

Sensitivity, especially Omitted Variable Bias

Stat 256, Spring 2026

Chad Hazlett

UCLA

`chazlett@ucla.edu`

Why sensitivity now

So far the workflow has been: pick an ID strategy (e.g. CI/SOO), defend the assumptions, estimate well.

“Defend the assumptions” is the fragile one!

Sensitivity analysis: Don't defend “confounding=0”, ask “How strong would confounding have to be to change my research conclusion?”

- Warning I will repeat: this is NOT “how much confounding is there”

Old idea. Cornfield et al. (1959) on smoking and lung cancer:

... an unobserved confounder ... would need to lead to a nine-fold increase in the odds of smoking to explain away the association ... and it was asserted that such a strong unobserved confounder was very unlikely to exist.

Effects of violence on attitudes: angry or weary?

Among Darfurian refugees in eastern Chad, did exposure to violence on average make people more willing to make concessions to achieve peace (“weary”), or more desiring of retribution (“angry”)?

Central ID claim: targeting of violence in 2003–2004 was based only on village and gender, otherwise highly indiscriminate. *Being directly harmed is “as-if random within village $\times$ gender strata.”*

- Aerial bombing was extremely crude
- *Janjaweed* militia attacks targeted on visible characteristics only (notably *female*)
- Purpose was to clear population, not target rebels

Strategy: SOO/CI conditioning on village and gender.

Darfur: estimate, and the worry

Setup: outcome *PeaceIndex* (0–1); treatment *DirectHarm* (0/1).

$$PeaceIndex_i = \alpha DirectHarm_i + \beta^T Village_i + \phi female_i + \epsilon_i.$$

Those directly harmed are on average **9.7 percentage points** more pro-peace (SE = 2.3).

But what if CI is “a little” wrong?

Questions we'd like sensitivity to help with:

- 1 If somebody names a confounder I don't have, how strong would it have to be to overturn the result?
- 2 What about *multiple* confounders, possibly acting non-linearly?
- 3 How can we report robustness compactly?
- 4 How can we make informed guesses about plausible confounding?

A landscape, and our choice

Several families of sensitivity analysis exist:

- **Matched-pair Γ bounds** (Rosenbaum 2002): a single odds-of-treatment parameter for matched designs.
- **Parametric** (Imbens 2003; Carnegie–Dorie–Hill 2016): fully specify a model for (D, Y, U) , sweep over parameters.
- **Partial identification** (Manski 1990; e.g. monotonicity bounds): assume less, get ranges.
- **OVB-based** (Frank 2000; Hosman–Hansen–Holland 2010; **Cinelli & Hazlett 2020**): use omitted-variable bias as the engine.

We focus on OVB-based: scale-free in R^2 , handles multiple/non-linear confounders, plugs into the regression workflow we just built.

A short detour: Frisch–Waugh–Lovell

Suppose we run OLS of Y on D and $\mathbf{X}$:

$$Y = \hat{\tau} D + \mathbf{X}\hat{\beta} + \hat{\varepsilon}.$$

The same $\hat{\tau}$ is given by a three-step recipe:

- ① Regress Y on $\mathbf{X}$, take residuals $Y^{\perp\mathbf{X}}$
- ② Regress D on $\mathbf{X}$, take residuals $D^{\perp\mathbf{X}}$
- ③ Regress $Y^{\perp\mathbf{X}}$ on $D^{\perp\mathbf{X}}$

In words: *multivariate OLS reduces to univariate OLS, after partialling out $\mathbf{X}$,*

$$\hat{\tau} = \frac{\text{cov}(D^{\perp\mathbf{X}}, Y^{\perp\mathbf{X}})}{\text{var}(D^{\perp\mathbf{X}})}.$$

This is the central tool for everything that follows.

Omitted variable bias: setup

Suppose an investigator wishes to run:

$$Y = \hat{\tau} D + \mathbf{X}\hat{\beta} + \hat{\gamma} Z + \hat{\epsilon}_{\text{full}}$$

But Z is unobserved, so they run the *restricted* regression:

$$Y = \hat{\tau}_{\text{res}} D + \mathbf{X}\hat{\beta}_{\text{res}} + \hat{\epsilon}_{\text{res}}$$

We ask: **how would $\hat{\tau}$ change if you could include Z ?**

Whether the “full” regression is wise, or its coefficient is the quantity of interest, is a separate question — here we just compute the consequence of omission.

