1. Potential outcomes and treatment effects

LPO 8852: Regression II

Sean P. Corcoran

What you learned in Regression I

The mechanics and properties of linear regression models:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + ... + \beta_k X_{ik} + u_i$$

- Model specification and interpretation
- Estimation (e.g., OLS, WLS)
- Inference: What is the *standard error* of your estimator? What is the estimator's *sampling distribution* in finite samples? In large samples? Knowledge of the sampling distribution is needed to construct *confidence intervals* and conduct *hypothesis tests*.

Model interpretation and statistical inference rely heavily on assumptions.

What you learned in Regression I

When I first learned econometrics, I often felt dissatisfied:

- Assumptions feel implausible
- How do we know the model is “correct”?
- There are always “omitted variables”!
- Causal interpretation feels like a pipe dream.

Regression II

Research designs for causal inference

- When can a regression be interpreted as causal?
- What does it mean for an estimator to have a causal interpretation?
- What research designs—which may or may not use regression—make a strong case for causal interpretation?

We will consider:

- Matching and weighting estimators
- Panel data models (e.g., fixed effects)
- Difference-in-differences
- Synthetic control methods
- Instrumental variables
- Regression discontinuity

What is a causal effect?

A **causal effect** is a change in some outcome (Y) that is the result of a change in some other (manipulable) factor (X).

For simplicity, assume the factor X is a binary "treatment." Example: the causal effect of taking an aspirin on headache pain, or the effect of getting a vaccine on contracting COVID-19.

Causal effects involve a **counterfactual** comparison between two different states of the world: e.g., Y whenever $X = 1$ versus Y whenever $X = 0$ (where all else is held constant).

Potential outcomes

The **potential outcomes framework** is useful for thinking about counterfactual comparisons and treatment effects. This approach is attributed to Neyman (1923) and Rubin, who later generalized the framework. It is often referred to as the **Neyman-Rubin causal model**.

Potential outcomes

Let D_i be a dichotomous indicator of a "treatment" where $D_i = 1$ means unit i is "treated" and $D_i = 0$ means i is "not treated." For every i there are two **potential outcomes**:

- $Y_i(1)$ or Y_{i1} = outcome when $D = 1$
- $Y_i(0)$ or Y_{i0} = outcome when $D = 0$

Note these are "all else equal" outcomes

These are called potential outcomes since units are not observed in more than one state at the same time. This is the *fundamental problem of causal inference* (Holland, 1986).

SUTVA

A common assumption invoked here is **SUTVA** (**stable unit treatment variable assignment**). What this says is that unit i's potential outcomes do not depend on the treatment assignment of other units. Cases in which this could be violated:

- Spillovers from treated to untreated (e.g., treatments for infectious disease, classroom peer effects, knowledge spillovers)
- "General equilibrium effects"

Violations of SUTVA create problems for what comes next. We'll ignore this possibility for now, but researchers should pay more attention to this.

Potential outcomes

The <u>observed</u> Y_i is either $Y_i(0)$ or $Y_i(1)$:

$$Y_i = D_i Y_i(1) + (1 - D_i) Y_i(0)$$

Call the above the **switching equation**.

A **counterfactual** is the outcome for the unit in the other (hypothetical, unobserved) state. E.g., the counterfactual for <u>treated</u> i would be $Y_i(0)$.

Example 1: job training program

Person	D_i	$Y_i(0)$	$Y_i(1)$	Y_i
1	1	10	14	14
2	1	8	12	12
3	1	12	16	16
4	1	8	12	12
5	1	6	10	10
6	1	4	8	8
7	0	4	8	4
8	0	6	10	6
9	0	8	12	8
10	0	4	8	4
11	0	10	14	10
12	0	8	12	8
13	0	2	6	2
14	0	1	5	1
Mean	0.429	6.5	10.5	8.2

Source: Jennifer Hill (2011) lecture notes. Assume Y_i is earnings and D_i indicates participation in job training program.

Treatment effects

The causal effect of D on Y for individual i (the **treatment effect**) is:

$$\tau_i = Y_i(1) - Y_i(0)$$

We can't know the τ_i for any individual, but we may be able to estimate an average of τ in some population, or some other information about the distribution of those τs.

