) of actual experiments – that is, experiments actually conducted with human subjects?
- Flaws in actual experiments can result in X and u being correlated (threats to internal validity).
- Some of these threats can be addressed using the regression estimation methods we have used so far: multiple regression, panel data, IV regression.

Idealized Experiments and Causal Effects
(SW Section 13.1)

- An ideal randomized controlled experiment randomly assigns subjects to treatment and control groups.
- More generally, the treatment level *X* is randomly assigned:

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

- If *X* is randomly assigned (for example by computer) then *u* and *X* are independently distributed and $E(u_i|X_i) = 0$, 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

Estimation of causal effects in an ideal randomized controlled experiment

- Random assignment of X implies that $E(u_i|X_i) = 0$.
- Thus the OLS estimator $\hat{\beta}_1$ is unbiased.
- When the treatment is binary, $\hat{\beta}_1$ is just the difference in mean outcome (Y) in the treatment vs. control group ($\overline{Y}^{treated} - \overline{Y}^{control}$).
- This differences in means is sometimes called the ***differences estimator*.**

Potential Problems with Experiments in Practice
(SW Section 13.2)

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 internal validity, ctd.

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

- some controls get the treatment
- some "treated" get controls
- "errors-in-variables" bias: corr(X,u) $\neq 0$
- Attrition (some subjects drop out)
- suppose the controls who get jobs move out of town; then corr(X,u) $\neq 0$

Threats to internal validity, ctd.

3. *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) $\neq 0$

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.

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)
4. Treatment v. eligibility effects (which is it you want to measure: effect on those who take the program, or the effect on those are eligible)

Regression Estimators of Causal Effects Using Experimental Data
(SW Section 13.3)

- Focus on the case that X is binary (treatment/control).
- Often you observe subject characteristics, $W_{1i},\ldots,W_{ri}$.
- Extensions of the differences estimator:
  - can improve efficiency (reduce standard errors)
  - can eliminate bias that arises when:
    - treatment and control groups differ
    - there is "conditional randomization"
    - there is partial compliance
- These extensions involve methods we have already seen – multiple regression, panel data, IV regression

Estimators of the Treatment Effect β_1 using Experimental Data (X = 1 if treated, 0 if control)

	Dep. vble	Ind. vble(s)	method	
differences	Y	X	OLS	
differences-in-differences	$\Delta Y = Y^{after} - Y^{before}$	X	OLS	adjusts for initial differences between treatment and control groups
differences with add'l regressors	Y	$X, W_1, \ldots, W_n$	OLS	controls for additional subject characteristics W

Estimators with experimental data, ctd.

	Dep. vble	**Ind. vble(s)**	**method**	
differences-in-differences with add'l regressors	$\Delta Y = Y^{after} - Y^{before}$	$X, W_1, \ldots, W_n$	OLS	adjusts for group differences + controls for subject char's W
Instrumental variables	Y	X	TSLS	Z = *initial* random assignment; eliminates bias from partial compliance

- TSLS with Z = initial random assignment also can be applied to the differences-in-differences estimator and the estimators with additional regressors (W's)

