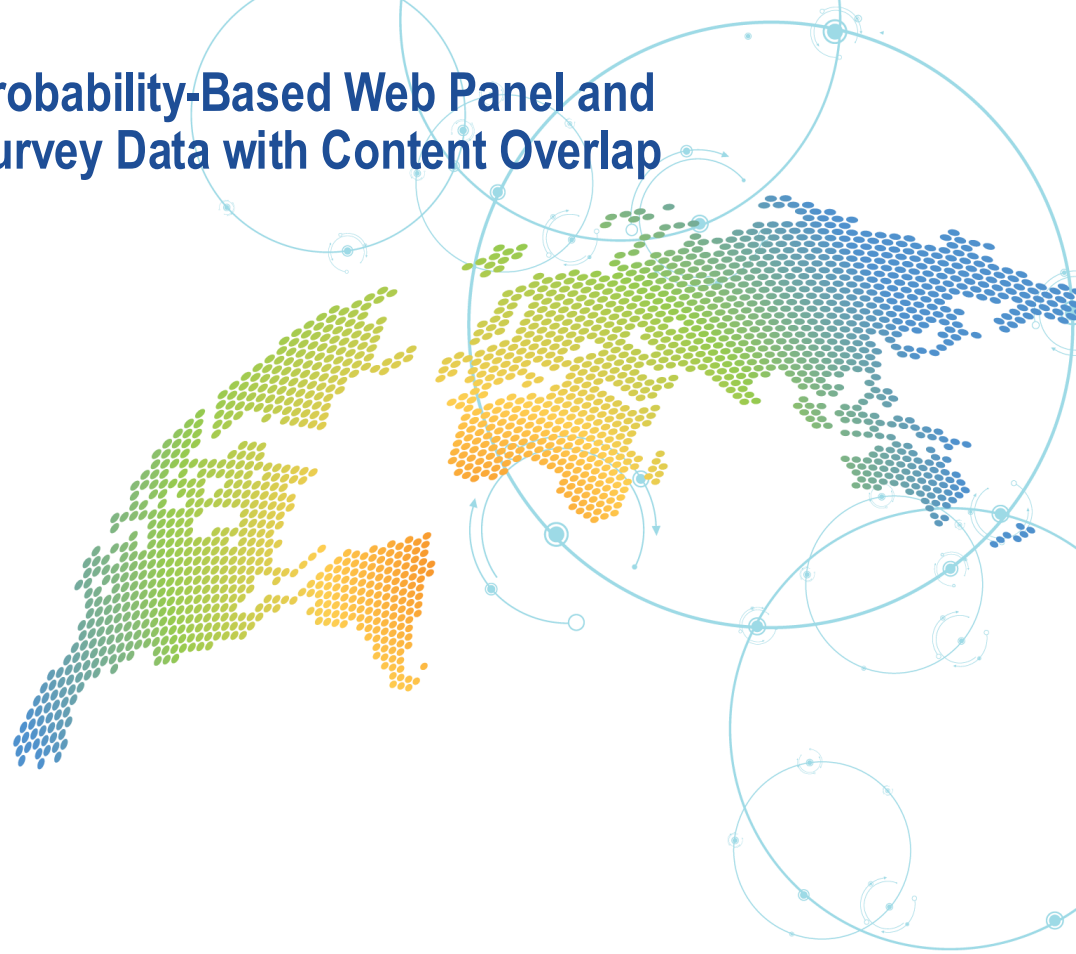


Multiple Imputation for Combining Probability-Based Web Panel and National Interviewer-Administered Survey Data with Content Overlap

Andy Peytchev, PhD

CIPHER, February 27, 2026

Washington, D.C.



Disclaimer and Acknowledgment

- The author has an advisory role on a contract for the National Center for Health Statistics' (NCHS) Rapid Survey System (RSS). The independent research in this presentation was not supported by this contract, used publicly-available data, and does not express the views of the Government.
- The author is grateful to NCHS for making these data available to researchers, making investigations like this possible and informing future data collection designs.