OVB via FWL: bias = $\hat{\gamma}\hat{\delta}$

Restating FWL,

$$\hat{\tau}_{\text{res}} = \frac{\text{cov}(D^{\perp \mathbf{X}}, Y^{\perp \mathbf{X}})}{\text{var}(D^{\perp \mathbf{X}})}.$$

Substitute the (true) full model $Y = \hat{\tau}D + \mathbf{X}\hat{\beta} + \hat{\gamma}Z + \hat{\epsilon}_{\text{full}}$, so

$$Y^{\perp \mathbf{X}} = \hat{\tau} D^{\perp \mathbf{X}} + \hat{\gamma} Z^{\perp \mathbf{X}} + \hat{\epsilon}_{\text{full}}^{\perp \mathbf{X}}.$$

Then:

$$\hat{\tau}_{\text{res}} = \hat{\tau} + \hat{\gamma} \underbrace{\frac{\text{cov}(D^{\perp \mathbf{X}}, Z^{\perp \mathbf{X}})}{\text{var}(D^{\perp \mathbf{X}})}}_{=: \hat{\delta}} = \hat{\tau} + \hat{\gamma}\hat{\delta}.$$

$\hat{\delta}$ is the coefficient on D in the auxiliary regression

$$Z = \hat{\delta}D + \mathbf{X}\hat{\psi} + \hat{\epsilon}_Z.$$

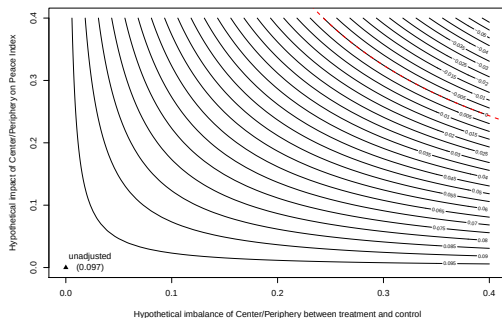
$$\widehat{\text{bias}}(\hat{\tau}_{\text{res}}) = \hat{\gamma}\hat{\delta} = \text{impact} \times \text{imbalance}$$

Consider a confounder:

Suppose a colleague says:

Those near the center of the village were both more likely to be injured and have different attitudes.

For each hypothesized pair $(\hat{\gamma}, \hat{\delta})$, the adjusted estimate is $\hat{\tau} = \hat{\tau}_{\text{res}} - \hat{\gamma}\hat{\delta}$.



Why reparameterize

Problems with the natural parameterization:

- “Impact” and “imbalance” are scale-dependent. *Center* happened to be binary but hard for abstract confounders (e.g. “political attitudes”).
- “Name a possible confounder” is too narrow. We want to handle multiple confounders, possibly acting non-linearly.

Fix: reparameterize using *partial* R^2 .

- Scale-free, intuitive, bounded in $[0, 1]$
- With this we will derive bias *and* variance, do bounds, handle all confounders jointly, and more.

What is the partial R^2 ?

The partial correlation is simply correlation *after* partialling out $\mathbf{X}$:

$$R_{Y \sim D | \mathbf{X}} = \text{cor}(Y^{\perp \mathbf{X}}, D^{\perp \mathbf{X}}).$$

Square it for the partial R^2 .

Or: “fraction of D ’s explained by Z beyond what $\mathbf{X}$ explained”:

$$R_{D \sim Z | \mathbf{X}}^2 = \frac{R_{D \sim \mathbf{X} + Z}^2 - R_{D \sim \mathbf{X}}^2}{1 - R_{D \sim \mathbf{X}}^2}.$$

And:

$$R_{Y \sim Z | D, \mathbf{X}}^2 = \frac{R_{Y \sim D + \mathbf{X} + Z}^2 - R_{Y \sim D + \mathbf{X}}^2}{1 - R_{Y \sim D + \mathbf{X}}^2}.$$

Direction-symmetric: $R_{D \sim Z | \mathbf{X}}^2 = R_{Z \sim D | \mathbf{X}}^2$

Deriving bias in the R^2 parameterization

By FWL,

$$\begin{aligned}
 \widehat{\text{bias}} &= \hat{\gamma} \hat{\delta} \\
 &= \left(\frac{\text{cov}(Y^{\perp X, D}, Z^{\perp X, D})}{\text{var}(Z^{\perp X, D})} \right) \left(\frac{\text{cov}(D^{\perp X}, Z^{\perp X})}{\text{var}(D^{\perp X})} \right) \\
 &= \text{cor}(Y^{\perp X, D}, Z^{\perp X, D}) \cdot \text{cor}(D^{\perp X}, Z^{\perp X}) \\
 &\quad \cdot \frac{\text{sd}(Y^{\perp X, D})}{\text{sd}(D^{\perp X})} \cdot \frac{\text{sd}(Z^{\perp X})}{\text{sd}(Z^{\perp X, D})}.
 \end{aligned}$$