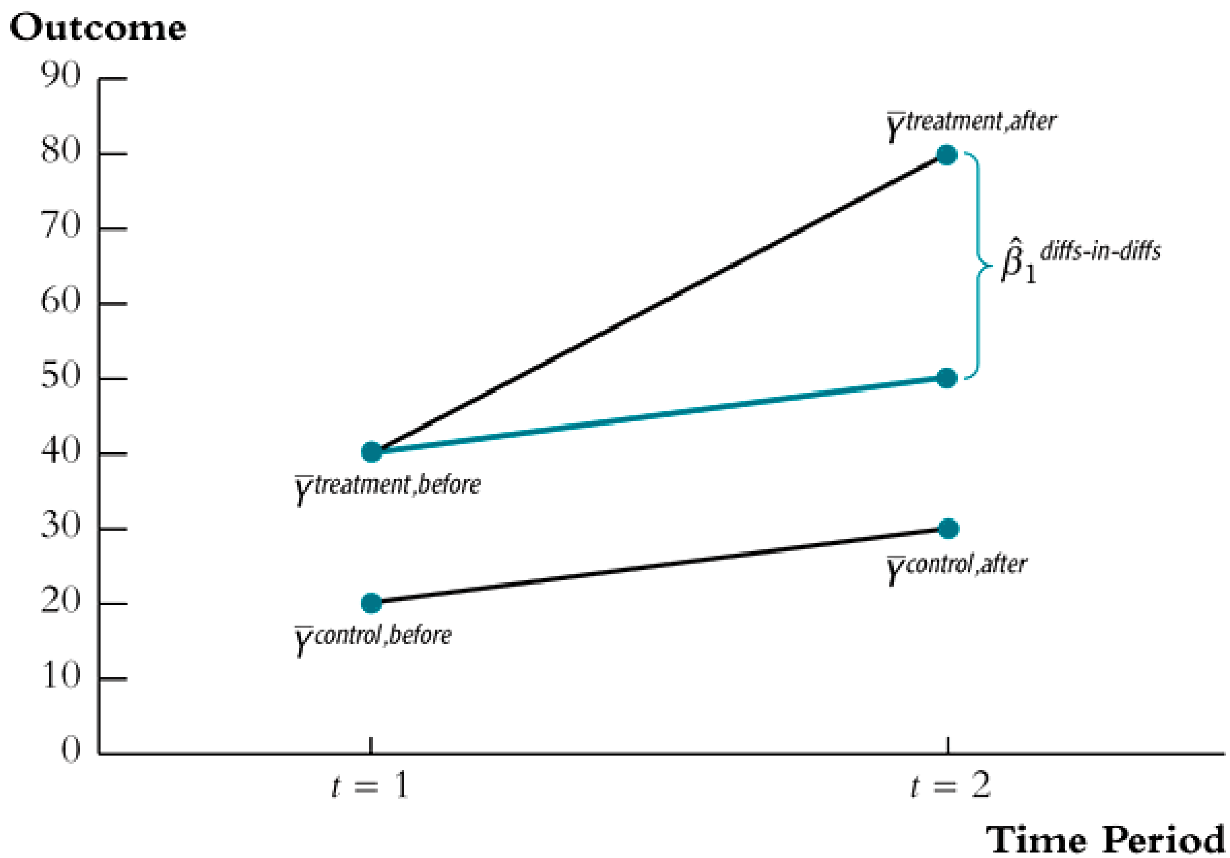
The differences-in-differences estimator

- Suppose the treatment and control groups differ systematically; maybe the control group is healthier (wealthier; better educated; etc.)
- Then X is correlated with u, and the differences estimator is biased.
- The differences-in-differences estimator adjusts for pre-experimental differences by subtracting off each subject's pre-experimental value of Y
  - Y_i^{before} = value of Y for subject i before the expt
  - Y_i^{after} = value of Y for subject i after the expt
  - $\Delta Y_i = Y_i^{after} - Y_i^{before}$ = *change* over course of expt

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

FIGURE 11.1 The Differences-in-Differences Estimator

The post-treatment difference between the treatment and control groups is 80 – 30 = 50, but this overstates the treatment effect because before the treatment $\bar{Y}$ was higher for the treatment than the control group by 40 – 20 = 20. The differences-in-differences estimator is the difference between the final and initial gaps, so that $\hat{\beta}_1^{diffs\text{-}in\text{-}diffs} =$ (80 – 30) – (40 – 20) = 50 – 20 = 30. Equivalently, the differences-in-differences estimator is the average change for the treatment group minus the average change for the control group, that is, $\hat{\beta}_1^{diffs\text{-}in\text{-}diffs} = \Delta\bar{Y}^{treatment} - \Delta\bar{Y}^{control} =$ (80 – 40) – (30 – 20) = 30.



The differences-in-differences estimator, ctd.

(1) "Differences" formulation:

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

where

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

$X_i = 1$ if treated, $= 0$ otherwise

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

The differences-in-differences estimator, ctd.

(2) Equivalent "panel data" version:

$$Y_{it} = \gamma_0 + \beta_1 X_{it} + \gamma_2 D_{it} + \gamma_3 G_{it} + v_{it},\ i = 1,\ldots,n$$

where

$t = 1$ (before experiment), 2 (after experiment)

$D_{it} = 0$ for $t = 1$, $= 1$ for $t = 2$

$G_{it} = 0$ for control group, $= 1$ for treatment group

$X_{it} = 1$ if treated, $= 0$ otherwise

$= D_{it} \times G_{it}$ = interaction effect of being in treatment group in the second period

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

Including additional subject characteristics (*W*'s)

- Typically you observe additional subject characteristics, $W_{1i},\ldots,W_{ri}$
- Differences estimator with add'l regressors:

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 W_{1i} + \ldots + \beta_{r+1} W_{ri} + u_i$$

- Differences-in-differences estimator with *W*'s:

$$\Delta Y_i = \beta_0 + \beta_1 X_i + \beta_2 W_{1i} + \ldots + \beta_{r+1} W_{ri} + u_i$$

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

Why include additional subject characteristics (*W*'s)?