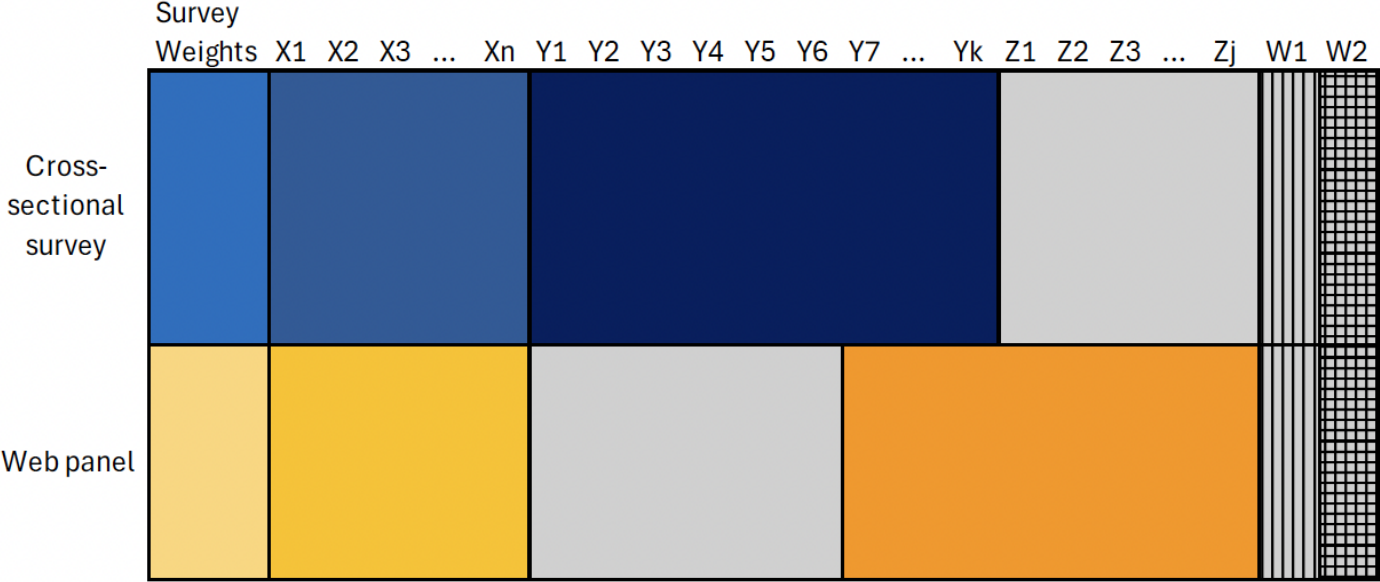
A New Paradigm

- Probability-based web panels offer an opportunity to collect data at lower cost (and faster) compared to common cross-sectional household survey designs
 - Collect data on new and emerging topics
 - Track changes
- Can also be used to augment a more costly cross-sectional survey:
 - Adding records (longer dataset)
 - Adding variables (wider dataset)
 - Adding some records with some variables

Key Research Questions

1. If we combine data collected with higher and lower cost protocols, does it affect inference, particularly from bi- and multivariate analyses?
 - Would weighting adjustments be sufficient?
 - Would it help if we calibrate on variables that are atypical for weighting?
2. Lower cost protocols may not be able to collect as much data from each respondent, creating a non-rectangular dataset. Given some results from (1) that show some differences in covariances, would it be better to impute data for lower cost protocol respondents?
 - For point estimates?
 - For variance estimates?

Data Integration Framework



Xn	Demographic variables
Yk	Substantive variables from core NHIS modules
Zj	Substantive variables not part of NHIS
	Multiple imputation
W1	Combined data survey weights, postratified to socio-demographic controls (X)
W2	Combined data survey weights, postratified to socio-demographic (X) and substantive controls (Y)

NCHS's Rapid Survey System (RSS) and NHIS

- Periodic surveys using two probability-based web panels
 - The instruments are designed to include questions from the National Health Interview Survey (NHIS)
 - Calibration
 - Evaluation
 - The second round (RSS2) was administered by web (n=6783) with very few by telephone (263), in late 2023
 - An overall (cumulative) response rate of 4% for both panels
- NHIS 2023, Quarter 3 (n=7,419)
 - Response rate for adults, 47.0%

Methods

- Identified 32 variables that are common to NHIS and RSS2
 - Of these, 20 are health-related and 4 are volunteering-related
- Weight 1. Recalibrated the RSS2 respondent survey weights to NHIS estimates for sex, age (appx. 10-year intervals), race/ethnicity, education, and region
- Weight 2. Added the 24 health- and volunteering-related variables
- Evaluated the impact on bivariate and multivariate associations before and after the more elaborate weighting using topic-related variables
 - Bivariate correlations
 - Logistic regression

Differences in Correlations between RSS2 and NHIS 2023 Q3, with Socio-Demographic Controls and with Added Health Controls