This information is useful for predicting what the effect might be for some other i (e.g., for policy and practice decisions)

Treatment effects

We are often interested in the population **average treatment effect (ATE)**:

$$ATE = E(\tau) = E[\underbrace{Y(1) - Y(0)}_{\text{not observed}}]$$

Or the **average treatment effect on the treated (ATT)**:

$$ATT = E(\tau|D=1) = E[Y(1)|D=1] - \underbrace{E[Y(0)|D=1]}_{\text{not observed}}$$

Treatment effects

Or the **average treatment effect on the untreated (ATU)**:

$$ATU = E(\tau|D=0) = \underbrace{E[Y(1)|D=0]}_{\text{not observed}} - E[Y(0)|D=0]$$

The ATE, ATT, and ATU are **estimands**—quantities of interest in the population. Researchers are often most interested in ATT or ATE.

Note the ATE is a weighted average of the ATT and ATU:

$$ATE = pATT + (1-p)ATU$$

where p is the proportion treated.

Example 1: job training program

Person	D_i	$Y_i(0)$	$Y_i(1)$	Y_i
1	1	10	14	14
2	1	8	12	12
3	1	12	16	16
4	1	8	12	12
5	1	6	10	10
6	1	4	8	8
7	0	4	8	4
8	0	6	10	6
9	0	8	12	8
10	0	4	8	4
11	0	10	14	10
12	0	8	12	8
13	0	2	6	2
14	0	1	5	1
Mean	0.429	6.5	10.5	8.2

In Example 1 there are constant treatment effects:

$$ATE = ATT = ATU = 4$$

Estimating treatment effects

Suppose we compare the mean observed Y for two groups, $D = 1$ and $D = 0$ (a "naïve" estimator):

$$E[Y(1)|D=1] - E[Y(0)|D=0] = E[Y(1)|D=1] - E[Y(0)|D=0] \\ - \underbrace{E[Y(0)|D=1] + E[Y(0)|D=1]}_{0}$$

$$E[Y(1)|D=1] - E[Y(0)|D=0] = ATT + \underbrace{E[Y(0)|D=1] - E[Y(0)|D=0]}_{\text{selection bias}}$$

Selection bias reflects differences in $Y(0)$ between the treated and untreated group ("baseline differences" or "unobserved heterogeneity").

Example 1: job training program

Person	D_i	Educ.	Age	$Y(0)$	$Y(1)$	Y
1	1	1	26	10	14	14
2	1	1	21	8	12	12
3	1	1	30	12	16	16
4	1	1	19	8	12	12
5	1	0	25	6	10	10
6	1	0	22	4	8	8
Mean ($D = 1$)	1	0.67	23.8	8	12	12
7	0	0	21	4	8	4
8	0	0	26	6	10	6
9	0	0	28	8	12	8
10	0	0	20	4	8	4
11	0	1	26	10	14	10
12	0	1	21	8	12	8
13	0	0	16	2	6	2
14	0	0	15	1	5	1
Mean ($D = 0$)	0	0.25	21.6	5.4	9.4	5.4

Estimating treatment effects

In Example 1, $ATT = 4$. But:

$$E[Y(1)|D=1] - E[Y(0)|D=0] = ATT + \underbrace{E[Y(0)|D=1] - E[Y(0)|D=0]}_{\text{selection bias}}$$

$$12.0 - 5.4 = 4.0 + \underbrace{8.0 - 5.4}_{\text{selection bias}} = 6.6$$

The treated group has a higher $Y(0)$ than the untreated group. This could be due to their higher average education and age (shown in the table), two things associated with higher earnings. Their Y would have been higher on average even in the absence of treatment.

Estimating treatment effects

Think of $Y_i(0)$ as shorthand for everything about unit i other than their treatment status. Comparing mean covariates can be revealing about differences in the treated and untreated groups.

Also, "when observed differences proliferate, so should our suspicions about unobserved differences" (*Mastering Metrics*).

Estimating treatment effects

Note the "naïve" estimator also generally fails to recover the ATE:

$$E[Y(1)|D=1] - E[Y(0)|D=0] = ATE + \underbrace{E[Y(0)|D=1] - E[Y(0)|D=0]}_{\text{selection bias}} + \underbrace{(1-p)(ATT - ATU)}_{\text{heterogeneous treatment effect bias}}$$

See *Mixtape* Potential Outcomes chapter for the algebra. In Example 1, ATT=ATU (constant treatment effect), so there is no heterogeneous treatment effect bias. (There is, however, selection bias).