1. *Efficiency*: more precise estimator of β_1 (smaller standard errors)
2. *Check for randomization*. If *X* is randomly assigned, then the OLS estimators with and without the *W*'s should be similar – if they aren't, this suggests that *X* wasn't randomly designed (a problem with the expt.)
    - *Note*: To check directly for randomization, regress *X* on the *W*'s and do a *F*-test.
3. *Adjust for conditional randomization* (*we'll return to this later…*)

Estimation when there is partial compliance

Consider diffs-in-diffs estimator, X = actual treatment

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

- Suppose there is partial compliance: some of the treated don't take the drug; some of the controls go to job training anyway
- Then X is correlated with u, and OLS is biased
- Suppose initial assignment, Z, is random
- Then (1) $\text{corr}(Z,X) \neq 0$ and (2) $\text{corr}(Z,u) = 0$
- Thus β_1 can be estimated by TSLS, with instrumental variable Z = initial assignment
- This can be extended to W's (included exog. variables)

Experimental Estimates of the Effect of Reduction: The Tennessee Class Size Experiment (SW Section 13.4)

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

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)

TABLE 11.1 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** (2.45)	29.78** (2.83)	19.39** (2.71)	15.59** (2.40)
Regular size with aide	0.31 (2.27)	11.96** (2.65)	3.48 (2.54)	−0.29 (2.27)
Intercept	918.04** (1.63)	1,039.39** (1.78)	1,157.81** (1.82)	1,228.51** (1.68)
Number of observations	5,786	6,379	6,049	5,967

The regressions were estimated using the Project STAR Public Access Data Set described in Appendix 11.1. The dependent variable is the student's combined score on the math and reading portions of the Stanford Achievement Test. Standard errors are given in parentheses under the coefficients. **The individual coefficient is statistically significant at the 1% significance level using a two-sided test.

TABLE 11.2 Project STAR: Differences Estimates with Additional Regressors for Kindergarten

Regressor	(1)	(2)	(3)	(4)
Small class	13.90** (2.45)	14.00** (2.45)	15.93** (2.24)	15.89** (2.16)
Regular size with aide	0.31 (2.27)	−0.60 (2.25)	1.22 (2.04)	1.79 (1.96)
Teacher's years of experience		1.47** (0.17)	0.74** (0.17)	0.66** (0.17)
Boy				−12.09** (1.67)
Free lunch eligible				−34.70** (1.99)
Black				−25.43** (3.50)
Race other than black or white				−8.50 (12.52)
Intercept	918.04** (1.63)	904.72** (2.22)		
School indicator variables?	no	no	yes	yes
$\bar{R}^2$	0.01	0.02	0.22	0.28
Number of observations	5,786	5,766	5,766	5,748

The regressions were estimated using the Project STAR Public Access Data Set described in Appendix 11.1. The dependent variable is the combined test score on the math and reading portions of the Stanford Achievement Test. The number of observations differ in the different regressions because of some missing data. Standard errors are given in parentheses under coefficients. The individual coefficient is statistically significant at the *5% level or **1% significance level using a two-sided test.

How big are these estimated effects?

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

TABLE 11.3 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** (0.03)	0.33** (0.03)	0.23** (0.03)	0.21** (0.03)
Regular size with aide	0.00 (0.03)	0.13** (0.03)	0.04 (0.03)	0.00 (0.03)
Sample standard deviation of test scores (s_Y)	73.70	91.30	84.10	73.30

The estimates and standard errors in the first two rows are the estimated effects in Table 11.1, 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 are given in parentheses under coefficients. **The individual coefficient is statistically significant at the 1% significance level using a two-sided test.

How do these estimates compare to those from the California, Mass. observational studies? (Ch. 4 – 7)

TABLE 11.4 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)

The estimated coefficient $\hat{\beta}_1$ for the STAR study is taken from column (1) of Table 11.2. The estimated coefficients for the California and Massachusetts studies are taken from the first column of Table 7.3. 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% level or **1% significance level using a two-sided test.

