

Mostly Harmless Econometrics

Session 2: What Does OLS Estimate?

Linear projection $\rightarrow$ Weighted contrasts $\rightarrow$ Weighted causal effects $\rightarrow$ Estimand choice

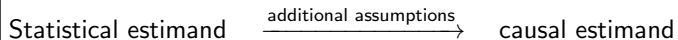
Kaiwen Luo

Washington University in St.Louis

October 16, 2026

What Does OLS Estimate? A Modern View

A useful way to think about OLS is:



OLS does **not** begin with a causal effect. It begins with a population linear projection.

As we impose additional structure, the same regression coefficient may admit increasingly informative representations.

Three questions

For any OLS coefficient, ask:

- 1 What is the coefficient *algebraically*?
- 2 Under what assumptions can it be written in terms of potential outcomes?
- 3 Which treatment effects does it average, and with what weights?

The Starting Point: No Causal Assumptions

Consider the population linear projection

$$(\alpha^{LP}, \beta^{LP}) = \arg \min_{\alpha, \beta} \mathbb{E} [(Y - \alpha - \beta D)^2] \quad \implies \quad \boxed{\beta^{LP} = \frac{\text{Cov}(D, Y)}{\text{Var}(D)}}$$

provided that $\text{Var}(D) > 0$.

This result requires:

- no potential-outcome model;
- no exogeneity;
- no linear conditional expectation;
- no homogeneous treatment effect.

Interpretation

β^{LP} is simply the slope of the best linear predictor of Y using D .

OLS is first a projection estimand, not a causal estimand.

The Assumption Ladder

It is useful to separate three layers.

Layer 1: Pure algebra

$$\beta^{LP} = \frac{\text{Cov}(D, Y)}{\text{Var}(D)}.$$

This is always the population OLS estimand.

Layer 2: Treatment support

Knowing whether D is

binary, ordered discrete, or continuous

allows us to rewrite the same covariance ratio as an average of different kinds of contrasts or slopes.

Layer 3: Identification assumptions

Assumptions such as

$$\mathbb{E}[Y(d) \mid D, X] = \mathbb{E}[Y(d) \mid X]$$

allow observed conditional contrasts to be interpreted causally.

Case 1: Binary Treatment, No Covariates

Suppose $D \in \{0, 1\}$. Then purely by algebra,

$$\beta^{OLS} = \mathbb{E}[Y \mid D = 1] - \mathbb{E}[Y \mid D = 0].$$

No causal assumption is needed for this equality.

With potential outcomes $Y = DY(1) + (1 - D)Y(0)$, we can write

$$\begin{aligned}\beta^{OLS} &= \mathbb{E}[Y(1) \mid D = 1] - \mathbb{E}[Y(0) \mid D = 0] \\ &= \underbrace{\mathbb{E}[Y(1) - Y(0) \mid D = 1]}_{ATT} + \underbrace{\mathbb{E}[Y(0) \mid D = 1] - \mathbb{E}[Y(0) \mid D = 0]}_{\text{selection bias}}.\end{aligned}$$

Binary Treatment: When Does OLS Become the ATE?

Suppose treatment is mean independent of potential outcomes:

$$\mathbb{E}[Y(d) \mid D] = \mathbb{E}[Y(d)], \quad d \in \{0, 1\}.$$

Then

$$\mathbb{E}[Y \mid D = 1] = \mathbb{E}[Y(1)] \quad \text{and} \quad \mathbb{E}[Y \mid D = 0] = \mathbb{E}[Y(0)],$$

so that $\beta^{OLS} = \mathbb{E}[Y(1) - Y(0)] = ATE$. Thus,

$$\frac{\text{Cov}(D, Y)}{\text{Var}(D)} \stackrel{D \in \{0,1\}}{=} \mathbb{E}[Y \mid 1] - \mathbb{E}[Y \mid 0] \stackrel{\text{exogeneity}}{=} ATE.$$

Key point

The first equality is algebra. The second equality is identification.