Heterogeneous treatment effects

In Example 1, $ATT = ATU = ATE$. In practice, ATT and ATU often differ from the ATE because units endogenously sort into treatments based on gains they expect from it. We might expect $ATT > ATE > ATU$.

Conditional and marginal treatment effects

Another treatment effect that is often of interest is a **conditional treatment effect** (ATE, ATT or ATU). That is, the average treatment effect conditional on something else being true. In Example 1, we might be interested the ATE conditional on $Education = 0$.

$$ATE|X = E(\tau|X) = E[Y(1) - Y(0)|X]$$

The experimental ideal

Under what conditions will selection bias be zero? When treatment assignment is **independent** of potential outcomes ("strong ignorability"):

$$(Y_i(1), Y_i(0)) \perp\!\!\!\perp D$$

One case where this holds is **randomization** to treatment. Under random assignment, $E[Y_i(0)|D_i = 1] = E[Y_i(0)|D_i = 0]$. The $D = 0$ and $D = 1$ groups are random draws from the same population. The untreated $D = 0$ can "stand in" as a counterfactual for the treated $D = 1$.

Note under random assignment, there is no heterogeneous treatment effect bias ($ATT = ATU$). So the mean difference in outcomes between $D = 0$ and $D = 1$ should give us the ATE, ATT, and ATU.

Conditional independence assumption

In the absence of randomization, it may be that treatment assignment is independent of potential outcomes *conditional* on some X:

$$(Y_i(1), Y_i(0)) \perp\!\!\!\perp D | X$$

In other words, i's with the same X have the same distribution of $Y(1)$ and $Y(0)$.

This is the **conditional independence assumption** (or again, strong ignorability). A big assumption, but may not be unreasonable in some circumstances. We'll come back to this.

Regression and causality

What does this have to do with regression? We often use regression to estimate average treatment effects. Suppose we estimate the following simple regression with the hope of estimating the causal effect of D:

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

When will this regression have a causal interpretation?

When it describes differences in average potential outcomes for a reference population of interest.

Regression and causality

Let's express β_1 in terms of mean potential outcomes. In large samples, you already know β_1 will consistently estimate:

$$\beta_1 = E[Y_i|D_i = 1] - E[Y_i|D_i = 0]$$

Which is the same as:

$$\beta_1 = E[Y(1)|D = 1] - E[Y(0)|D = 0]$$

Is this a parameter we care about? Does it represent differences in average potential outcomes for a population of interest? Does it estimate a treatment effect for a population of interest?

Regression and causality

Earlier we saw:

$$E[Y(1)|D = 1] - E[Y(0)|D = 0] = ATT + \underbrace{E[Y(0)|D = 1] - E[Y(0)|D = 0]}_{\text{selection bias}}$$

and:

$$E[Y(1)|D = 1] - E[Y(0)|D = 0] = ATE + \underbrace{E[Y(0)|D = 1] - E[Y(0)|D = 0]}_{\text{selection bias}} + \underbrace{(1 - p)(ATT - ATU)}_{\text{heterogeneous treatment effect bias}}$$

So, <u>no</u>: β_1 will not generally give us a treatment effect we care about!

Regression and causality

We also saw that:

- If D_i is randomly assigned, this difference in population means corresponds to the ATE: $E[Y(1) - Y(0)]$ (and the ATT).
- Under this condition, the regression does reveal a difference in potential outcomes for a population of interest.
- Without random assignment this is not generally true.

The name of the game: under what condition(s) does your regression/ estimator/research design provide a treatment effect of interest? Do those conditions hold in your case? When is your treatment effect **identified**?

- Also referred to as **internal validity**.

Regression and causality

As another illustration, suppose there are constant treatment effects, so for every i, $Y_i(1) = Y_i(0) + \delta$. We don't observe potential outcomes, but rather the observed Y_i:

$$Y_i = D_i Y_i(1) + (1 - D_i) Y_i(0)$$

Plugging in for $Y_i(1)$, this can be written:

$$Y_i = \underbrace{E[Y(0)]}_{\beta_0} + \underbrace{\delta}_{\beta_1} D_i + \underbrace{Y_i(0) - E[Y(0)]}_{\text{residual}}$$

Note the residual is the deviation of $Y_i(0)$ from the population mean $Y(0)$. With random assignment, D_i is uncorrelated with this residual. If there is *selection bias*—e.g., the treated tend to have higher baseline outcomes—then there is omitted variables bias.