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

What is the Difference Between a Control Variable and the Variable of Interest? (SW App. 13.3)

Example: "free lunch eligible" in the STAR regressions

- Coefficient is large, negative, statistically significant
- Policy interpretation: *Making students ineligible for a free school lunch will improve their test scores.*
- Is this really an estimate of a causal effect?
- Is the OLS estimator of its coefficient unbiased?
- Can it be that the coefficient on "free lunch eligible" is biased but the coefficient on *SmallClass* is not?

TABLE 11.2 Project STAR: Differences Estimates with Additional Regressors for Kindergarten

Regressor	(1)	(2)	(3)	(4)
Small class	13.90** (2.45)	14.00** (2.45)	15.93** (2.24)	15.89** (2.16)
Regular size with aide	0.31 (2.27)	−0.60 (2.25)	1.22 (2.04)	1.79 (1.96)
Teacher's years of experience		1.47** (0.17)	0.74** (0.17)	0.66** (0.17)
Boy				−12.09** (1.67)
Free lunch eligible				−34.70** (1.99)
Black				−25.43** (3.50)
Race other than black or white				−8.50 (12.52)
Intercept	918.04** (1.63)	904.72** (2.22)		
School indicator variables?	no	no	yes	yes
$\overline{R}^2$	0.01	0.02	0.22	0.28
Number of observations	5,786	5,766	5,766	5,748

The regressions were estimated using the Project STAR Public Access Data Set described in Appendix 11.1. The dependent variable is the combined test score on the math and reading portions of the Stanford Achievement Test. The number of observations differ in the different regressions because of some missing data. Standard errors are given in parentheses under coefficients. The individual coefficient is statistically significant at the *5% level or **1% significance level using a two-sided test.

Example: "free lunch eligible," ctd.

- Coefficient on "free lunch eligible" is large, negative, statistically significant
- Policy interpretation: *Making students ineligible for a free school lunch will improve their test scores.*
- Why (precisely) can we interpret the coefficient on *SmallClass* as an unbiased estimate of a causal effect, but not the coefficient on "free lunch eligible"?
- This is not an isolated example!
  - Other "control variables" we have used: gender, race, district income, state fixed effects, time fixed effects, city (or state) population,…
- What is a "control variable" anyway?

Simplest case: one *X*, one control variable *W*

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

For example,

- *W* = free lunch eligible (binary)
- *X* = small class/large class (binary)
- Suppose random assignment of *X* depends on *W*
  - for example, 60% of free-lunch eligibles get small class, 40% of ineligibles get small class)
  - note: *this wasn't the actual STAR randomization procedure* – this is a hypothetical example
- Further suppose *W* is correlated with *u*

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

Suppose:

- The control variable W is correlated with u
- Given $W = 0$ (ineligible), X is randomly assigned
- Given $W = 1$ (eligible), X is randomly assigned.

Then:

- Given the value of W, X is randomly assigned;
- That is, controlling for W, X is randomly assigned;
- Thus, controlling for W, X is uncorrelated with u
- Moreover, $E(u|X,W)$ doesn't depend on X
- That is, we have ***conditional mean independence:***

$$E(u|X,W) = E(u|W)$$

Implications of conditional mean independence

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

Suppose $E(u|W)$ is linear in W (*not restrictive – could add quadratics etc.*): then,

$$E(u|X,W) = E(u|W) = \gamma_0 + \gamma_1 W_i \qquad (*)$$

so

$$\begin{aligned} E(Y_i|X_i,W_i) &= E(\beta_0 + \beta_1 X_i + \beta_2 W_i + u_i|X_i,W_i) \\ &= \beta_0 + \beta_1 X_i + \beta_2 W_i + E(u_i|X_i,W_i) \\ &= \beta_0 + \beta_1 X_i + \beta_2 W_i + \gamma_0 + \gamma_1 W_i \qquad \text{by } (*) \\ &= (\gamma_0+\beta_0) + \beta_1 X_i + (\gamma_1+\beta_2) W_i \end{aligned}$$

Implications of conditional mean independence:

- The conditional mean of Y given X and W is

$$E(Y_i|X_i,W_i) = (\gamma_0+\beta_0) + \beta_1 X_i + (\gamma_1+\beta_2)W_i$$

- The effect of a change in X under conditional mean independence is the desired causal effect:

$$E(Y_i|X_i = x+\Delta x,W_i) - E(Y_i|X_i = x,W_i) = \beta_1\Delta x$$

or

$$\beta_1 = \frac{E(Y_i \mid X_i = x+\Delta x, W_i) - E(Y_i \mid X_i = x, W_i)}{\Delta x}$$

- If X is binary (treatment/control), this becomes:

$$\beta_1 = \frac{E(Y_i \mid X_i = 1, W_i) - E(Y_i \mid X_i = 0, W_i)}{\Delta x}$$

which is the desired treatment effect.