Recognise:

- $\text{cor}(Y^{\perp X, D}, Z^{\perp X, D})^2 = R_{Y \sim Z|D, X}^2$
- $\text{cor}(D^{\perp X}, Z^{\perp X})^2 = R_{D \sim Z|X}^2$
- $\frac{\text{var}(Z^{\perp X, D})}{\text{var}(Z^{\perp X})} = 1 - R_{D \sim Z|X}^2$

$$|\widehat{\text{bias}}| = \sqrt{\frac{R_{Y \sim Z|D, X}^2 R_{D \sim Z|X}^2}{1 - R_{D \sim Z|X}^2}} \cdot \frac{\text{sd}(Y^{\perp X, D})}{\text{sd}(D^{\perp X})}.$$

Bias anatomy

An OLS variance fact (for now and later):

$$\widehat{\text{se}}(\hat{\tau}_{\text{res}}) = \frac{\text{sd}(Y^{\perp \mathbf{X}, D})}{\text{sd}(D^{\perp \mathbf{X}}) \sqrt{\text{df}}}.$$

Nice to rewrite bias then as:

$$|\widehat{\text{bias}}| = \sqrt{\frac{R_{Y \sim Z|D, \mathbf{X}}^2 R_{D \sim Z|\mathbf{X}}^2}{1 - R_{D \sim Z|\mathbf{X}}^2}} \cdot \frac{\text{sd}(Y^{\perp \mathbf{X}, D})}{\text{sd}(D^{\perp \mathbf{X}})} = \widehat{\text{se}}(\hat{\tau}_{\text{res}}) \cdot \sqrt{\underbrace{\frac{R_{Y \sim Z|D, \mathbf{X}}^2 R_{D \sim Z|\mathbf{X}}^2}{1 - R_{D \sim Z|\mathbf{X}}^2}}_{\text{BF}^2}} \cdot \text{df}.$$

Consider the “relative bias”, i.e. the proportion of $\hat{\tau}_{\text{res}}$ that is bias. Substitute $|\hat{\tau}_{\text{res}}| = |t_D| \cdot \widehat{\text{se}}(\hat{\tau}_{\text{res}})$ in the denominator ($|t_D|$ = absolute t-stat for D):

$$\frac{|\widehat{\text{bias}}|}{|\hat{\tau}_{\text{res}}|} = \frac{\widehat{\text{se}}(\hat{\tau}_{\text{res}}) \cdot \sqrt{\text{BF}^2 \cdot \text{df}}}{|t_D| \cdot \widehat{\text{se}}(\hat{\tau}_{\text{res}})} = \frac{\sqrt{\text{df}} \cdot \text{BF}}{|t_D|}.$$

Bias anatomy:

One last touch:

$$\frac{\sqrt{df} \cdot BF}{|t_D|} = \frac{BF}{|f_{Y \sim D|\mathbf{X}}|}$$

where $f_{Y \sim D|\mathbf{X}} = R_{Y \sim D|\mathbf{X}} / \sqrt{1 - R_{Y \sim D|\mathbf{X}}}$ is the (partial) Cohen's f .

Key insights

- Numerator: BF = essential expression of confounding strength.
- Denominator: $f_{Y \sim D|\mathbf{X}}$ is essentially how much Y is explained uniquely by D and is *all* you need from the data to understand sensitivity. It is what confounding is competing to explain.
- And $f_{Y \sim D|\mathbf{X}}$ is not function of sample size: more data doesn't help you fight confounding.