Regression and causality

Now continue with constant treatment effects (δ) but suppose that potential outcomes depend linearly on X_i (and random noise u_i):

$$Y_i(0) = \alpha_0 + \alpha_1 X_i + u_i$$
$$Y_i(1) = \alpha_0 + \alpha_1 X_i + \delta + u_i$$

and that there is selection into treatment, such that D_i and X_i are related:

$$X_i = \gamma_0 + \gamma_1 D_i + w_i$$

Regression and causality

The observed Y_i (using the switching equation) is:

$$\begin{aligned} Y_i &= D_i Y_i(1) + (1 - D_i) Y_i(0) \\ &= D_i(\alpha_0 + \alpha_1 X_i + \delta + u_i) + (1 - D_i)(\alpha_0 + \alpha_1 X_i + u_i) \\ &= \alpha_0 + \delta D_i + \alpha_1 X_i + u_i \end{aligned}$$

If we estimate a naïve simple regression, $\alpha_1 X_i$ is in the residual:

$$Y_i = \beta_0 + \beta_1 D_i + \underbrace{v_i}_{\alpha_1 X_1 + u_i}$$

This is not a problem if X_i is uncorrelated with D_i, but in this case it is. There is omitted variables bias. X is related to potential outcomes <u>and</u> the $D = 1$ and $D = 0$ groups differ in their mean X.

Regression and causality

If we plug in what we know about how X_i is related to D_i:

$$\begin{aligned} Y_i &= \alpha_0 + \delta D_i + \alpha_1(\gamma_0 + \gamma_1 D_i + w_i) + u_i \\ &= \alpha_0 + (\delta + \alpha_1\gamma_1)D_i + \alpha_1\gamma_0 + \alpha_1 w_i + u_i \end{aligned}$$

The slope coefficient is $\delta + \alpha_1\gamma_1$. The latter is the omitted variables bias.

Regression and causality

Another way to see this. We know our estimator of β_1 in the simple regression will provide:

$$\begin{aligned} \beta_1 &= E[Y_i|D_i = 1] - E[Y_i|D_i = 0] \\ &= \alpha_0 + \alpha_1 E[X_i|D_i = 1] + \delta - \alpha_0 - \alpha_1 E[X_i|D_i = 0] \\ &= \delta + \underbrace{\alpha_1(E[X_i|D_i = 1] - E[X_i|D_i = 0])}_{\text{selection bias}} \\ &= \delta + \underbrace{\alpha_1(\gamma_0 + \gamma_1 - \gamma_0)}_{\text{selection bias}} \\ &= \delta + \underbrace{\alpha_1\gamma_1}_{\text{selection bias}} \end{aligned}$$

Conditional independence assumption

This is a pretty simple case where estimating a regression that *conditions on* (controls for) X would eliminate the selection bias. Here, the only reason treated and untreated units differ in their potential outcomes (on average) is that they have different levels of X.

$$Y_i = \beta_0 + \beta_1 D_i + \alpha_1 X_i + u_i$$

The conditional independence assumption holds here. Holding X constant, there is no association between treatment and potential outcomes.

Example 2: private colleges

Does attending a selective private college result in higher earnings?

TABLE 2.3
Private school effects: Average SAT score controls

	No selection controls			Selection controls		
	(1)	(2)	(3)	(4)	(5)	(6)
Private school	.212 (.060)	.152 (.057)	.139 (.043)	.034 (.062)	.031 (.062)	.037 (.039)
Own SAT score ÷ 100		.051 (.008)	.024 (.006)		.036 (.006)	.009 (.006)
Log parental income			.181 (.026)			.159 (.025)
Female			−.398 (.012)			−.396 (.014)
Black			−.003 (.031)			−.037 (.035)
Hispanic			.027 (.052)			.001 (.054)
Asian			.189 (.035)			.155 (.037)
Other/missing race			−.166 (.118)			−.189 (.117)
High school top 10%			.067 (.020)			.064 (.020)
High school rank missing			.003 (.025)			−.008 (.023)
Athlete			.107 (.027)			.092 (.024)
Average SAT score of schools applied to ÷ 100				.110 (.024)	.082 (.022)	.077 (.012)
Sent two applications				.071 (.013)	.062 (.011)	.058 (.010)
Sent three applications				.093 (.021)	.079 (.019)	.066 (.017)
Sent four or more applications				.139 (.024)	.127 (.023)	.098 (.020)

Notes: This table reports estimates of the effect of attending a private college or university on earnings. Each column shows coefficients from a regression of log earnings on a dummy for attending a private institution and controls. The sample size is 14,238. Standard errors are reported in parentheses.