Implications of conditional mean independence, ctd.

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

Conditional mean independence says:

$$E(u|X,W) = E(u|W)$$

which, with linearity, implies:

$$E(Y_i|X_i,W_i) = (\gamma_0+\beta_0) + \beta_1 X_i + (\gamma_1+\beta_2)W_i$$

Then:

- The OLS estimator $\hat{\beta}_1$ is unbiased.
- $\hat{\beta}_2$ is not consistent and not meaningful
- The usual inference methods (standard errors, hypothesis tests, etc.) apply to $\hat{\beta}_1$.

So, what is a control variable?

A ***control variable*** W is a variable that results in X satisfying the conditional mean independence condition:

$$E(u|X,W) = E(u|W)$$

- Upon including a control variable in the regression, X ceases to be correlated with the error term.
- The control variable itself can be (in general will be) correlated with the error term.
- The coefficient on X has a causal interpretation.
- The coefficient on W does *not* have a causal interpretation.

Example: Effect of teacher experience on test scores

More on the design of Project STAR:

- Teachers didn't change school because of the expt.
- Within their normal school, teachers were randomly assigned to small/regular/reg+aide classrooms.
- What is the effect of X = years of teacher education?

The design implies ***conditional*** *mean independence*:

- W = school binary indicator
- Given W (school), X is randomly assigned
- That is, $E(u|X,W) = E(u|W)$
- W is plausibly correlated with u (nonzero school fixed effects: some schools are better/richer/etc than others)

TABLE 11.2 Project STAR: Differences Estimates with Additional Regressors for Kindergarten

Regressor	(1)	(2)	(3)	(4)
Small class	13.90** (2.45)	14.00** (2.45)	15.93** (2.24)	15.89** (2.16)
Regular size with aide	0.31 (2.27)	−0.60 (2.25)	1.22 (2.04)	1.79 (1.96)
Teacher's years of experience		1.47** (0.17)	0.74** (0.17)	0.66** (0.17)
Boy				−12.09** (1.67)
Free lunch eligible				−34.70** (1.99)
Black				−25.43** (3.50)
Race other than black or white				−8.50 (12.52)
Intercept	918.04** (1.63)	904.72** (2.22)		
School indicator variables?	no	no	yes	yes
$\overline{R}^2$	0.01	0.02	0.22	0.28
Number of observations	5,786	5,766	5,766	5,748

The regressions were estimated using the Project STAR Public Access Data Set described in Appendix 11.1. The dependent variable is the combined test score on the math and reading portions of the Stanford Achievement Test. The number of observations differ in the different regressions because of some missing data. Standard errors are given in parentheses under coefficients. The individual coefficient is statistically significant at the *5% level or **1% significance level using a two-sided test.

Example: teacher experience, ctd.

- Without school fixed effects (2), the estimated effect of an additional year of experience is 1.47 (*SE* = .17)
- "*Controlling for the school*" (3), the estimated effect of an additional year of experience is .74 (*SE* = .17)
- Direction of bias makes sense:
  - less experienced teachers at worse schools
  - years of experience picks up this school effect
- OLS estimator of coefficient on years of experience is biased up without school effects; with school effects, OLS yields unbiased estimator of causal effect
- School effect coefficients don't have a causal interpretation (effect of student changing schools)

Quasi-Experiments
(SW Section 13.5)

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

Two cases:

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

(b) A variable (Z) that influences treatment (X) is "as if" randomly assigned (IV)

Two types of quasi-experiments

(a) Treatment (*X*) is "as if" randomly assigned (perhaps conditional on some control variables *W*)

- *Ex*: Effect of marginal tax rates on labor supply
  - *X* = marginal tax rate (rate changes in one state, not another; state is "as if" randomly assigned)

(b) A variable (*Z*) that influences treatment (*X*) is "as if" randomly assigned (IV)

- Effect on survival of cardiac catheterization

  X = cardiac catheterization;

  Z = differential distance to CC hospital

Econometric methods

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

Diffs-in-diffs estimator using panel data methods:

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \beta_2 D_{it} + \beta_3 G_{it} + u_{it},\ i = 1,\ldots,n$$

where

$t = 1$ (before experiment), 2 (after experiment)

