LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Weight and See? Investigating the Strategic (Mis)Use of Survey Weights in Social Research

Daniel Krähmer | 27 October 2025

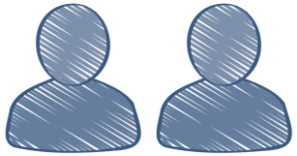
Leibniz Open Science Day 2025

Funded through the DFG Priority Program META-REP (Project No. 464507200).

A Motivating Example...

We show that *X* affects politicians' attitudes.

Our replication efforts suggest that the original findings may be **sensitive to the use of survey weights**.



This is an honest question: Isn't it **standard practice to use weights** when using this data? In any case, happy to adapt if there is a clear consensus on this.

Paraphrased from Chu & Recchia (2022); N.N. (2023)

…and the Larger Problem

"Top-notch empirical scholars make conflicting choices about whether and how to weight and often provide little or no rationale for their choices"

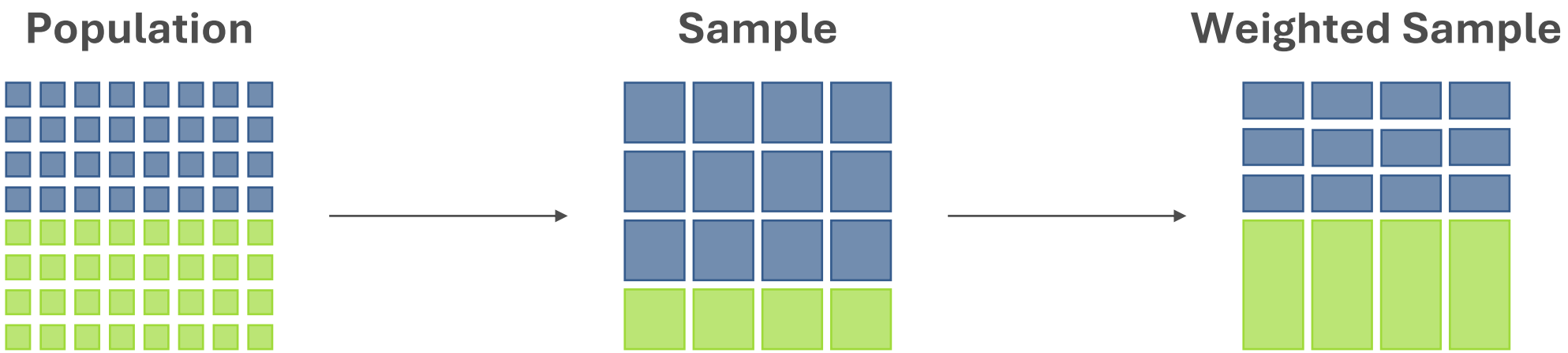
(Solon et al. 2015: 301)

"Some always weight and others never do"

(Bollen et al. 2016: 390)