Source: *Mastering Metrics ch. 2*

Example 2: private colleges

Column (1): attendance at a private college is not randomly assigned; we should be concerned that the coefficient on private school does not describe differences in average potential outcomes for any population of interest. It may be that students attending selective private colleges are better qualified on a number of dimensions than students not attending such colleges. The causal effect is *not identified*.

Another example: class size and student achievement

Omitted variables bias

Suppose that potential outcomes (log earnings) are governed by:

$$Y_i(0) = \alpha + \gamma A_i + u_i$$
$$Y_i(1) = \alpha + \gamma A_i + \beta + u_i$$

A_i is a measure of "ability" and u_i is a random error term. $P_i = 1$ is an indicator for private college attendance ("treatment"). The switching equation gives us:

$$Y_i = \alpha + \beta P_i + \gamma A_i + u_i$$

Call this the "long" regression. Relabel the coefficients:

$$Y_i = \alpha^{\ell} + \beta^{\ell} P_i + \gamma A_i + u_i^{\ell}$$

Omitted variables bias

Suppose instead we estimated the "short" regression (as in column (1) above):

$$Y_i = \alpha^s + \beta^s P_i + u_i^s$$

We know the true model is the "long" regression ($\gamma \neq 0$), so there may be *omitted variables bias* if A_i is related to P_i. The error term in the short regression is: $u_i^s = \gamma A_i + u_i^\ell$.

Omitted variables bias formula

There is a formal (and mechanical) link between β^s and β^ℓ:

$$\beta^s = \beta^\ell + \pi_1 \gamma$$

Where:

- γ comes from the long regression: it is the relationship between Y_i and A_i (conditional on P_i).
- π_1 comes from an "auxiliary" regression of the omitted variable (A_i) on the included variable (P_i).

$$A_i = \pi_0 + \pi_1 P_i + v_i$$

Example 2

Table 2.5: auxiliary regressions where A_i is SAT score (in hundreds)

TABLE 2.3
Private school effects: Average SAT score controls

	No selection controls			Selection controls		
	(1)	(2)	(3)	(4)	(5)	(6)
Private school	.212	.152	.139	.034	.031	.037
	(.060)	(.057)	(.043)	(.062)	(.062)	(.039)
Own SAT score ÷ 100		.051	.024		.036	.009
		(.008)	(.006)		(.006)	(.006)
Log parental income			.181			.159
			(.026)			(.025)
Female			−.398			−.396
			(.012)			(.014)
Black			−.003			−.037
			(.031)			(.035)
Hispanic			.027			.001
			(.052)			(.054)
Asian			.189			.155
			(.035)			(.037)
Other/missing race			−.166			−.189
			(.118)			(.117)
High school top 10%			.067			.064
			(.020)			(.020)
High school rank missing			.003			−.008
			(.025)			(.023)
Athlete			.107			.092
			(.027)			(.024)
Average SAT score of schools applied to ÷ 100				.110	.082	.077
				(.024)	(.022)	(.012)
Sent two applications				.071	.062	.058
				(.013)	(.011)	(.010)
Sent three applications				.093	.079	.066
				(.021)	(.019)	(.017)
Sent four or more applications				.139	.127	.098
				(.024)	(.023)	(.020)

Notes: This table reports estimates of the effect of attending a private college or university on earnings. Each column shows coefficients from a regression of log earnings on a dummy for attending a private institution and controls. The sample size is 14,238. Standard errors are reported in parentheses.