$D_{it} = 0$ for $t = 1$, $= 1$ for $t = 2$

$G_{it} = 0$ for control group, $= 1$ for treatment group

$X_{it} = 1$ if treated, $= 0$ otherwise

$= D_{it} \times G_{it}$ = interaction effect of being in treatment group in the second period

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

The panel data diffs-in-diffs estimator simplifies to the "changes" diffs-in-diffs estimator when T = 2

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \beta_2 D_{it} + \beta_3 G_{it} + u_{it},\ i = 1,\ldots,n \quad (*)$$

For $t = 1$: $D_{i1} = 0$ and $X_{i1} = 0$ (nobody treated), so

$$Y_{i1} = \beta_0 + \beta_3 G_{i1} + u_{i1}$$

For $t = 2$: $D_{i2} = 1$ and $X_{i2} = 1$ if treated, = 0 if not, so

$$Y_{i2} = \beta_0 + \beta_1 X_{i2} + \beta_2 + \beta_3 G_{i2} + u_{i2}$$

so

$$\Delta Y_i = Y_{i2} – Y_{i1} = (\beta_0+\beta_1 X_{i2}+\beta_2+\beta_3 G_{i2}+u_{i2}) – (\beta_0+\beta_3 G_{i1}+u_{i1})$$

$$= \beta_1 X_i + \beta_2 + (u_{i1} – u_{i2}) \qquad \text{(since } G_{i1} = G_{i2}\text{)}$$

or

$$\Delta Y_i = \beta_2 + \beta_1 X_i + v_i, \text{ where } v_i = u_{i1} – u_{i2} \quad (**)$$

Differences-in-differences with control variables

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \beta_2 D_{it} + \beta_3 G_{it} + \beta_4 W_{1it} + \ldots + \beta_{3+r} W_{rit} + u_{it},$$

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

= $G_{it} \times D_{it}$ (= 1 for treatment group in second period)

- If the treatment (X) is "as if" randomly assigned, given W, then u is conditionally mean indep. of X:

$$E(u|X,D,G,W) = E(u|D,G,W)$$

- OLS is a consistent estimator of β_1, the causal effect of a change in X
- In general, the OLS estimators of the other coefficients do *not* have a causal interpretation.

(b) A variable (Z) that influences treatment (X) is "as if" randomly assigned (IV)

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \beta_2 D_{it} + \beta_3 G_{it} + \beta_4 W_{1it} + \ldots + \beta_{3+r} W_{rit} + u_{it},$$

X_{it} = 1 if the treatment is received, = 0 otherwise
= $G_{it} \times D_{it}$ (= 1 for treatment group in second period)

Z_{it} = variable that influences treatment but is uncorrelated with u_{it} (given W's)

TSLS:

- X = endogenous regressor
- $D, G, W_1, \ldots, W_r$ = included exogenous variables
- Z = instrumental variable

Potential Threats to Quasi-Experiments
(SW Section 13.6)

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

4. ***Failure to randomize*** (imperfect randomization)
   Is the "as if" randomization really random, so that X (or Z) is uncorrelated with u?
5. ***Failure to follow treatment protocol & attrition***
6. ***Experimental effects*** (not applicable)
7. ***Instrument invalidity*** (relevance + exogeneity)
   (Maybe healthier patients *do* live closer to CC hospitals –they might have better access to care in general)

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

5. Nonrepresentative sample
6. 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 that study used data from the early 90's – do its findings apply to CC usage today?

Experimental and Quasi-Experiments Estimates in Heterogeneous Populations (SW Section 13.7)

- We have discussed "the" treatment effect
- But the treatment effect could vary across individuals:
  - Effect of job training program probably depends on education, years of education, etc.
  - Effect of a cholesterol-lowering drug could depend other health factors (smoking, age, diabetes,…)
- If this variation depends on observed variables, then this is a job for interaction variables!
- But what if the source of variation is unobserved?

Heterogeneity of causal effects

When the causal effect (treatment effect) varies among individuals, the population is said to be ***heterogeneous***.

When there are heterogeneous causal effects that are not linked to an observed variable:

- What do we *want* to estimate?
  - Often, the average causal effect in the population
  - But there are other choices, for example the average causal effect for those who participate (effect of treatment on the treated)
- What do we *actually* estimate?
  - using OLS? using TSLS?

Population regression model with heterogeneous causal effects:

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

- β_{1i} is the causal effect (treatment effect) for the i^{th} individual in the sample
- For example, in the JTPA experiment, β_{1i} could be zero if person *i* already has good job search skills
- What do we want to estimate?
    - effect of the program on a randomly selected person (the "*average causal effect*") – our main focus
    - effect on those most (least?) benefited
    - effect on those who choose to go into the program?

