



Evaluating a Web/Mail Alternative to a National Face-to-Face Survey: Initial Results from the American Family Health Study (AFHS)

BRADY WEST, MICK COUPER, BILL AXINN,
JAMES WAGNER, REBECCA GATWARD,
HTAY-WAH SAW, AND SHIYU ZHANG

[INSTITUTE FOR SOCIAL RESEARCH, UNIVERSITY OF
MICHIGAN-ANN ARBOR]

*Thank you to the NIH for support
of the American Family Health
Study (AFHS) via NIH Grant
R01HD095920 (PI: West).*



American Family Health Study

afhs.isr.umich.edu

What's The Problem?

Rising costs of large-scale face-to-face data collections, combined with rising survey nonresponse rates and reluctance of the public to participate in time-consuming in-person data collections, threaten the future of national face-to-face surveys.

The pandemic has further eroded researchers' reliance on face-to-face surveys.

These trends create an urgent need for the evaluation of **innovative web-based data collection methods** that are convenient for the general public and yield high-quality scientific information for population researchers.

What's The Solution?

The development of an alternative data collection methodology is especially important for research teams without the resources of larger government agencies.

The **web mode** is particularly appealing because it is **relatively inexpensive** and affords a **high level of privacy and confidentiality** when correctly implemented.

Today, we present initial results from a **sequential mixed-mode web/mail data collection approach** that was implemented on a **national probability sample** in 2020-2021, and compare them to results from a benchmark face-to-face national survey of population health: the 2017-2019 National Survey of Family Growth (NSFG).

Study Objectives

1. Evaluate the **features of the respondents recruited** using this approach and the population that they would represent prior to any weighting adjustments
2. **Compare survey estimates** based on this web/mail approach to those computed from NSFG, **measuring the same health content** in a similar time frame
3. Evaluate **design effects on survey estimates** due to the complex probability sampling designs employed in each case
4. **Compare the data collection costs** per completed survey associated with these two approaches

AFHS Design Summary

The **American Family Health Study (AFHS)** recently completed data collection with a national address-based probability sample of 19,000 U.S. addresses (see afhs.isr.umich.edu for details).

Sequential mixed-mode protocols involving **push-to-web** and mail follow-up were applied at both the screening (target population = 18-49 years old) and main data collection stages.

We also experimented with modular design; for purposes of this analysis, data from the modular and full surveys were combined, as few differences were observed.

AFHS Design Summary, cont'd

Screening Protocol:

1. Invitation letter with web link and \$2 pre-paid incentive
2. Reminder postcard with web link one week later
3. Reminder letter with paper screener included two weeks later
4. A random subsample of 5,000 nonrespondents received a fourth priority mailing with \$5 included

AFHS Design Summary, cont'd

Main Protocol:

1. Invitation letter with web link to complete full survey or module; promised incentive of up to \$70
2. Reminder postcard / email (if provided) two weeks later
3. Shorter paper version of questionnaire / module mailed four weeks later, with web reminder
4. Telephone reminders to complete the survey six weeks later (phone numbers from completed screeners or linked MSG data)

Response Rates

The AFHS obtained a response rate in the screening stage of 15.0% (n = 2,556) and a conditional AAPOR RR4 response rate of 66.0% in the main stage.

For individuals assigned to the modular condition, completing at least two sections of the questionnaire in the first module was counted as a partial response.

These two rates resulted in a net AAPOR RR4 response rate of 9.9% (n = 998).

We found that 89% of AFHS respondents participated via the web mode, while the remaining 11% returned the shortened paper questionnaire.

Objective 1: Representation

The AFHS recruited more respondents that were **non-Hispanic White** and **higher-educated** (for both males and females; $p < 0.01$); the male respondents in the AFHS also tended to be younger than in the NSFG, although this difference was weaker ($p = 0.031$).

These results are not unique to the AFHS, and imply that nonresponse adjustments to the base sampling weights may be needed to correct for potential biases in estimates due to race/ethnicity and education.

