Spurious Precision in Meta-Analysis

(of Observational Research)

Z. Irsova[1] P. Bom[2] T. Havranek[1,3] H. Rachinger[4]

[1]Charles University, Prague
[2]University of Deusto, Bilbao
[3]Centre for Economic Policy Research, London
[4]University of the Balearic Islands, Palma

7 September 2023

Meta-Research Innovation Center at Stanford

In a nutshell

Main Message

1. Researchers may report artificially high precision.
2. Then standard meta-analysis methods fail.
3. Solution: Meta-Analysis Instrumental Variable Estimator (MAIVE).

Project Website

meta-analysis.cz/maive

STAR experiment: class size and learning

Table: Regression of individual test scores on class size

	(1)	(2)	(3)	(4)
Small class (treatment)	4.82	5.37	5.36	5.37
	(2.19)	(1.26)	(1.21)	(1.19)
White/Asian			8.35	8.44
			(1.35)	(1.36)
Girl			4.48	4.39
			(0.63)	(0.63)
Free lunch			-13.15	-13.07
			(0.77)	(0.77)
White teacher				-0.57
				(2.1)
Teacher experience				0.26
				(0.10)
Master's degree				-0.51
				(1.06)
School intercepts	No	Yes	Yes	Yes
Sample	5,861	5,861	5,861	5,861

Notes: Adapted from Krueger (1999). Dependent variable: test score percentile.
Standard errors in parentheses.

Many ways to compute the standard error

Table: Regression of individual test scores on class size

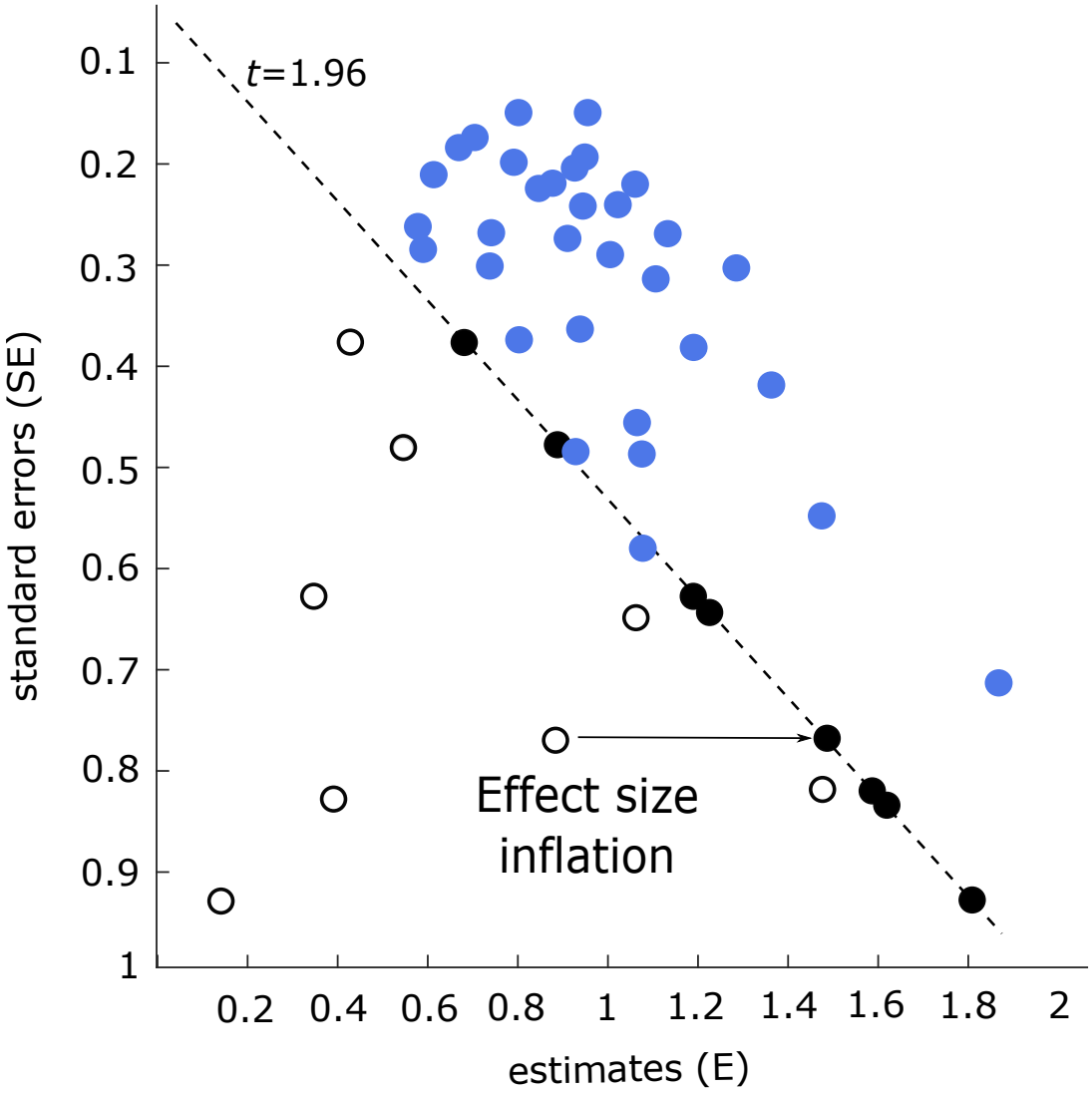
	Krueger (1999)	Replications using different computations of SE				
	(1)	(2) Bootstrap of class clusters	(3) Class clusters	(4) School clusters	(5) Huber-White SE	(6) Plain vanilla SE
Small class	4.82	4.71	4.71	4.71	4.71	4.71
	(2.19)	(2.00)	(1.88)	(1.38)	(0.79)	(0.76)
Sample	5,861	5,743	5,743	5,743	5,743	5,743