Adding Covariates: Two Different Objects

Consider the familiar additive regression $Y = \alpha + \beta D + X'\gamma + u$. By Frisch–Waugh–Lovell,

$$\beta^{OLS} = \frac{\mathbb{E}[\tilde{D}Y]}{\mathbb{E}[\tilde{D}^2]}, \quad \tilde{D} = D - L(D \mid 1, X),$$

where $L(D \mid 1, X)$ is the *linear projection* of D on X .

This is generally different from residualizing by the true conditional mean, $D - \mathbb{E}[D \mid X]$.

Important distinction

In general,

$$L(D \mid 1, X) \neq \mathbb{E}[D \mid X].$$

Therefore, ordinary additive OLS with a finite-dimensional set of controls does not automatically inherit the clean causal weighting formulas that arise under saturated or flexible adjustment.

A Clean Benchmark: Flexible Partialling Out

Consider instead the partially linear projection

$$(\beta^*, g^*) = \arg \min_{\beta, g} \mathbb{E} \left[(Y - \beta D - g(X))^2 \right],$$

where $g(X)$ is unrestricted. Define $\mu_D(X) = \mathbb{E}[D | X]$. Then

$$\beta^* = \frac{\mathbb{E}[(D - \mu_D(X))Y]}{\mathbb{E}[(D - \mu_D(X))^2]}.$$

Using iterated expectations,

$$\beta^* = \frac{\mathbb{E}[\text{Cov}(D, Y | X)]}{\mathbb{E}[\text{Var}(D | X)]}.$$

This provides a useful benchmark for all treatment types.

Case 2: Binary Treatment with Flexible Covariate Adjustment

Let

$$D \in \{0, 1\}, \quad e(X) = \Pr(D = 1 \mid X).$$

Then

$$\text{Var}(D \mid X) = e(X)\{1 - e(X)\}.$$

Also,

$$\text{Cov}(D, Y \mid X) = e(X)\{1 - e(X)\} [\mathbb{E}(Y \mid D = 1, X) - \mathbb{E}(Y \mid D = 0, X)].$$

Hence

$$\beta^* = \frac{\mathbb{E}[e(X)\{1 - e(X)\}\Delta_Y(X)]}{\mathbb{E}[e(X)\{1 - e(X)\}]},$$

where

$$\Delta_Y(X) = \mathbb{E}[Y \mid D = 1, X] - \mathbb{E}[Y \mid D = 0, X].$$

Binary Treatment + Conditional Ignorability

Suppose conditional mean independence holds together with overlap:

$$\mathbb{E}[Y(d) \mid D, X] = \mathbb{E}[Y(d) \mid X], \quad d \in \{0, 1\}, \quad 0 < e(X) < 1.$$

Define the conditional average treatment effect

$$\tau(X) = \mathbb{E}[Y(1) - Y(0) \mid X].$$

Then $\Delta_Y(X) = \tau(X)$, and therefore

$$\beta^* = \frac{\mathbb{E}[e(X)\{1 - e(X)\}\tau(X)]}{\mathbb{E}[e(X)\{1 - e(X)\}]}.$$

Interpretation

The coefficient is an **overlap-weighted average treatment effect**, not generally the population ATE.

Binary Treatment: Why the OLS Weight Matters

The weight is proportional to $e(X)\{1 - e(X)\}$. It is largest when $e(X) = \frac{1}{2}$, and approaches zero when $e(X) \rightarrow 0$ or $e(X) \rightarrow 1$.

Thus the estimand emphasizes covariate cells in which both treated and untreated observations are common.

In general,

$$\beta^* \neq ATE = \mathbb{E}[\tau(X)].$$

Equality holds under additional conditions, for example:

- homogeneous treatment effects, $\tau(X) = \tau$;
- randomized treatment with constant propensity, $e(X) = p$;
- or other special restrictions that make the OLS weights unrelated to treatment-effect heterogeneity.

Case 3: Ordered Discrete Treatment, No Covariates

Suppose

$$D \in \{d_0, d_1, \dots, d_J\}, \quad d_0 < d_1 < \dots < d_J,$$

and define the conditional mean $m(d) = \mathbb{E}[Y \mid D = d]$.

