Lecture 10: Meta-Research

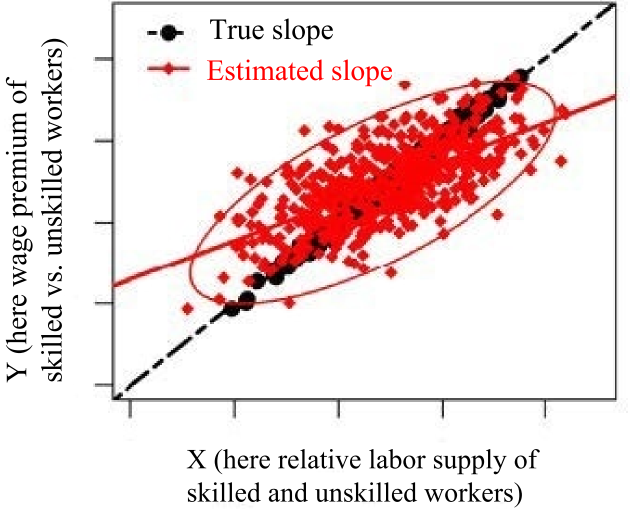
Tomas Havranek

Charles University, CEPR, METRICS

Research Synthesis in Economics and Finance,
University of Canterbury

March 18, 2025

Example 1: Attenuation bias, aka regression dilution

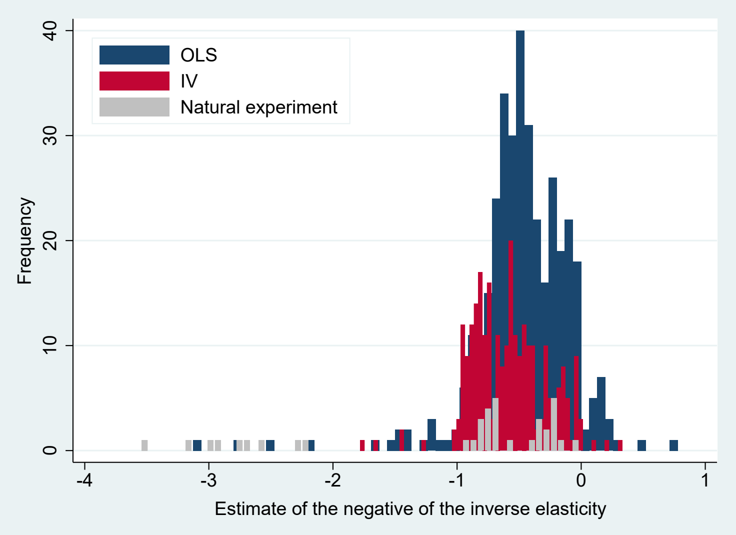


Important complication

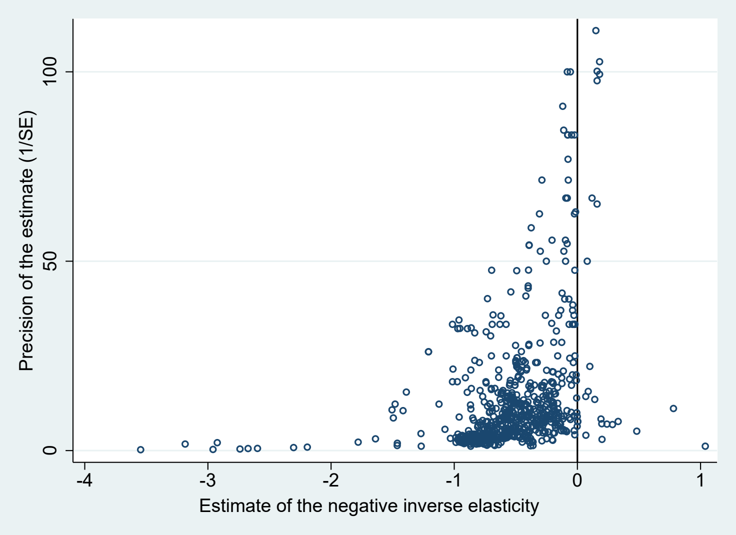
Researchers estimate negative inverse elasticity: they regress skill premium on relative labor supply (not vice versa!)

