Lecture 2: Applied Meta-Analysis Examples

Tomas Havranek

Charles University, CEPR, METRICS

Research Synthesis in Economics and Finance,
University of Canterbury

February 18, 2025

Example 1: Daylight saving time

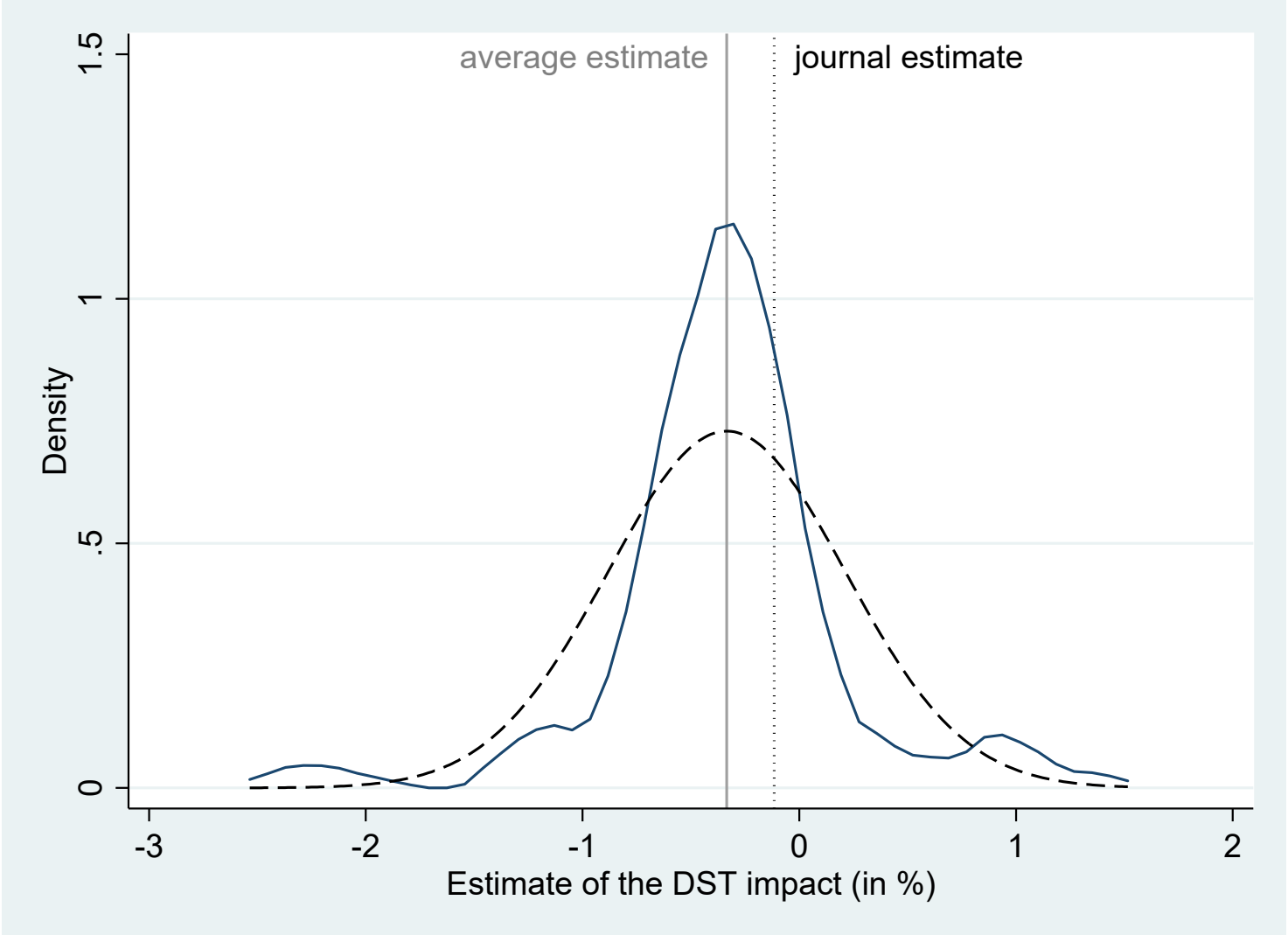
When using econometric analysis, researchers typically estimate the following model:

$$ln\text{Consumption}_t = \alpha + DST \cdot \text{Treatment effect}_t + \text{Controls}_t + \epsilon.$$

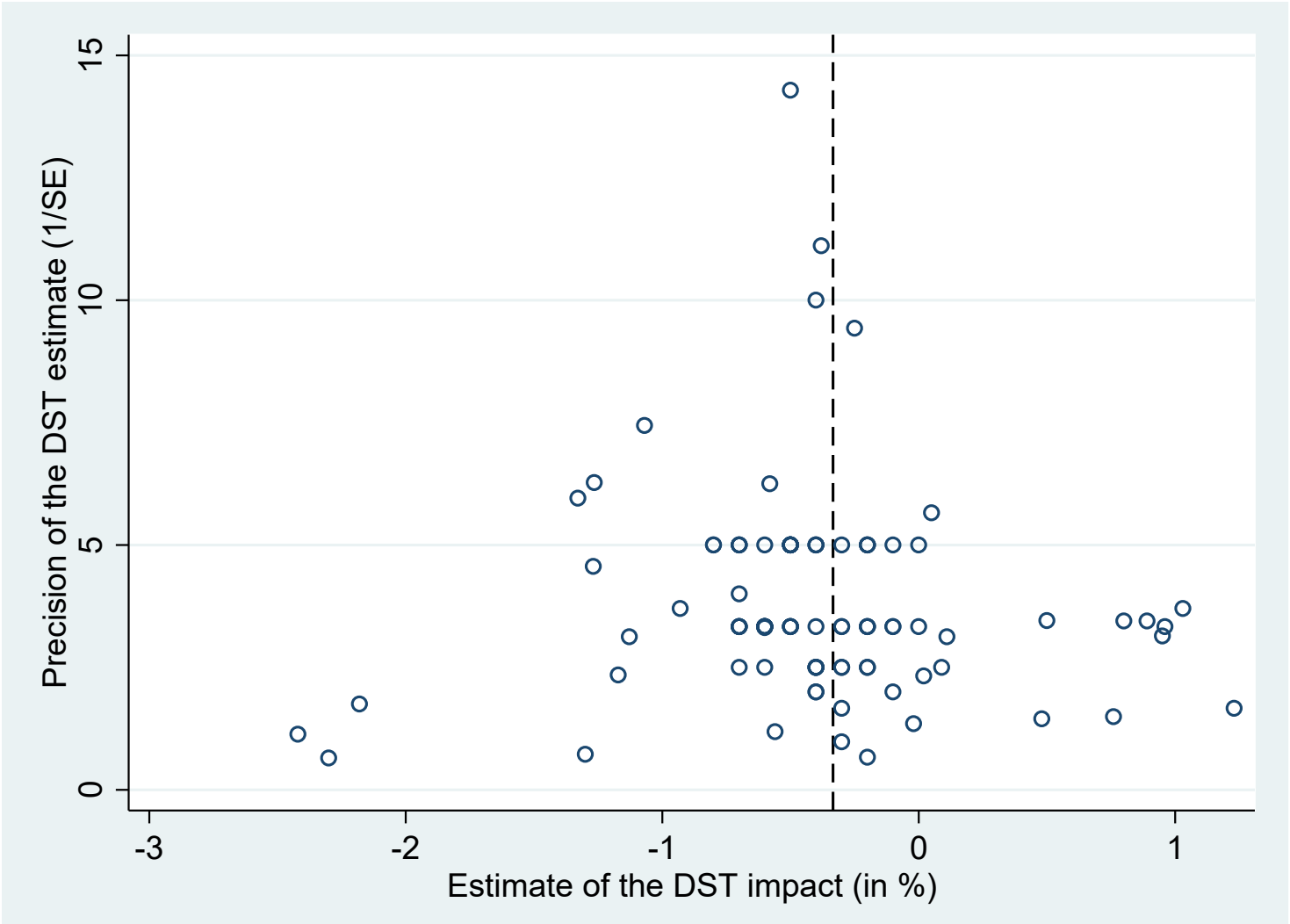
- *Consumption*: the average energy consumption during time *t* for a given hour, day, and year.
- *Treatment effect*: equals 1 for all hours when daylight saving time applies.
- *Controls*: seasonality and holidays, weather, the intensity of sunlight, heterogeneity among consumption units, etc.

No consensus in the literature

Journals report smaller savings



Publication bias? Probably not



Funnel asymmetry tests

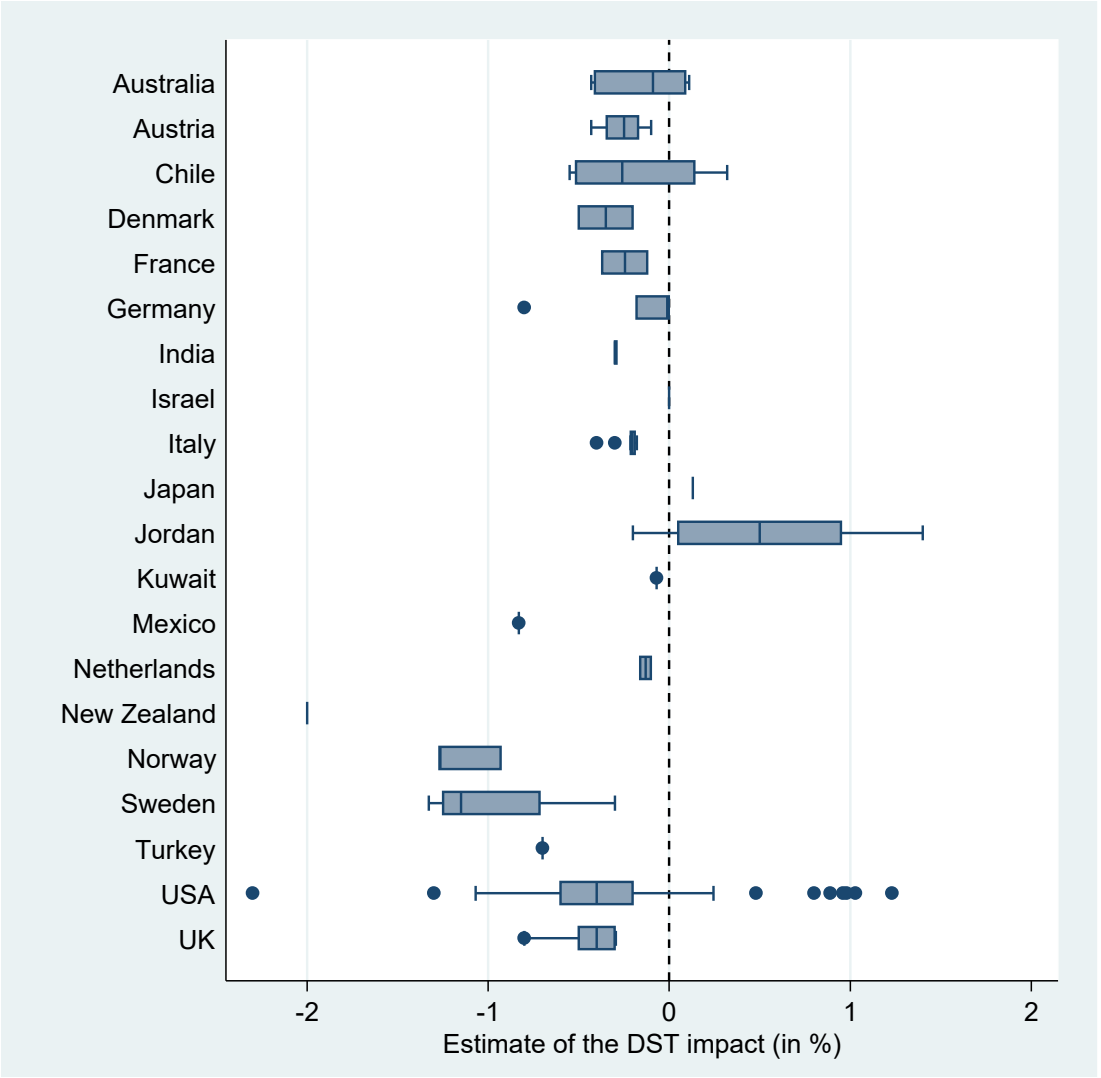
If the ratio of the point estimate to its standard error has a *t*-distribution, then the two quantities should be independent:

$$\text{estimate}_i = \underbrace{\beta}_{\text{true effect}} + \underbrace{\beta_0 SE_i}_{\text{publication bias}} + \mu_i.$$