Notes: Dependent variable: test score percentile. Standard errors (SE) in parentheses.

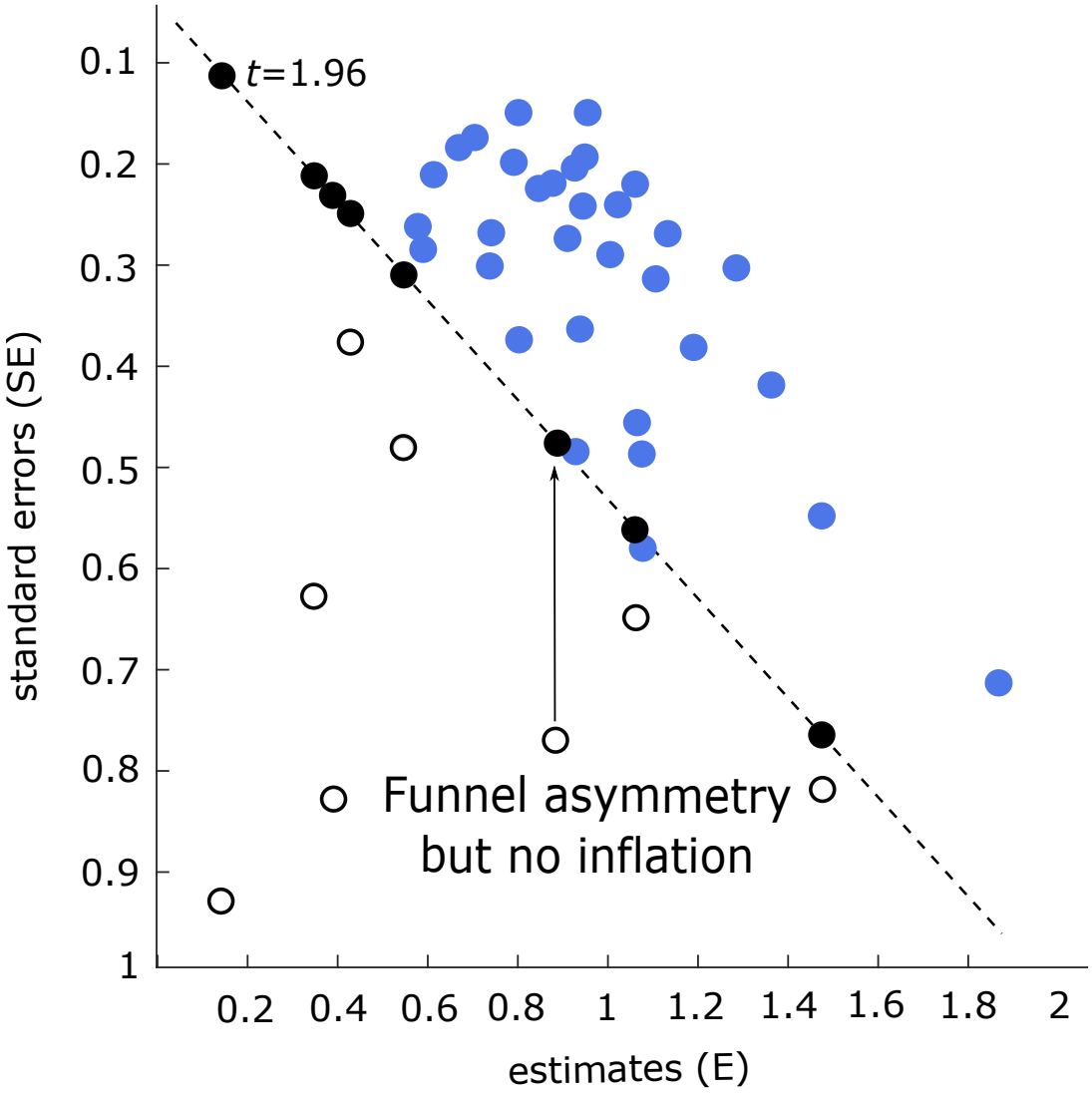
Meta-analysis of the class size effect

Available at meta-analysis.cz/class. Spoiler: effect size tiny.