Since $\mathbb{E}[Y - m(D) \mid D] = 0$, the residual $Y - m(D)$ is uncorrelated with D , and we obtain

$$\beta^{OLS} = \frac{\text{Cov}(D, m(D))}{\text{Var}(D)}.$$

Implication

OLS only uses the conditional mean function $m(\cdot)$. The question becomes: how does the covariance ratio summarize the shape of $m(\cdot)$ across the support of D ?

Case 3: Ordered Discrete Treatment as Weighted Slopes

Define the adjacent slope

$$s_j = \frac{m(d_j) - m(d_{j-1})}{d_j - d_{j-1}}, \quad j = 1, \dots, J.$$

Then

$$\beta^{OLS} = \sum_{j=1}^J w_j s_j, \quad w_j = \frac{(d_j - d_{j-1}) \text{Cov}(D, \mathbf{1}\{D \geq d_j\})}{\text{Var}(D)}.$$

Moreover,

$$w_j \geq 0, \quad \sum_{j=1}^J w_j = 1.$$

Interpretation

OLS is a convex combination of adjacent slopes of the conditional mean function. Margins near the middle of the distribution of D typically receive the largest weights.

Ordered Discrete Treatment: Causal Interpretation

Suppose $\mathbb{E}[Y(d) \mid D] = \mathbb{E}[Y(d)]$ for every d . Then $m(d) = \mathbb{E}[Y(d)]$. Therefore,

$$\beta^{OLS} = \sum_{j=1}^J w_j \frac{\mathbb{E}[Y(d_j) - Y(d_{j-1})]}{d_j - d_{j-1}}.$$

Interpretation

OLS is a weighted average of causal effects associated with *different treatment margins*:

$$d_0 \rightarrow d_1, \quad d_1 \rightarrow d_2, \quad \dots, \quad d_{J-1} \rightarrow d_J.$$

Unlike the binary case, there is no longer a unique treatment margin.

Related result: Yitzhaki (1996), *Journal of Business & Economic Statistics*.

Ordered Discrete Treatment with Covariates

Under flexible partialling out,

$$\beta^* = \frac{\mathbb{E}[\text{Var}(D | X)\beta(X)]}{\mathbb{E}[\text{Var}(D | X)]}, \quad \beta(X) = \frac{\text{Cov}(D, Y | X)}{\text{Var}(D | X)}.$$

Within each covariate cell, with $m(d, X) = \mathbb{E}[Y | D = d, X]$,

$$\beta(X) = \sum_{j=1}^J w_j(X) \frac{m(d_j, X) - m(d_{j-1}, X)}{d_j - d_{j-1}}.$$

Thus OLS contains **two layers of weighting**:

$\underbrace{\text{Var}(D X)}_{\text{across covariate cells}}$	$\times$	$\underbrace{w_j(X)}_{\text{across treatment margins}}$	$.$
--	----------	---	-----

Ordered Discrete Treatment + Conditional Ignorability

Suppose $\mathbb{E}[Y(d) \mid D, X] = \mathbb{E}[Y(d) \mid X]$. Then $m(d, X) = \mathbb{E}[Y(d) \mid X]$.

Define the conditional causal slope

$$\tau_j(X) = \frac{\mathbb{E}[Y(d_j) - Y(d_{j-1}) \mid X]}{d_j - d_{j-1}}.$$

Then

$$\beta^* = \frac{\mathbb{E} \left[\text{Var}(D \mid X) \sum_{j=1}^J w_j(X) \tau_j(X) \right]}{\mathbb{E}[\text{Var}(D \mid X)]}.$$

Interpretation

A variance-weighted average across covariate cells, within which OLS further averages across treatment margins.

A Caveat: Unordered Discrete Treatments

Suppose instead that treatment categories have no natural ordering: $D \in \{A, B, C\}$.

Assigning arbitrary numerical codes such as $A = 1$, $B = 2$, $C = 3$ and estimating a single regression slope is generally not meaningful. Changing the coding changes the estimand.