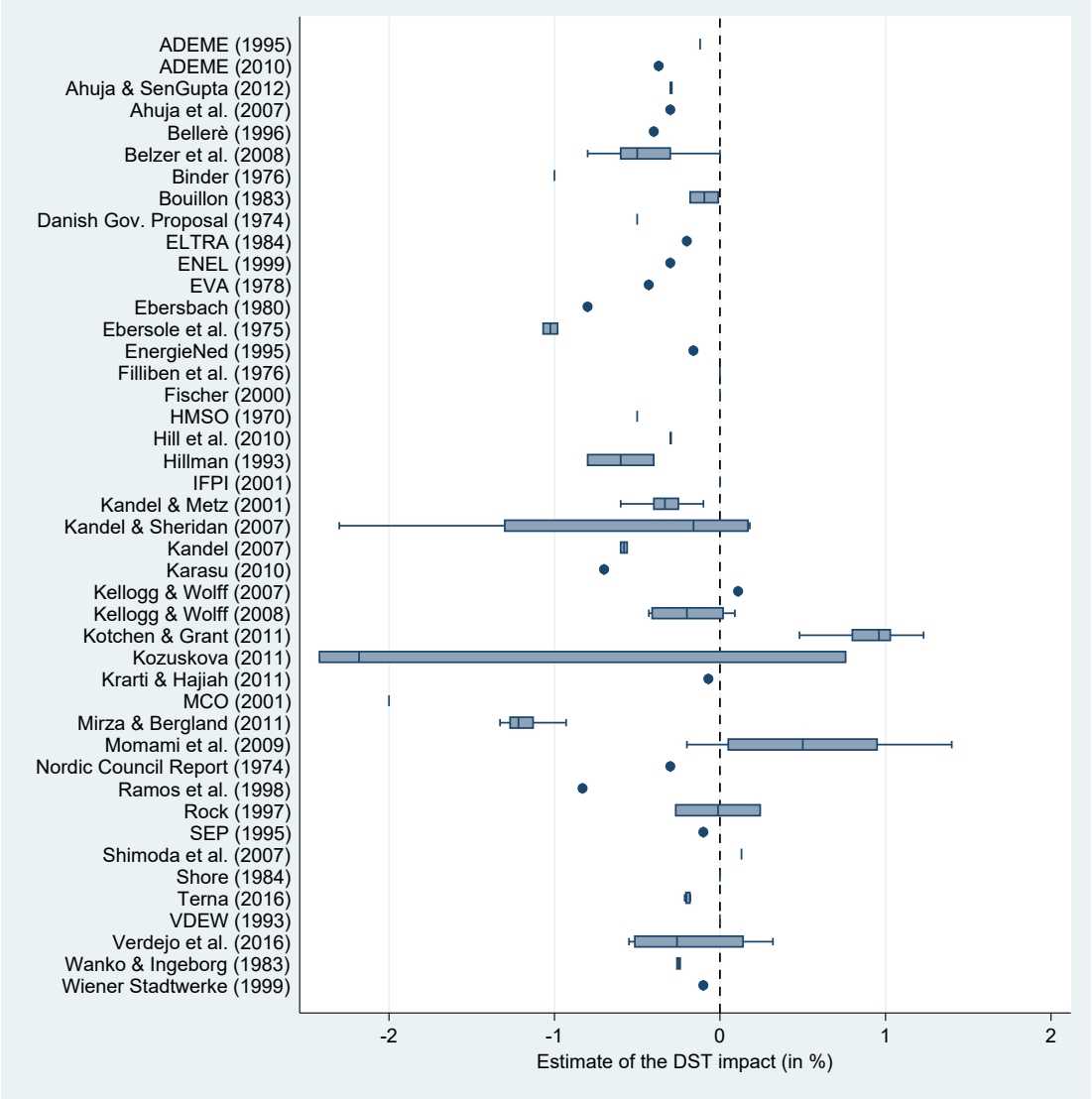
The no. of obs. can be used as an instrument for SE.

	OLS	FE	BE	Country	ME	IV
SE (*publication bias*)	-0.410	-1.217	-0.410	-0.496	-0.449	0.226
	(0.265)	(0.790)	(0.757)	(0.805)	(0.688)	(1.088)
Constant (*true effect*)	-0.293***	-0.222***	-0.294***	-0.278***	-0.291***	-0.445*
	(0.000778)	(0.0700)	(0.00812)	(0.0459)	(0.00731)	(0.243)
Observations	101	101	101	101	101	90

Country heterogeneity



Method heterogeneity (just studies for the US)



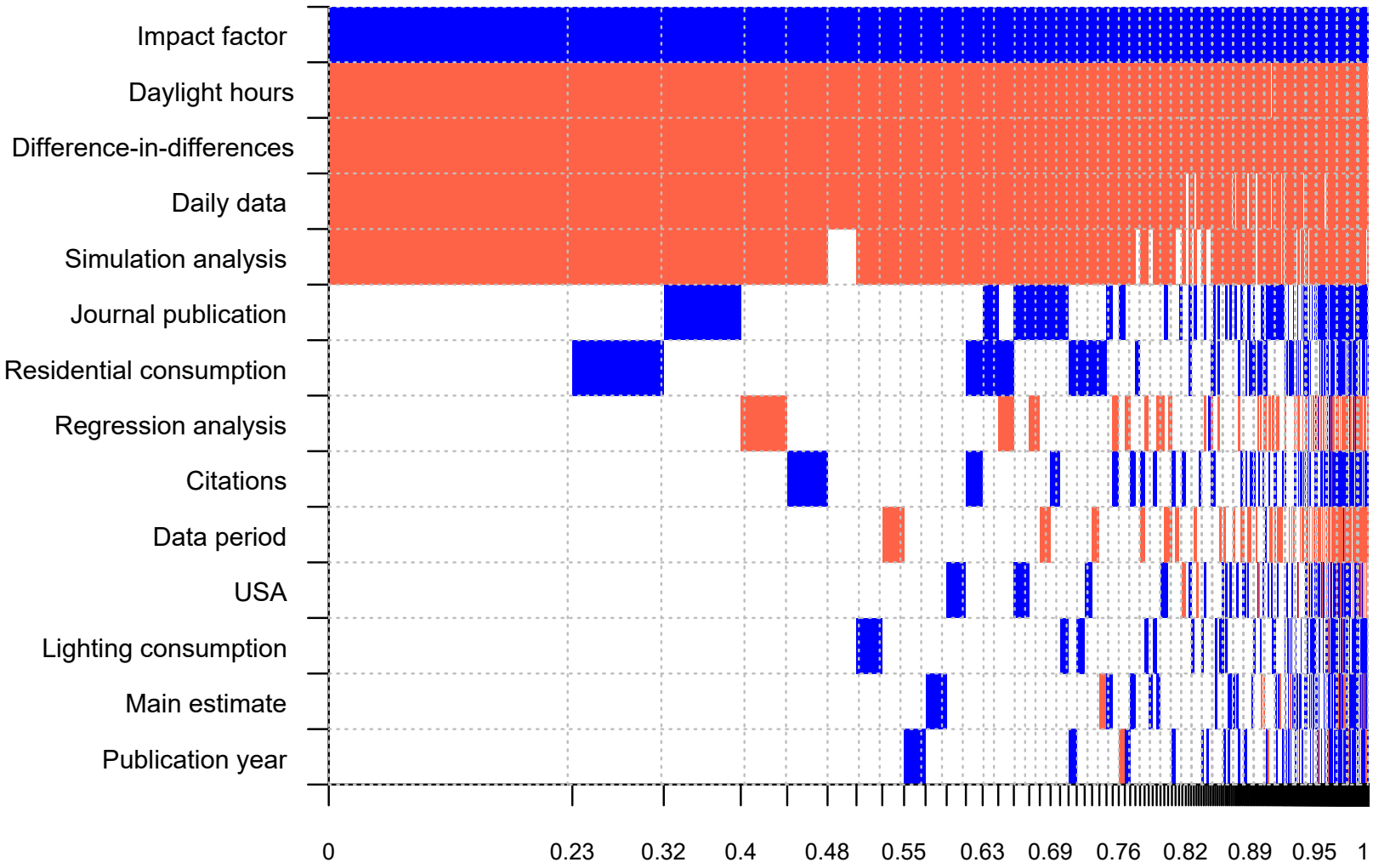
What explains the differences in conclusions? (1)

Data characteristics	
Data period	The number of years used in the estimation.
Main estimate	= 1 if the estimate is preferred by the authors of the study.
Hourly data	= 1 if the data are examined on hourly or higher than hourly granularity.
Daily data	= 1 if the data are examined on a daily basis.
Daylight hours	Average time between sunrise and sunset on the longest day for the country or region under examination.
Europe	= 1 if European countries are examined.
USA	= 1 if US data are examined.
Design of the analysis	
Regression analysis	= 1 if the primary study is based on regression analysis.
Simulation analysis	= 1 if the study is based on simulation.
Difference-in-diff.	= 1 if the difference-in-differences approach is employed.
Residential cons.	= 1 if only residential consumption is examined.
Lighting cons.	= 1 if total energy savings are reported as a result of lighting reduction.

What explains the differences in conclusions? (2)

Publication characteristics	
Publication year	The publication year of the study (base = 1970).
Journal article	= 1 if the study was published in a peer-reviewed journal.
Impact factor	The recursive RePEc impact factor of the outlet.
Citations	The logarithm of the total number of citations of the study in Google Scholar.

Bayesian model averaging



Main Findings

1. The mean reported estimate suggests energy savings of 0.34% due to DST.
2. The results are driven by the method employed (simulation vs. regression).
3. Energy savings are larger for countries farther away from the equator.

