OVB, Interactions, and the Bridge to OLS

What omitting a variable does to the BLP

Gov 2001 · Scott Cunningham · Spring 2026

Duncan (1961): income and education both predict prestige — but which is driving it?

45 US occupations, 1950 census

- Prestige: % of respondents rating the occupation "good" or better
- Income: % in the occupation earning > \$3,500/year
- Education: % in the occupation with a high school diploma

$$\text{Cor}(\text{Income, Education}) = 0.73$$

High-income occupations also tend to require education.
Which variable is really driving prestige?

The bivariate and long BLPs give very different income slopes

Bivariate BLP (omits Education):

$$\widehat{\text{Prestige}} = 2.46 + \underbrace{\mathbf{1.080}}_{\delta_{\text{inc}}}\ \text{Inc}$$

$$R^2 = 0.708$$

Long BLP (includes Education):

$$\widehat{\text{Prestige}} = -6.07 + \underbrace{\mathbf{0.599}}_{\beta_{\text{inc}}}\ \text{Inc} + 0.546\ \text{Educ}$$

$$R^2 = 0.828$$

Income slope drops from 1.080 to 0.599 once we control for education

FWL Step 1 on Duncan data: partial income out of education

Auxiliary BLP — regress Income on Education:

$$\widehat{\text{Inc}} = 10.60 + 0.595\,\text{Educ}$$

Residual — income with education's influence removed:

$$\widetilde{\text{Inc}} \;=\; \text{Inc} - \widehat{\text{Inc}}$$

By construction: $\text{Cov}(\widetilde{\text{Inc}}, \text{Educ}) = 0$

FWL Step 2: regress Prestige on residual Income — slope matches the long BLP exactly

Bivariate BLP of Prestige on $\widetilde{\text{Inc}}$:

$$\widehat{\text{Prestige}} = 47.69 + 0.599\,\widetilde{\text{Inc}}$$

$\boxed{\text{Step 2 slope} = 0.599 = \beta_{\text{inc}} \text{ from the long BLP} \quad \text{(exact to 4 decimal places)}}$

FWL is not an approximation. It is an algebraic identity.

Three regressions, two distinct slopes — FWL explains why they match

	Bivariate	**Long BLP**	**FWL**
	Prestige ~ Inc	Prestige ~ Inc + Educ	Prestige ~ $\widetilde{\text{Inc}}$
$\hat{\beta}_{\text{income}}$	**1.080**	**0.599**	**0.599**
$\hat{\beta}_{\text{education}}$	—	0.546	—
R^2	0.708	0.828	0.102

Bivariate ≠ long: Income and Education are correlated ($r = 0.73$), so the bivariate slope absorbs Education's effect.
FWL = long: $\widetilde{\text{Inc}}$ is orthogonal to Education by construction.

The multivariate BLP raises three questions — today we answer them

1. What does β_k measure?

Each coefficient is a *partial* effect — the slope on X_k after removing everything the other regressors explain.

Frisch-Waugh-Lovell

2. What if you leave a variable out?

The remaining coefficients absorb the omitted variable's effect — and the gap is exactly quantifiable.

Omitted variable bias formula

3. What if the effect of X_1 depends on X_2?

Interaction terms let the BLP slope on one variable shift with the value of another.

Interactions and polynomials

All three are BLP concepts. Wednesday: OLS as the sample analog.

We continue from last time: the multivariate BLP

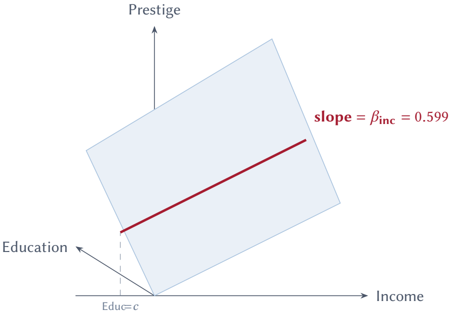
$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + U$$

- β_j: slope on X_j in the best linear predictor of Y given all regressors
- Defined by the $k+1$ first-order conditions — one per coefficient
- $U = Y - \text{BLP}(X_1, \ldots, X_k)$: prediction error, orthogonal to all regressors by construction

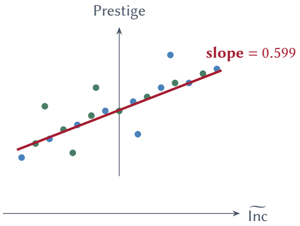
Last time: why the BLP exists and what the FOCs pin down.
Today: what each β_j actually *measures*.