The Average Causal Effect

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

- The ***average causal effect*** (or ***average treatment effect***) is the mean value of β_{1i} in the population.
- We can think of β_1 as a random variable: it has a distribution in the population, and drawing a different person yields a different value of β_1 (just like X and Y)
- For example, for person #34 the treatment effect is not random – it is her true treatment effect – but before she is selected at random from the population, her value of β_1 can be thought of as randomly distributed.

The average causal effect, ctd.

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

- The average causal effect is $E(\beta_1)$.
- What does OLS estimate:
  (a) When the conditional mean of u given X is zero?
  (b) Under the stronger assumption that X is randomly assigned (as in a randomized experiment)?
    In this case, OLS is a consistent estimator of the average causal effect.

OLS with Heterogeneous Causal Effects

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

(a) Suppose $E(u_i|X_i) = 0$ so $\text{cov}(u_i,X_i) = 0$.

- If X is binary (treated/untreated), $\hat{\beta}_1 = \bar{Y}^{treated} - \bar{Y}^{control}$ estimates the causal effect among those who receive the treatment.
- Why? For those treated, $\bar{Y}^{treated}$ reflects the effect of the treatment *on them*. But we don't know how the untreated would have responded had they been treated!

The math: suppose X is binary and $E(u_i|X_i) = 0$.

Then

$$\hat{\beta}_1 = \bar{Y}^{treated} - \bar{Y}^{control}$$

For the treated:

$$E(Y_i|X_i=1) = \beta_0 + E(\beta_{1i}X_i|X_i=1) + E(u_i|X_i=1)$$

$$= \beta_0 + E(\beta_{1i}|X_i=1)$$

For the controls:

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

$$= \beta_0$$

Thus:

$$\hat{\beta}_1 \xrightarrow{p} E(Y_i|X_i=1) - E(Y_i|X_i=0) = E(\beta_{1i}|X_i=1)$$

= average effect of the treatment **on the treated**

OLS with heterogeneous treatment effects: general *X* with $E(u_i|X_i) = 0$

$$\hat{\beta}_1 = \frac{s_{XY}}{s_X^2} \xrightarrow{p} \frac{\sigma_{XY}}{\sigma_X^2} = \frac{\text{cov}(\beta_0 + \beta_{1i}X_i + u_i, X_i)}{\text{var}(X_i)}$$

$$= \frac{\text{cov}(\beta_0, X_i) + \text{cov}(\beta_{1i}X_i, X_i) + \text{cov}(u_i, X_i)}{\text{var}(X_i)}$$

$$= \frac{\text{cov}(\beta_{1i}X_i, X_i)}{\text{var}(X_i)} \quad \text{(because cov}(u_i,X_i) = 0)$$

- If *X* is binary, this simplifies to the "effect of treatment on the treated"
- Without heterogeneity, $\beta_{1i} = \beta_1$ and $\hat{\beta}_1 \xrightarrow{p} \beta_1$
- In general, the treatment effects of individuals with large values of *X* are given the most weight

(b) Now make a stronger assumption: that X is randomly assigned (experiment or quasi-experiment). Then what does OLS actually estimate?

- X_i is randomly assigned, it is distributed independently of β_{1i}, so there is no difference between the population of controls and the population in the treatment group
- Thus the effect of treatment on the treated = the average treatment effect in the population.

The math:

$$\hat{\beta}_1 \xrightarrow{p} \frac{\operatorname{cov}(\beta_{1i}X_i, X_i)}{\operatorname{var}(X_i)} = E\left\{E\left[\frac{\operatorname{cov}(\beta_{1i}X_i, X_i)}{\operatorname{var}(X_i)} \mid \beta_{1i}\right]\right\}$$

$$= E\left[\beta_{1i}\frac{\operatorname{cov}(X_i, X_i)}{\operatorname{var}(X_i)}\right] = E\left[\beta_{1i}\frac{\operatorname{var}(X_i)}{\operatorname{var}(X_i)}\right]$$

$$= E(\beta_{1i})$$

Summary

- If X_i and β_{1i} are independent (X_i is randomly assigned), OLS estimates the average treatment effect.
- If X_i is not randomly assigned but $E(u_i|X_i) = 0$, OLS estimates the effect of treatment on the treated.
- *Without heterogeneity, the effect of treatment on the treated and the average treatment effect are the same*

IV Regression with Heterogeneous Causal Effects

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.