The Weighting Controversy

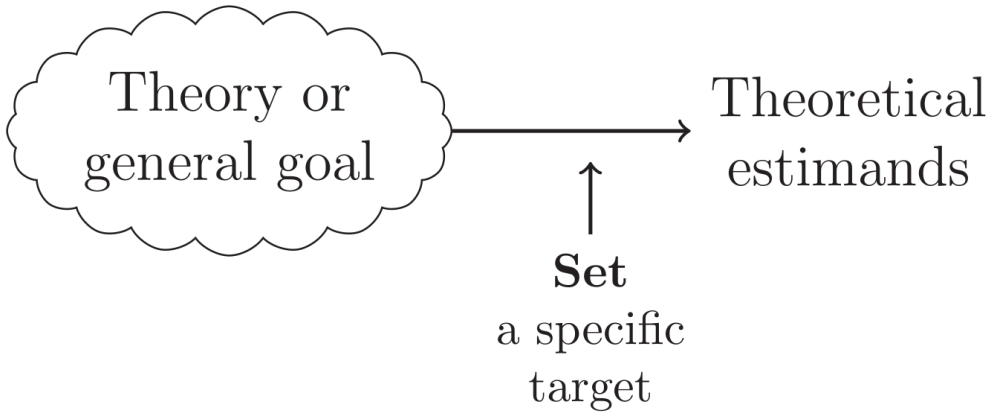
Pro: Weighting improves representation



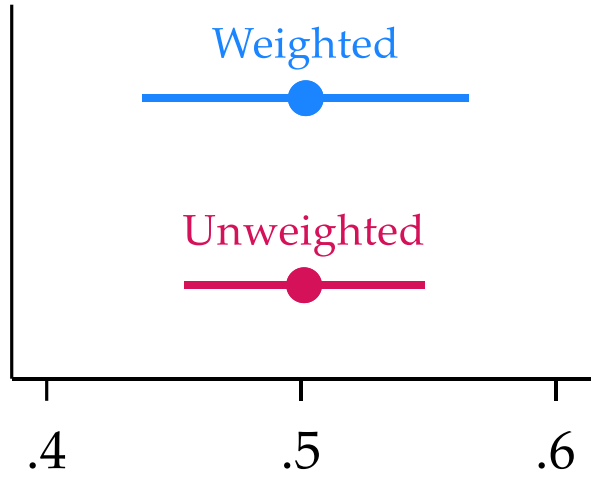
Solon et al. (2015); Kalton (1989)

The Weighting Controversy

Con: Weighting is difficult and costly



⚡ **Estimand?** ⚡



⚡ **Precision?** ⚡

Lundberg at al. (2021); Solon et al. (2015); Bollen et al. (2016)

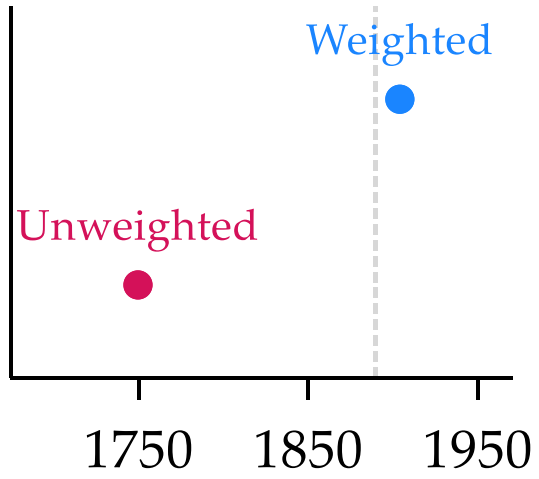
Why It's a Problem

Weights matter.

They can impact:

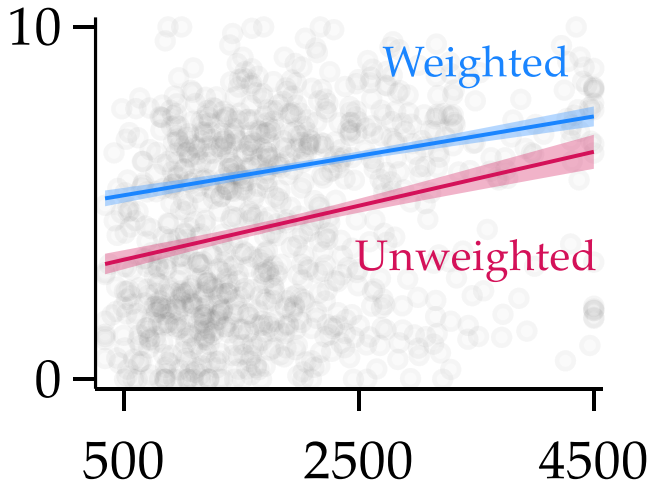
Univariate

point estimates

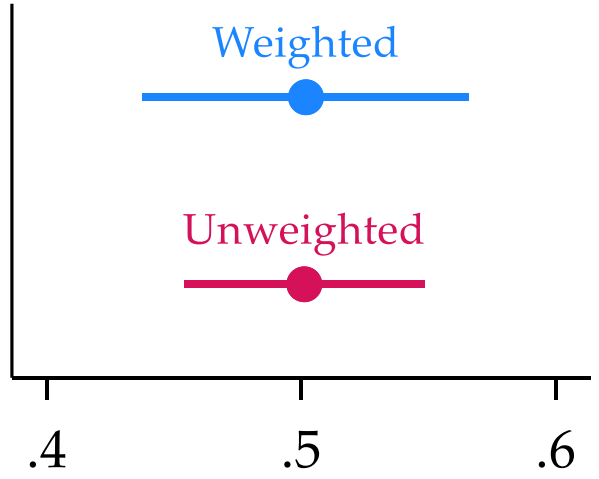


Multivariate

point estimates



Standard errors



What Makes a Good *p*-Hack?