TABLE 2.5
Private school effects: Omitted variables bias

	Dependent variable					
	Own SAT score ÷ 100			Log parental income		
	(1)	(2)	(3)	(4)	(5)	(6)
Private school	1.165	1.130	.066	.128	.138	.028
	(.196)	(.188)	(.112)	(.035)	(.037)	(.037)
Female		−.367			.016	
		(.076)			(.013)	
Black		−1.947			−.359	
		(.079)			(.019)	
Hispanic		−1.185			−.259	
		(.168)			(.050)	
Asian		−.014			−.060	
		(.116)			(.031)	
Other/missing race		−.521			−.082	
		(.293)			(.061)	
High school top 10%		.948			−.066	
		(.107)			(.011)	
High school rank missing		.556			−.030	
		(.102)			(.023)	
Athlete		−.318			.037	
		(.147)			(.016)	
Average SAT score of schools applied to ÷ 100			.777			.063
			(.058)			(.014)
Sent two applications			.252			.020
			(.077)			(.010)
Sent three applications			.375			.042
			(.106)			(.013)
Sent four or more applications			.330			.079
			(.093)			(.014)

Notes: This table describes the relationship between private school attendance and personal characteristics. Dependent variables are the respondent's SAT score (divided by 100) in columns (1)–(3) and log parental income in columns (4)–(6). Each column shows the coefficient from a regression of the dependent variable on a dummy for attending a private institution and controls.

Omitted variables bias: example

Assessing omitted variables bias in column (1) vs. (2):

- $\hat{\beta}^s = 0.212$
- $\beta^s = \beta^\ell + \pi_1 \gamma$
- What do you think the signs of π_1 and γ are?
- The estimated $\widehat{\pi_1} = 1.165$ (the difference in SAT scores between private and public college students) and $\hat{\gamma} = 0.051$
- So, $0.212 = \beta^\ell + (1.165 * 0.051)$. Our estimator of β using β_s is likely biased upward.
- $\hat{\beta}^\ell = 0.152$ (compare to column (2))

Omitted variables bias

Table 10.3. The omitted variable bias formula helps us think about whether failing to control for a confounder results in an over- or under-estimate of the causal effect.

	Omitted Variable Positively Correlated with Treatment $\pi > 0$	**Omitted Variable Negatively Correlated with Treatment** $\pi < 0$
Omitted Variable Positively Correlated with Outcome $\gamma > 0$	Positive bias $\pi \cdot \gamma > 0$	Negative bias $\pi \cdot \gamma < 0$
Omitted Variable Negatively Correlated with Outcome $\gamma < 0$	Negative bias $\pi \cdot \gamma < 0$	Positive bias $\pi \cdot \gamma > 0$

Source: Bueno de Mesquita & Fowler (2021)

Example 2

Of course, a model with two explanatory variables is probably not sufficient in this example: it alone is unlikely to describe differences in average potential outcomes. Column (3) of Table 2.3 includes additional student covariates, such as log parental income, gender, race/ethnicity, athlete, and HS top 10%. The reduction in $\hat{\beta}$ suggests the estimator used in column (2) was still biased upward.

In a setting like this, one should still be concerned about *unobserved*, possibly *unobservable* omitted variables.

The "unobservables"

Example 2

In a further attempt to address unobserved omitted variables, columns (4) - (6) in Table 2.3 represent what might be called a "self-revelation" model. They include the number and characteristics of schools to which students *applied*. This behavior might proxy for unobserved differences that are related to both private college attendance and earnings.

Example 2

Table 2.3
Private school effects: Average SAT score controls

	No selection controls			Selection controls		
	(1)	(2)	(3)	(4)	(5)	(6)
Private school	.212	.152	.139	.034	.031	.037
	(.060)	(.057)	(.043)	(.062)	(.062)	(.039)
Own SAT score ÷ 100		.051	.024		.036	.009
		(.008)	(.006)		(.006)	(.006)
Log parental income			.181			.159
			(.026)			(.025)
Female			−.398			−.396
			(.012)			(.014)
Black			−.003			−.037
			(.031)			(.035)
Hispanic			.027			.001
			(.052)			(.054)
Asian			.189			.155
			(.035)			(.037)
Other/missing race			−.166			−.189
			(.118)			(.117)
High school top 10%			.067			.064
			(.020)			(.020)
High school rank missing			.003			−.008
			(.025)			(.023)
Athlete			.107			.092
			(.027)			(.024)
Average SAT score of schools applied to ÷ 100				.110	.082	.077
				(.024)	(.022)	(.012)
Sent two applications				.071	.062	.058
				(.013)	(.011)	(.010)
Sent three applications				.093	.079	.066
				(.021)	(.019)	(.017)
Sent four or more applications				.139	.127	.098
				(.024)	(.023)	(.020)

Notes: This table reports estimates of the effect of attending a private college or university on earnings. Each column shows coefficients from a regression of log earnings on a dummy for attending a private institution and controls. The sample size is 14,238. Standard errors are reported in parentheses.