The BLP plane slice and FWL both recover slope $\beta_{\text{inc}} = 0.599$ — from the same partial variation

Long BLP plane: fix Education, vary Income



FWL: Prestige vs $\widetilde{\text{Inc}}$



Both red lines have slope $\beta_{\text{inc}} = 0.599$ — the BLP plane slice and FWL are the same object

Last week: fixing Education at any value c gives a line through the BLP plane with slope β_{inc}. FWL extracts that slope from the 2D scatter by removing Education's variation.

FWL Step 1: regress X_1 on X_2 to extract what X_2 predicts about X_1

Auxiliary BLP:

$$X_1 = \pi_0 + \pi_1 X_2$$

Fitted value — the component of X_1 explained by X_2:

$$\hat{X}_1 = \hat{\pi}_0 + \hat{\pi}_1 X_2$$

How much of the variation in X_1 is "driven by" X_2?
The fitted value captures exactly that.

The residual $\tilde{X}_1$ is the part of X_1 that X_2 cannot explain

$$\tilde{X}_1 \;=\; X_1 - \hat{X}_1$$

$$\boxed{\tilde{X}_1 \perp X_2 \qquad \text{by construction of the BLP}}$$

Whatever is left in $\tilde{X}_1$ is genuinely unrelated to X_2.
This is the “pure” variation in X_1.

FWL Step 2: regress Y on $\tilde{X}_1$ — slope equals β_1 from the long BLP

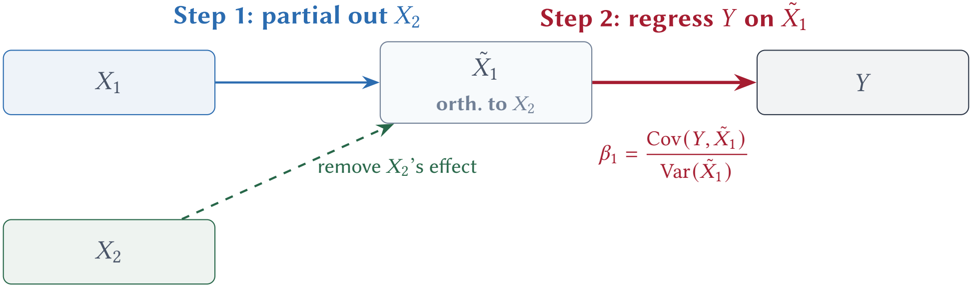
Bivariate BLP of Y on $\tilde{X}_1$:

$$Y = \gamma_0 + \gamma_1 \tilde{X}_1 \qquad \gamma_1 = \frac{\mathrm{Cov}(Y, \tilde{X}_1)}{\mathrm{Var}(\tilde{X}_1)}$$

$\gamma_1 = \beta_1$ (Filoso 2013 — an exact algebraic identity, not an approximation)

The Frisch-Waugh-Lovell theorem: the bivariate slope on $\tilde{X}_1$ equals the partial slope β_1 from the full multivariate BLP.

Frisch-Waugh-Lovell generalizes: every coefficient in the multivariate BLP is a partial-out



Every coefficient in a multivariate BLP is the result of this process for its own variable — simultaneously, not sequentially

$\beta_1 = \gamma_1$: different regression, identical slope — the Frisch-Waugh-Lovell theorem

Long BLP		**FWL bivariate BLP**
$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2$	=	$Y = \gamma_0 + \gamma_1 \tilde{X}_1$
slope on $X_1 = \beta_1$		**slope on** $\tilde{X}_1 = \gamma_1$

$\beta_1 = \gamma_1$ (not an approximation — an exact algebraic identity)

FWL wraps up multiple regression — now: what is U, and what happens when you omit a variable?

The BLP always comes with a prediction error:

$$Y = \underbrace{\beta_0 + \beta_1 X_1 + \beta_2 X_2}_{\text{BLP}(X_1, X_2)} + U$$

U is the population prediction error

$U = Y - \text{BLP}(X_1, X_2)$

Always exists. Always defined.

FOC $\Rightarrow$ tautologies:

$\mathbb{E}[U] = 0$

$\text{Cov}(X_1, U) = 0$

$\text{Cov}(X_2, U) = 0$

These are **not** assumptions — they hold by construction for every joint distribution of (Y, X_1, X_2)

U in BLP world vs. ε in structural world — same symbol, completely different meaning

BLP world: U

$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + U$

U = population prediction error

$\text{Cov}(X, U) = 0$ is a **tautology**

Cannot fail — U is defined to be orthogonal to X

Structural world: ε

$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon$

ε = unobserved causes of Y

$\text{Cov}(X, \varepsilon) = 0$ is an **assumption**

Can fail — this is what identification is about

OVB is **not** about $\text{Cov}(X, U) \neq 0$ — that can never happen in BLP world.
OVB is about two different population objects answering different questions.

Frisch-Waugh-Lovell recovers β_1 from the long model — what if we omit Z entirely?