Ambiguity ✅

Impact (✅)

- Simulations (Cai 2012; Becker & Ismail 2016)
- Stylized examples (Korn & Graubard 1995; Pfeffermann 1996)
- Case studies (e.g., Boto-García 2024; Hahs-Vaughn 2005)

→ Are survey weights misused strategically?

Auspurg & Hinz (2011), Feigenbaum & Levy 1993, Abele-Brehm & Bühner (2016), Simmons et al. (2011); Credit: FlatIcon

Research Design

Data:

- 69 studies, 2015 – 2020
- European Social Survey (ESS)
- 360 reproducible empirical results
- Software: Stata

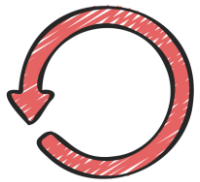
ESS Weighting Guide 2015: "The most accurate estimates will be obtained only after weighting."

Weight	Corrects	Usage
1) dweight	sampling design	country-round analysis
2) pspwght	non-response	
3) pweight	population size	pooled country analysis

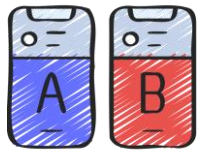
Procedure:

check original

repeat analysis

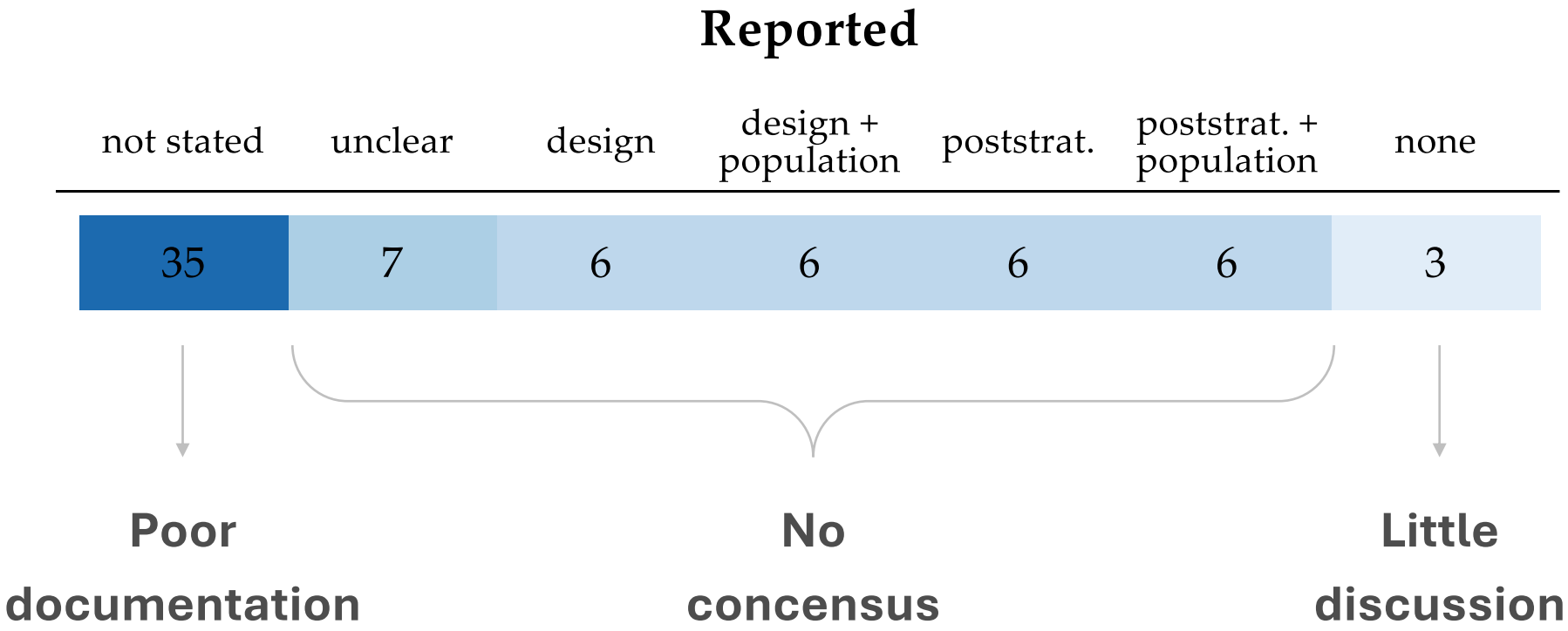


compare results

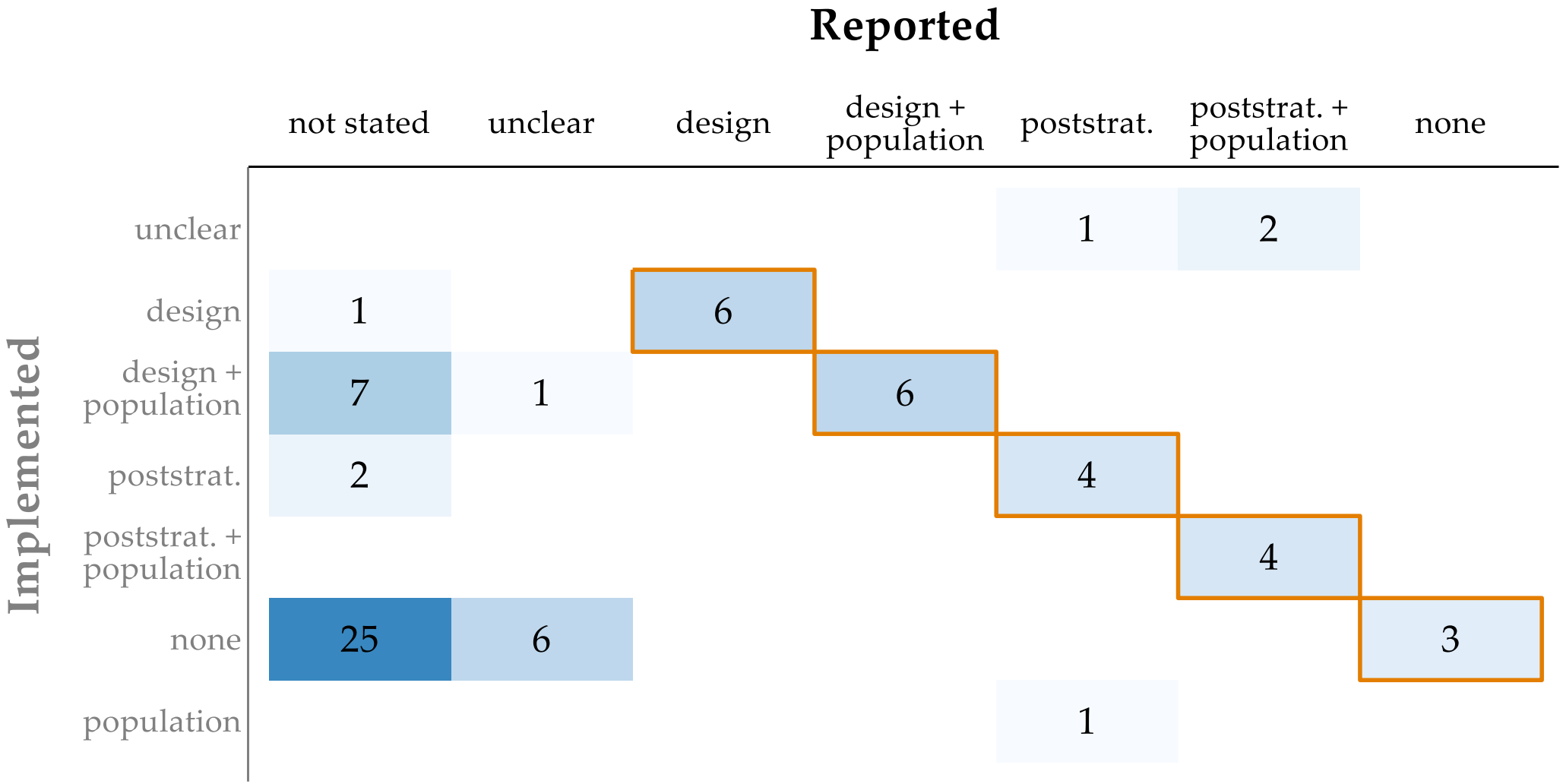


Credit: FlatIcon

Result 1: Reported Weighting Practices

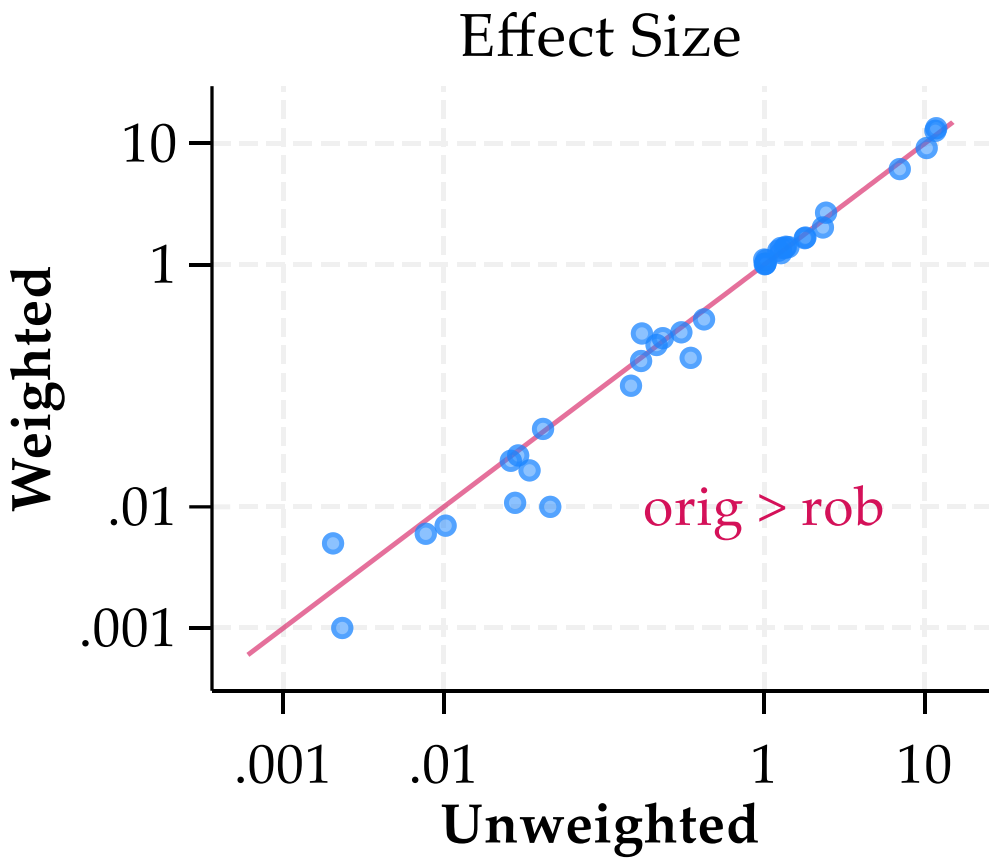


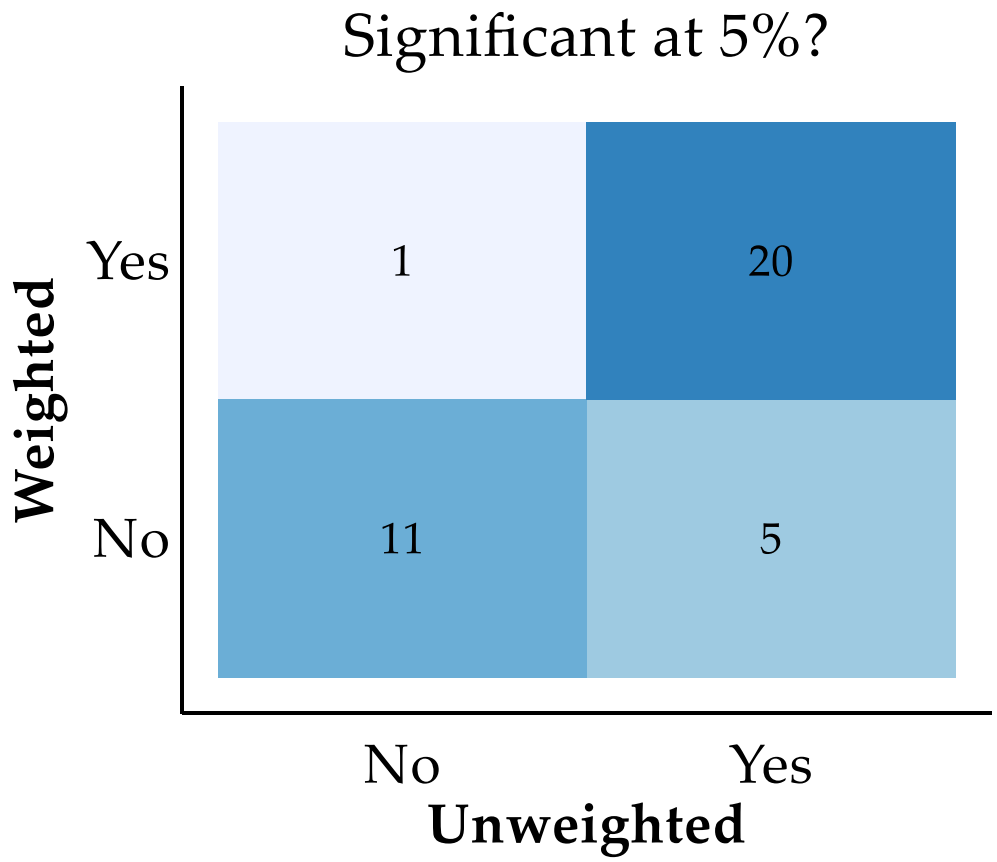
Result 2: Implemented Weighting Practices



Result 3: Impact of Weights (poststrat. + population)

Preliminary:





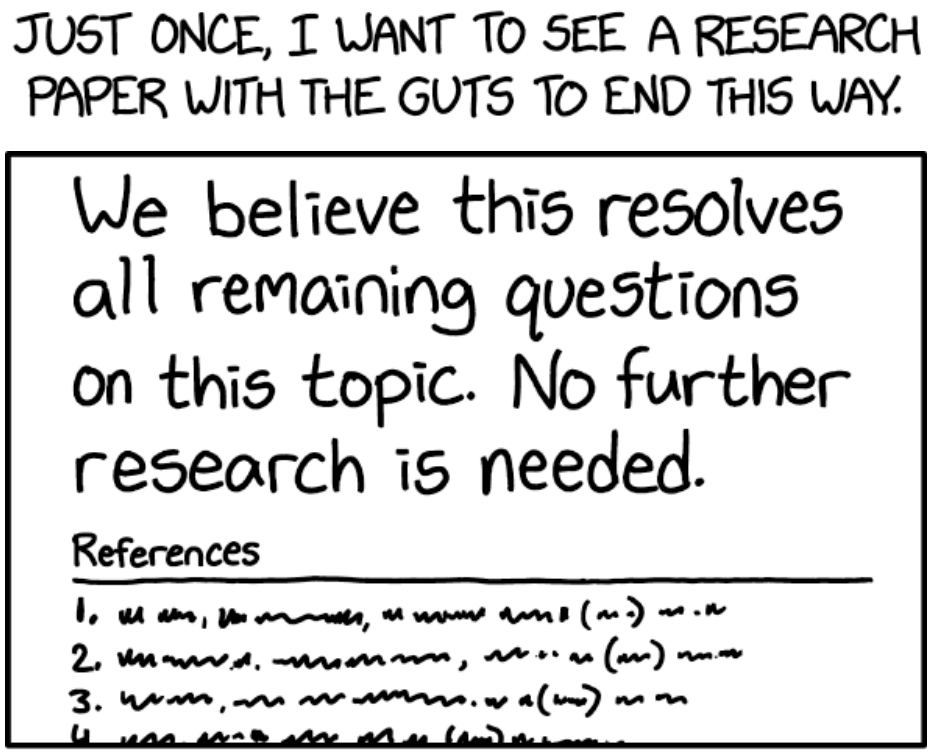
Summary & Discussion

Main findings:

- Confusion & inconsistencies surrounding weights
- Preliminary:
  - Little impact on results
  - No systematic bias

Impliactions & Limitations:

- Generalizability?
- We need more...
  - Transparency in documentation
  - Clarity in estimands



Credit: XKCD (modified)

THANKS!

✉ Daniel.Kraehmer@soziologie.uni-muenchen.de

⌨ www.danielkraehmer.com