IV with heterogeneous causal effects, ctd.

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

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

Intuition:

- Suppose π_{1i}'s were known. If for some people π_{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.

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:

- β_{1i} and π_{1i} are distributed independently of (u_i,v_i,Z_i)
- $E(u_i|Z_i) = 0$ and $E(v_i|Z_i) = 0$
- $E(\pi_{1i}) \neq 0$

  Then $\hat{\beta}_1^{TSLS} \xrightarrow{p} \dfrac{E(\beta_{1i}\pi_{1i})}{E(\pi_{1i})}$ (derived in SW App. 11.4)

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

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 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 treatment effect for those most influenced by Z_1 might differ from the treatment effect for those most influenced by Z_2

When does TSLS estimate the average causal effect?

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

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

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

- TSLS estimates the average causal effect (that is, $\hat{\beta}_1^{TSLS} \xrightarrow{p} E(\beta_{1i})$) if:
  - If β_{1i} and π_{1i} are independent
  - If $\beta_{1i} = \beta_1$ (no heterogeneity in equation of interest)
  - If $\pi_{1i} = \pi_1$ (no heterogeneity in first stage equation)
- But in general $\hat{\beta}_1^{TSLS}$ does *not* estimate $E(\beta_{1i})$!

Example: 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}?

Equation of interest:

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

First stage (*linear probability model*):

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

- 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? ("We might as well go to the CC hospital, its not too much farther")
- This is one 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.
- When these differences depend on observable variables, heterogeneous causal effects can be estimated using interactions (nothing new here).
- When these differences are unobserved (β_{1i}) the average causal (or treatment) effect is the average value in the population, $E(\beta_{1i})$.
- When causal effects are heterogeneous, OLS and TSLS estimate….

OLS with Heterogeneous Causal Effects

***X* is:**	**Relation between X_i and u_i:**	**Then OLS estimates:**
binary	$E(u_i \mid X_i) = 0$	effect of treatment on the treated: $E(\beta_{1i} \mid X_i=1)$
	X randomly assigned (so X_i and u_i are independent)	average causal effect $E(\beta_{1i})$
general	$E(u_i \mid X_i) = 0$	weighted average of β_{1i}, placing most weight on those with large $\lvert X_i - \mu_X \rvert$
	X randomly assigned	average causal effect $E(\beta_{1i})$

Without heterogeneity, $\beta_{1i} = \beta_1$ and $\hat{\beta}_1 \xrightarrow{p} \beta_1$ in all these cases.

*T*SLS with Heterogeneous Causal Effects

- TSLS estimates the causal effect for those individuals for whom *Z* is most influential (those with large π_{1i}).
- What TSLS estimates depends on the choice of *Z*!!
- In CC example, these were the individuals for whom the decision to drive to a CC lab was heavily influenced by the extra distance (those patients for whom the EMT was otherwise "on the fence")
- Thus TSLS also estimates a causal effect: the average effect of treatment on those most influenced by the instrument
  - In general, this is neither the average causal effect nor the effect of treatment on the treated

Summary: Experiments and Quasi-Experiments
(SW Section 13.8)

Experiments:

- Average causal effects are defined as expected values of ideal randomized controlled experiments
- Actual experiments have threats to internal validity
- These threats to internal validity can be addressed (in part) by:
  - panel methods (differences-in-differences)
  - multiple regression
  - IV (using initial assignment as an instrument)

Summary, ctd.

Quasi-experiments:

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

Summary, ctd.

Two additional subtle issues:

- What is a control variable?
  - A variable *W* for which *X* and *u* are uncorrelated, given the value of *W* (conditional mean independence: $E(u_i|X_i,W_i) = E(u_i|W_i)$)
  - *Example*: STAR & effect of teacher experience
    - within their school, teachers were randomly assigned to regular/reg+aide/small class
    - OLS provides an unbiased estimator of the causal effect, but only after controlling for school effects.

Summary, ctd.

- What do OLS and TSLS estimate when there is unobserved heterogeneity of causal effects?
- In general, weighted averages of causal effects:
  - If X is randomly assigned, then OLS estimates the average causal effect.
  - If X_i is not randomly assigned but $E(u_i|X_i) = 0$, OLS estimates the average effect of treatment on the treated.
  - If $E(u_i|Z_i) = 0$, TSLS estimates the average effect of treatment on those most influenced by Z_i.