Weight 1 (Demographic Characteristics)	poor	CHR	FCHR	CCHR	ACHR	C SOC	S SOC	S PAY	B PAY	P ACC	F ACC	I ACC	F PAI	FRIMM	S IMM	S MED	F MED	F MTL	M MTL	M CIG	S CIG	CIVIC	CIV	V CIG	M CIV	V CIV
Poor or fair health	0.000	-0.088	-0.080	0.012	-0.050	-0.055	-0.061	0.005	0.030	-0.060	-0.050	-0.099	-0.041	-0.077	-0.047	-0.067	-0.006	0.047	0.020	-0.006	0.085	-0.025	0.007	-0.068	0.012	0.020
Ever been told you had hypertension	-0.088	0.000	-0.011	0.001	-0.005	0.004	0.038	0.016	0.034	-0.009	-0.046	-0.027	-0.055	-0.028	-0.013	-0.006	0.039	0.035	0.057	0.027	0.056	0.012	0.020	0.012	0.020	0.020
Ever told you had high cholesterol	-0.080	-0.011	0.000	0.008	-0.002	0.000	-0.026	-0.015	-0.038	0.033	0.017	-0.006	-0.019	-0.023	0.002	-0.016	-0.011	0.064	0.010	0.027	0.045	0.041	0.022	0.029	0.034	0.029
Ever had asthma	-0.012	0.001	-0.008	0.000	-0.009	0.013	-0.007	-0.013	-0.029	-0.022	0.018	0.007	-0.007	0.007	-0.022	-0.013	-0.009	0.032	-0.024	-0.023	-0.044	-0.011	-0.004	-0.037	0.012	-0.037
Ever been told you had cancer	-0.050	-0.005	-0.002	0.009	0.000	0.038	0.020	-0.013	-0.022	0.007	-0.010	0.040	0.002	-0.005	0.020	0.002	-0.009	0.006	0.001	-0.007	0.034	0.005	0.014	0.031	0.012	0.031
Difficulty doing errands alone	-0.055	-0.064	-0.090	0.013	-0.038	0.000	0.033	0.069	0.118	-0.113	-0.034	-0.051	0.040	-0.090	-0.102	-0.067	0.085	0.113	0.063	0.025	0.086	-0.043	0.004	-0.104	0.004	-0.104
Work limited due to health problem	-0.061	0.004	-0.026	-0.007	0.020	0.033	0.000	0.038	0.047	-0.019	-0.004	-0.043	0.000	-0.051	-0.073	-0.001	0.060	0.130	0.044	0.018	-0.036	-0.021	0.034	-0.039	0.000	-0.039
Problems paying medical bills, past 12m	0.005	0.038	-0.015	-0.013	-0.013	0.089	0.038	0.000	0.075	-0.033	0.001	-0.001	0.035	0.003	0.021	0.007	0.078	0.080	0.036	-0.004	-0.037	-0.038	0.007	-0.037	0.000	-0.037
Get sick or have accident, worry about paying m	0.030	-0.016	-0.039	-0.029	-0.027	0.118	0.047	0.075	0.000	-0.014	0.007	-0.013	0.009	-0.016	-0.030	0.000	0.048	0.063	0.040	0.004	-0.027	-0.025	0.045	-0.032	0.000	-0.032
Time since last saw doctor	-0.060	0.034	0.033	-0.022	0.007	-0.113	-0.019	-0.033	-0.014	0.000	0.064	-0.026	-0.010	0.040	0.005	0.061	0.017	0.078	-0.014	0.034	0.022	0.029	-0.034	0.043	0.000	0.043
Have a usual place for care	-0.050	-0.009	0.017	0.018	-0.010	-0.034	-0.004	0.001	0.007	0.064	0.000	0.004	-0.034	0.016	0.019	0.055	0.056	0.060	0.012	0.016	-0.008	-0.002	-0.045	0.014	0.000	0.014
Hospitalized overnight, past 12 months	-0.099	-0.046	-0.006	0.007	0.040	-0.051	-0.043	-0.001	-0.013	-0.026	0.004	0.000	-0.030	-0.029	0.019	-0.020	0.069	0.005	0.021	-0.036	0.011	0.027	0.041	0.024	0.000	-0.024