Project Website

www.meta-analysis.cz/dst

For further reading

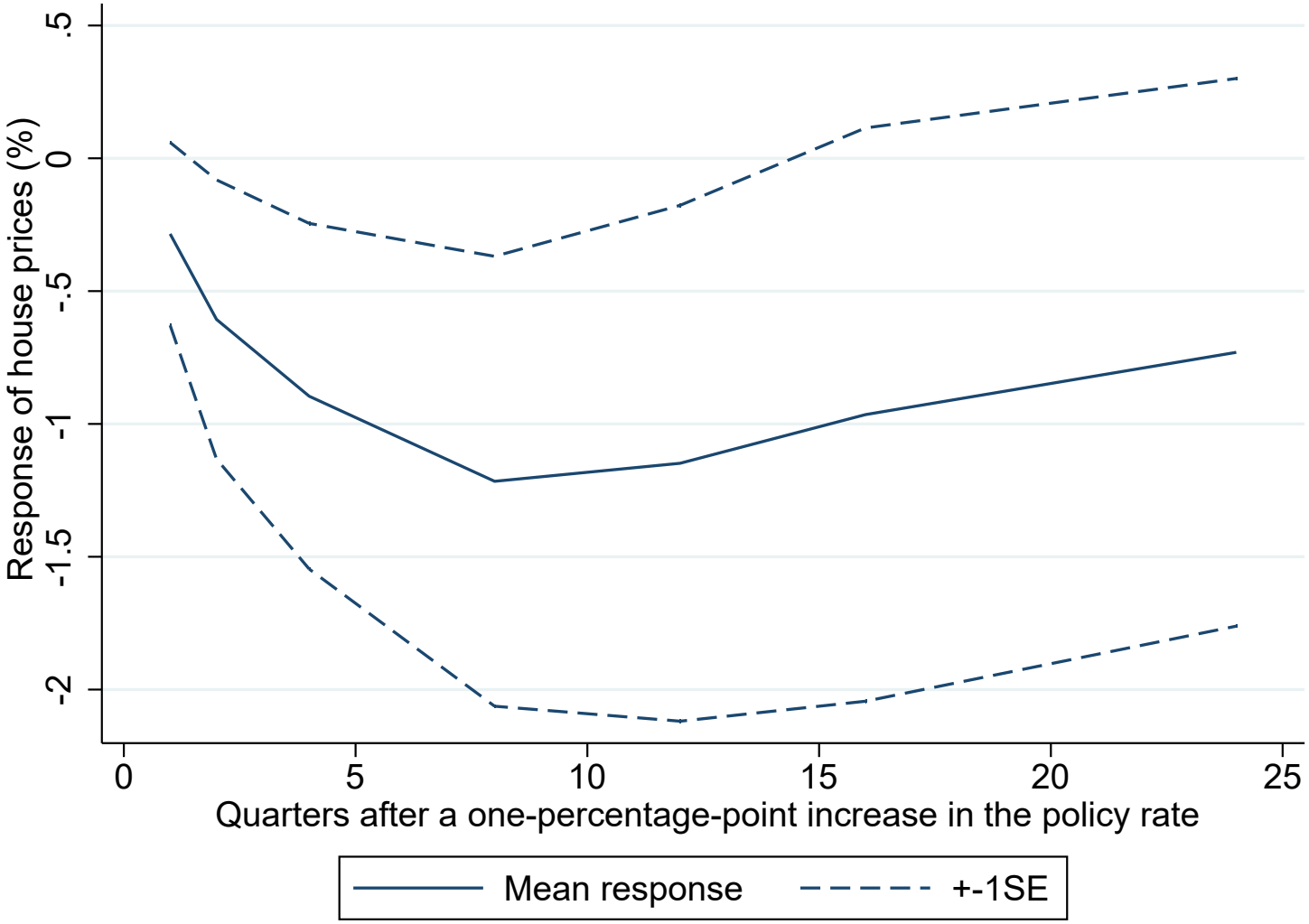

Stanley, T. D. & C. Doucouliagos (2012): Meta-Regression Analysis in Economics and Business.
Routledge, 1st. edition.


Brodeur, A., M. Sangnier & Y. Zylberg (2016): Star Wars: The Empirics Strike Back.
AEJ Applied 8(1): pp. 1–32.


Havranek, T. (2015): Measuring Intertemporal Substitution: The Importance of Method Choices and Publication Bias.
JEEA 13(6): pp. 1180–1204.

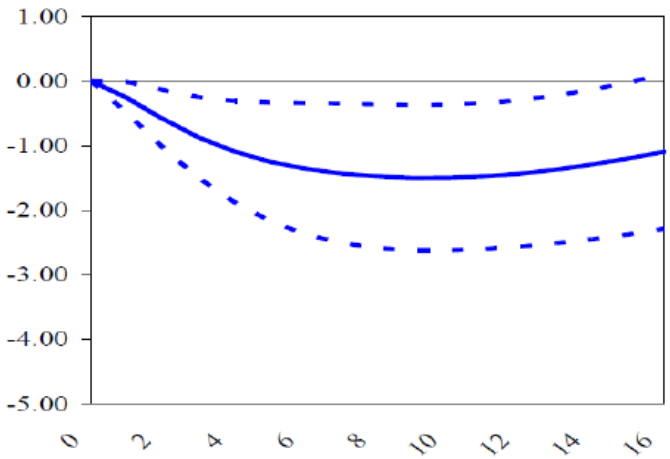
Reading list on RePEc: Google “meta-analysis in economics.”

Example 2: Interest rates and house prices



Converting graphs to data

We measure pixel coordinates using WebPlotDigitizer:

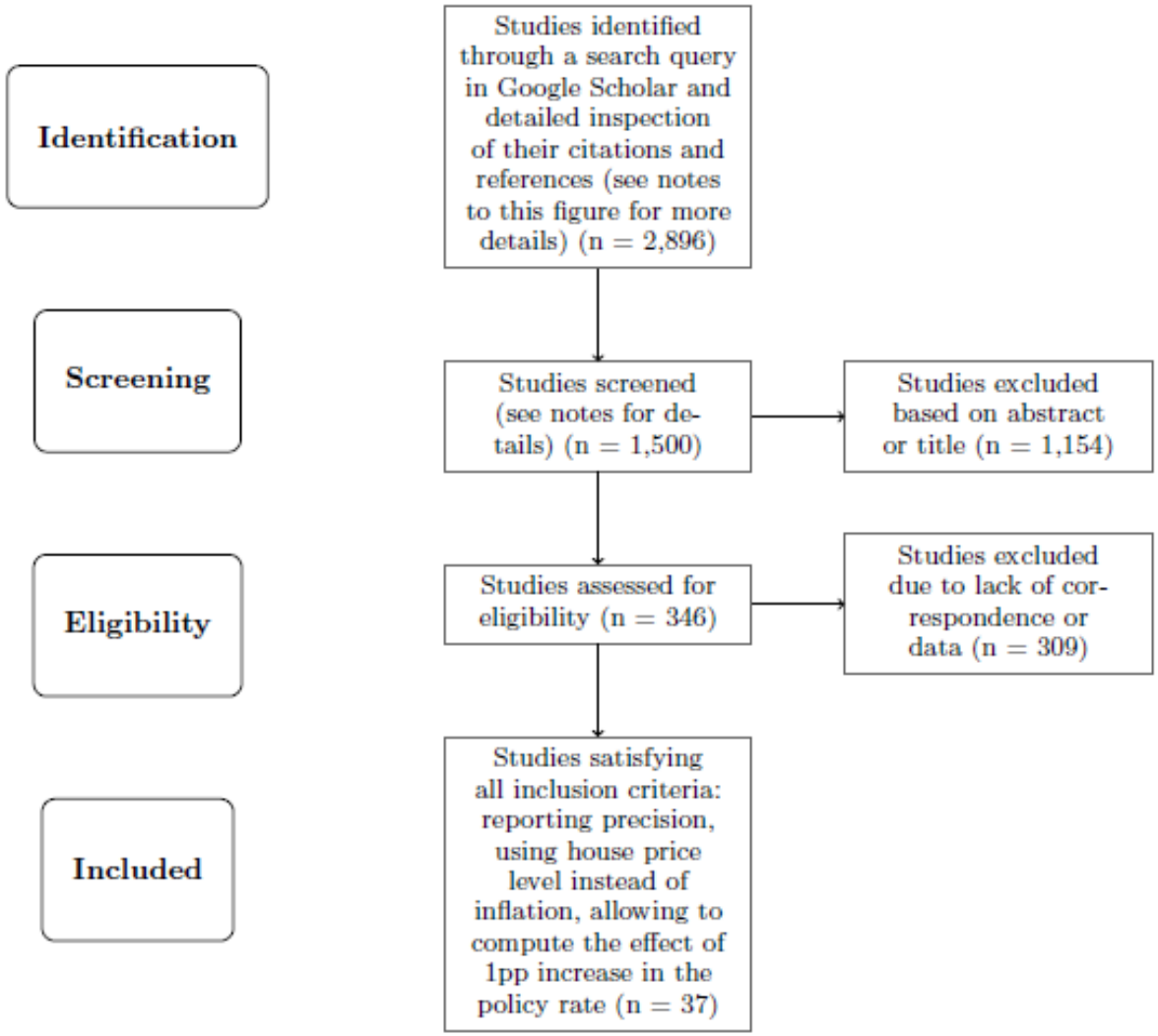


Canada

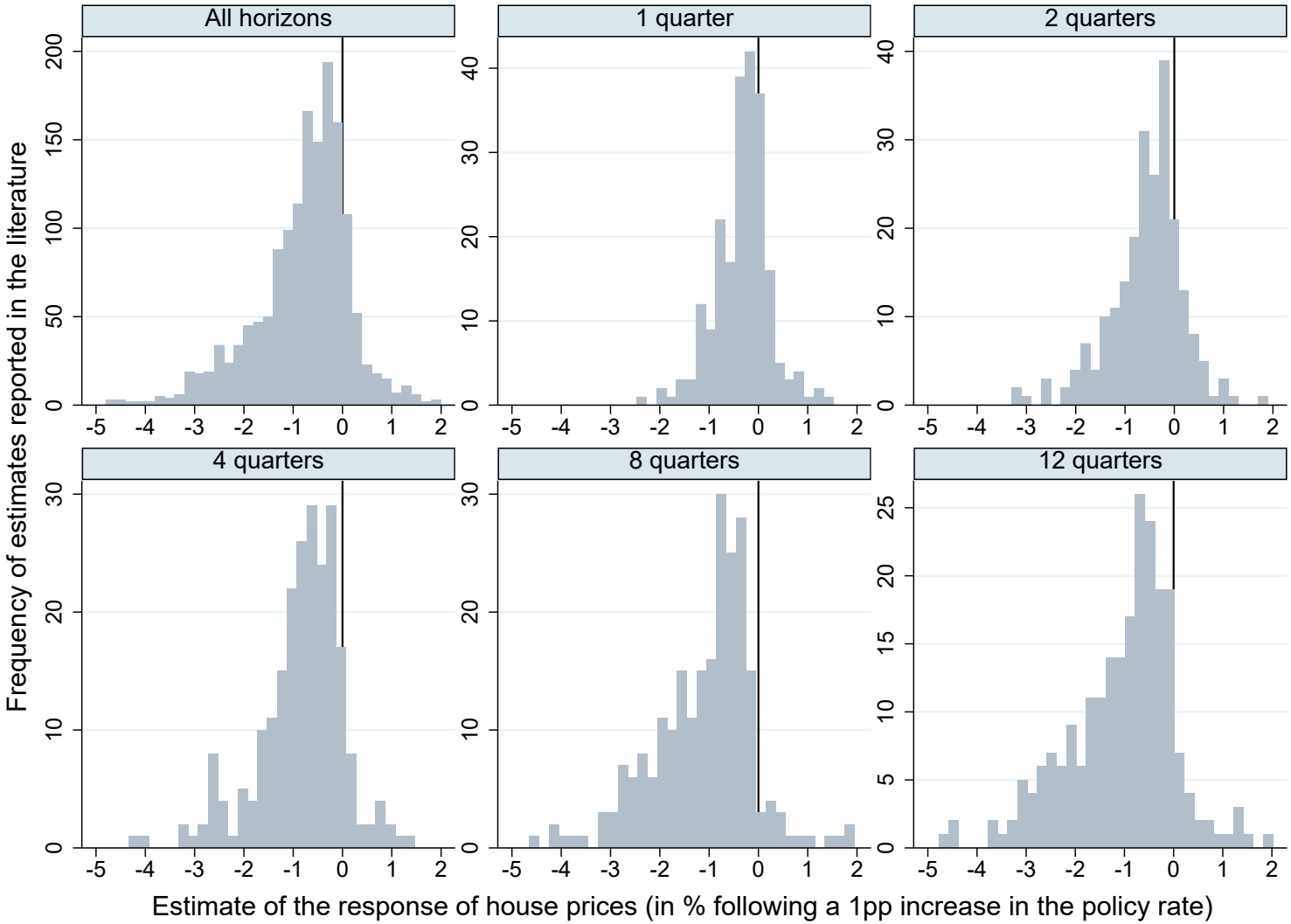
horizon	response	confidence interval
1	-0.231	(-0.407, -0.016)
2	-0.578	(-1.012, -0.135)
4	-1.076	(-1.853, -0.299)
8	-1.454	(-2.554, -0.359)
12	-1.414	(-2.534, -0.287)
16	-1.084	(-2.259, 0.131)
max	n.a.	n.a.

Obtained from: page 50, Figure 2
Frequency: Quarterly
Note: Responses are in %.