FWL:

Two ways to get β_1 from the *long* BLP

Direct regression and partialing out give the same answer

New question:

What does the *short* BLP give?

$$m_S(X) = \delta_0 + \delta_1 X$$

Z omitted entirely — not partialed out, just gone

The question: does omitting Z change the slope? By how much?

The OVB formula answers this exactly — pure algebra, no simulation needed

The short and long BLPs are two different population objects

Short BLP (omits Z):

$$m_S(X) = \delta_0 + \delta_1 X$$

Estimand: $\delta_1 = \dfrac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(X)}$

What we get from the bivariate projection

Long BLP (includes Z):

$$m_L(X,Z) = \beta_0 + \beta_1 X + \gamma Z$$

Estimand: β_1 from the multivariate FOC

What we get when we hold Z fixed

Auxiliary projection:

$$Z = \pi_0 + \pi_1 X + \nu \qquad \Longrightarrow \qquad \pi_1 = \frac{\mathrm{Cov}(X,Z)}{\mathrm{Var}(X)}$$

π_1: co-movement of Z and X in the population

Derive the OVB formula: substitute the auxiliary regression into the long BLP

Long BLP: $Y = \beta_0 + \beta_1 X + \gamma Z$ γ = coefficient on Z in the long model

Auxiliary: $Z = \pi_0 + \pi_1 X + v, \quad v \perp X$

Substitute Z into the long BLP:

$$\begin{aligned} Y &= \beta_0 + \beta_1 X + \gamma(\pi_0 + \pi_1 X + v) \\ &= \underbrace{(\beta_0 + \gamma\pi_0)}_{\text{new intercept}} + \underbrace{(\beta_1 + \gamma\pi_1)}_{\delta_1} X + \gamma v \end{aligned}$$

$\gamma v \perp X$ by construction: the short BLP slope picks up $\beta_1 + \gamma\pi_1$

$$\boxed{\delta_1 = \beta_1 + \pi_1 \cdot \gamma}$$

OVB formula: the short coefficient equals the long coefficient plus a bias term

$$\underbrace{\delta_1}_{\text{short BLP}} = \underbrace{\beta_1}_{\text{long BLP}} + \underbrace{\pi_1}_{\substack{\text{regression of}\\ Z \text{ on } X}} \cdot \underbrace{\gamma}_{\substack{\text{long coeff.}\\ \text{on } Z}}$$

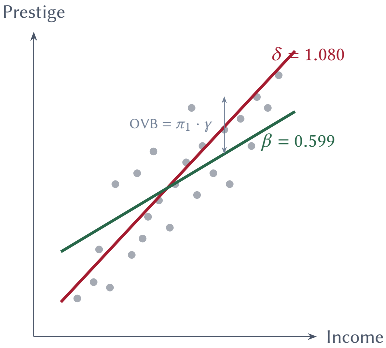
This is pure algebra

- Both δ_1 and β_1 are population BLP coefficients
- No estimation, no sampling error

The term $\pi_1 \cdot \gamma$

- $\gamma \neq 0$: Z predicts Y given X
- $\pi_1 \neq 0$: Z is correlated with X
- Both required for $\delta_1 \neq \beta_1$

Omitting education inflates the income slope — the gap is OVB in the Duncan data



Short BLP (omits Education):

$\delta_{\text{inc}} = 1.080$

Absorbs Education's effect via the Income–Education correlation ($r = 0.73$)

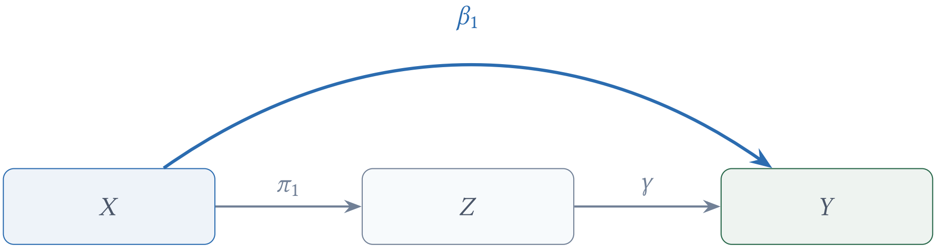
Long BLP (fixes Education):

$\beta_{\text{inc}} = 0.599$

Pure income effect, holding Education constant

$$\text{Gap} = \underbrace{0.595}_{\pi_1} \times \underbrace{0.546}_{\gamma} \approx 0.325$$

OVB is zero under two conditions — each kills one channel



$\pi_1 \cdot \gamma = 0$ iff $\pi_1 = 0$ **or** $\gamma = 0$

$\gamma = 0$: Z does not predict Y once X is in the model

$\pi_1 = 0$: Z is orthogonal to X in the population
$\delta_1 = \beta_1$

$\delta_1 \neq \beta_1$ is algebra — calling it "bias" is a causal claim

What the OVB formula says:

$$\delta_1 - \beta_1 = \pi_1 \cdot \gamma$$

Two different population BLP coefficients answering different questions

What "bias" additionally requires:

- β_1 is the *true causal effect* of X on Y
- δ_1 is a *contaminated* version
- Neither follows from algebra alone

What makes β_1 causal — the conditional independence assumption

What makes β_1 the causal effect of X:

$$(Y_{0i}, Y_{1i}) \perp X \mid Z \quad \textbf{(CIA)}$$

Requires a design argument — not more algebra

Today: $\delta_1 \neq \beta_1$ is algebra. Whether that gap is "bias" is a causal question.