Whether there would be bias in estimates due to these differentials of course depends on the associations of race/ethnicity and education with the measures of substantive interest collected in the AFHS.

Objective 2: Key Estimates

We identified 42 and 89 key measures in the male and female surveys, respectively, capturing data on critical domains of family reproductive and health behaviors.

Both AFHS and NSFG estimates were weighted by the final survey weights, and design-adjusted standard errors were computed for the weighted estimates accounting for the complex sampling features inherent to each study.

We assessed the similarity of estimates based on their standard errors and confidence intervals, looked at changes in estimates across the four stages of the main data collection protocol, and examined whether the web and mail respondents in AFHS varied in response distributions.

Objective 2: Key Estimates

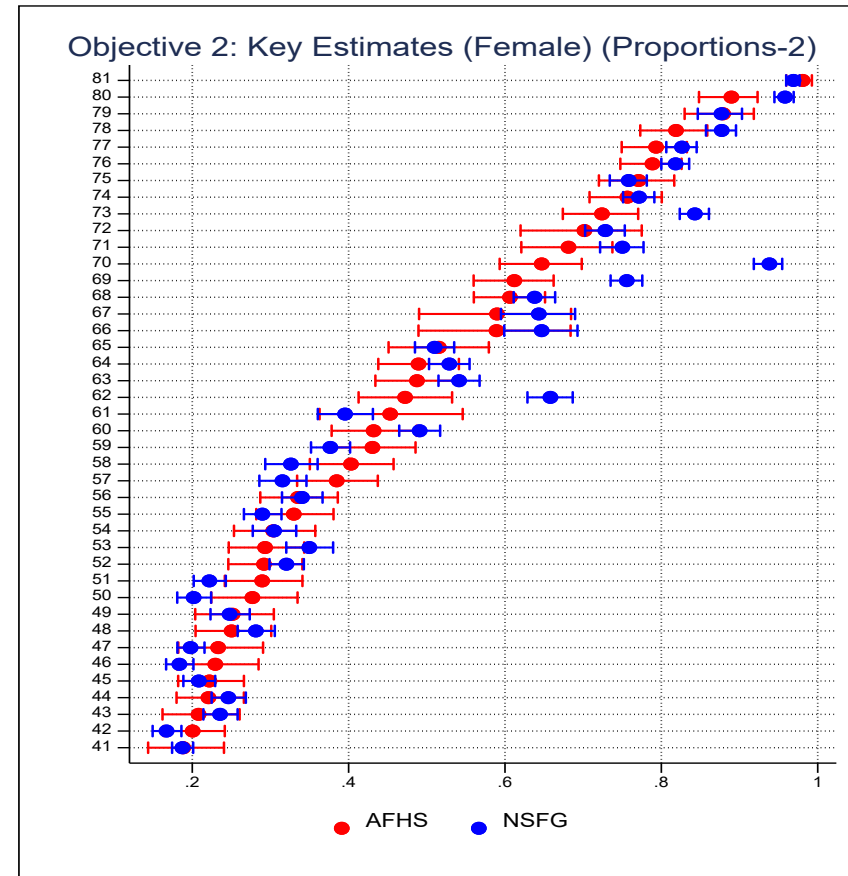
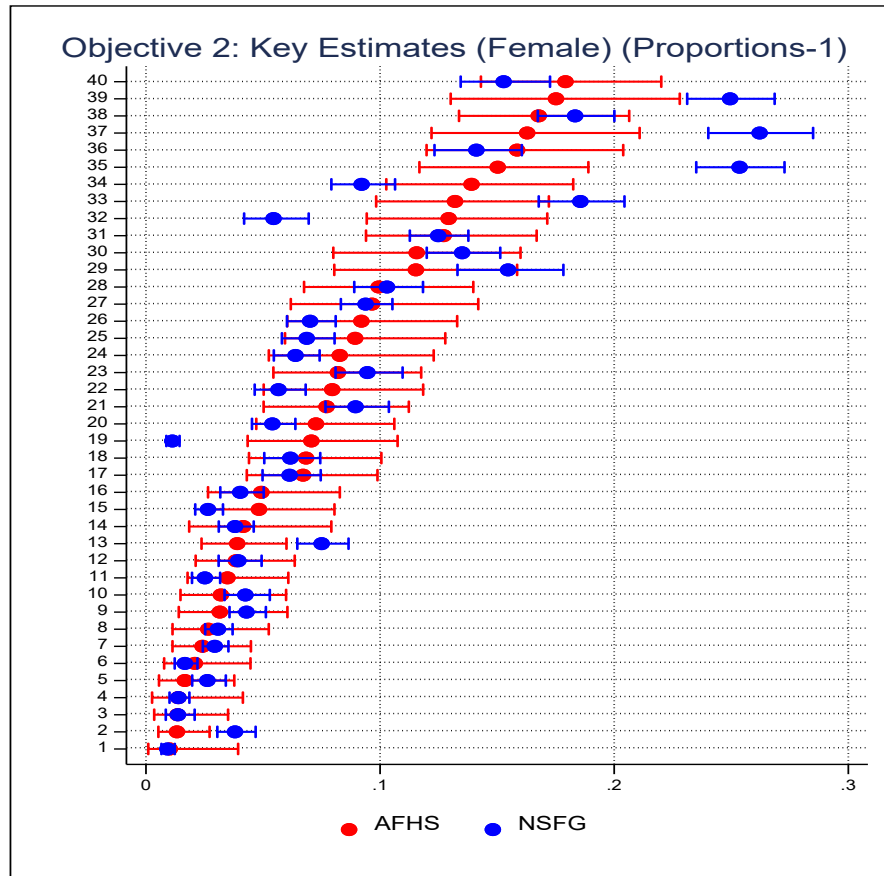
71% of the AFHS estimates had confidence intervals that covered the NSFG point estimates, and for only 24% of estimates was the standardized difference of the estimates between the two studies more than two pooled standard errors away from zero.

