

Method for selecting calibration weights in a non-probability epidemiological survey

Black JC¹, Forber A¹, Rockhill KM¹, Amioka EC¹, May KM¹, Dart RC¹

¹Rocky Mountain Poison & Drug Center - Denver Health, Denver, CO

Introduction

- The Survey of Non-Medical Use of Prescription Drugs (NMURx) Program is a study that collects information on drug use and abuse using an online survey instrument.
- In rapidly-changing fields such as the study of drug use, accurate and timely data is essential to properly inform policy and intervention decisions.
- Validity of responses is often a concern for anonymous online surveys, and NMURx is further challenged by collecting information on rare outcomes related to behaviors that may encounter social stigmas.
- In order to improve data quality and key outcome estimates, a method was developed that analyzed multiple weighting schemes and selected the best improvement in key estimates.

Methods

The Researched Abuse, Diversion and Addiction-Related Surveillance (RADARS[®]) System NMURx Program deploys a general population survey via an online panel company. The goal of this program is to provide accurate and timely estimates of prescription drug non-medical use and associated motivations and behaviors among the general population. The NMURx Program employs a cross-sectional, opt-in online survey that utilizes survey panels to contact respondents. The questionnaire is a self-administered, anonymous survey completed by respondents on their own time, in their own setting, and on their own device.

The survey was deployed from 28 Sept. 2018 to 21 Nov. 2018. There were 148,274 invitations sent to panelists; the participation rate was 20.2%. A total of 29,841 respondents were utilized in this analysis after exclusion criteria were applied.

Eight weighting variables were tested:

- Sex
- Census region
- Age category
- Household size
- Household income
- Self-assessed health status
- Cigarette smoking status
- Limitation in daily activity

Sex, census region, and age category were included in all schemes. The remaining variables were included in all possible combinations, resulting in 33 total schemes tested.

For each scheme, the relative difference was calculated individually for 26 benchmark estimates from probability surveys. Benchmarks included health estimates (e.g. illicit drug use, hospital stays) and non-health estimates (e.g. race, employment status). All relative differences were averaged into a single measure of difference between NMURx estimates and probability survey estimates. The standard error was also calculated for each benchmark, and standard errors averaged into a single measure.

Results

Key Results

- The best scheme reduced the average relative difference in benchmarks by 31.2%, from 36.1% (unweighted) to 24.8% (weighted)
- Weighting improved health characteristic estimates relative to benchmarks of the sample and lowered drug use estimates to become more in line with benchmarks
- Weighted estimates of drug use were comparable to national probability survey estimates; these estimates were not used in determining the weighting scheme