Variance anatomy

Same OLS variance fact, applied to the *full* regression:

$$\widehat{\text{se}}(\hat{\tau}) = \frac{\text{sd}(Y^{\perp \mathbf{X}, D, Z})}{\text{sd}(D^{\perp \mathbf{X}, Z}) \sqrt{\text{df} - 1}}.$$

Two ratios of variances are partial R^2 's:

$$\frac{\text{sd}^2(Y^{\perp \mathbf{X}, D, Z})}{\text{sd}^2(Y^{\perp \mathbf{X}, D})} = 1 - R_{Y \sim Z | \mathbf{X}, D}^2, \quad \frac{\text{sd}^2(D^{\perp \mathbf{X}, Z})}{\text{sd}^2(D^{\perp \mathbf{X}})} = 1 - R_{D \sim Z | \mathbf{X}}^2.$$

Combine, factoring out $\widehat{\text{se}}(\hat{\tau}_{\text{res}})$:

$$\widehat{\text{se}}(\hat{\tau}) = \widehat{\text{se}}(\hat{\tau}_{\text{res}}) \cdot \sqrt{\underbrace{(1 - R_{Y \sim Z | \mathbf{X}, D}^2)}_{\text{VRF}} \underbrace{\frac{1}{1 - R_{D \sim Z | \mathbf{X}}^2}}_{\text{VIF}} \underbrace{\frac{\text{df}}{\text{df} - 1}}_{\text{df adj.}}}$$

- **VRF**: confounder shrinks residual Y variance.
- **VIF**: confounder's correlation with D inflates the SE.

Two simple statistics for routine reporting

Robustness Value (RV): the level of (equal) association with D and Y that would eliminate the effect.

- RV_q : the equal-association level needed to bring the estimate down by proportion q
- RV_α : the equal-association level needed to bring the t-stat to the threshold for level α

Partial R^2 of D with Y : $R^2_{Y \sim D | X}$

- If confounding explained *all* residual Y -variance, this is how much it would have to explain of D (residual variance) to bring the estimate to zero.

Both come from t_D and df alone — you can compute them from a published regression table.

Augmented regression table: Darfur

Outcome: *Peace Index*

Treatment:	Est.	SE	t	$R^2_{Y \sim D X}$	RV	$RV_{\alpha=0.05}$
<i>Directly Harmed</i>	0.097	0.023	4.18	2.2%	13.9%	7.6%
df = 783	<i>Bound (1× female): $R^2_y = 12\%$, $R^2_d = 1\%$</i>					

Cold-call check:

- What does this RV mean?
- What does $RV_{\alpha=0.05}$ mean?
- What does $R^2_{Y \sim D|X}$ mean?

Bounding by comparison to observed covariates

“Benchmarking” interpretation: For an observed covariate X_j , define

$$k_D := \frac{R_{D \sim Z | \mathbf{x}_{-j}}^2}{R_{D \sim X_j | \mathbf{x}_{-j}}^2}, \quad k_Y := \frac{R_{Y \sim Z | \mathbf{x}_{-j}, D}^2}{R_{Y \sim X_j | \mathbf{x}_{-j}, D}^2}.$$

“ Z is at most k_D times as predictive of D , k_Y times as predictive of Y , as X_j is.” Given $\{X_j, k_D, k_Y\}$, this *bounds* both $R_{Y \sim Z | \mathbf{x}, D}^2$ and $R_{D \sim Z | \mathbf{x}}^2$ ¹

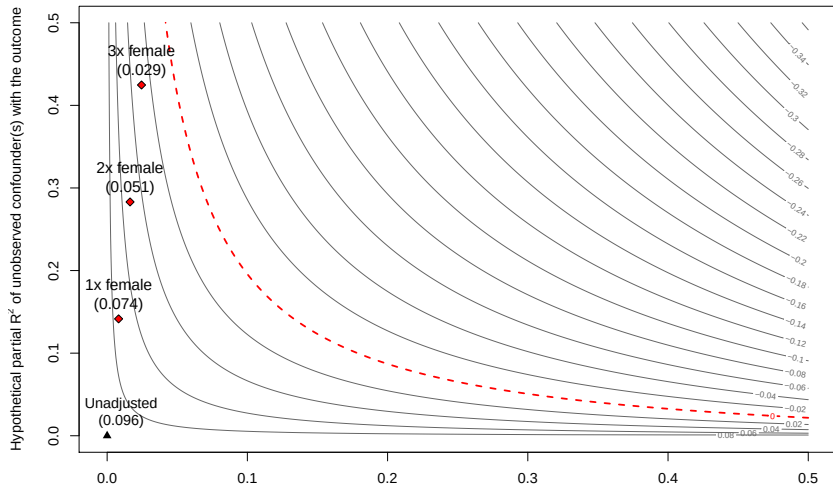
Yields a worst-case bias and so a worst-case adjusted estimate.
Conservative for multiple, non-linear confounders.