Adding an interaction term changes the estimand — the slope on X_1 becomes a function of X_2

Additive BLP:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2$$

Slope on X_1: **constant**

$$\frac{\partial \text{BLP}}{\partial X_1} = \beta_1$$

Same effect of X_1 at every level of X_2

Interaction BLP:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2$$

Slope on X_1: **a function of X_2**

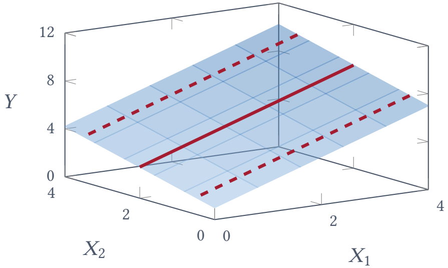
$$\frac{\partial \text{BLP}}{\partial X_1} = \beta_1 + \beta_3 X_2$$

Effect of X_1 depends on level of X_2

β_3: how much the effect of X_1 changes per unit increase in X_2

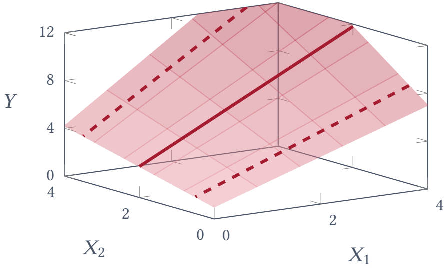
The additive BLP is a flat plane — the interaction BLP is a twisted surface

Additive: flat plane



All red lines: slope = $\beta_1 = 1.5$

Interaction: twisted surface



Red lines fan out: slope = $\beta_1 + \beta_3 X_2$

β_1 in an interaction model is the slope when $X_2 = 0$

Interaction model: $$\frac{\partial \text{BLP}}{\partial X_1} = \beta_1 + \beta_3 X_2$$

β_1 is the effect of X_1 when $X_2 = 0$

If 0 is not in the support of X_2: β_1 has no direct interpretation.

Fix: Center X_2. Then β_1 = effect of X_1 at the *mean* of X_2.

Centering X_2 also eliminates OVB from dropping the interaction

OVB from omitting X_1X_2:

The bias on β_1 from dropping the interaction term is $\pi_1 \cdot \beta_3$,
where $\pi_1 = \text{Cov}(X_1X_2, X_1) \,/\, \text{Var}(X_1)$.

If X_2 is mean-centered: $\pi_1 = 0$, and β_1 from the additive model = β_1 from the interaction model

Saturated model: when $X_1, X_2 \in \{0, 1\}$, the interaction BLP *is* the CEF

Four cells, four parameters:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2 + U$$

	$X_2 = 0$	$X_2 = 1$
$X_1 = 0$	β_0	$\beta_0 + \beta_2$
$X_1 = 1$	$\beta_0 + \beta_1$	$\beta_0 + \beta_1 + \beta_2 + \beta_3$

One parameter per cell

BLP = CEF exactly

4 parameters, 4 cells
No approximation error

$\mathbb{E}[U \mid X_1, X_2] = 0$ everywhere

For binary covariates: every regression with the full interaction is saturated — the CEF is linear, not approximately linear

Everything so far is a population object — OLS is the sample analog

Estimands covered in Lectures 10–11a:

- **Multivariate BLP:** $\beta_0 + \beta_1 X_1 + \cdots + \beta_k X_k$ — partial slopes, FWL
- **OVB formula:** $\delta_1 = \beta_1 + \pi_1 \cdot \gamma$ — pure algebra
- **Interaction BLP:** $\partial\text{BLP}/\partial X_1 = \beta_1 + \beta_3 X_2$ — effect heterogeneity
- **Saturated model:** BLP = CEF exactly for binary X

Population
$\beta_0 + \beta_1 X_1 + \cdots + \beta_k X_k$
multivariate BLP

Lecture 11b →
sample analog:
$(X'X)^{-1}X'Y$

Sample
$\hat{\beta}_0 + \hat{\beta}_1 X_1 + \cdots + \hat{\beta}_k X_k$
OLS