In general, we found that the weighting adjustments applied to the AFHS estimates were more effective at shifting the estimates closer to the NSFG benchmarks than the sequential mixed-mode design, although weighting tended to increase the standard errors (SE).

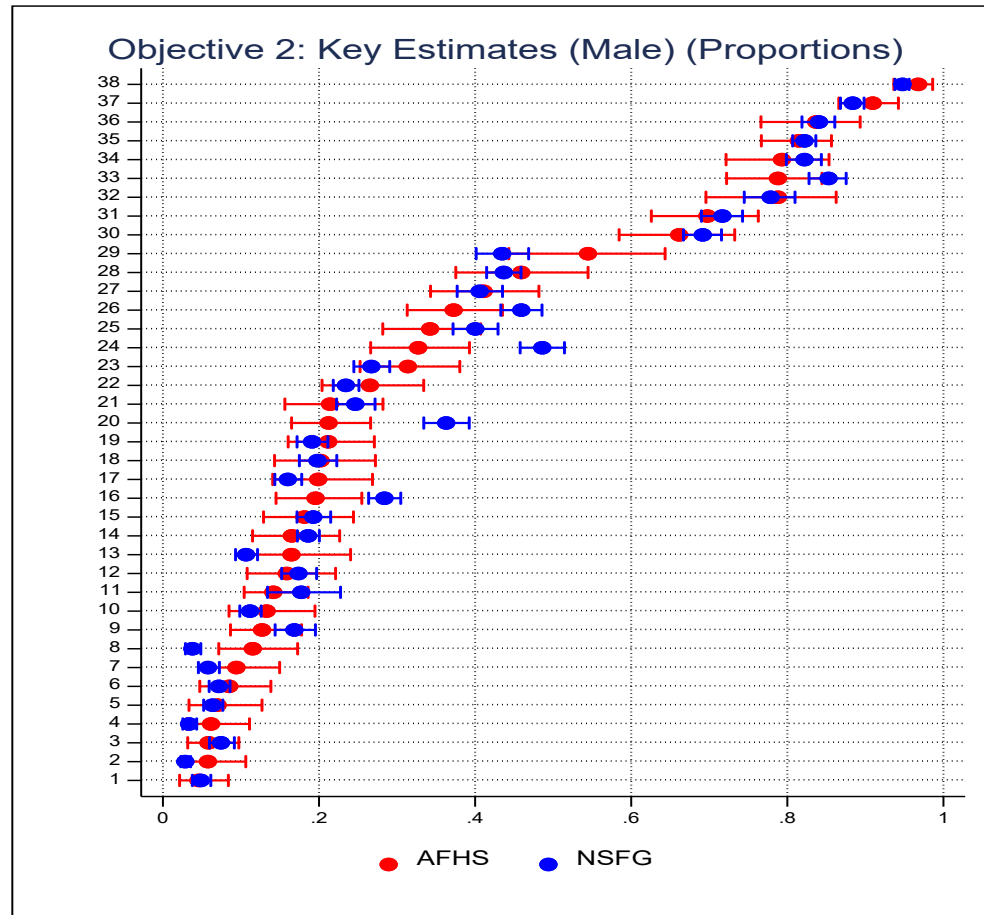
The weighted estimates based on the full sample were quite similar to those based on the web subsample.