How often had pain, past 3 months	-0.041	-0.027	-0.019	-0.007	0.002	0.040	0.000	0.035	0.009	-0.010	-0.034	-0.030	0.000	-0.025	-0.039	-0.064	0.133	0.072	-0.029	-0.007	-0.041	-0.029	-0.014	-0.019	0.000	-0.019
Ever had pneumonia shot	-0.077	-0.055	-0.023	0.007	-0.005	-0.090	-0.051	0.003	-0.016	0.040	0.016	-0.029	-0.025	0.000	-0.018	0.000	-0.007	0.038	0.043	0.001	0.016	0.051	0.020	-0.016	0.000	0.028
Flu vaccine, past 12 months	-0.047	-0.028	0.002	-0.022	0.020	-0.102	-0.073	0.021	0.030	0.005	0.019	0.019	-0.039	-0.018	0.000	-0.032	-0.011	0.085	0.004	0.009	0.035	0.023	0.000	0.028	0.000	0.028
Took prescription medication, past 12 months	-0.067	-0.013	-0.016	-0.013	0.002	-0.067	-0.001	0.007	-0.007	0.061	0.055	-0.020	-0.064	0.000	-0.032	0.000	0.047	0.154	-0.005	0.030	-0.060	-0.034	-0.052	-0.031	0.000	-0.031
Needed prescription medication but did not get	-0.006	-0.006	-0.011	-0.009	-0.009	0.085	0.060	0.078	0.048	0.017	0.056	0.069	0.013	-0.007	-0.011	0.047	0.000	0.105	0.035	-0.014	-0.035	-0.013	0.013	-0.056	0.000	-0.056
Took medication for other emotions/concentra	0.047	0.039	0.064	0.032	0.006	0.113	0.130	0.080	0.063	0.078	0.060	0.005	0.072	0.038	0.085	0.154	0.105	0.000	0.213	0.015	-0.005	-0.048	-0.004	0.015	0.000	0.015
Received counseling/therapy from mental hea	0.020	0.035	0.010	-0.024	-0.001	0.063	0.044	0.036	0.040	0.014	0.012	0.021	-0.029	0.043	0.004	-0.005	0.035	0.213	0.000	-0.003	0.074	-0.061	-0.016	-0.058	0.000	-0.058
Ever smoked 100 cigarettes	-0.006	0.057	0.027	-0.023	-0.007	0.025	0.018	-0.004	0.004	0.034	-0.016	-0.036	-0.007	0.001	0.009	0.030	0.014	0.015	-0.003	0.000	-0.008	-0.003	0.002	0.011	0.000	0.011
Civic engagement including volunteering	-0.065	0.027	0.045	-0.044	0.034	-0.086	-0.036	-0.037	-0.027	0.022	-0.008	0.011	-0.041	0.016	0.035	-0.060	-0.035	-0.005	-0.074	-0.008	0.000	0.041	0.004	-0.001	0.000	-0.001
Volunteer for organization or association	-0.025	0.056	0.041	-0.011	0.005	-0.043	-0.021	-0.038	-0.025	0.029	-0.002	0.027	-0.029	0.051	0.023	-0.034	-0.013	-0.048	-0.061	0.003	0.041	0.000	0.004	0.000	0.004	0.084
Attend public meeting	-0.007	0.012	0.022	-0.004	-0.014	0.007	0.045	0.034	-0.007	-0.045	0.041	-0.014	0.020	0.000	-0.052	0.013	-0.004	-0.016	0.002	0.004	-0.004	0.000	-0.027	0.000	0.000	-0.027
Vote in last local elections	-0.066	0.020	0.029	-0.037	0.031	-0.104	-0.039	-0.032	0.043	0.114	-0.024	-0.019	-0.016	0.028	-0.031	-0.056	0.015	-0.058	0.011	-0.001	0.084	-0.027	0.000	0.000	0.000	-0.027
Weight 2 (Added Health Variables)	poor	CHR	FCHR	CCHR	ACHR	C SOC	S SOC	S PAY	B PAY	P ACC	F ACC	I ACC	F PAI	FRIMM	S IMM	S MED	F MED	F MTL	M MTL	M CIG	S CIG	CIVIC	CIV	V CIG	M CIV	V CIV
Poor or fair health	0.000	-0.011	-0.091	-0.018	-0.015	-0.027	0.012	0.047	0.070	-0.140	-0.037	-0.162	0.030	-0.053	0.001	-0.056	-0.031	-0.007	-0.037	0.085	-0.057	-0.035	-0.008	-0.035	0.000	-0.035