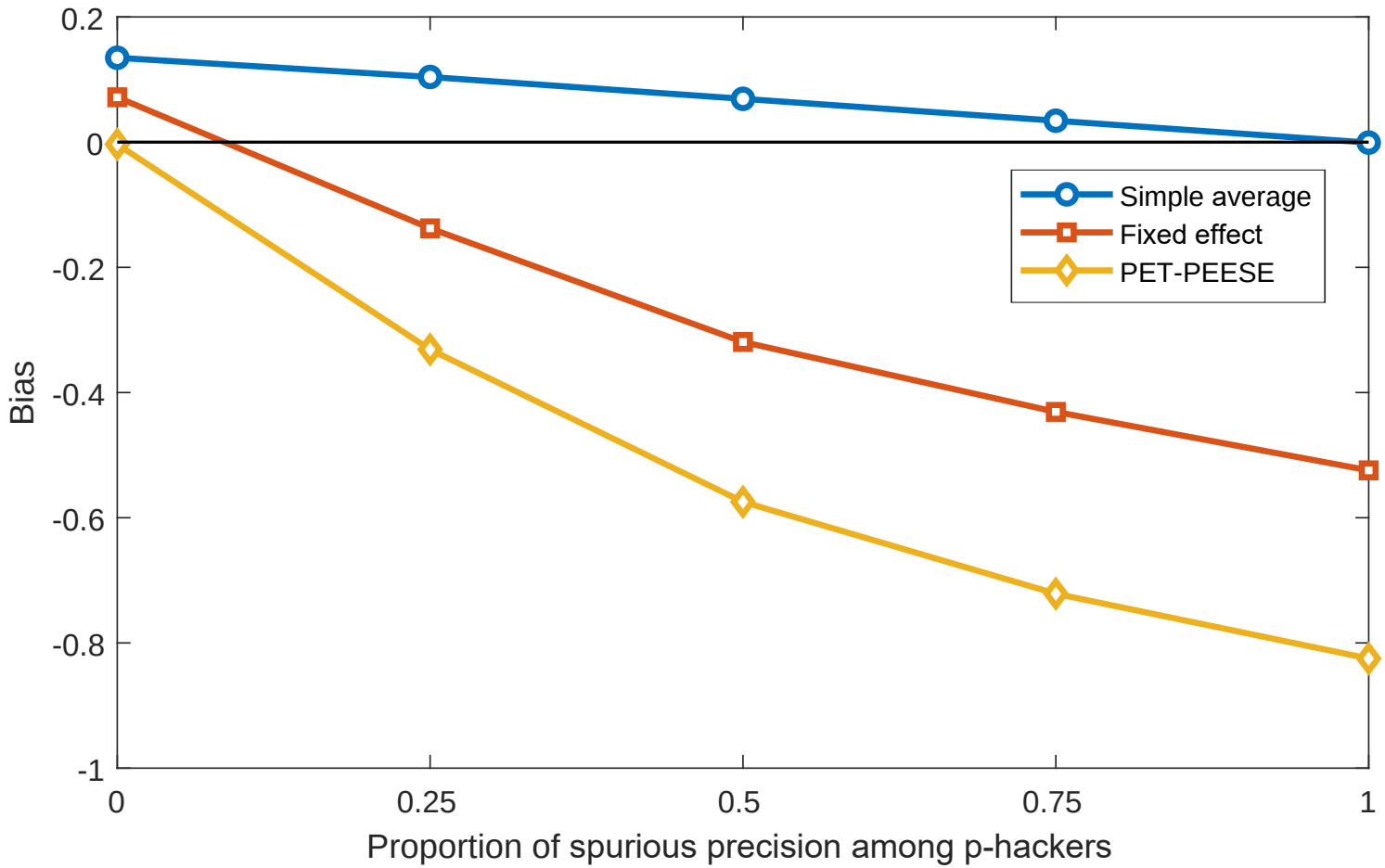
Conventional p-hacking



Spurious precision



Repercussions: bias



Meta-Analysis Instrumental Variable Estimator

Take the PEESE regression without inverse-variance weigths:

$$\hat{E}_i = E_0 + \beta SE(\hat{E})_i^2 + v_i, \tag{1}$$

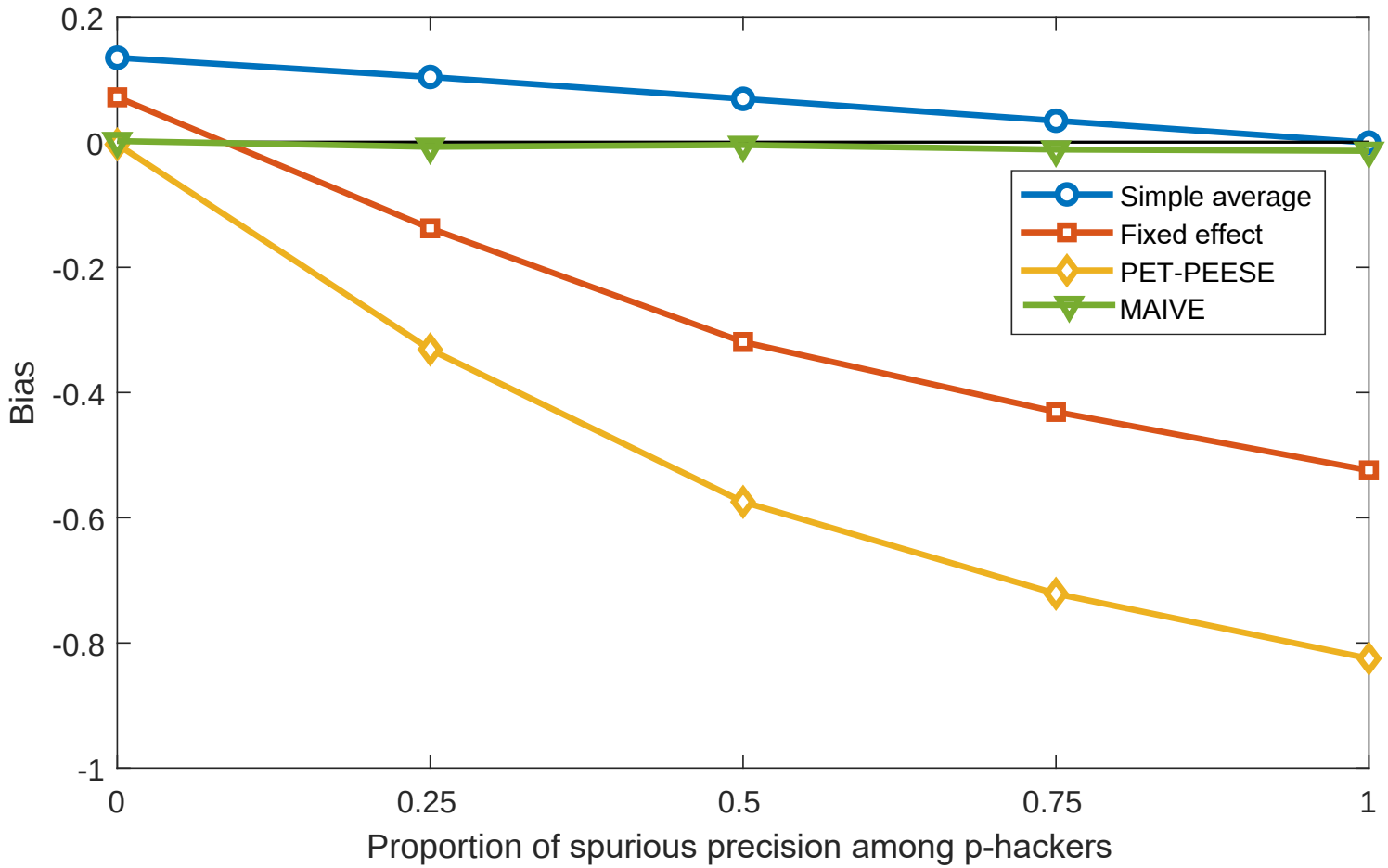
But instead of reported SE, use the portion of SE that can be explained by sample size $[\hat{\alpha_0} + \hat{\alpha_1}(1/N_i)]$:

$$SE(\hat{E})_i^2 = \alpha_0 + \alpha_1(1/N_i) + \pi_i, \tag{2}$$

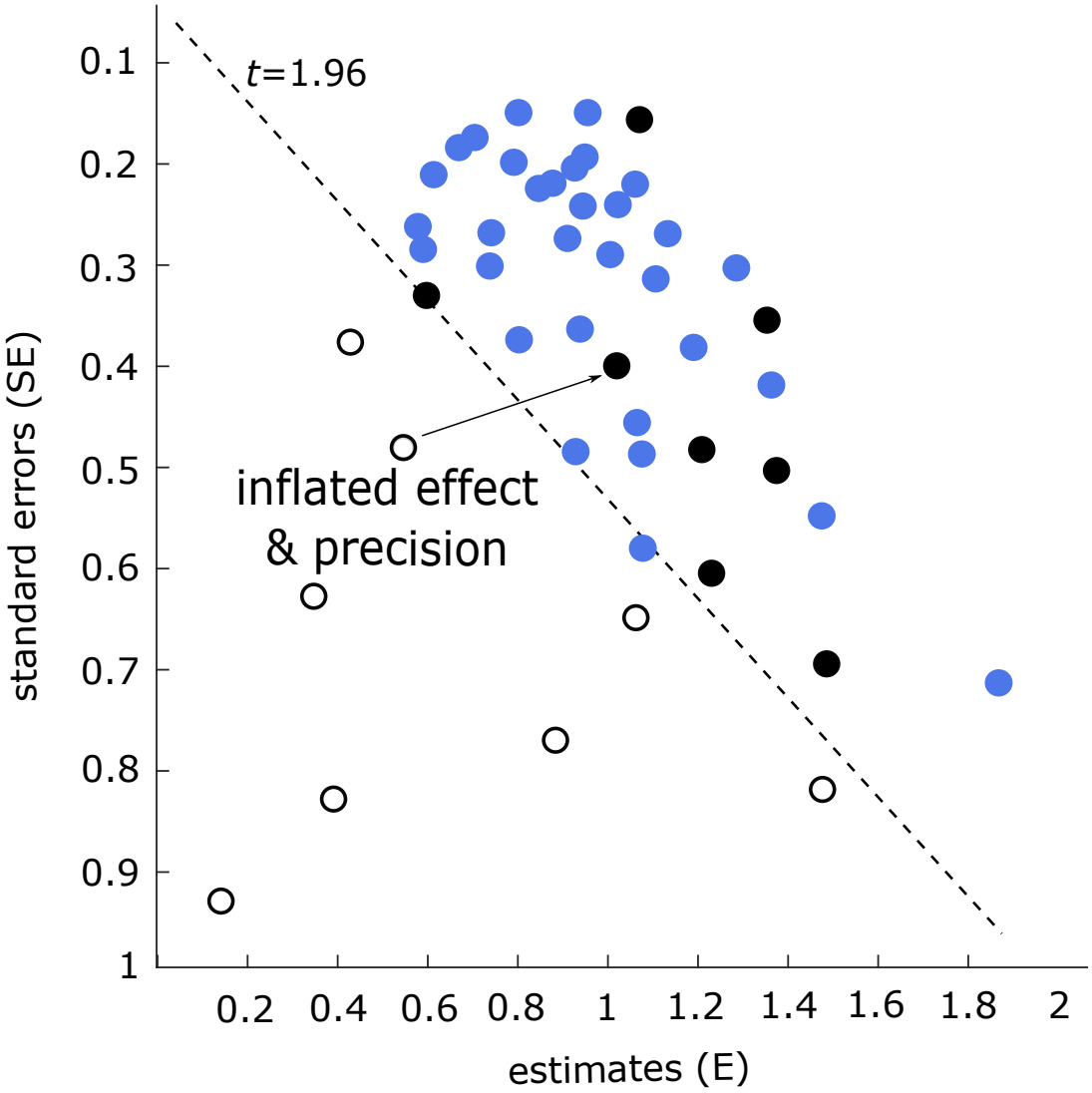
Why not simply replace SE^2 with 1/N?

1. SE influenced by many things. More instruments possible.
2. N imperfect predictor of SE $\rightarrow$ need correct meta SEs.
3. Replacement a special case of the instrumental approach.