This shows the implications of an assumption — you have to argue the assumption itself is believable.

¹This is more complicated than it sounds: the observed R^2 of observed X are themselves biased, including by collider bias from the inclusion of D . But you can construct bounds; see Cinelli & Hazlett 2020.

R^2 contour with bounds: Darfur

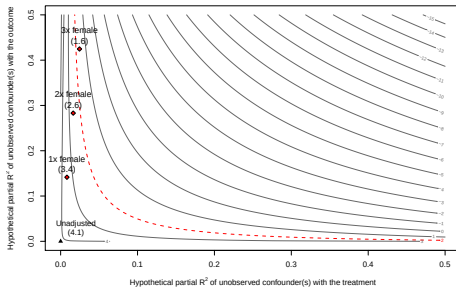
Within village, it is hard to argue *any* basis for targeting could matter more than *female*. So $X_j = \text{female}$, with various k .



Point estimate robust to a confounder $1\times$, $2\times$, even $3\times$ as strong as *female*.

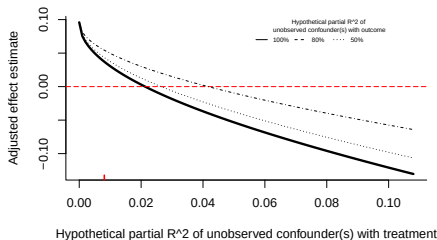
Inference and extreme scenarios

t -statistic contour



t remains significant against a confounder up to $\sim 2 \times$ female; loses sig. at $3 \times$.

Extreme scenario



Assume $R^2_{Y \sim Z | X, D} = 1$. Then Z would still need to explain $\geq 2.2\%$ of D to zero the effect. Pay a real price when you can't bound.

How to argue it, and what readers should demand

Occasionally you might be able to argue in support of a claim, e.g. in Darfur:

*“... aircraft dropped makeshift and unguided bombs over villages, and militia raided without concern for who they would attack — the only known major exception, due to sexual assaults, was targeting gender, which is also one of the most visually apparent characteristics. **A confounder twice as strong as female would be surprising.**”*

Flipside: dismissing a result requires arguing for confounding of a given strength. (And all you need is t and df to get started)

More broadly, remember:

- No magic RV threshold: refine the debate, not hide behind a number.
- This is never “how much feasible confounding is there” — it’s “how much confounding would it take to overturn the conclusion”
- The result might be informative or might show that we can’t say much, but either way it is the honest conversation.

Resources

Papers (all linked from chadhazlett.com/publications):

- Cinelli & Hazlett (2020). *Making sense of sensitivity: Extending omitted variable bias*. JRSS B. — the framework
- Hazlett & Parente (2023). *From “Is it unconfounded?” to “How much confounding would it take?”* JoP. — applied / how-to-argue companion
- Cinelli, Ferwerda & Hazlett (2024). *sensemakr: Sensitivity Analysis Tools for OLS in R and Stata*. Observational Studies. — software paper
- Cinelli & Hazlett (2025). *An OVB framework for sensitivity analysis of IV*. Biometrika. — IV extension (cf. extras)

Software & tutorials:

- `sensemakr` R/Stata + worked vignettes: carloscinelli.com/sensemakr
- Carlos Cinelli's site (talks, tutorials, videos): carloscinelli.com
- Software list on my site: chadhazlett.com/software

Rosenbaum's Γ for matched pairs

- A different sensitivity tradition (Rosenbaum 2002), with a single parameter $\Gamma \geq 1$ quantifying departure from zero confounding.
- Consider matched pairs i and j with $X_i = X_j$. Under SOO, both have the same $\pi(X)$.
- A binary unobserved U may make their odds of treatment differ:

$$\frac{1}{\Gamma} \leq \frac{\pi_i(1 - \pi_j)}{(1 - \pi_i)\pi_j} \leq \Gamma.$$

Formal sensitivity tests (Γ)

- $\Gamma = 1$: no hidden bias. $\Gamma = 2$: unit i has twice the odds of treatment as j despite identical X .
- “Primal” version: only the odds of *treatment* are varied; effect of U on the outcome is taken to be infinite (worst-case).
- At any chosen Γ , run a permutation-style test where assignment odds respect Γ .
- Try increasing Γ until significance is lost or the sign flips.
- R: `psens()`, `hlsens()` in `rbounds`.

Rosenbaum-style sensitivity: smoking example

TABLE 4.1. Sensitivity Analysis for Hammond's Study of Smoking and Lung Cancer: Range of Significance Levels for Hidden Biases of Various Magnitudes.

