**Supplemental Appendix: A rigorous evaluation of a method to adjust BMI for self-report bias**

Zachary J. Ward, PhD1\*, Steven L. Gortmaker2, PhD

1. Center for Health Decision Science, Harvard T.H. Chan School of Public Health, Boston, MA
2. Department of Social and Behavioral Sciences, Harvard T.H. Chan School of Public Health, Boston, MA

\*Corresponding author:

Zachary J. Ward, MPH, PhD

Center for Health Decision Science

Harvard T.H. Chan School of Public Health

718 Huntington Ave, Boston MA, 02115

Email: [zward@hsph.harvard.edu](mailto:zward@hsph.harvard.edu)

Tel: 617-432-2019

Fax: 617-432-0190

**Methods: NHANES Cross-validation**

We obtained data from NHANES 1999-2018 for 45,959 adults with no missing variables of interest. We performed 1,000 iterations of cross-validation in which we randomly sampled half of NHANES respondents into a ‘training set’ and the other half into a ‘testing set’. We adjusted self-reported BMI in the ‘testing set’ using the measured BMI values in the ‘training set’, and compared the adjusted BMI for each individual to their measured BMI. We used the NHANES sample weights in our primary analysis (as is recommended), but also repeated the analysis without sample weights to explore the impact of weighted vs unweighted bias-correction. We estimated the mean error and 95% confidence intervals (CI, calculated as the 2.5 and 97.5 percentiles of the simulated results) of self-reported and adjusted mean BMI and obesity prevalence for the total adult population, and for various subgroups: sex (male, female), race/ethnicity (White non-Hispanic, Black non-Hispanic, Hispanic, Other non-Hispanic), household income (<$20,000, $20,000-<$55,000, $55,000+), smoking status (never smoker, current smoker, former smoker), and age group (18-39, 40-64, 65+). We applied our bias-correction method by sex, and also by sex and age group. Our full R code to replicate this analysis is available at: <https://github.com/zward/BMI-bias-correction>.

We find that unweighted bias-correction generally resulted in larger errors by subgroup, especially for current smokers, highlighting the importance of using sample weights when analyzing complex survey data.