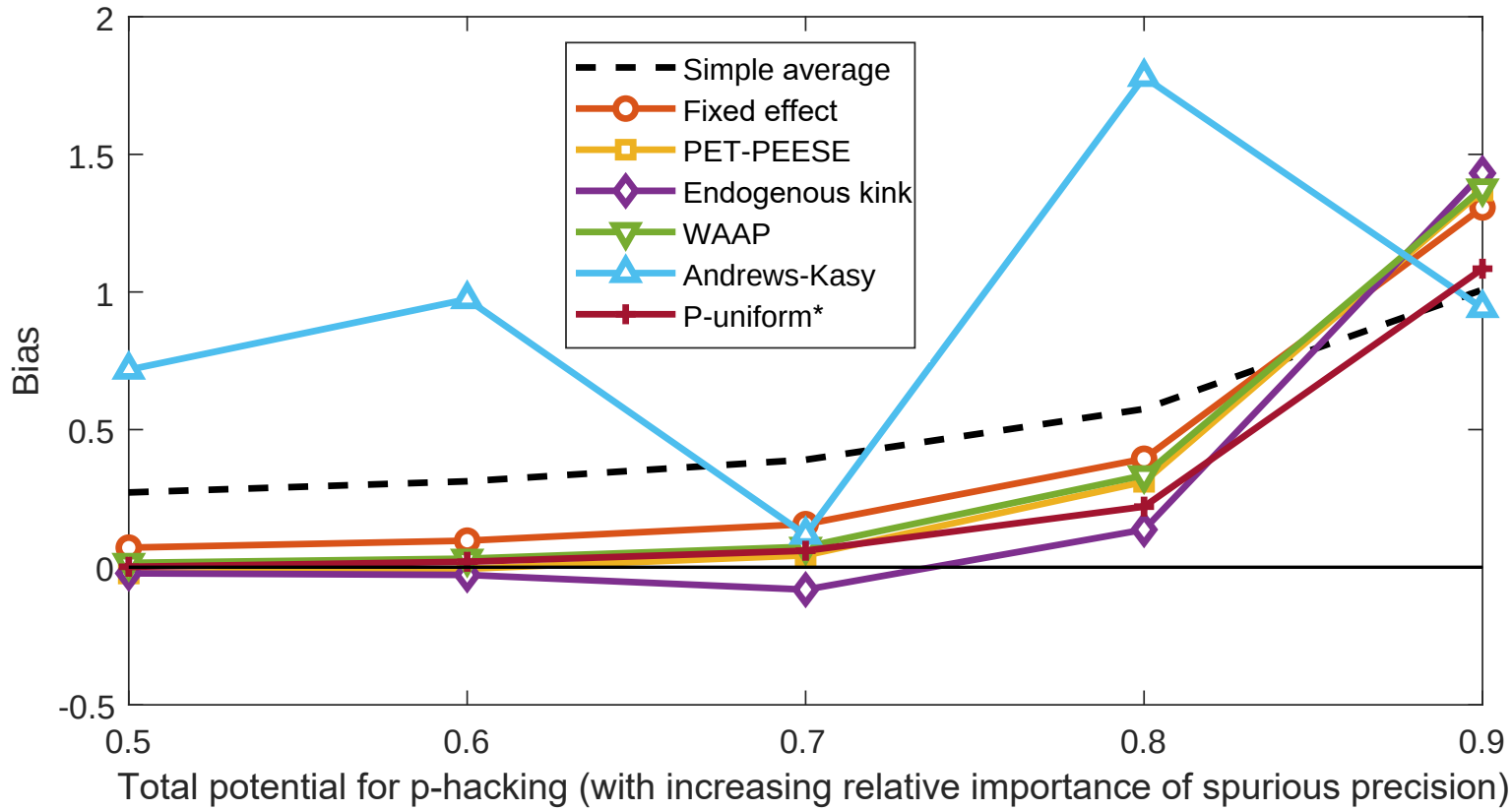
MAIVE solves the problem



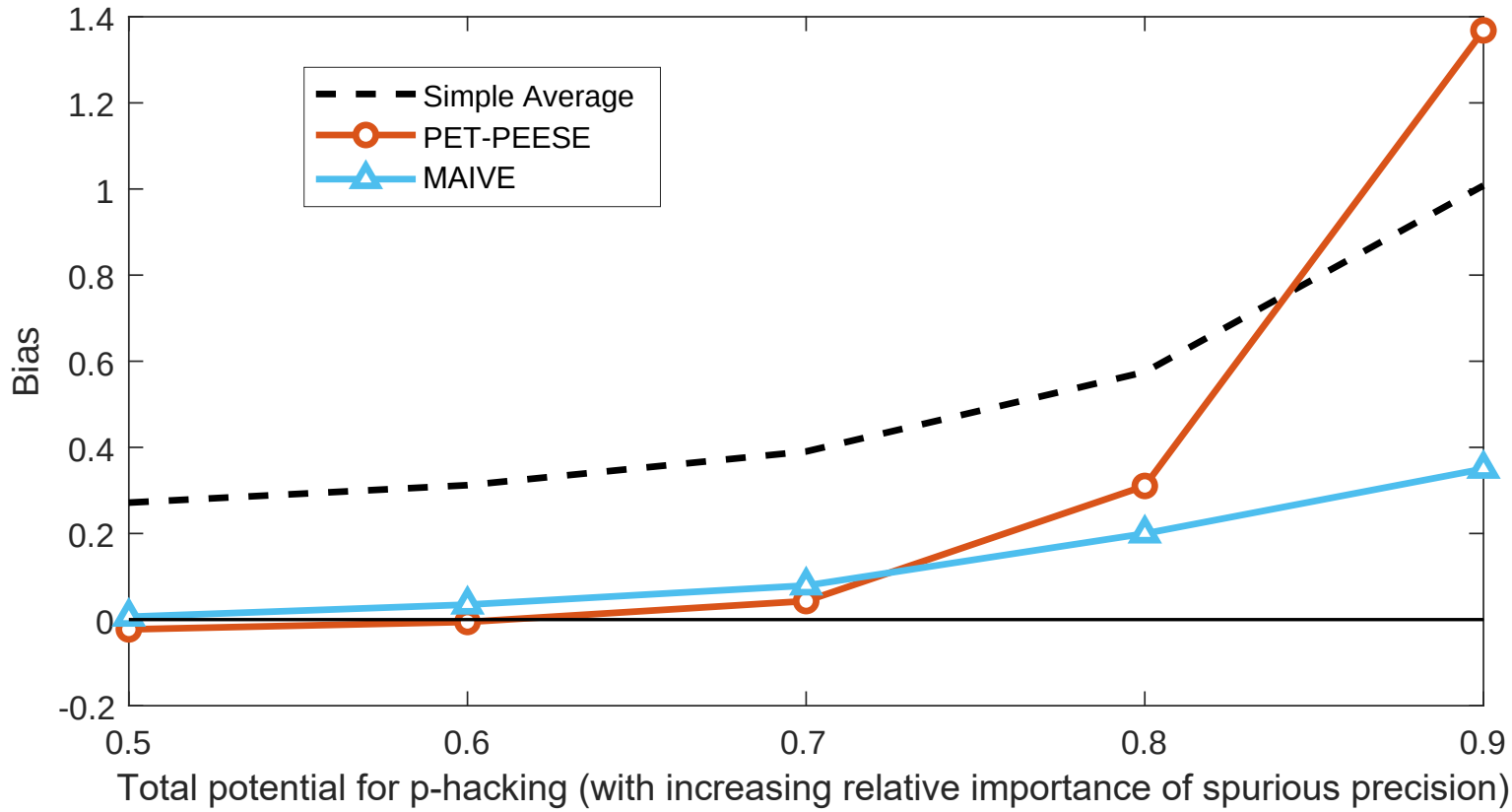
What if both types of p-hacking interact?



Standard estimators still biased



MAIVE still works (but imperfectly)



Next steps

Working on (revision & sequel):

1. Add more complexities to simulations.
2. Check whether MAIVE helps explain differences between meta-analyses and multilab replications.
3. Compare the extent of spurious precision across fields.

Meta-research group in Prague

meta-analysis.cz

Our team's recent meta-analyses

Matousek, J., T. Havranek, & Z. Irsova (2022): Individual Discount Rates: A Meta-Analysis of Experimental Evidence.
Experimental Economics 25: 318-358.

Elminejad, A., T. Havranek, & Z. Irsova (2023): Relative Risk Aversion: A Meta-Analysis.
meta-analysis.cz/risk

Cala, P., T. Havranek, Z. Irsova, J. Matousek (2023): Financial Incentives and Performance: A Meta-Analysis of Economics Evidence.
meta-analysis.cz/incentives