In general, the AFHS and NSFG estimates were well-aligned (see figures to follow).

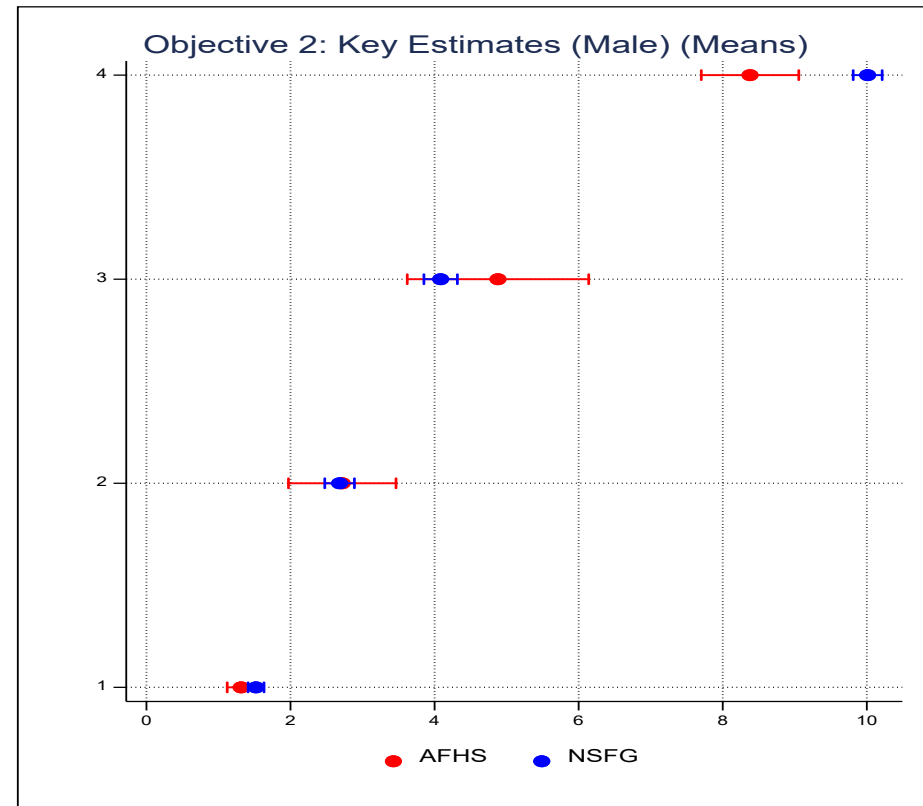
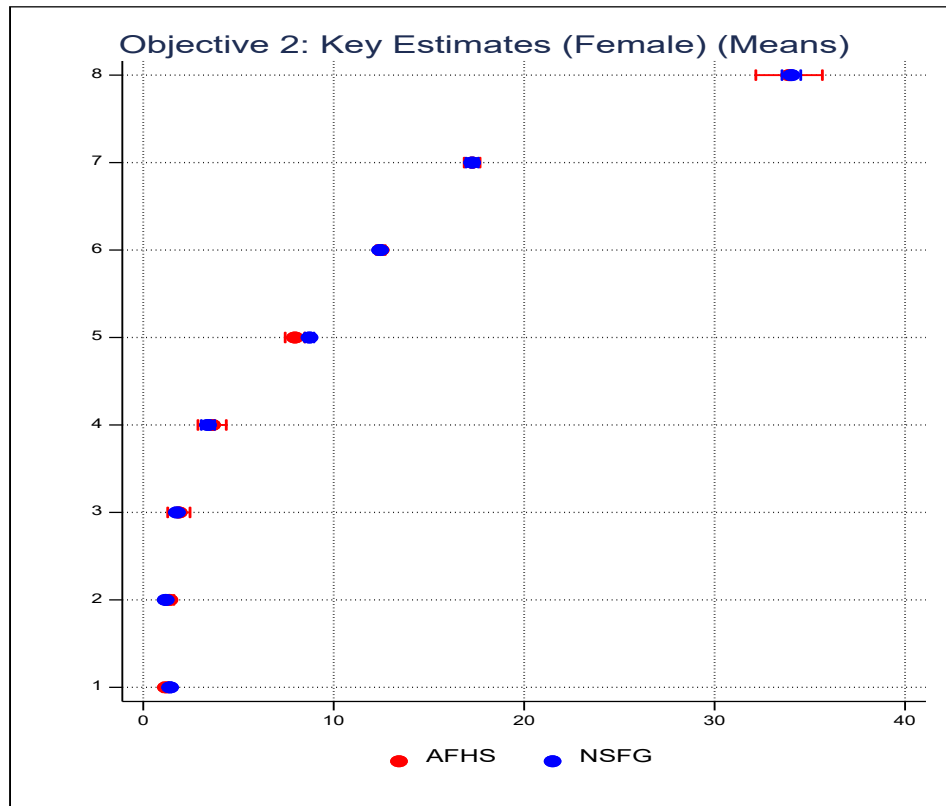
Objective 2: Key Estimates