Figure 1: Relative Difference and Standard Error

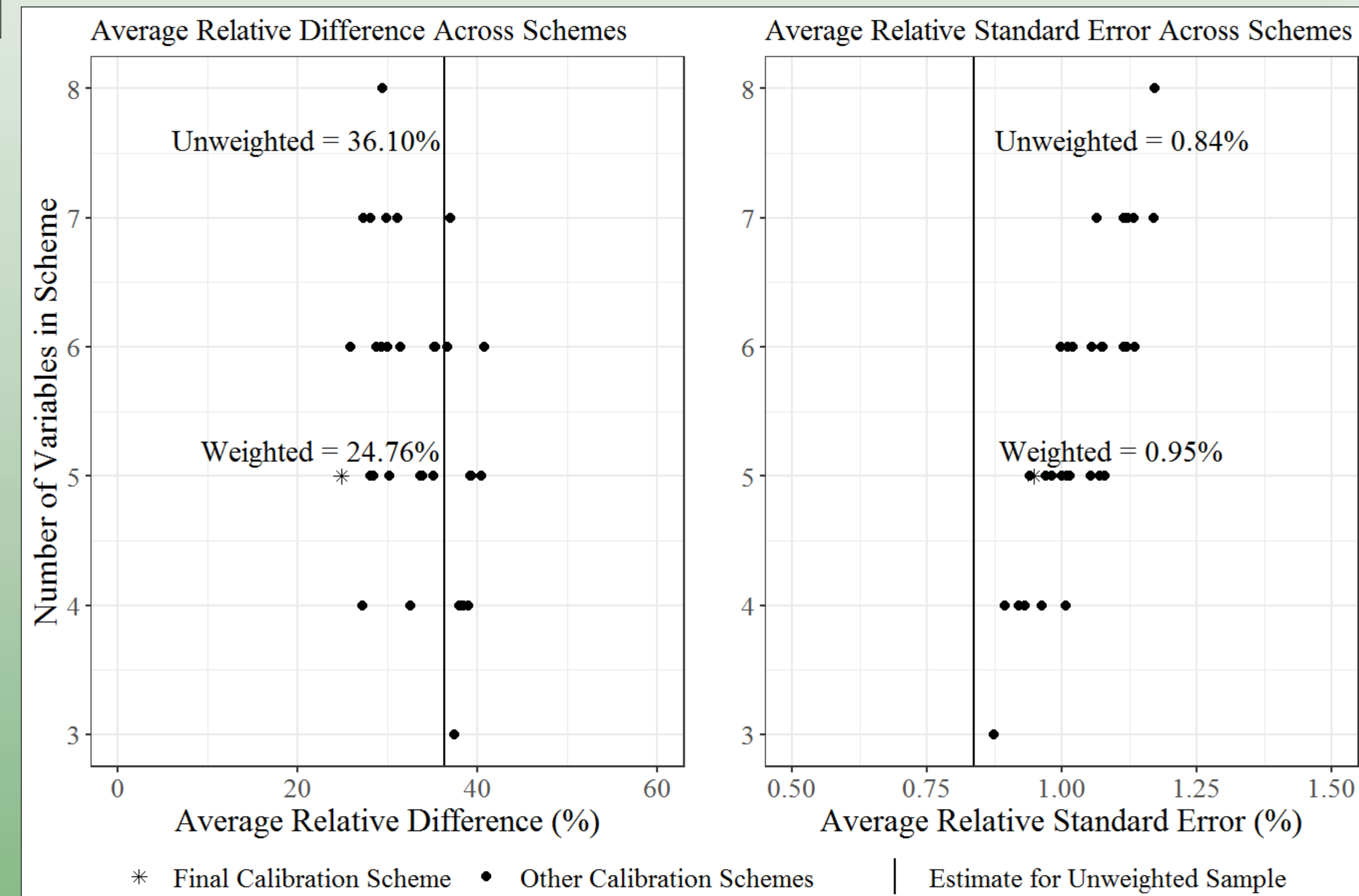
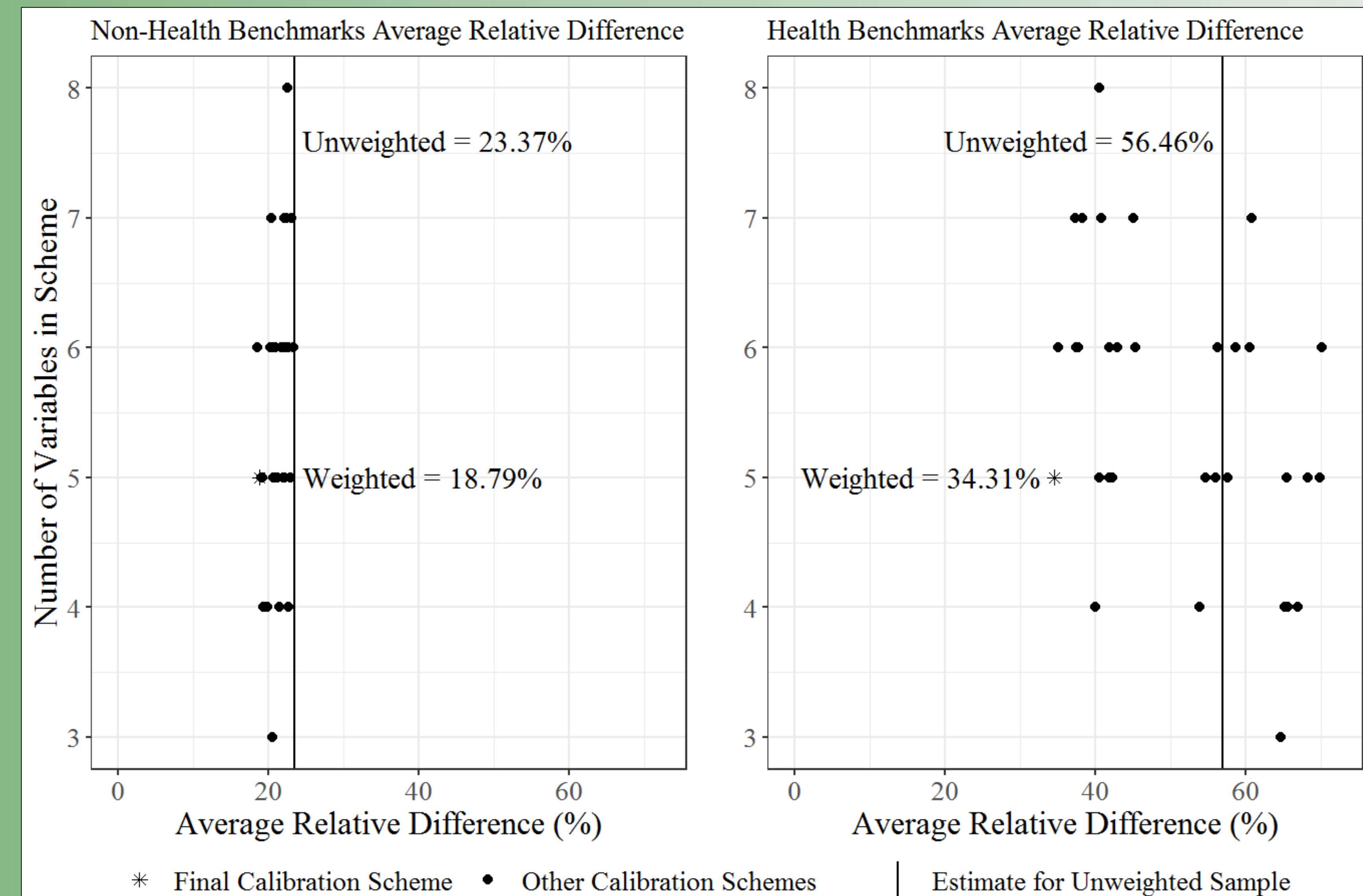
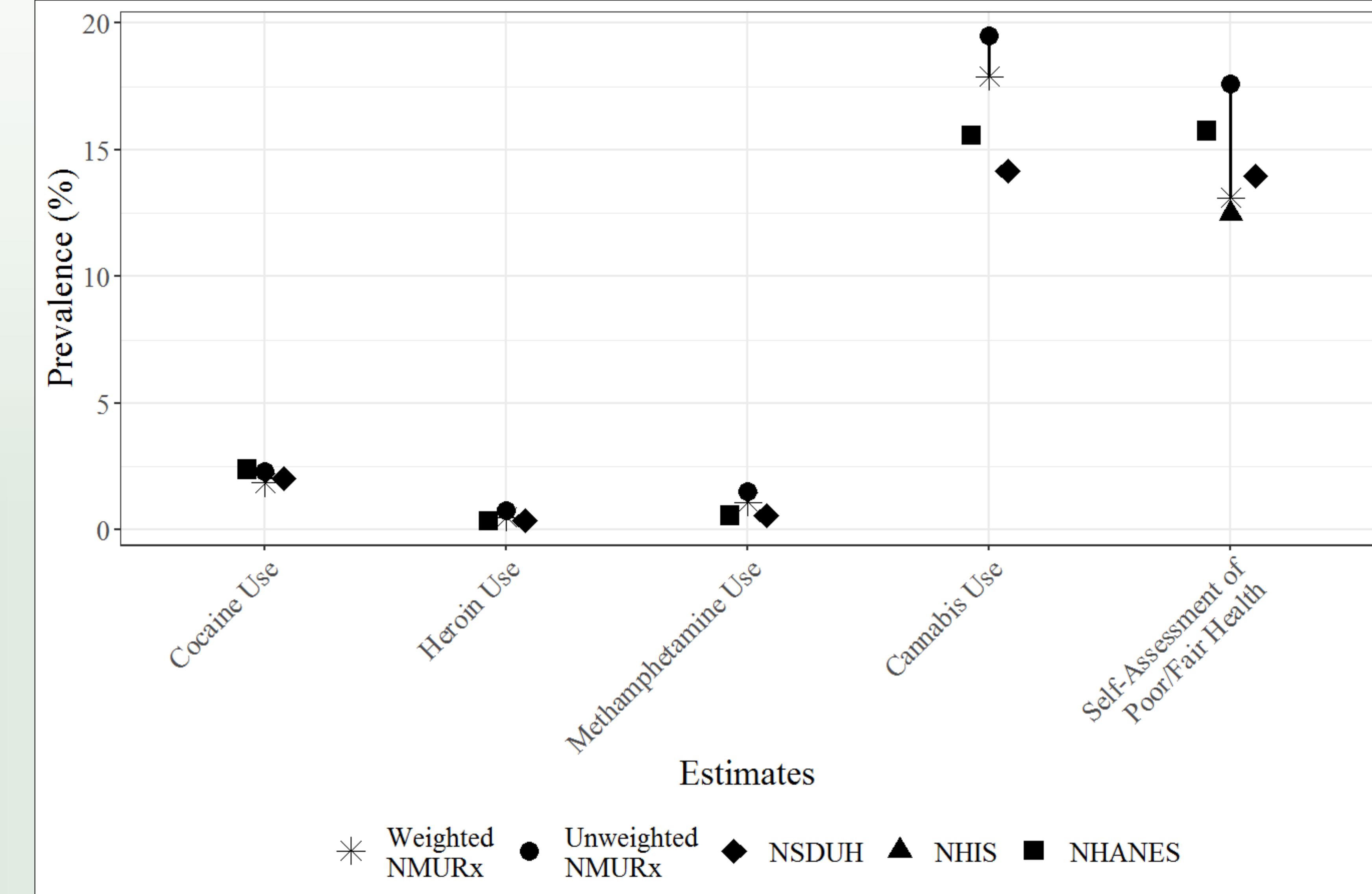


Figure 2: Relative Difference Stratified by Benchmark Type



Results (cont.)

Figure 3: Comparison to Probability Surveys



Conclusions

- The weighting scheme that included sex, census region, age category, cigarette smoking status, and limitation in daily activity produced the smallest average relative difference
- Use of health-related weighting variables had the most influence on drug use estimates
- Differences in health benchmarks varied widely across schemes and were heavily influenced by the inclusion of health-related weighting variables
- Convergent validity was shown against probability surveys in a more timely manner than large governmental survey efforts
- 95% confidence intervals represent precision within the sampling framework

Limitations

- 10% of Americans do not use the internet, which represents a gap in the sampling frame because sampling was taken from internet users
- 95% CIs derived from this method do not represent true 95% probabilities because exact selection probability from the sampling frame into the sample is unknown.
- Due to complete anonymity of the survey, information on non-responding panelists was not available and therefore, non-response adjustment was not included.
- Questions and timeframes are not identical between the NMURx survey and the probability surveys; therefore, estimates are not expected to be identical