Ever been told you had hypertension	-0.011	0.000	-0.001	-0.011	-0.060	-0.009	0.053	0.124	0.029	-0.093	-0.019	-0.039	-0.001	-0.090	-0.055	-0.001	0.018	0.016	-0.022	0.130	-0.002	0.013	-0.042	-0.004	0.000	-0.004
Ever told you had high cholesterol	-0.091	-0.001	0.000	0.006	-0.026	-0.113	-0.031	-0.021	-0.009	-0.004	0.005	-0.032	-0.027	-0.076	-0.009	-0.026	0.010	0.045	0.024	0.036	-0.011	0.018	-0.012	-0.016	0.000	-0.016
Ever had asthma	-0.018	-0.011	0.006	0.000	-0.008	-0.049	-0.029	-0.020	-0.029	-0.006	0.007	0.010	-0.012	-0.023	-0.052	-0.028	-0.004	-0.011	-0.082	0.002	-0.084	-0.017	-0.023	-0.058	0.000	-0.058
Ever been told you had cancer	-0.015	-0.060	-0.026	0.008	0.000	0.013	0.008	-0.017	0.020	-0.024	-0.023	0.021	0.012	-0.040	-0.019	-0.028	-0.009	0.026	0.034	0.010	0.012	0.002	0.018	0.009	0.000	0.009
Difficulty doing errands alone	-0.027	-0.009	-0.113	-0.049	0.013	0.000	-0.006	0.100	0.081	-0.178	-0.041	-0.092	0.054	-0.048	-0.062	-0.081	0.009	-0.007	-0.076	0.045	-0.009	-0.014	0.000	-0.018	0.000	-0.018
Work limited due to health problem	0.012	0.053	-0.031	-0.029	0.008	-0.006	0.000	0.052	0.013	-0.070	0.011	0.006	0.072	-0.014	-0.057	-0.011	-0.021	0.022	-0.006	0.112	0.005	-0.021	0.017	0.028	0.000	0.028
Problems paying medical bills, past 12m	0.047	0.124	-0.021	-0.020	-0.017	-0.100	0.052	0.000	0.059	-0.018	-0.014	-0.057	0.046	0.038	0.052	0.012	0.014	-0.018	-0.072	0.082	-0.009	-0.085	0.000	0.007	0.000	0.007
Get sick or have accident, worry about paying m	-0.070	0.029	-0.009	-0.029	0.020	0.081	0.013	0.059	0.000	-0.048	-0.017	-0.052	0.006	0.037	-0.002	0.001	0.028	0.009	0.018	0.066	-0.035	-0.063	0.038	0.034	0.000	-0.034
Time since last saw doctor	-0.144	-0.093	-0.004	-0.006	-0.024	-0.179	-0.070	-0.128	-0.048	0.000	0.038	-0.016	-0.093	-0.090	-0.059	-0.080	0.005	0.007	0.012	-0.011	-0.061	0.004	-0.042	-0.039	0.000	-0.039
Have a usual place for care	-0.037	-0.019	0.005	-0.007	-0.023	-0.041	0.011	-0.014	-0.017	0.038	0.000	-0.022	-0.073	-0.045	-0.014	0.006	0.043	-0.013	0.011	0.010	-0.045	-0.006	-0.089	-0.008	0.000	-0.008
Hospitalized overnight, past 12 months	-0.162	-0.039	-0.032	0.010	0.021	-0.092	0.006	-0.057	-0.052	-0.116	-0.022	0.000	-0.036	-0.051	-0.022	-0.009	0.031	-0.005	0.019	-0.090	-0.010	-0.006	0.003	-0.016	0.000	-0.016
How often had pain, past 3 months	0.030	-0.001	-0.027	-0.012	0.012	0.054	0.072	0.046	0.006	-0.093	-0.073	-0.036	0.000	-0.022	-0.043	-0.043	-0.025	0.003	-0.008	0.048	-0.082	-0.050	-0.016	-0.037	0.000	-0.037
Ever had pneumonia shot	-0.053	-0.090	-0.076	-0.023	-0.040	-0.048	0.014	0.038	0.037	-0.090	-0.045	-0.051	-0.022	0.000	-0.058	-0.072	-0.035	-0								