Γ	Minimum	Maximum
1	< 0.0001	< 0.0001
2	< 0.0001	< 0.0001
3	< 0.0001	< 0.0001
4	< 0.0001	0.0036
5	< 0.0001	0.03
6	< 0.0001	0.1

If a hidden confounder gave the unit with the worse outcome $6\times$ higher odds of smoking, the adjusted estimate would lose significance.

Imbens (2003): parametric approach

- Highly parametric:
 - Assume $Y_1, Y_0 \perp\!\!\!\perp D \mid X, U$ with $U \sim \text{Bernoulli}(0.5)$, $U \perp X$.
 - Logistic propensity: $P(D = 1 \mid X, U) = \frac{\exp(X\theta + \gamma U)}{1 + \exp(X\theta + \gamma U)}$. Here γ governs U – D strength.
 - Outcome conditionally normal: $Y \mid X, U \sim N(\alpha D + X\beta + \delta U, \sigma^2)$. δ governs U – Y strength.
 - Choose (γ, δ) , MLE for $\hat{\alpha}(\gamma, \delta)$.
- Reparameterize (γ, δ) as partial R^2 's with D and with Y for interpretability.

Trades stronger assumptions for more direct probabilistic interpretation.
 Carnegie–Dorie–Hill (2016) extends with BART for $\mu_0(X)$.

OVB with instrumental variables

We can apply the same OVB machinery to IV. Walks through, applied to the Card (1993) returns-to-education example.

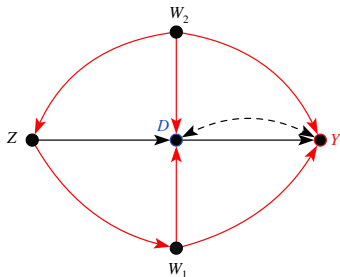
Running example: returns to education (Card 1993)

- Estimating returns to education has long been a challenge.
- OLS with $N = 3010$, covariates $X = \{\text{race, experience, regional}\}$ suggests 7.5% higher earnings per year of education.
- Confounding may be severe: ability, motivation, family/social.
- Card uses *proximity to college* as instrument.

These assumptions may fail!

The IV estimate can be biased — worse than OLS — if *Proximity* is not a valid instrument due to uncontrolled:

- “side-effects” (e.g. high-wage jobs near colleges)
- “common-cause confounders” (e.g. unmeasured SES)



How “bad” would such violations need to be to change the conclusion?

IV regression: three estimators

1. Indirect least squares:

Reduced form: $\text{Earnings} = \hat{\lambda}_{\text{res}} \text{Proximity} + \mathbf{X}\hat{\beta}_{\text{res}} + \hat{\epsilon}_{y,\text{res}}$

First stage: $\text{Education} = \hat{\theta}_{\text{res}} \text{Proximity} + \mathbf{X}\hat{\psi}_{\text{res}} + \hat{\epsilon}_{d,\text{res}}$

2. Two-stage least squares:

2nd stage: $\text{Earnings} = \hat{\tau}_{2\text{SLS},\text{res}} \widehat{\text{Education}} + \mathbf{X}\hat{\beta}_{2\text{SLS},\text{res}} + \hat{\epsilon}_{2\text{SLS},\text{res}}$

Same answer:

$$\hat{\tau}_{2\text{SLS}} = \hat{\tau}_{\text{ILS}} := \frac{\hat{\lambda}_{\text{res}}}{\hat{\theta}_{\text{res}}} = \frac{0.042}{0.319} \approx 0.132.$$

IV regression: Anderson–Rubin

3. Anderson–Rubin (Fieller):

$$\text{Earnings} - \tau_0 \text{Education} = \hat{\phi}_{\tau_0, \text{res}} \text{Proximity} + \mathbf{X} \hat{\beta}_{\tau_0, \text{res}} + \hat{\epsilon}_{\tau_0, \text{res}}$$

- τ_0 that makes $\hat{\phi}_{\tau_0, \text{res}} = 0$ is the point estimate
- $\{\tau_0 : \text{do not reject } \hat{\phi}_{\tau_0, \text{res}} = 0\}$ is the CI
- Correct coverage regardless of instrument strength

All give 13.2% higher earnings per year of education.

How would these change had we included W ?