Table 2.5
Private school effects: Omitted variables bias

	Dependent variable					
	Own SAT score ÷ 100			Log parental income		
	(1)	(2)	(3)	(4)	(5)	(6)
Private school	1.165	1.130	.066	.128	.138	.028
	(.196)	(.188)	(.112)	(.035)	(.037)	(.037)
Female		−.367			.016	
		(.076)			(.013)	
Black		−1.947			−.359	
		(.079)			(.019)	
Hispanic		−1.185			−.259	
		(.168)			(.050)	
Asian		−.014			−.060	
		(.116)			(.031)	
Other/missing race		−.521			−.082	
		(.293)			(.061)	
High school top 10%		.948			−.066	
		(.107)			(.011)	
High school rank missing		.556			−.030	
		(.102)			(.023)	
Athlete		−.318			.037	
		(.147)			(.016)	
Average SAT score of schools applied to ÷ 100			.777			.063
			(.058)			(.014)
Sent two applications			.252			.020
			(.077)			(.010)
Sent three applications			.375			.042
			(.106)			(.013)
Sent four or more applications			.330			.079
			(.093)			(.014)

Notes: This table describes the relationship between private school attendance and personal characteristics. Dependent variables are the respondent's SAT score (divided by 100) in columns (1)–(3) and log parental income in columns (4)–(6). Each column shows the coefficient from a regression of the dependent variable on a dummy for attending a private institution and controls. The sample size is 14,238. Standard errors are reported in parentheses.

Example 2

In columns (4) - (6) the estimated coefficient on private school shrinks and becomes statistically insignificant.

Interestingly, the correlation between *own* SAT score and private school enrollment is eliminated once application behavior has been controlled for (the self-revelation model). See column (3) of Table 2.5.

Randomized controlled trials

Given their ability to eliminate selection bias, **randomized controlled trials (RCTs)** are considered the "gold standard" design for estimating treatment effects.

- Since treated and untreated are random draws from the same population, they should be balanced on observables and unobservables.
- This implies they are balanced on potential outcomes (and thus on individual treatment effects).
- It is important to collect baseline data so these assumptions can be tested. (While we can never test for differences in unobserved variables, differences in observed variables can be indicative of a problem).

Randomized controlled trials

RCTs come with their own pitfalls and challenges, however. For example:

- Imperfect compliance: subjects may not comply with their treatment assignment. Estimating the ATE is not possible, but could estimate an **intent-to-treat (ITT) effect**, which may also be of interest. This represents the impact of being *offered* treatment.
- **Spillovers** of treatment effect onto untreated (SUTVA violation). Randomization at the *group* level may help.
- **Attrition** from the study, particularly if non-random. Important to compare attrition rates for treated and untreated, and covariates for those remaining in the sample. **Bounds analyses** can probe the potential impact of attrition.

Randomized controlled trials

- **Randomization bias**: when behavior changes as a result of randomization (e.g., "Hawthorne effects")
- **Ethical questions**: is it ethical to withhold a treatment that has a likelihood of success?
- **External validity**: can the causal effect estimated from an RCT be generalized to other populations, places, times?

Randomized controlled trials

Treatment effects are often estimated from RCTs using regression, although (given balance) one could estimate them from a difference in means (e.g., t-test). Benefits of regression:

- Convenient way to obtain standard errors and test for statistical significance.
- Can include controls to improve precision. (These controls should not be affected by the treatment).
- Can include strata dummy variables randomization is within strata.
- Can estimate treatment effects for *subgroups*

Randomized controlled trials

For some simple RCT examples, see:

- Class exercise 1.4 on private school vouchers.
- RAND Health Insurance Experiment and Oregon Health Insurance Experiment in *Mastering Metrics* chapter 1.
- Gennetian et al., (2022) "Unconditional Cash and Family Investments in Infants: Evidence from a Large-Scale Cash Transfer Experiment in the U.S."