Interim Conclusion of Relevance

- Caution should be used if “data stacking” is considered for adding records to a more rigorous survey at a lower cost

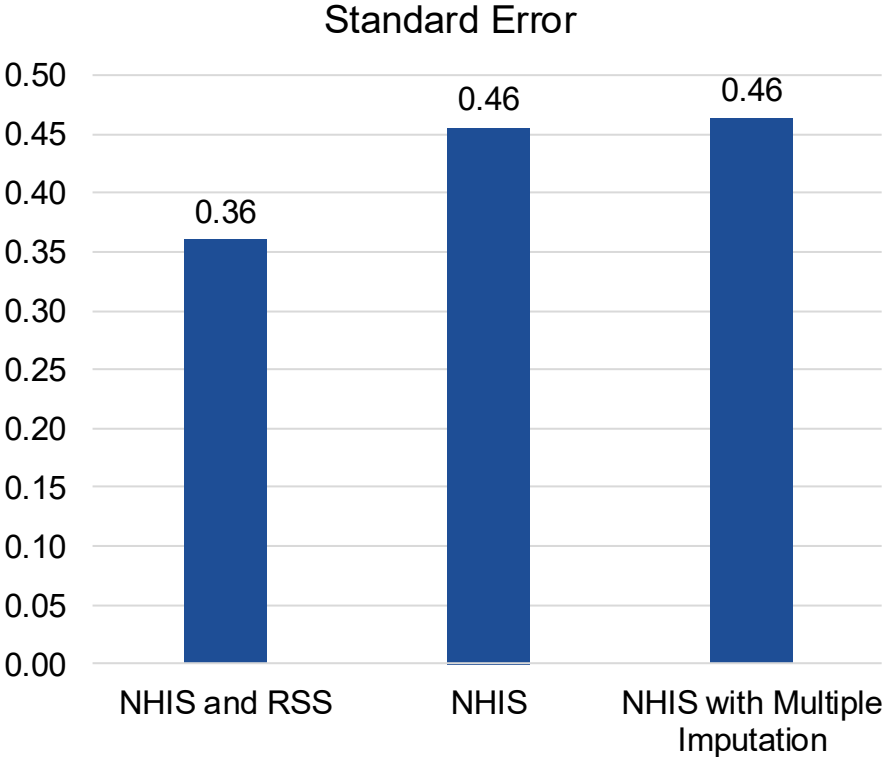
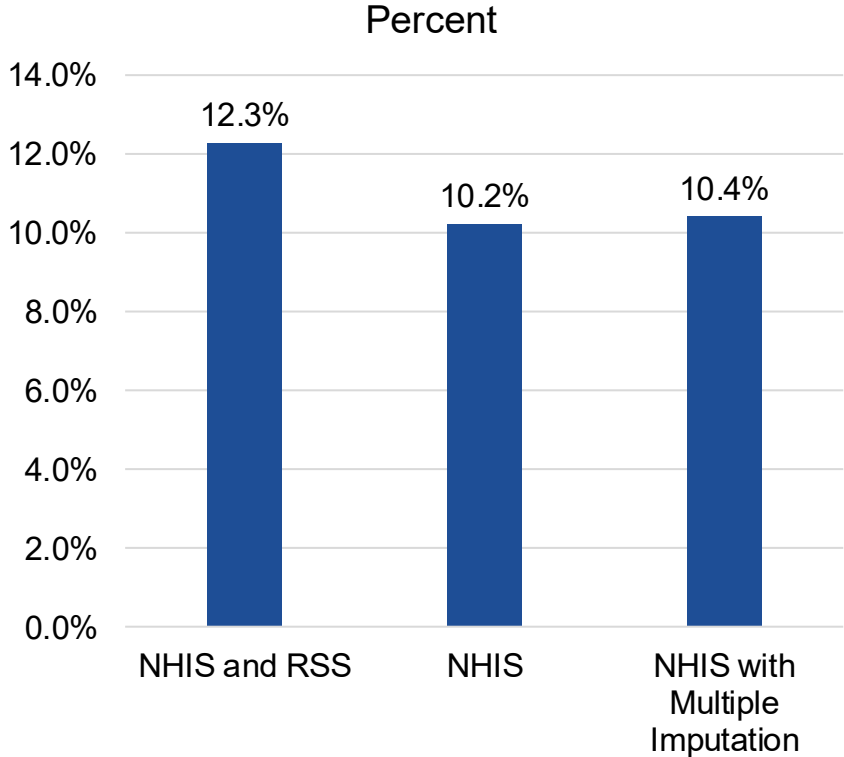
Collect Less from the Panel, Impute More

- Reduction in Web Panel (RSS) Survey Length
- Impute data not collected from the web panel
- Multiple imputation to account for uncertainty in imputed values