Literature search



Histograms for different horizons



Funnel asymmetry test (Card & Krueger)

In the absence of publication bias the funnel should be symmetrical:

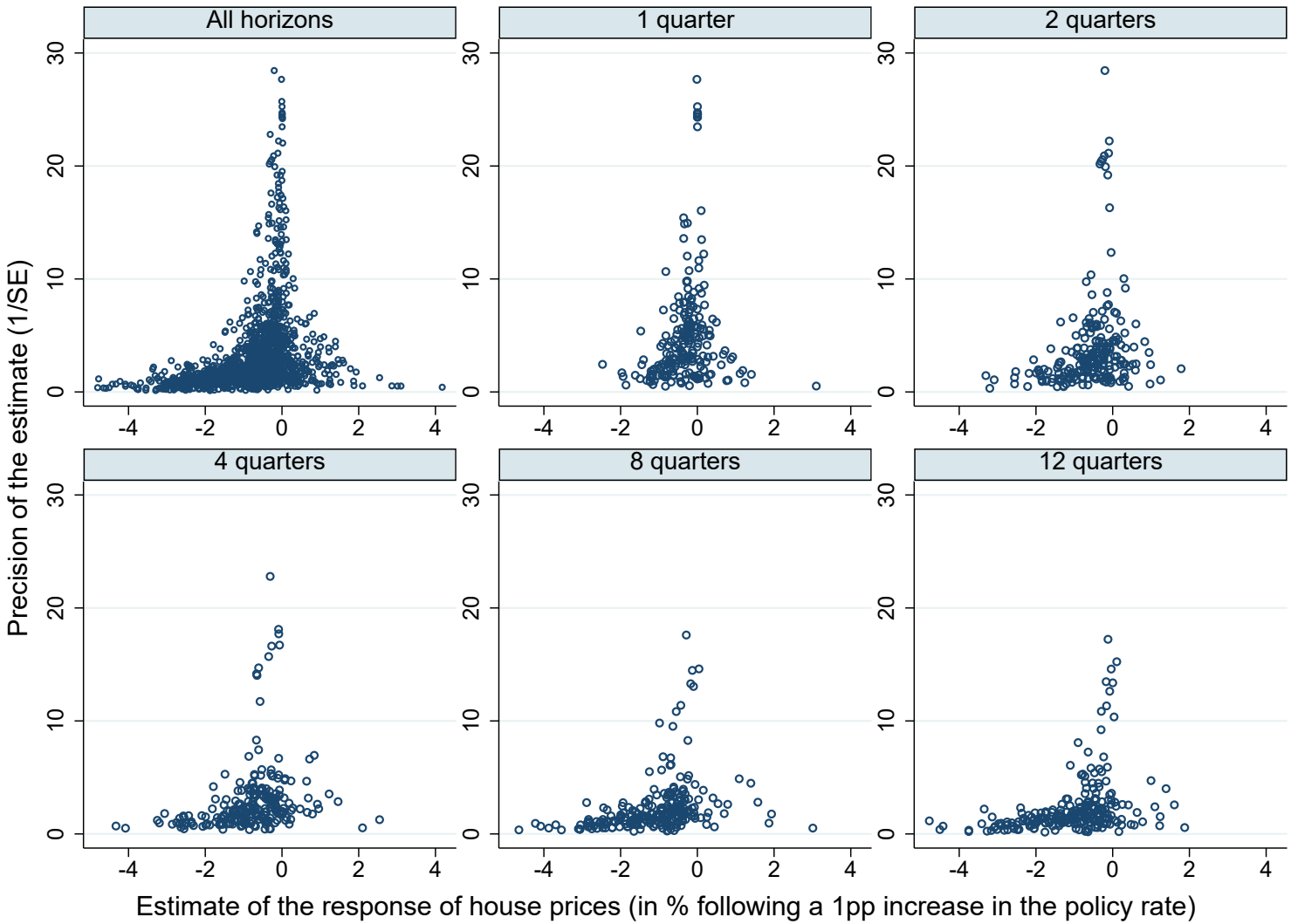
$$\text{estimate}_i = \underbrace{\beta}_{\text{true effect}} + \underbrace{\beta_0 SE_i}_{\text{publication bias}} + \mu_i.$$

The equation is heteroscedastic. Weighted least squares yield

$$t_i = \beta_0 + \beta \left(\frac{1}{SE_i} \right) + \vartheta_i.$$

The no. of obs. can be used as an instrument for SE.

Publication bias for longer horizons?



Corroborated by funnel asymmetry tests

Note that the slope is negative, the intercept is smaller than the reported mean effect:

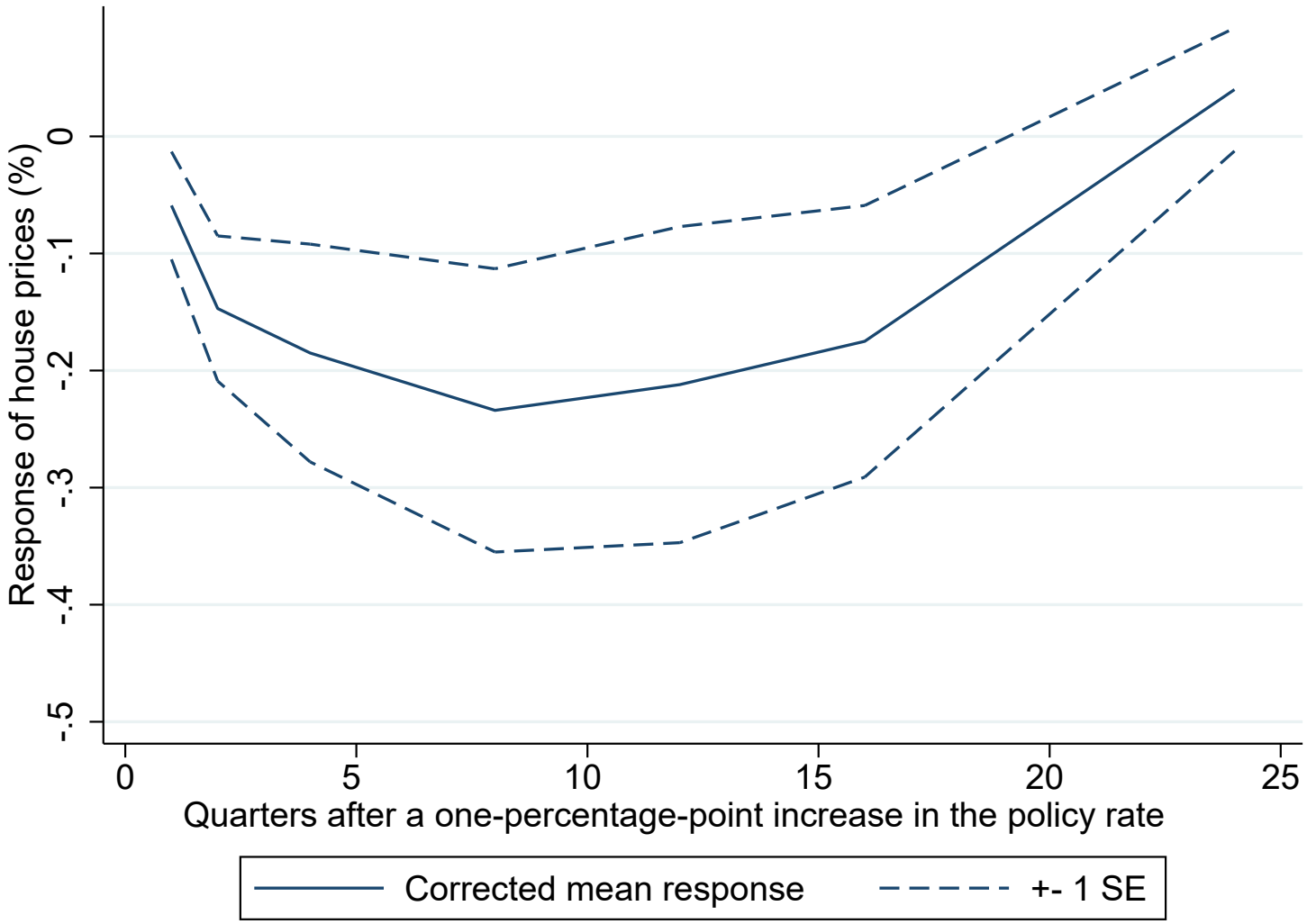
Time after a monetary policy shock:	1 quarter	2 quarters	4 quarters	8 quarters	12 quarters	16 quarters
PANEL A: Linear models						
Regression of reported estimates on their standard errors, ordinary least squares						
Standard error (publication bias)	-0.815*	-1.117***	-1.353***	-1.160***	-0.667**	-0.375*
	(0.463)	(0.367)	(0.427)	(0.308)	(0.317)	(0.215)
	[-1.220, -.015]	[-1.987, -.351]	[-2.373, -.410]	[-2.051, -.150]	[-1.542, .009]	[-1.281, .163]
Constant (corrected mean effect)	-0.014	-0.020	-0.015	-0.233	-0.501**	-0.560***
	(0.168)	(0.185)	(0.248)	(0.219)	(0.252)	(0.201)
	[-.213, .145]	[-.417, .479]	[-.576, .739]	[-.790, .444]	[-1.072, .154]	[-.968, -.039]
Regression of reported estimates on their standard errors, weighted by inverse variance						
Standard error (publication bias)	-0.705***	-0.874***	-1.092***	-1.160***	-0.964***	-0.732***
	(0.179)	(0.147)	(0.203)	(0.204)	(0.234)	(0.199)
	[-1.228, -.393]	[-1.207, -.586]	[-1.555, -.730]	[-1.629, -.710]	[-1.534, -.497]	[-1.343, -.337]
Constant (corrected mean effect)	-0.059	-0.147**	-0.185**	-0.234*	-0.212	-0.175
	(0.046)	(0.062)	(0.093)	(0.121)	(0.135)	(0.116)
	[-.058, .014]	[-.303, .081]	[-.360, .130]	[-.496, .044]	[-.521, -.030]	[-.533, -.010]

Nonliner tests also imply bias

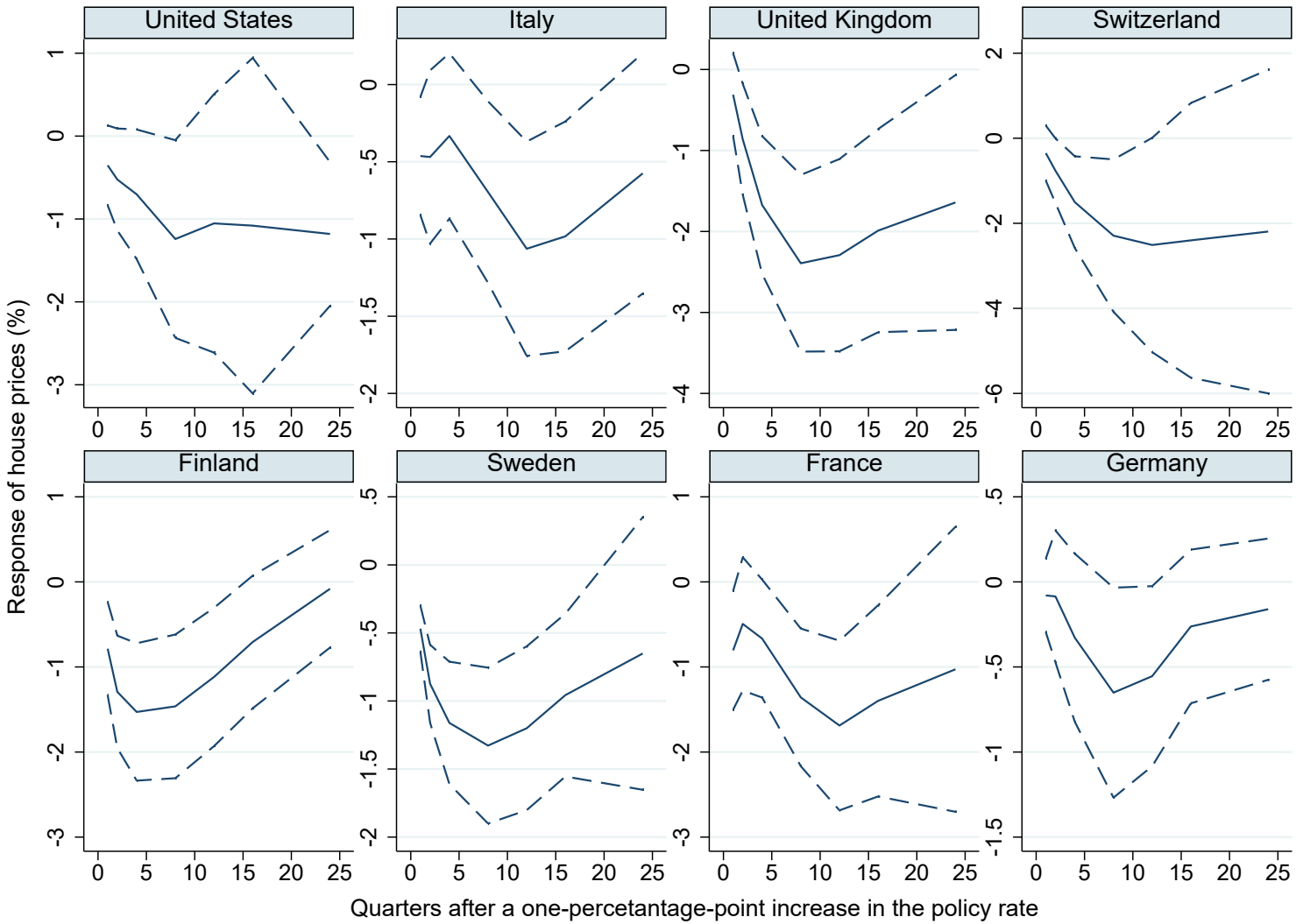
Note that the corrected effect is smaller than the reported mean effect:

Time after a monetary policy shock:	1 quarter	2 quarters	4 quarters	8 quarters	12 quarters	16 quarters
PANEL B: Nonlinear models						
Stem-based method (Furukawa, 2021)						
Corrected mean effect	-0.017	-0.192***	-0.318***	-0.324*	-0.172	-0.146**
	(0.053)	(0.074)	(0.116)	(0.185)	(0.145)	(0.071)
Selection model (Andrews & Kasy, 2019)						
Corrected mean effect, break at $t = 1.645$	-0.004***	-0.155*	-0.361***	-0.404**	-0.205	-0.065
	(0.002)	(0.094)	(0.079)	(0.190)	(0.187)	(0.115)
Corrected mean effect, break at $t = 1$	-0.023	-0.008	-0.213**	-0.149	-0.182**	-0.030
	(0.027)	(0.234)	(0.130)	(0.194)	(0.110)	(0.114)
P-uniform (van Aert & van Assen, 2021)*						
Corrected mean effect, break at $t = 1.645$	-0.133***	-0.121***	-0.140***	-0.134***	-0.118***	-0.095***
Corrected mean effect, break at $t = 1$	-0.072***	-0.091***	-0.103***	-0.098***	-0.093***	-0.075***

Mean impulse response beyond publication bias



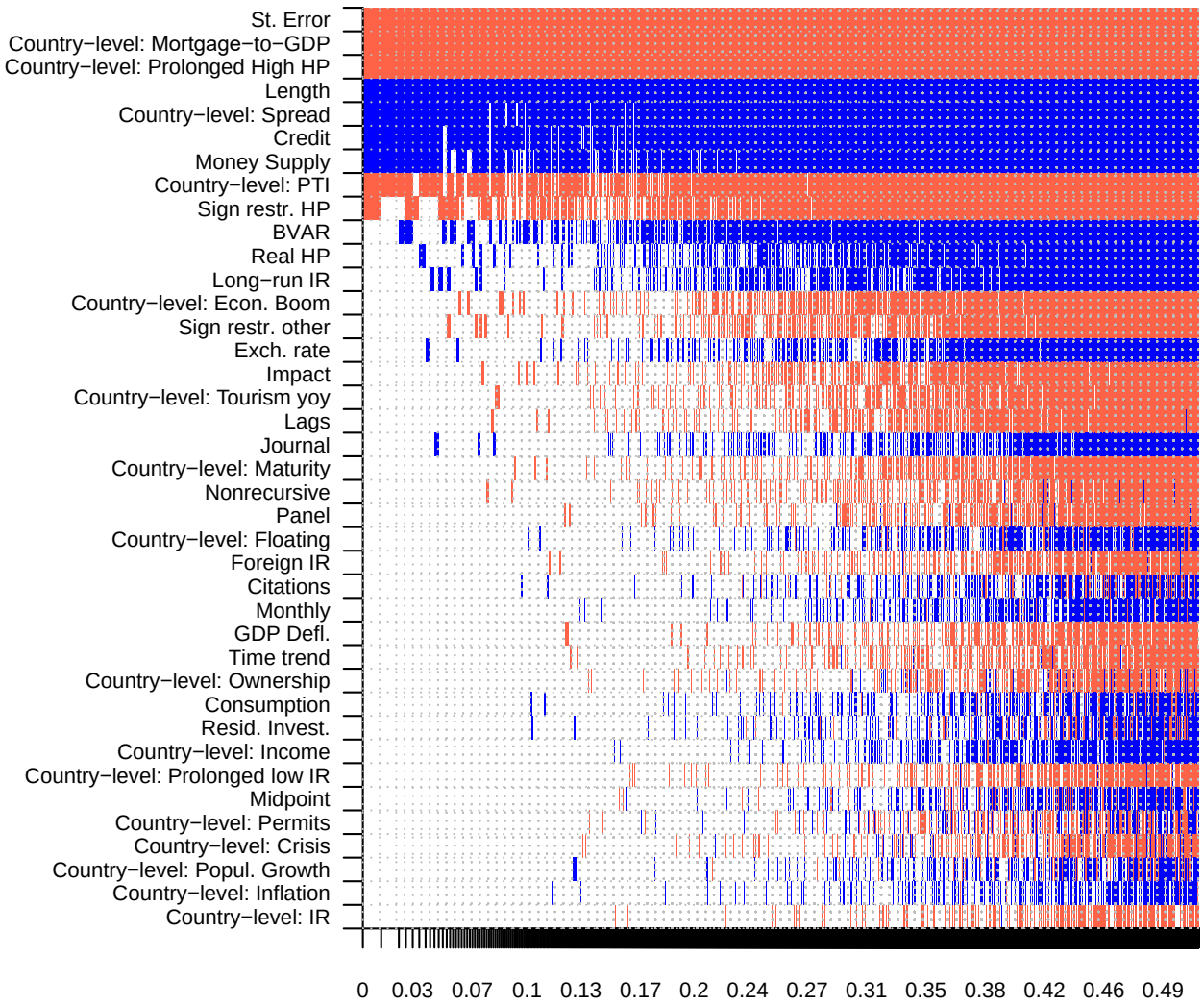
Cross-country heterogeneity



Estimation context

- **Data characteristics**: monthly, panel, length, age
- **Specification characteristics**: foreign IR, credit, consumption, res. invest., money supply, exch. rate, long-run IR, lags
- **Estimation characteristics**: BVAR, sign restr. HP, sign restr. other, nonrecursive
- **Publication characteristics**: citations, impact, published
- **Country characteristics**: mortgage-to-GDP, floating, maturity, IR, spread, prolonged high HP, PTI, crisis, prolonged low IR, tourism, income, inflation, popul. growth, permits, ownership

Bayesian model averaging

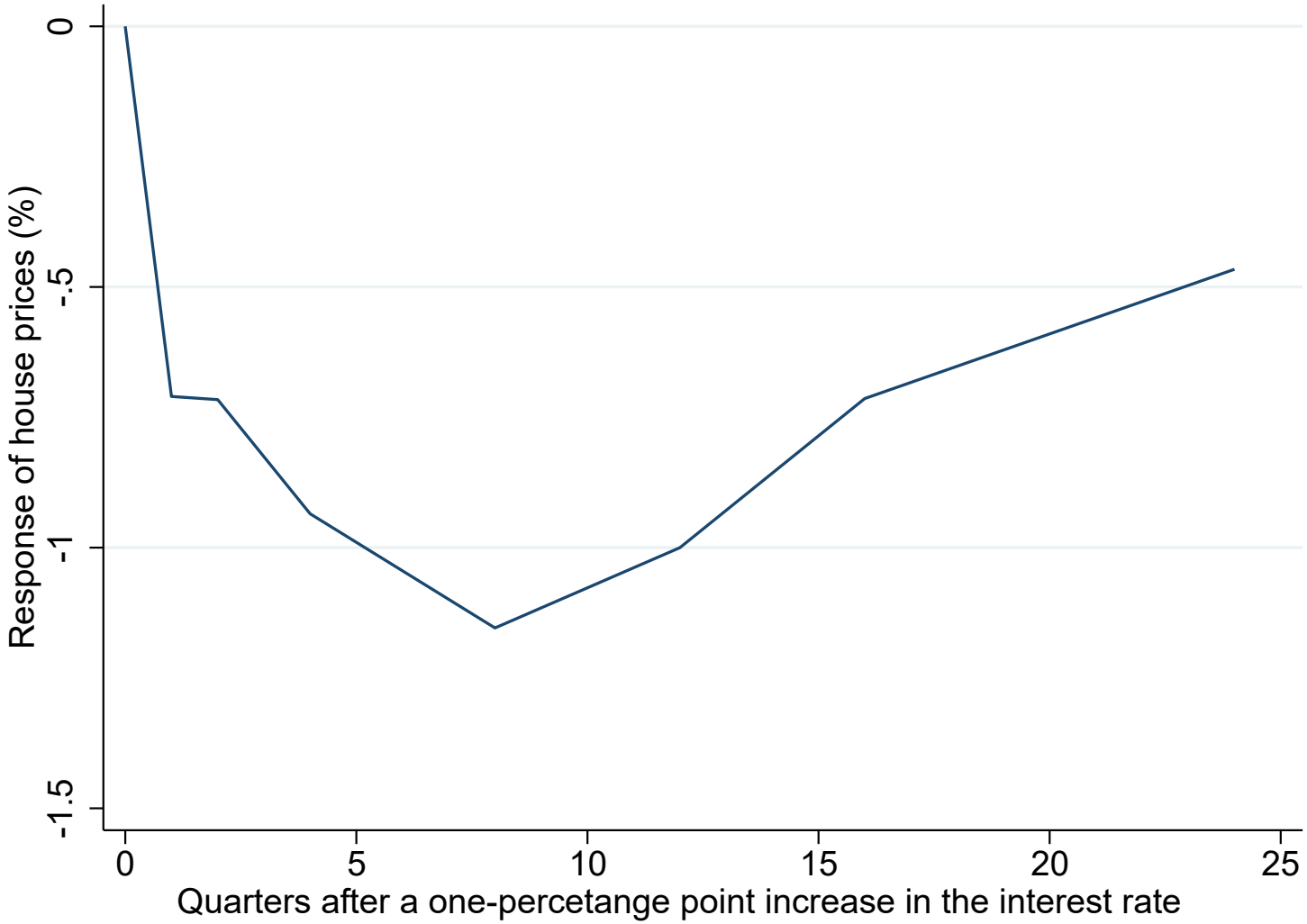


Best practice

We construct a hypothetical ideal study: no publication bias (SE = 0), recent data, long series, nonrecursive identification, published, good journal, highly cited.

Horizon of the corresponding impulse response:	1Q	2Q	4Q	8Q	12Q	16Q
Agnostic on control variables (baseline)	-0.710	-0.716	-0.935	-1.154	-1.000	-0.714
Control for credit, money supply, and long-run IR	-0.016	-0.023	-0.241	-0.461	-0.306	-0.021
Excluding credit, money supply, and long-run IR	-0.920	-0.927	-1.145	-1.365	-1.210	-0.925
Finland	-0.576	-0.583	-0.801	-1.021	-0.866	-0.581
France	-1.477	-1.484	-1.702	-1.922	-1.767	-1.482
Germany	-0.497	-0.503	-0.722	-0.941	-0.787	-0.501
Italy	0.093	0.087	-0.132	-0.352	-0.197	0.088
United Kingdom	-1.344	-1.351	-1.569	-1.789	-1.634	-1.349
United States	-0.910	-0.916	-1.135	-1.354	-1.200	-0.914

Impulse response implied by best practice



Main Findings

1. Among 237 impulse responses from 37 studies, the mean decrease in house prices is 1.2% two years after a 1 pp. increase in the policy rate.
2. The mean is exaggerated because of publication bias.
3. Stronger transmission for i) more developed markets, ii) flatter yield curve, iii) house price boom.