Start with the reduced form

We wanted to run

$$Y = \hat{\lambda} Z + \mathbf{X}\hat{\beta} + \hat{\gamma} W + \hat{\epsilon}_y.$$

But W was not observed, so we ran

$$Y = \hat{\lambda}_{\text{res}} Z + \mathbf{X}\hat{\beta}_{\text{res}} + \hat{\epsilon}_{y,\text{res}}.$$

How would including W change inference about λ ?

Same OVB analysis, different variable names

- $R^2_{Y \sim W|Z, \mathbf{X}}$: W 's share of outcome variance
- $R^2_{Z \sim W|\mathbf{X}}$: W 's share of *instrument* variance

$$\begin{aligned} |\widehat{\text{bias}}| &= \sqrt{\frac{R^2_{Y \sim W|Z, \mathbf{X}} R^2_{Z \sim W|\mathbf{X}}}{1 - R^2_{Z \sim W|\mathbf{X}}}} \text{ df} \cdot \widehat{\text{se}}(\hat{\lambda}_{\text{res}}) \\ &= \text{BF} \sqrt{\text{df}} \cdot \widehat{\text{se}}(\hat{\lambda}_{\text{res}}) \end{aligned}$$

$$\begin{aligned} \widehat{\text{se}}(\hat{\lambda}) &= \sqrt{\frac{1 - R^2_{Y \sim W|Z, \mathbf{X}}}{1 - R^2_{Z \sim W|\mathbf{X}}}} \frac{\text{df}}{\text{df} - 1} \cdot \widehat{\text{se}}(\hat{\lambda}_{\text{res}}) \\ &= \text{SeF} \sqrt{\text{df} / (\text{df} - 1)} \cdot \widehat{\text{se}}(\hat{\lambda}_{\text{res}}) \end{aligned}$$

An embarrassingly simple sensitivity analysis

Change the critical t rather than the point estimate.

Familiar $1 - \alpha$ CI:

$$\hat{\lambda}_{\text{res}} \pm t_{\alpha, \text{df}}^* \cdot \widehat{\text{se}}(\hat{\lambda}_{\text{res}}).$$

Adjust for confounding (limited by R^2):

$$\hat{\lambda}_{\text{res}} \pm t_{\alpha, \text{df} - 1, R^2}^\dagger \cdot \widehat{\text{se}}(\hat{\lambda}_{\text{res}}),$$

where

$$t_{\alpha, \text{df} - 1, R^2}^\dagger := \text{SeF} \sqrt{\text{df} / (\text{df} - 1)} \cdot t_{\alpha, \text{df} - 1}^* + \text{BF} \sqrt{\text{df}}.$$

Routine reporting: (extreme) robustness values

What would it take to reduce the coefficient until we no longer reject (with confidence $1 - \alpha$) that it is $100q\%$ as large?

Using the RF coefficient:

- $RV_{q,\alpha}(\lambda)$: minimal strength of association with both Z and Y needed.
- $XR_{V_{q,\alpha}}(\lambda)$: minimal strength with Z alone, even if W explains *all* residual outcome variance.

Sensitivity of RF and FS

What can we learn about IV from sensitivity of RF and FS? Recall $\tau = \lambda/\theta$.

Sensitivity of reduced form: for $H_0 : \tau = 0$, sensitivity of RF *is exactly* sensitivity of IV.

“If it ain’t in the reduced form, it ain’t there.” (Angrist & Pischke)

Sensitivity of first stage: If the sign of the FS isn’t guaranteed, you cannot rule out unbounded estimates / CIs.
AR is consistent with this logic.

Sensitivity with the AR regression

$$Y_{\tau_0} = \hat{\phi}_{\tau_0, \text{res}} Z + \mathbf{X} \hat{\beta}_{\tau_0, \text{res}} + \hat{\epsilon}_{\tau_0, \text{res}}$$

The AR confidence interval:

$$\text{CI}_{1-\alpha}(\tau) = \left\{ \tau_0 : t_{\hat{\phi}_{\text{res}, \tau_0}}^2 \leq (t_{\alpha, \text{df}-1}^*)^2 \right\}.$$

Replace $t_{\alpha, \text{df}-1}^*$ with $t_{\alpha, \text{df}-1, \mathbf{R}^2}^{\dagger \max}$:

$$\text{CI}_{1-\alpha, \mathbf{R}^2}^{\max}(\tau) = \left\{ \tau_0 : t_{\hat{\phi}_{\text{res}, \tau_0}}^2 \leq (t_{\alpha, \text{df}-1, \mathbf{R}^2}^{\dagger \max})^2 \right\}.$$

Analytical solution via quadratic inequality. Yields $\text{RV}_{q, \alpha}(\tau)$, $\text{XRV}_{q, \alpha}(\tau)$ and $\text{RV}_{\geq q, \alpha}(\tau)$, $\text{XRV}_{\geq q, \alpha}(\tau)$.