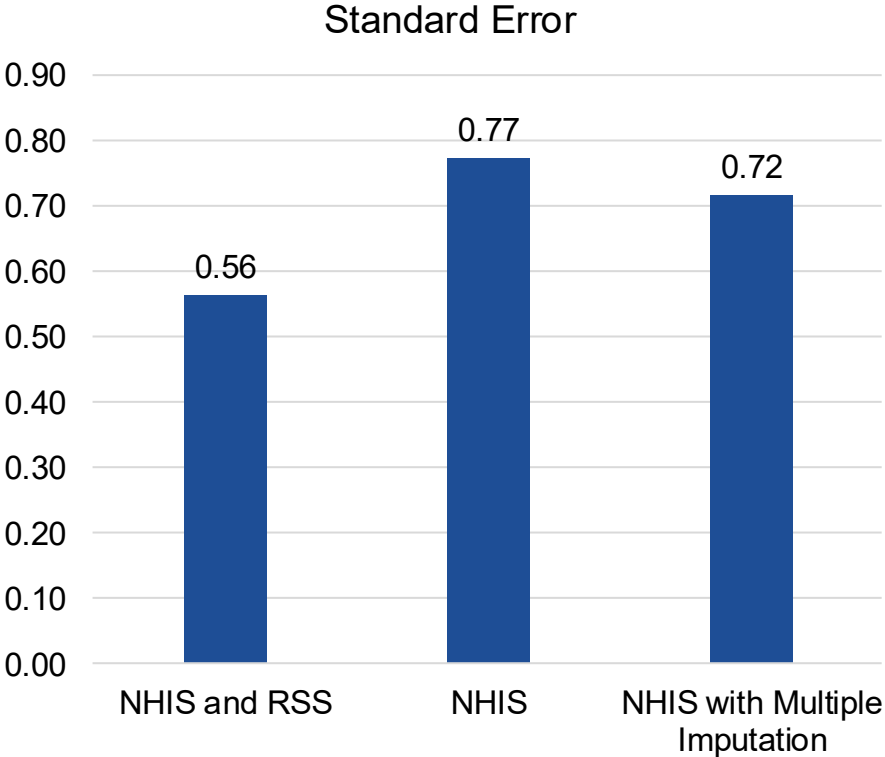
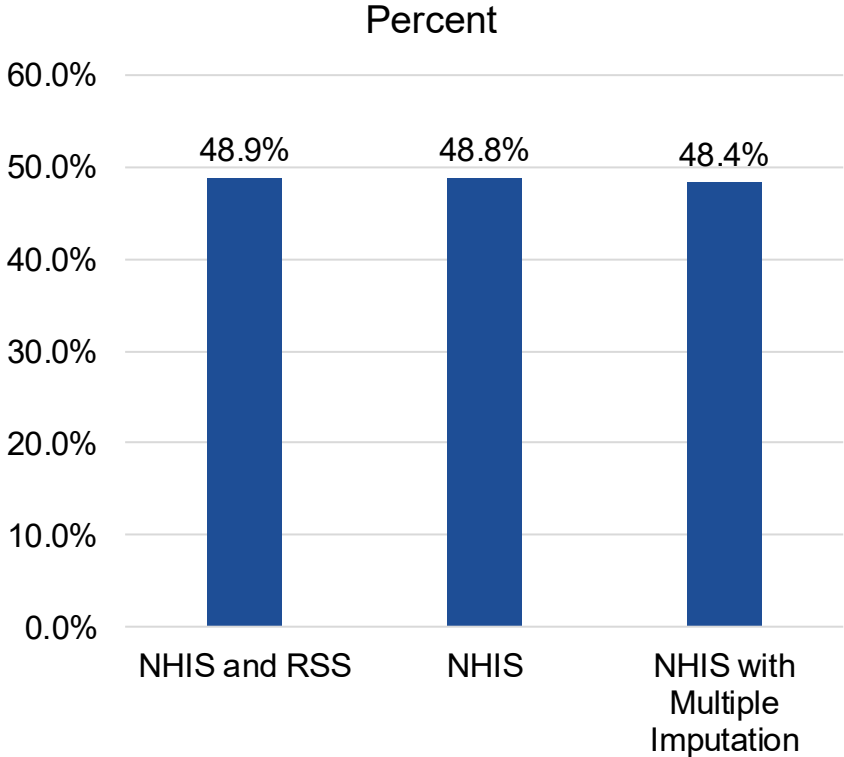
Design

- Delete RSS panel survey data for three variables, as if never asked
- Impute the deleted panel data
 - Stack NHIS and RSS data with the Wt1 from the the previous step
 - Multiple imputation in SAS, fully conditional specification (chained equations) method with discriminant function
 - M=45 datasets
- Key comparisons
 - NHIS vs. NHIS + Imputed RSS (also include NHIS + Observed RSS)
 - Bias and variance