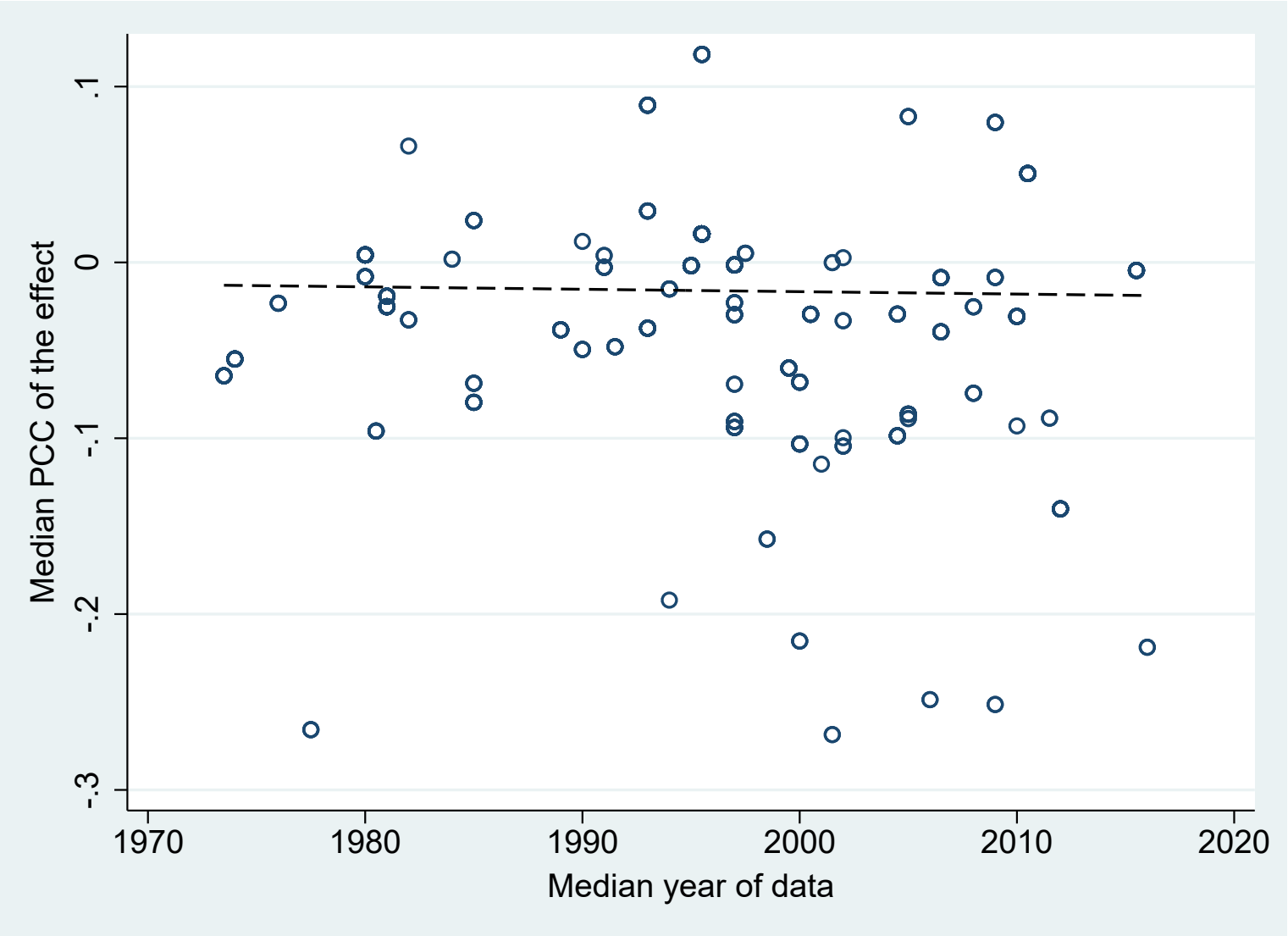
Project Website

www.meta-analysis.cz/house_prices

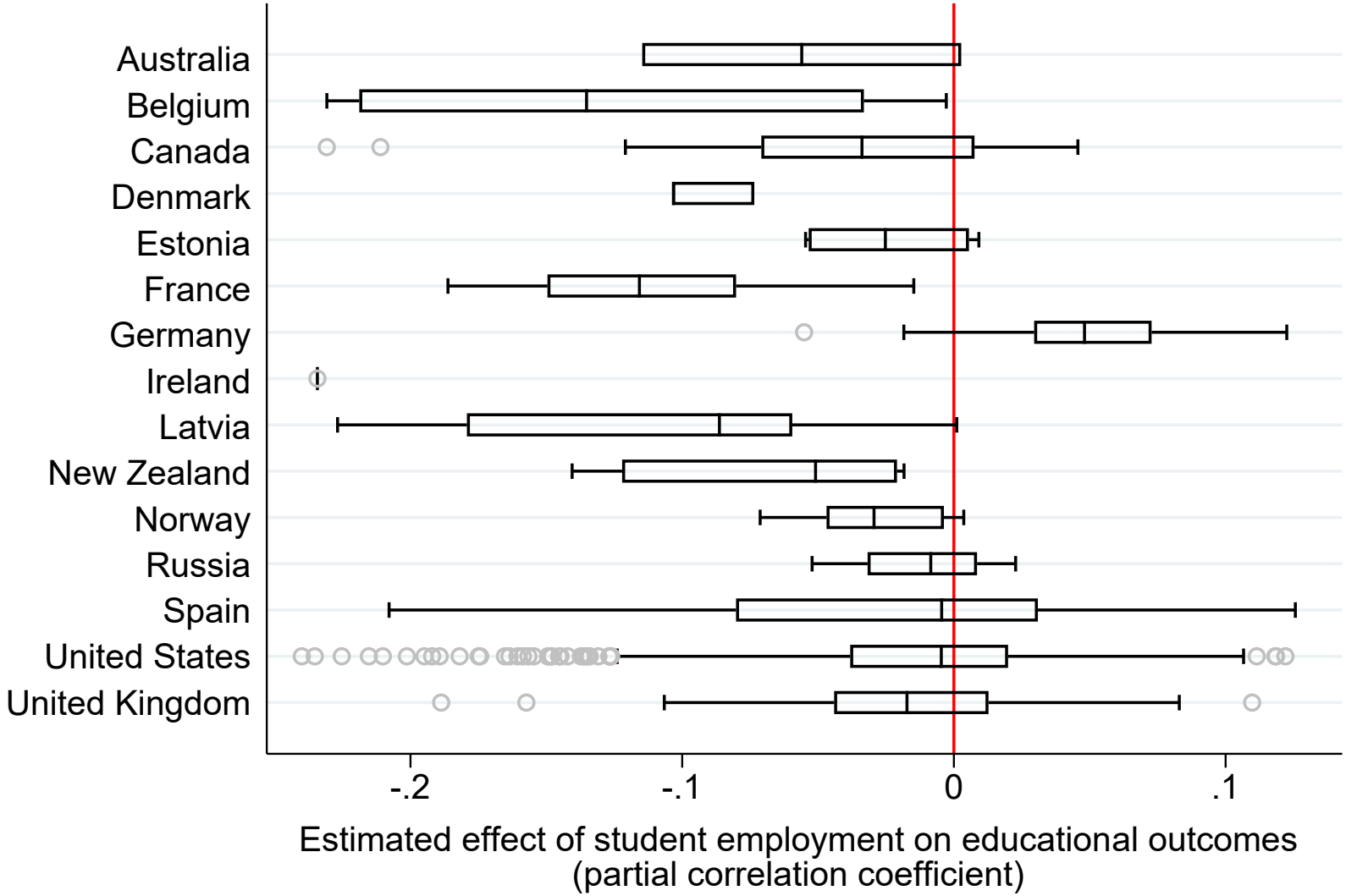
Example 3: Working while in school

- Hard to influence commonly cited determinants of educational outcomes: ability, ethnicity, gender, parental education and affluence
- But most students can decide whether or not to work
- How does that influence educational outcomes?
- Should you motivate your kids to work while in school?
- Mixed empirical results: from large negative to positive estimates
- 861 estimates from 69 studies