The natural approach is to use treatment indicators:

$$Y = \alpha + \beta_B \mathbf{1}\{D = B\} + \beta_C \mathbf{1}\{D = C\} + u.$$

Without covariates,

$$\beta_B = \mathbb{E}[Y \mid D = B] - \mathbb{E}[Y \mid D = A],$$

and similarly for β_C .

Conclusion

A scalar “OLS slope” is intrinsically meaningful for an *ordered scalar treatment*, but not for arbitrarily coded unordered categories.

Case 4: Continuous Treatment, No Covariates

Suppose D is continuously distributed on $\mathcal{D} = [a, b]$, and define $m(d) = \mathbb{E}[Y \mid D = d]$. Again,

$$\beta^{OLS} = \frac{\text{Cov}(D, m(D))}{\text{Var}(D)}.$$

If $m(d)$ is differentiable, then

$$\beta^{OLS} = \int_a^b \omega(d) m'(d) dd,$$

$$\omega(d) = \frac{\text{Cov}(D, \mathbf{1}\{D \geq d\})}{\text{Var}(D)}.$$

The weights satisfy

$$\omega(d) \geq 0, \quad \int_a^b \omega(d) dd = 1.$$

Interpretation

The continuous analogue of the ordered discrete case: adjacent slopes become local derivatives $m'(d)$.

Continuous Treatment: What Does the Slope Average?

The OLS coefficient is therefore a weighted average derivative:

$$\beta^{OLS} = \int \omega(d) \frac{\partial}{\partial d} \mathbb{E}[Y \mid D = d] dd.$$

This should be distinguished from the ordinary average derivative

$$ADE = \mathbb{E}[m'(D)] = \int m'(d) f_D(d) dd.$$

In general,

$$\omega(d) \neq f_D(d) \quad \implies \quad \beta^{OLS} \neq ADE.$$

Key lesson

With a nonlinear conditional mean, a linear regression coefficient is one particular weighted summary of local slopes.

Related result: Yitzhaki (1996).

Continuous Treatment: Causal Interpretation

Suppose $\mathbb{E}[Y(d) \mid D] = \mathbb{E}[Y(d)]$ for every d . Then $m(d) = \mathbb{E}[Y(d)]$. Hence

$$\beta^{OLS} = \int \omega(d) \frac{\partial}{\partial d} \mathbb{E}[Y(d)] dd.$$

Thus the coefficient can be interpreted as a weighted average of marginal causal responses at different treatment levels.

But this does **not** imply that OLS estimates a uniquely privileged “average causal effect.”

Important

For continuous treatments, different scientifically meaningful summaries include:

dose-response curves, shift effects, average derivative effects, weighted derivative effects.

There is no unique scalar analogue of the binary-treatment ATE.

Case 5: Continuous Treatment with Covariates

Under flexible partialling out,

$$\beta^* = \frac{\mathbb{E}[\text{Var}(D | X)\beta(X)]}{\mathbb{E}[\text{Var}(D | X)]}, \quad \beta(X) = \frac{\text{Cov}(D, Y | X)}{\text{Var}(D | X)}.$$

For a differentiable conditional mean $m(d, x) = \mathbb{E}[Y | D = d, X = x]$, we can write

$$\beta(x) = \int_{\mathcal{D}(x)} \omega(d | x) \frac{\partial m(d, x)}{\partial d} dd,$$

with

$$\omega(d | x) = \frac{\text{Cov}(D, \mathbf{1}\{D \geq d\} | X = x)}{\text{Var}(D | X = x)}.$$

Continuous Treatment + Conditional Ignorability

Suppose $\mathbb{E}[Y(d) \mid D, X] = \mathbb{E}[Y(d) \mid X]$. Then $m(d, x) = \mathbb{E}[Y(d) \mid X = x]$. Therefore,

$$\beta^* = \frac{\mathbb{E} [\text{Var}(D \mid X) \int \omega(d \mid X) \frac{\partial}{\partial d} \mathbb{E}[Y(d) \mid X] dd]}{\mathbb{E}[\text{Var}(D \mid X)]}.$$

Again there are two layers of weighting:

$$\underbrace{\text{Var}(D \mid X)}_{\text{which covariate cells?}} \times \underbrace{\omega(d \mid X)}_{\text{which treatment margins?}}.$$

Interpretation

A weighted average of conditional marginal causal effects.