Problems paying medical bills in past 12 months

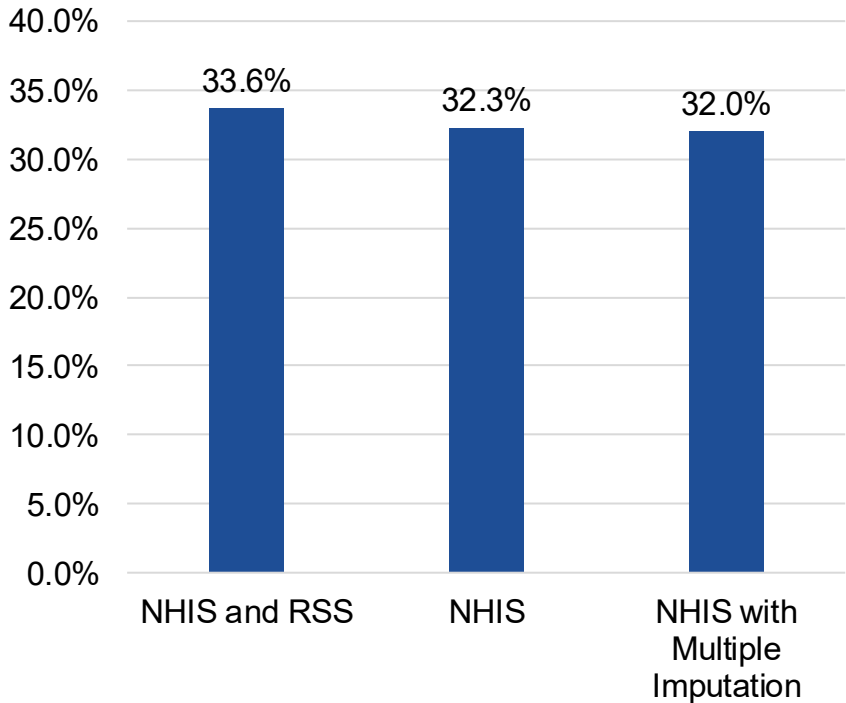


Flu vaccine in past 12 months

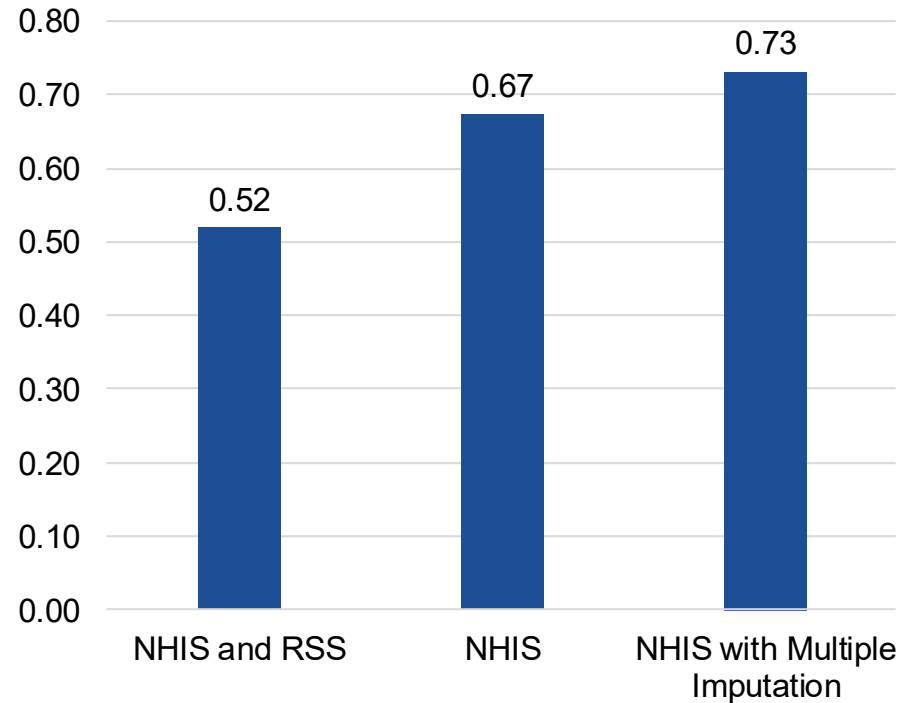


Ever smoked 100 cigarettes

Percent



Standard Error



Conclusions

- Some bivariate relationships are the same, but others can differ across the lower-cost and rigorous survey design
- Weighting to more variables does not seem to help with this
- Imputation seems to help with point estimates likely linked to differences founded in covariance structure
- Multiple imputation leads to lower variance for some estimates, yet higher for others
 - MI also provides a rectangular dataset for data users

Contact

Andy Peytchev
Fellow and Senior Survey Methodologist
RTI International

apectchev@rti.org