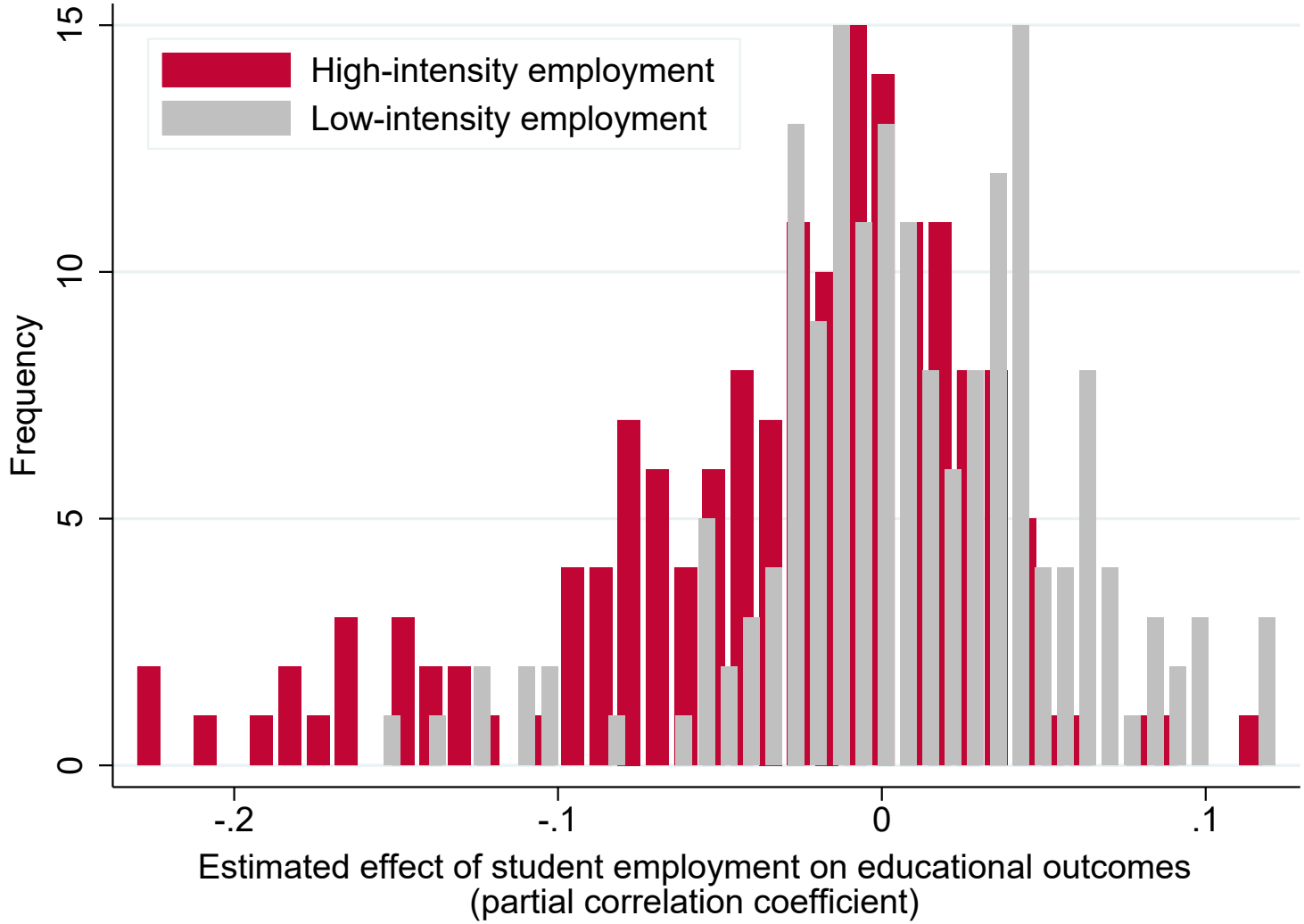
No consensus in the literature



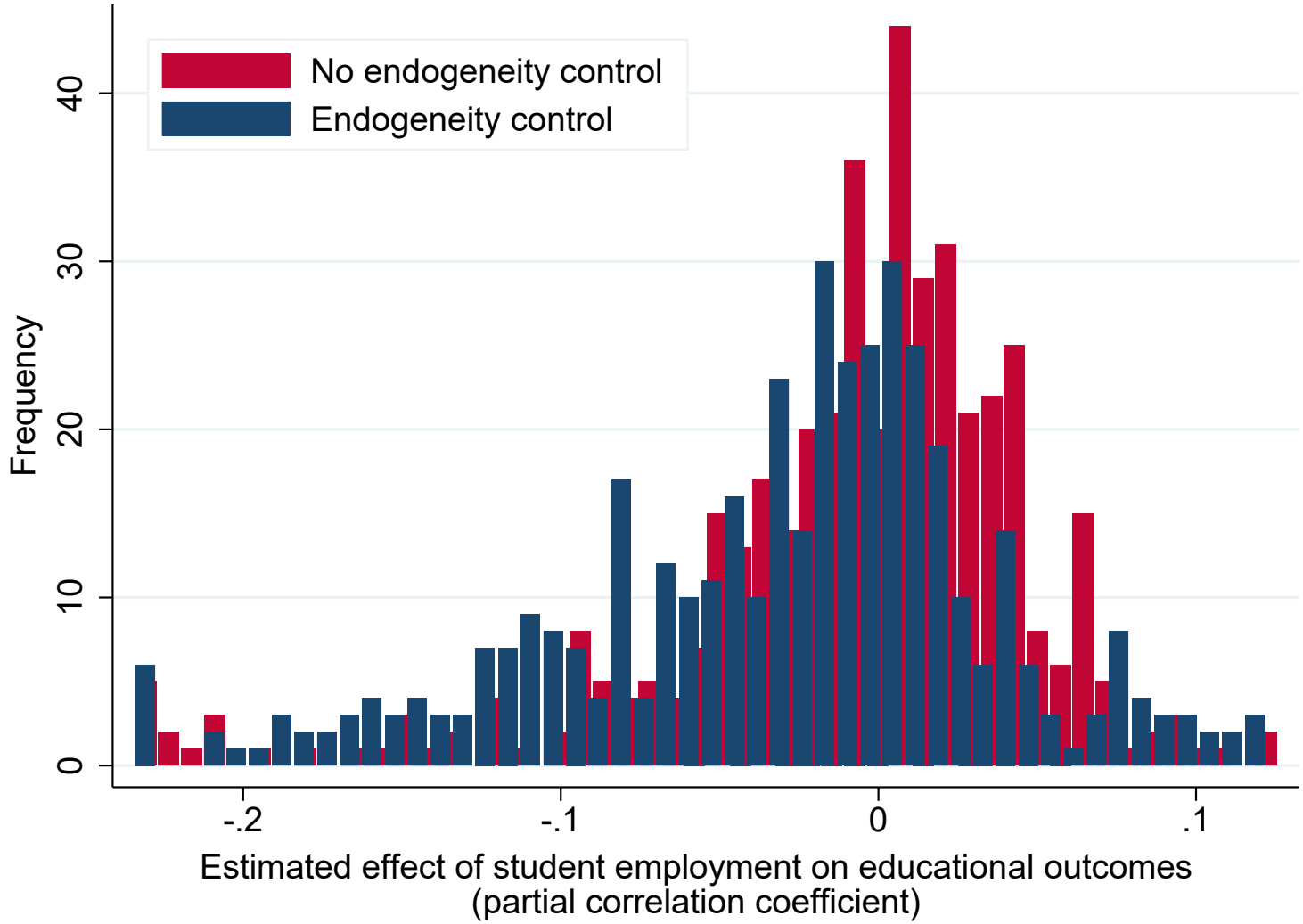
Negative for all countries except Germany



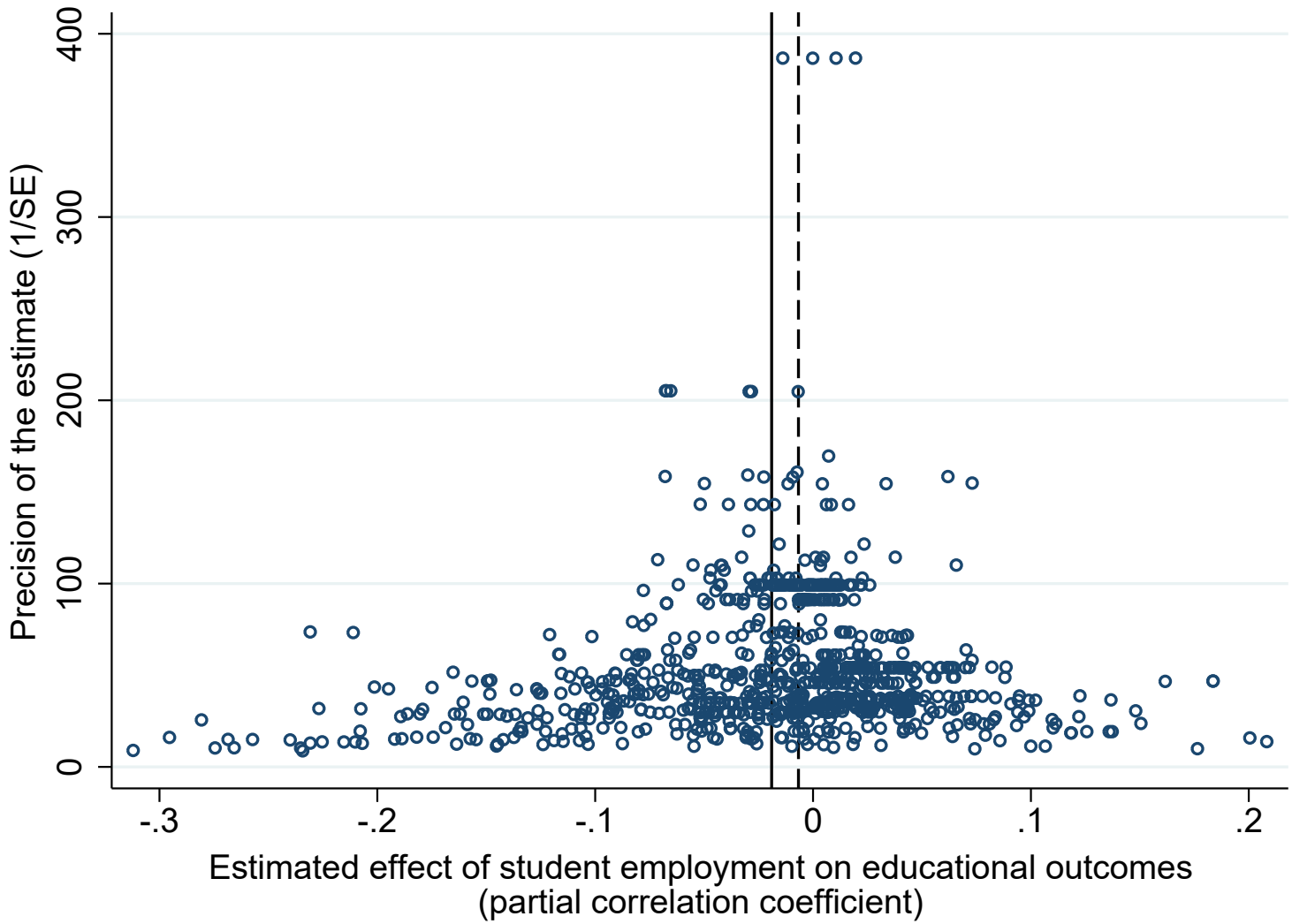
Work intensity clearly matters



Endogeneity control is important



Little publication bias on average



No bias in studies addressing endogeneity

[Block 2] Studies trying to take endogeneity into account					
Panel A: Linear techniques					
	OLS	IV	Study	Precision	
Standard error *(Publication bias)*	-0.311 (0.347) [-1.174, 0.710]	-0.311 (0.392) [-1.306, 0.829]	-0.244 (0.431) [-1.556, 0.987]	-0.449 (0.387) [-1.3, 0.489]	
Constant *(Effect beyond bias)*	-0.0189** (0.00932) [-0.040, -0.0001]	-0.0185* (0.00984) [-0.039, -0.0002]	-0.0218 (0.0159) [-0.057, 0.011]	-0.0151** (0.00751) [-0.034, 0.002]	
Observations	425	425	425	425	
Panel B: Between- and within-study variation					
	BE	FE	RE		
Standard error *(Publication bias)*	-1.334*** (0.457)	1.041*** (0.373)	0.235 (0.287)		
Constant *(Effect beyond bias)*	0.00317 (0.0169)	-0.0556*** (0.0101)	-0.0431*** (0.0126)		
Observations	425	425	425		
Panel C: Nonlinear techniques					
	WAAP	Stem method	Kinked model	Selection model	p-uniform*
Effect beyond bias	-0.0200*** (0.00687)	0.00996 (0.0294)	-0.0138*** (0.00369)	-0.0260*** (0.008)	-0.0319*** (0.0106)
Observations	425	425	425	425	425

Strong bias in studies ignoring endogeneity

[Block 1] Studies ignoring endogeneity

Panel A: Linear techniques

	OLS	IV	Study	Precision
Standard error	-1.858***	-1.945***	-2.420***	-0.959**
(Publication bias)	(0.457)	(0.483)	(0.545)	(0.397)
	[-3.189, -0.866]	[-3.23, -0.921]	[-3.83, -1.17]	[-2.058, -0.24]
Constant	0.0405**	0.0425**	0.0591***	0.0171
(Effect beyond bias)	(0.0172)	(0.0176)	(0.0189)	(0.0112)
	[0.0001, 0.084]	[0.002, 0.086]	[0.0003, 0.105]	[-0.001, 0.067]
Observations	436	436	436	436

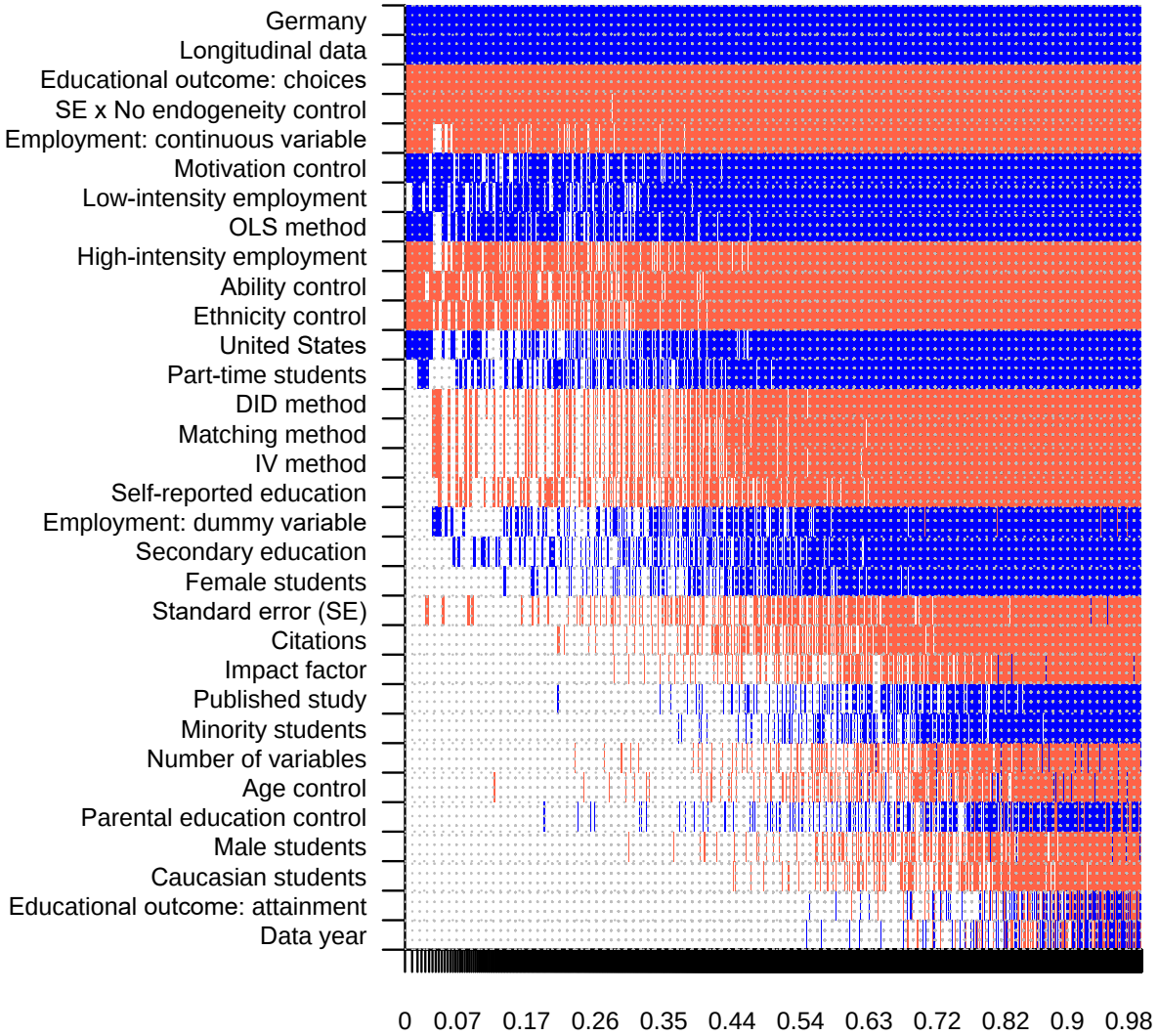
Panel B: Between- and within-study variation