Card example: sensitivity of RF and FS

1. Reduced form: $H : \tau = 0$.

Outcome: *Earnings* (log)

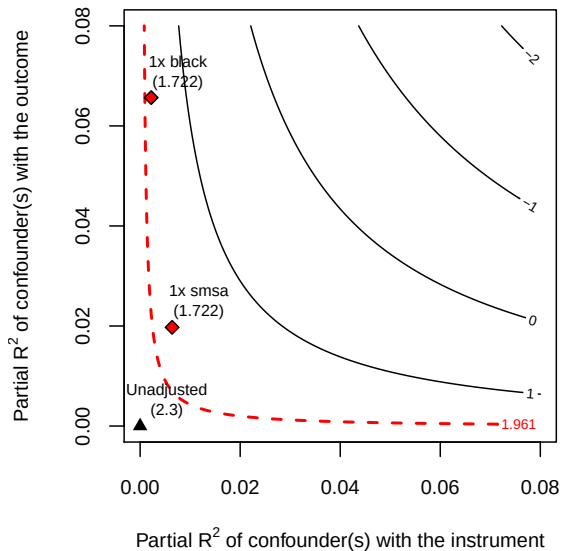
Instrument	Est.	SE	t	$R^2_{Y \sim Z X}$	$XRV_{q^*, \alpha}$	$RV_{q^*, \alpha}$
<i>Proximity</i>	0.042	0.018	2.33	0.18%	0.05%	0.67%
<i>Bound</i> ($1 \times smsa$): $R^2_{Y \sim W Z, X} = 2\%$, $R^2_{Z \sim W X} = 0.6\%$, $t^\dagger = 2.55$ df = 2994, $q^* = 1$, $\alpha = 0.05$						

2. First stage: stability.

Outcome: *Education* (years)

Instrument	Est.	SE	t	$R^2_{D \sim Z X}$	$XRV_{q^*, \alpha}$	$RV_{q^*, \alpha}$
<i>Proximity</i>	0.32	0.088	3.64	0.44%	0.31%	3.02%
<i>Bound</i> ($1 \times smsa$): $R^2_{D \sim W Z, X} = 0.5\%$, $R^2_{Z \sim W X} = 0.6\%$, $t^\dagger = 2.26$						

Card example: reduced form contours



Card example: sensitivity of the IV (AR)

Outcome: *Earnings* (log)

Treatment	Est.	LL	UL	t	$XRV_{\geq q^*, \alpha}$	$RV_{\geq q^*, \alpha}$
<i>Education</i>	0.132	0.025	0.285	2.33	0.05%	0.67%

Bound ($1 \times \text{smsa}$): $R^2_{Y_0 \sim W|Z, X} = 2\%$, $R^2_{W \sim Z|X} = 0.6\%$, $t^\dagger = 2.55$

df = 2994, $q^* = 1$, $\alpha = 0.05$

Table: Minimal sensitivity reporting of IV (AR)

- Identical to RF results because the null is $\tau = 0$.
- More interesting under other null hypotheses.

Card example: CI-limit contours

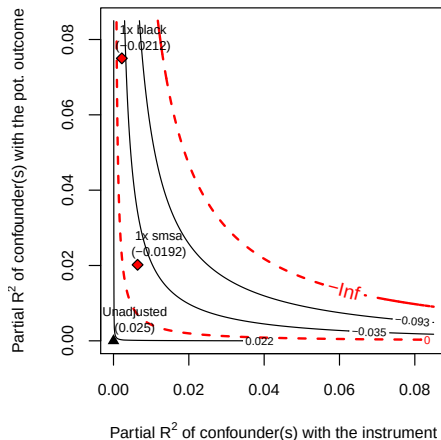


Figure: Sensitivity contours: lower 95% limit

IV sensitivity: conclusions

- OLS sensitivity to confounding can be as simple as adjusting your critical value.
- For null of zero in IV, sensitivity of the reduced form *is all you need*.
- General sensitivity for IV is also easy: in the AR framework, explore how postulated confounding moves point estimates, CI limits, RVs, and bounds.