- Measures substitutability between skilled (college) and unskilled (high school) workers
- Important for wage inequality, explaining cross-country differences in productivity, . . .
- 682 estimates from 77 studies, "consensus" elasticity = 1.5 (negative inverse = -0.67)

Three estimation frameworks



Strong publication bias (or *p*-hacking?)



OLS: no effect (but attenuation & endogeneity biases)

Panel A: OLS estimates					
	FE	BE	IV	EK	AK
Publication bias	-5.804***	-4.277***	-6.962***	-5.465***	P=0.468
	(1.999)	(1.266)	(1.694)	(0.540)	(0.139)
			[-11.770, -2.494]		
			{-11.972, -3.133}		
Effect beyond bias	-0.0207	-0.0965	0.0103	-0.0361**	-0.289**
	(0.103)	(0.0627)	(0.104)	(0.0191)	(0.113)
			[-0.331, 0.214]		
First-stage robust F-stat			46.17		
Observations	347	347	251	347	347

IV: strong effect (no biases beyond publication)

Panel B: IV estimates					
	FE	BE	IV	EK	AK
Publication bias	-2.287**	-0.923	-0.553	-1.485***	P=0.336
	(0.843)	(1.365)	(0.681)	(0.268)	(0.093)
			[-1.913, 1.078]		
			{-1.991, 0.748}		
Effect beyond bias	-0.149	-0.297**	-0.400***	-0.252***	-0.333***
	(0.109)	(0.115)	(0.114)	(0.0246)	(0.058)
			[-0.719, 0.175]		
First-stage robust F-stat			69.98		
Observations	264	264	212	264	264

Natural experiments: no effect (but attenuation bias)

Panel C: Natural experiment estimates					
	FE	BE	IV	EK	AK
Publication bias	-3.557***	-1.874*	-3.176***	-3.115***	P=0.187
	(0.0178)	(0.682)	(0.853)	(0.343)	(0.075)
			[-4.854, -1.407]		
			{-4.653, -1.444}		
Effect beyond bias	0.0496***	-0.121	-0.00307	0.00302	-0.009
	(0.00246)	(0.0824)	(0.0297)	(0.0280)	(0.066)
			[NA, NA]		
First-stage robust F-stat			260.41		
Observations	40	40	40	40	40

Publication bias $>$ attenuation bias

- $|\text{IV}| > |\text{OLS}| \rightarrow$ attenuation or other endogeneity biases important
- OLS = Natural experiments $\rightarrow$ other endogeneity biases negligible

$\rightarrow$ $|\text{IV} - \text{OLS}|$ measures attenuation bias, which is strong

- Uncorrected inverted elasticity = -0.67
- Inverted elasticity corrected for pub and attenuation bias (IV) = -0.25

$\rightarrow$ Pub bias dominates attenuation

$\rightarrow$ Implied elasticity = 4

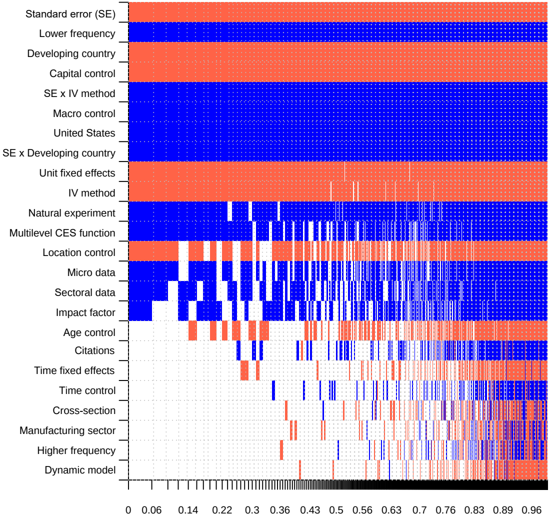
Assumptions of Andrews & Kasy not satisfied