Where the Clean Results Stop

The previous covariate-adjusted formulas used the benchmark $D - \mathbb{E}[D \mid X]$. Ordinary additive OLS instead uses $D - L(D \mid 1, X)$.

If $L(D \mid 1, X) \neq \mathbb{E}[D \mid X]$ and treatment effects are heterogeneous, the coefficient generally does **not** reduce to the simple weighting formulas derived above.

Status of the literature

There are important characterizations for particular settings, especially for binary treatment. However, for an arbitrarily misspecified additive OLS regression with heterogeneous effects, there is no single universal interpretation such as

“OLS always equals a particular ATE.”

The interpretation depends on the specification and the source of treatment variation.

See, e.g., Słoczyński (2022) and Lee and Han (2024).

Binary Treatment: Ordinary Additive OLS Can Reweight Effects

For binary treatment, modern work shows that when treatment effects are heterogeneous, the coefficient from

$$Y = \alpha + \beta D + X'\gamma + u$$

can place implicit weights on different treatment-effect components that differ substantially from their population shares.

Therefore,

conditional ignorability alone $\nrightarrow$ ordinary additive OLS coefficient = *ATE*.

The problem is not necessarily confounding. The problem may instead be the estimand generated by the regression's implicit weighting scheme.

Modern interpretation

Even after identification has been solved, one still needs to ask:

Which causal effects does this regression average?

A More Modern Strategy: Choose the Estimand First

Rather than asking “What causal parameter does this regression happen to estimate?”, a modern approach is:

Scientific question $\rightarrow$ causal estimand $\rightarrow$ identification $\rightarrow$ estimator.

Examples include:

- Binary treatment:

ATE , ATT , ATU , overlap-weighted ATE .

- Multi-valued treatment: $\mathbb{E}[Y(d_j) - Y(d_k)]$ or prespecified averages of treatment contrasts.
- Continuous treatment:

dose-response curve, shift effect, ADE , weighted ADE .

OLS may estimate one useful summary, but the regression coefficient need not define the scientific target.

An Alternative Modern Object: Average Linear Regression Function

Graham and Pinto propose working with the conditional linear projection within each covariate cell. Let

$$b_0(x) = \frac{\text{Cov}(D, Y \mid X = x)}{\text{Var}(D \mid X = x)}.$$

Instead of the pooled OLS weighting, consider

$$\frac{\mathbb{E}[\text{Var}(D \mid X)b_0(X)]}{\mathbb{E}[\text{Var}(D \mid X)]} \quad \text{vs.} \quad \boxed{\beta_{ALRF} = \mathbb{E}[b_0(X)]}.$$

Why this matters

The choice between these two objects is itself an estimand choice.

Pooled regression automatically gives more weight to cells with greater treatment variance.

The average linear regression function averages conditional regression slopes over the covariate distribution instead.

Graham and Pinto (2022), *Journal of Econometrics*.

Continuous Treatments: The Frontier Is About the Target

For continuous treatment, the algebraic interpretation of OLS is well understood:

OLS = a particular weighted average derivative.

What is **not** uniquely determined is whether this is the scientifically appropriate causal summary.

Recent work studies alternative targets such as

$$ADE = \mathbb{E} \left[\frac{\partial}{\partial d} \mathbb{E}[Y(d) | X] \Big|_{d=D} \right],$$

as well as more general weighted average derivative effects.

Frontier status

There is no single universally preferred scalar causal estimand for a general continuous treatment.

The appropriate target depends on the intervention and the scientific question.

See Graham and Pinto (2022) and Hines, Diaz-Ordaz, and Vansteelandt (2026).