	BE	FE	RE
Standard error	-2.535***	-0.903	-1.310***
(Publication bias)	(0.489)	(0.794)	(0.255)
Constant	0.0331	0.0156	-0.0111
(Effect beyond bias)	(0.0217)	(0.0207)	(0.0146)
Observations	436	436	436

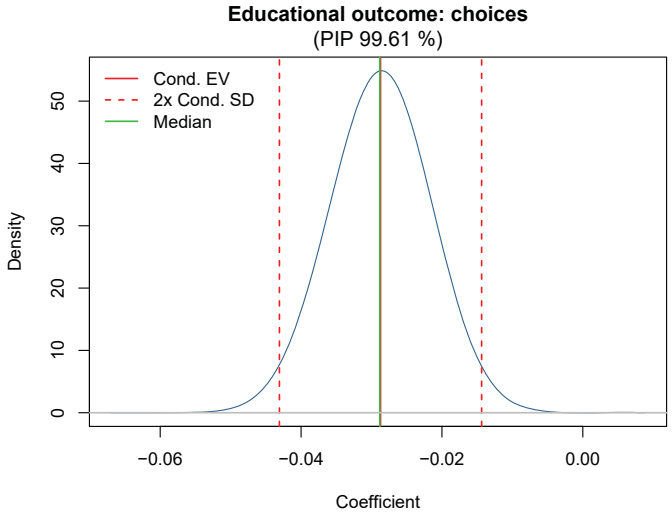
Panel C: Nonlinear techniques

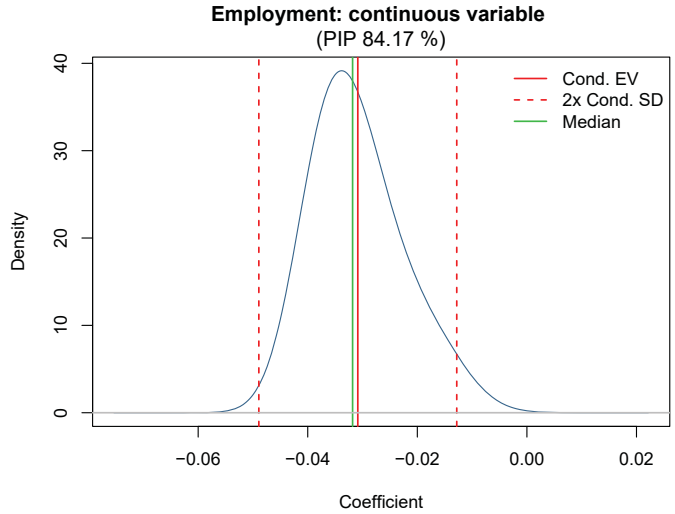
	WAAP	Stem method	Kinked model	Selection model	p-uniform*
Effect beyond bias	·	-0.00959**	0.00187	0.000	·
	(·)	(0.00432)	(0.00386)	(0.00500)	(·)
Observations	436	436	436	436	436

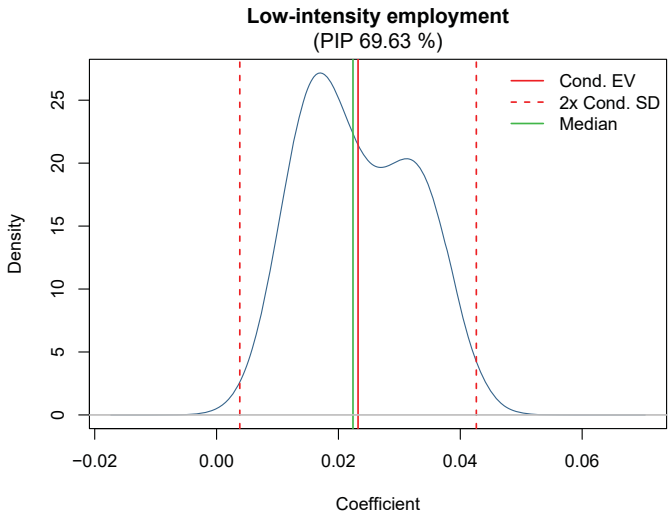
Model uncertainty; endogeneity control is key

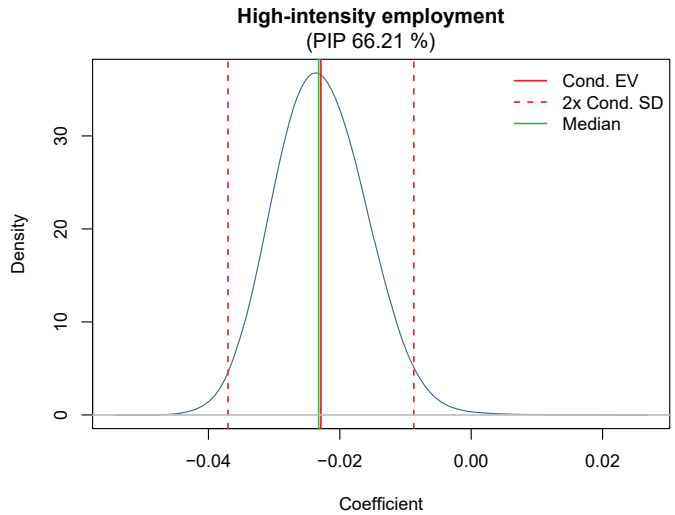


Closer look: work intensity, measurement

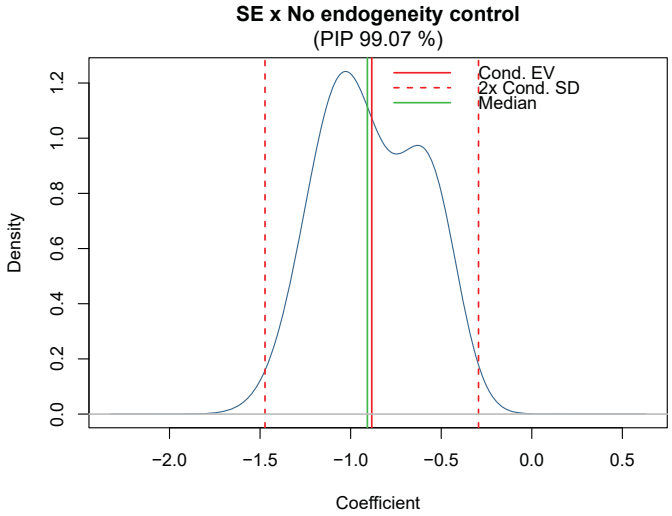


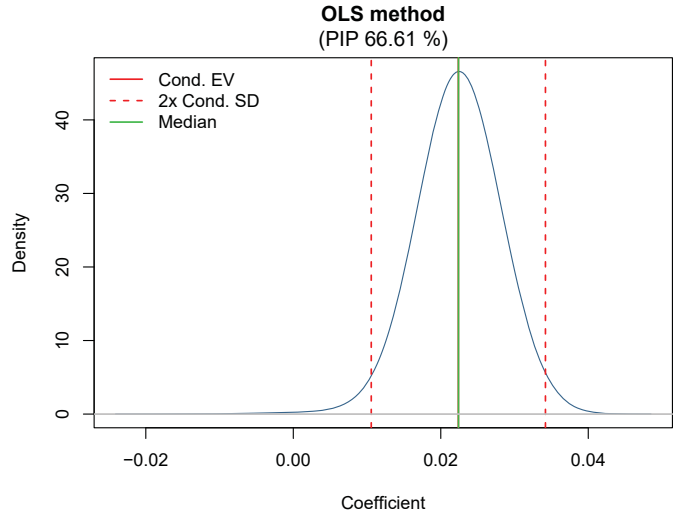


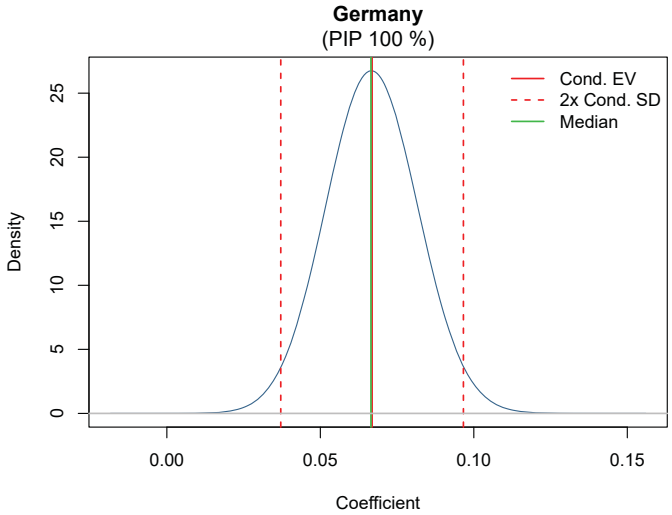


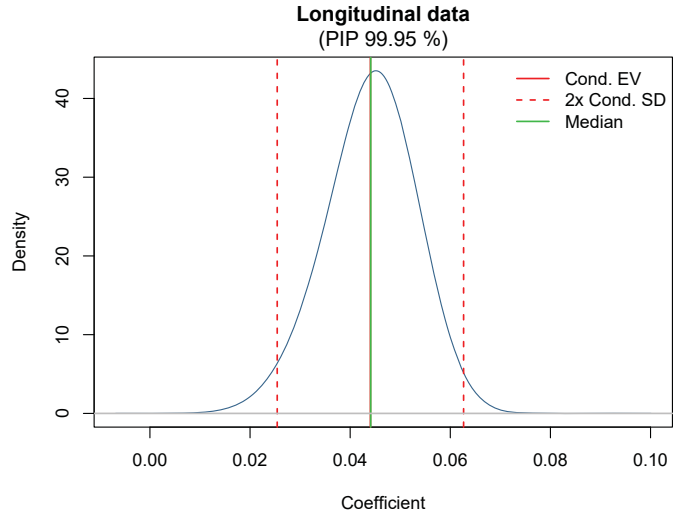


Closer look: endogeneity, panel data, Germany









Best practice estimates

	Mean	95% conf. int.	
USA	-0.039	-0.078	0.001
Germany	0.019	-0.046	0.083
Other countries	-0.048	-0.104	0.008
Male students	-0.039	-0.083	0.005
Female students	-0.037	-0.079	0.005
Part-time students	-0.025	-0.069	0.020
Low-intensity employment	-0.023	-0.068	0.022
High-intensity employment	-0.054	-0.098	-0.010
Educational outcome: choices	-0.062	-0.107	-0.017
Educational outcome: test scores	-0.033	-0.075	0.009
Educational outcome: attainment	-0.033	-0.080	0.014
Overall effect	-0.038	-0.079	0.003

Main Findings

1. Publication and endogeneity biases interact.
2. The corrected effect is negative but economically zero.
3. Effects negative for high-intensity work, positive for Germany.

Project Website

www.meta-analysis.cz/students