Table C7: Specification test for the Andrews & Kasy (2019) model

	All inverse	OLS method	IV method	Natural experiment	Developed country
Correlation	0.515	0.265	0.635	0.947	0.558
	[0.436, 0.605]	[0.160, 0.392]	[0.499, 0.743]	[0.928, 0.976]	[0.462, 0.667]
Observations	654	347	264	40	418
	Developing country	**Country estimate**	**Region estimate**	**One-level CES function**	**Multilevel CES function**
Correlation	0.330	0.856	0.476	0.199	0.592
	[0.159, 0.482]	[0.775, 0.95]	[0.384, 0.568]	[0.047, 0.342]	[0.498, 0.709]
Observations	151	555	93	198	444

Notes: Following Kranz & Putz (2022), the table shows, for various subsets of the literature, the correlation coefficient between the logarithm of the absolute value of the estimated inverse elasticity and the logarithm of the corresponding standard error, weighted by the inverse publication probability estimated by the Andrews & Kasy (2019) model. If the assumptions of the model hold, the correlation is zero. Bootstrapped 95% confidence interval in parentheses.

Model uncertainty; biases confirmed



Best practice estimate around 4

Table 5: Implied elasticities

	Subjective best practice	Autor (2014)	Card (2009)	Carneiro *et al.* (2022)
All countries	-0.27	-0.13	-0.24	0.05
	(-0.48, -0.05)	(-0.24, -0.02)	(-0.39, -0.09)	(-0.12, 0.23)
	$\sigma = 3.7$	$\sigma = 7.7$	$\sigma = 4.2$	$\sigma = -18.4$
USA	-0.16	-0.02	-0.13	0.16
	(-0.38, 0.06)	(-0.12, 0.07)	(-0.28, 0.02)	(-0.02, 0.34)
	$\sigma = 6.3$	$\sigma = 45.0$	$\sigma = 7.8$	$\sigma = -6.2$
Developing countries	-0.47	-0.33	-0.44	-0.15
	(-0.70, -0.24)	(-0.47, -0.19)	(-0.60, -0.27)	(-0.33, 0.04)
	$\sigma = 2.1$	$\sigma = 3.0$	$\sigma = 2.3$	$\sigma = 6.8$

Main Findings

1. Attenuation bias is important ($|\text{IV}| > |\text{OLS}| = |\text{natural experiments}|$).
2. Publication bias is more important (inverse elasticity: simple mean -0.67 ; corrected IV mean -0.25).
3. Implied elasticity around 4 (compared to consensus 1.5).

Project Website

www.meta-analysis.cz/skill

Example 2: Too much heterogeneity?

Common measure: estimate the relative heterogeneity of underlying effects *(Higgins and Thompson, 2002)*

$$I^2 \equiv \frac{\text{underlying variance of effects}}{\text{observed variance of estimates}}$$

- $I^2 = 0.75 \sim 1.0$ . . . *"considerable heterogeneity" in medicine (Cochrane Review Guideline)*
- $I^2 \simeq 0.9$ . . . common in economics meta-analyses (e.g., *Vivalt, 2021*)

Unresolved concern:

- no universally agreed criteria for "too much" heterogeneity
- undesirable pattern: with more precise studies, higher I^2

This Paper

1. Propose a new test for external validity in meta-analysis:
   - Based on the tail index of the underlying distribution, without contamination by estimate precision.
   - Test whether the implied variance of underlying effect is infinite (i.e., too heterogeneous).

2. Show that the implied variance is infinite in meta-analysis of 126 nudge RCTs *(DellaVigna and Linos, 2020)*:
   - **Nudge:** “choice architecture that alters people’s behavior in a predictable way without forbidding any options or significantly changing their economic incentives.” *(Thaler and Sunstein, 2008)*
   - The possibility of infinite variance can be important in real-world applications.

Pareto tail:

$$F(b) = 1 - Cb^{-\alpha}, \quad b > b_{\min} \quad \text{and} \quad C > 0$$

- e.g. earthquakes, word frequency, city size, wealth distribution
- If $\alpha < 2$, then $Var(b) = \infty$ because

$$\int_{b_{\min}}^{\infty} b^2 f(b)\, db \propto \int_{b_{\min}}^{\infty} b^2 b^{-3}\, db = \int_{b_{\min}}^{\infty} b^{-1}\, db = \log b \Big|_{b_{\min}}^{\infty} = \infty$$

Tail Index Test of External Validity:

- **Criteria 1:** if $\hat{\alpha} > 2$, then pass the test
- **Criteria 2:** if $\mathbb{P}(\hat{\alpha} \leq 2) < p$, then pass the test

Estimation Methods (i) Inner Problem of Tail Index

Following *Hill (1975)*, use Maximum Likelihood Estimation for $\hat{\alpha}_{MLE}$: given $b_{\min}$,

$$\hat{\alpha}_{MLE} = \arg\max_{\alpha} \sum_{i=1}^{n} \ln l\left(\alpha \mid \beta_i, SE_i, b_{\min}\right),$$

where

$$l\left(\alpha \mid \beta_i, SE_i, b_{\min}\right) \equiv \int_{\theta_{\min}}^{\infty} \frac{\phi\left(\frac{\beta_i - \theta}{SE_i}\right)}{1 - \Phi\left(\frac{b_{\min} - \theta}{SE_i}\right)} \alpha b_{\min}^{\alpha} b^{-(\alpha+1)}\, db.$$

- Numerical integration

Estimation Methods (ii) Outer Problem of Cut-off

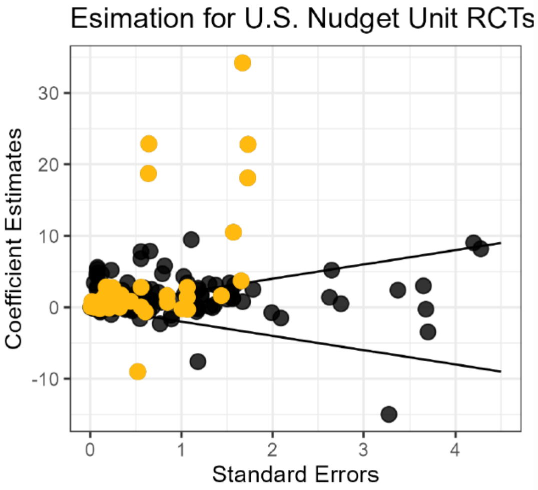
Following *Clauset et al. (2006)*, use the goodness-of-fit metric to choose $b^*_{\min}$:

$$b^*_{\min} \in \arg\min_{b_{\min}} \tilde{K}S\,(b_{\min} \mid \hat{\alpha}_{MLE})\,,$$

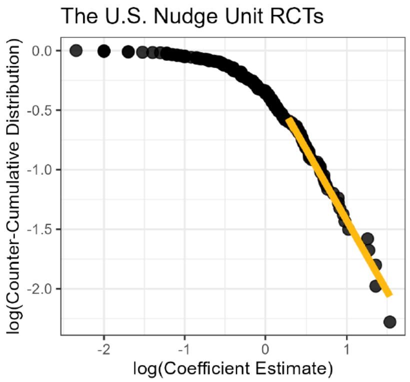
where the reweighted Kolmogorov-Smirnov statistic is

$$\tilde{K}S\,(b_{\min} \mid \hat{\alpha}_{MLE}) \equiv \sup_{\beta} \frac{\left| G(\beta \mid \beta \geq b_{\min}) - \hat{G}(\beta \mid \cdot) \right|}{\sqrt{\hat{G}(\beta \mid \cdot)\left[1 - \hat{G}(\beta \mid \cdot)\right]}},$$

and the estimated distribution is $\hat{G}(\beta \mid \hat{\alpha}_{MLE}, \beta \geq b_{\min})$ and the empirical distribution is $G(\beta \mid \beta \geq b_{\min})$.


Esimation for U.S. Nudget Unit RCTs
30
20
10
0
-10
Coefficient Estimates
0
1
2
3
4
Standard Errors

Nudges



$$\log\left(1-G\left(\beta\right)\right)_i = -\underbrace{1.19}_{=\hat{\alpha}} \log \beta_i - \log \underbrace{1.90}_{=\hat{b}_{\text{min}}}$$

Summary

1. Proposed the concept of *"infinite variance"* as representing the persistent criticism that *"too much" heterogeneity* underlies the overall mean estimate.

2. Developed an estimation method by extending the maximum likelihood estimation and goodness-of-fit statistics commonly used in tail estimation.

3. Applied the estimation method and its implied test to show that the influential nudge meta-analysis exhibits the infinite variance.