Objective 2: Key Estimates



Objective 2: Key Estimates



Objective 2: Key Estimates

We find that significant differences between the NSFG and the AFHS generally arise for measures that were likely affected by the pandemic.

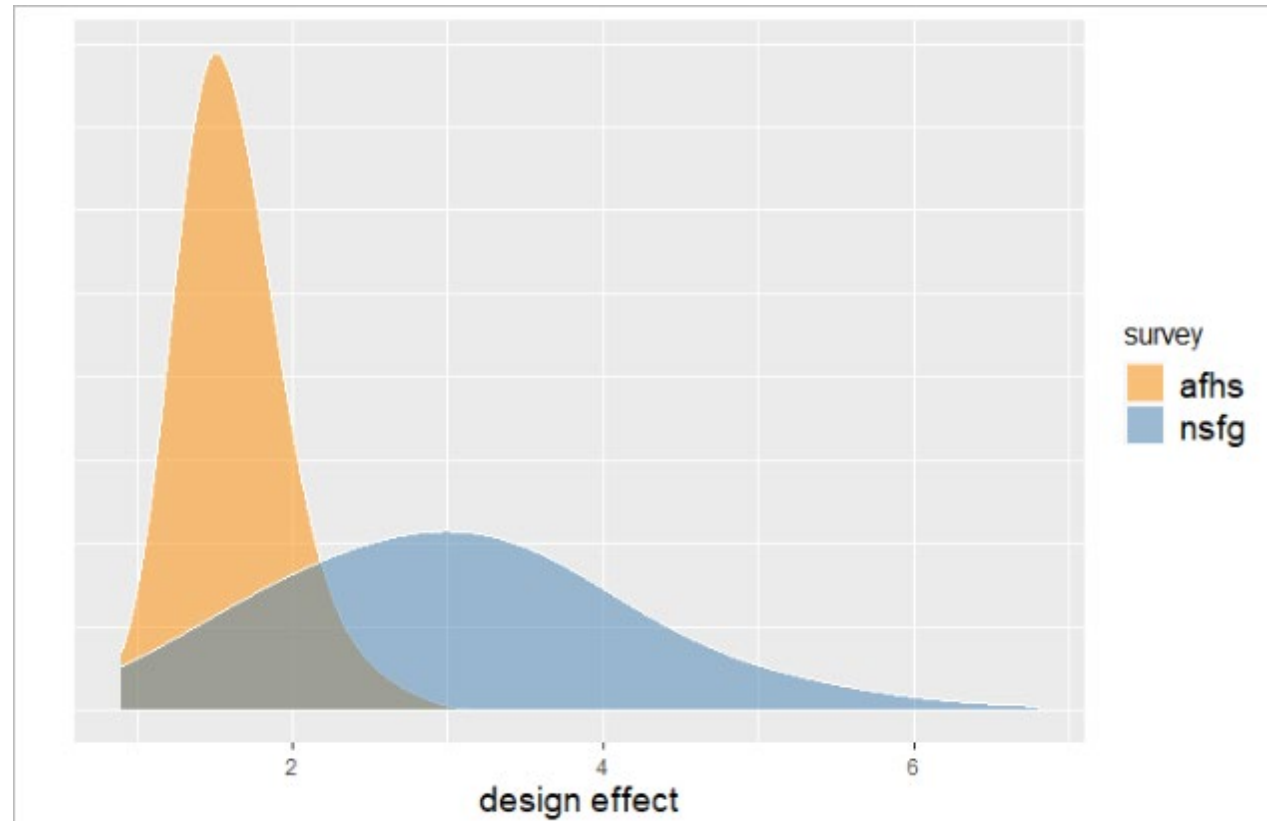
For example, among males, the mean number of months working for pay in the past year was noticeably lower in the AFHS; was this a function of the so-called “Great Resignation”?

Use of birth control and emergency contraception among females also declined significantly in the AFHS, possibly due to changing social circumstances introduced by the pandemic.

Reports of “excellent” health and ever being tested for HIV also declined significantly for both males and females in the AFHS, possibly reflecting poorer health during the pandemic and less in-person medical visits for clinical testing. **Also: was this simply more honest reporting via web?**

Objective 3: Design Effects

The lack of cluster sampling (only stratification) in the AFHS sample design proved incredibly important!



Objective 4: Costs

Based on data from the **last four quarters** of the 2017-2019 NSFG, the AFHS produced significant cost savings per completed survey (\$300).

Considering the average design effects of 3 and 1.5, the estimated costs per effective n are $\$717 \times 3 = \$2,151$ (NSFG) and $\$417 \times 1.5 = \626 (AFHS).

	NSFG (Q29-Q32)	AFHS (2020-2021)
Completed Surveys	5731	998
Cost per Completed Survey	\$717	\$417

Summary of Findings

1. The web/mail approach recruits more non-Hispanic White and higher-educated individuals; careful weighting approaches can compensate for this potential source of nonresponse bias
2. About three-quarters of the estimates were statistically similar to those produced by the NSFG, with the remaining one-quarter likely having shifts introduced by the pandemic
3. Design effects on the variances of estimates due to complex sampling are a fraction of those found in the NSFG, largely owing to the absence of area cluster sampling in the AFHS
4. The cost per completed survey was \$300 less in the web/mail approach compared to the NSFG when considering all data collection activities

Thinking About the Future

Can clever adaptive design approaches for web/mail surveys (e.g., SMART designs) lessen the role that weights need to play in adjusting estimates?

Comparisons of estimates for key socio-demographic subgroups are critical as well, which will require larger sample sizes.

- We just finished data collection for our second replicate (now, $n = 2,373$), which will help.

Analyses of paradata generated from this approach will also be critical to understand whether aspects of the approach work better or worse for certain subgroups, with future adaptive designs in mind.

Thank You!

Please email Brady West at bwest@umich.edu with any questions, and visit the AAPOR 77th library after the conference for access to all of these materials.