Summary I: Without Covariates

Binary D

$$\beta^{OLS} = \mathbb{E}[Y \mid 1] - \mathbb{E}[Y \mid 0].$$

Under exogeneity:

$$\beta^{OLS} = ATE.$$

Ordered discrete D

$$\beta^{OLS} = \sum_j w_j \frac{m(d_j) - m(d_{j-1})}{d_j - d_{j-1}}.$$

Under exogeneity: a weighted average of **causal treatment-margin slopes**.

Continuous D

$$\beta^{OLS} = \int \omega(d) m'(d) dd.$$

Under exogeneity: a weighted average of **marginal causal responses**.

Same covariance ratio; the support of D determines what it averages.

Summary II: With Flexible Covariate Adjustment

The common formula is

$$\beta^* = \frac{\mathbb{E}[\text{Var}(D | X)\beta(X)]}{\mathbb{E}[\text{Var}(D | X)]}.$$

Binary treatment

$$\beta(X) = \mathbb{E}[Y | 1, X] - \mathbb{E}[Y | 0, X].$$

Under conditional ignorability:

$$\beta^* = \text{overlap-weighted CATE.}$$

Ordered discrete treatment

$$\beta(X) = \text{weighted average across discrete treatment margins.}$$

Continuous treatment

$$\beta(X) = \text{weighted average of conditional derivatives.}$$

The Main Lesson

The classical question is: “What is the coefficient on D ?” The modern question is:

What population variation does the coefficient use,
and which heterogeneous effects does it average?

OLS can be viewed as an **averaging machine**:

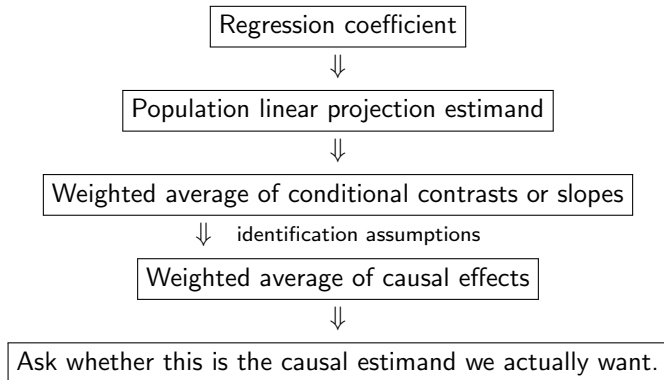
OLS = selection of identifying variation + implicit weighting.

For binary treatment without covariates, this complexity is largely hidden because there is only one treatment margin. For multivalued or continuous treatment, the weighting problem becomes explicit.

Whose effect? At which treatment margin? With what weight?

From “Regression First” to “Estimand First”

A useful intellectual progression is



This last step is the central modern shift:

Define the estimand first; choose the estimator second.

Selected References

- Graham, B. S., and C. C. de Xavier Pinto (2022). "Semiparametrically Efficient Estimation of the Average Linear Regression Function." *Journal of Econometrics*, 226(1), 115–138.
- Hines, O. J., K. Diaz-Ordaz, and S. Vansteelandt (2026). "Parameterizing the Effect of a Continuous Treatment Using Average Derivative Effects." *Biometrika*, 113(2).
- Ishimaru, S. (2024). "Empirical Decomposition of the IV–OLS Gap with Heterogeneous and Nonlinear Effects." *Review of Economics and Statistics*, 106(2), 505–520.
- Lee, M.-J., and C. Han (2024). "Ordinary Least Squares and Instrumental-Variables Estimators for Any Outcome and Heterogeneity." *Stata Journal*, 24(1), 72–92.
- Słoczyński, T. (2022). "Interpreting OLS Estimands When Treatment Effects Are Heterogeneous: Smaller Groups Get Larger Weights." *Review of Economics and Statistics*, 104(3), 501–509.
- Yitzhaki, S. (1996). "On Using Linear Regressions in Welfare Economics." *Journal of Business & Economic Statistics*, 14(4), 478–486.