Investigating the Bias of Alternative Statistical Inference Methods in Mixed-Mode Surveys

Z. Tuba Suzer-Gurtekin[1], Steven G. Heeringa[2], Richard Valliant [3]
[1,2]Michigan Program in Survey Methodology, Institute for Social Research, Ann Arbor, MI
[3]Program in Survey Methodology, Institute for Social Research, Ann Arbor, Michigan and Joint Program in Survey Methodology, University of Maryland, College Park, MD

Abstract
Mixed-mode surveys combine different data collection modes to reduce non-observational survey errors under certain cost constraints. In this survey design, usually there is no control over who responds by which mode. As a result, data are obtained by a nonrandom mixture of survey administration modes. Without adjusting for this nonrandom mixture of modes, the standard method of estimation that combines responses from different modes has a bias that depends on both mode effects and the mix of respondents that choose each mode. Unless mode effects are zero, data should be adjusted for both nonresponse and nonrandom mixture of modes. We present alternative methods that account for both nonresponse and the nonrandom mixture of modes. Although in principle the separate mode effects are not estimable in a mixed-mode survey design, the alternative estimators do allow estimation of the difference in average mode effects. In addition, the bias properties of alternative methods can be better understood when compared to the standard estimation method. The alternative methods use models to impute each respondent's values for each counterfactual response mode—e.g. a telephone response value of in-person respondents. Combining the observed values with the imputations results in a "completed" data set for each mode. Alternative estimators are then used to combine these mode-specific "completed" data sets in an attempt to reduce bias associated with confounded and nonrandom influences of mode choice and mode effects. This paper presents some results for empirical comparisons of mean personal income and percent health insurance coverage based on the alternative methods and standard method. The public-use 2012 Current Population Survey (CPS) March data are used for empirical evaluations.
Key Words: Mixed mode surveys, Mode effects, Imputation, Current Population Survey (CPS)

1. Introduction

Early in the history of survey research, mixed-mode surveys were proposed to decrease non-observational survey errors under constrained survey budgets (Hansen & Hurwitz, 1946; Hochstim, 1967). Inference in the earlier studies implicitly assumed ignorable mode effects; that is, all survey modes generate values close to true values for all the members of the population. Later, theoretical frameworks were developed to discuss the factors that may yield nonignorable mode effects for different subgroups in the population (De Leeuw, 1992, 2005; Groves et al., 2009; Schwarz, Strack, Hippler, & Bishop, 1991; Tourangeau, Rips, & Rasinski, 2000a). But empirical work could only study parts of the frameworks and was conditional on specific survey designs. With a few exceptions (Aquilino, 1994; Beland & St-Pierre, 2008; Soulakova, Hartman, Gibson, & Davis, 2009), the focus of the empirical work was on estimates of full population quantities. In particular, the theory and the empirical results emphasized the possible differences between the self- and interviewer-administered surveys, audio and visual channel dependent presentations and the variable dependent nature of mode effects.

Recently, pressing issues of increasing non-observational survey error and survey costs influenced survey researchers to adapt many variations of mixed-mode surveys (Brick & Lepkowski, 2008; Couper, 2011; De Leeuw, 2005). Inference in later mixed-mode survey designs, generally adopted the early assumption that mode effects could be ignored and did not challenge that assumption with any empirical work. In sequential or concurrent mixed-mode survey inference, in which data are collected via multiple response modes, responses from multiple modes have been combined without adjusting for any measurement error.

In practice, survey modes are not randomly assigned in mixed-mode surveys. As an implication of the assumption of ignorable mode effects, this nonrandom assignment of modes is ignored. This paper defines this nonrandom assignment as *mode choice*. Recent methods have been developed to unconfound the mode choice and the mode effects (Camillo & D'Attoma, 2011; Jäckle, Roberts, & Lynn, 2010; Lugtig, Lensvelt-Mulders, Frerichs, & Greven, 2011; Vannieuwenhuyze, Loosveldt, & Molenberghs, 2010, 2012). These methods challenge the general notion of ignorable mode choice and mode effects and motivate a more systematic approach to evaluate mode effects. Buelens & Van den Brakel (2011) also propose a mode calibrated method for estimating changes in means over time.

This paper extends the single survey mode statistical error model to a mixed-mode survey context. We investigate the bias properties of the standard method of estimation, which ignores mode choice and mode effects, and proposed alternative methods, which adjust for mode choice and mode effects based on unit covariates. The alternative methods use imputation models. In particular, the respondent data for a given mode and phase are used to create completed data sets for a given sample. Then, the completed data sets are used to compute mode-specific survey means. The survey means are then combined to produce one survey estimate. The ways in which the mean estimates can be combined (CM1) a simple average, (CM2) a minimum variance combination, and (CM3) a minimum mean square error combination. Figure 1 shows a schematic chart for the proposed inference method when there are two modes. This approach can be extended to three or more modes. In this paper complete response is assumed, but the methods can also be extended to account for unit nonresponse.

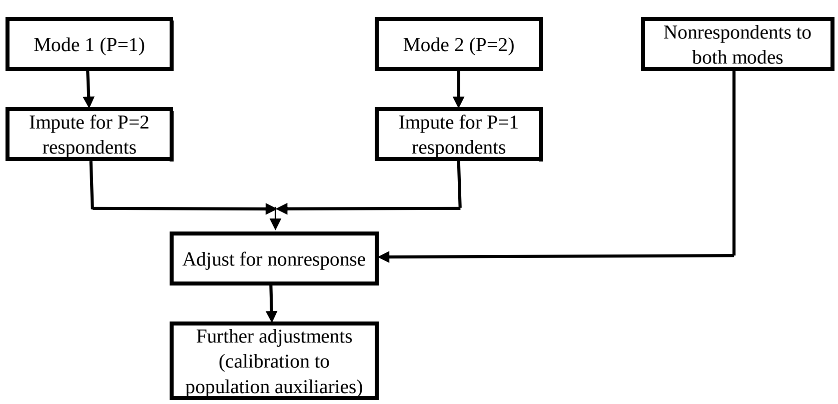

Figure 1: Schematic Chart for the Proposed Mixed-Mode Survey Inference Method

The first empirical/simulation study focused on a variable of interest, wage and salary income, for which measurement complexities are minimal (Suzer-Gurtekin, Heeringa, & Valliant, 2012). The data used in this first study include benchmark values, which may not be the usual case. In contrast, this paper analyzes 2012 Current Population Survey (CPS) March Data mean personal income and percent health insurance coverage for which no benchmark values are available. The Current Population Survey (CPS), a monthly rotating panel survey, implements a mixed-mode survey design. The CPS rotating panel design employs a 4-8-4 cycle for a selected household. Interviews are conducted for two sets of four consecutive waves that are eight months apart. A majority of the CPS interviews are conducted by telephone, except for the first and fifth wave interviews. In this paper, the CPS data structure is conceptualized as an example of a mixed-mode survey design in which two response modes are available for a given survey period.

Suzer-Gurtekin, Heeringa and Valliant (2012) evaluated the alternative methods using a subset of 1973 CPS Match Data[1], which includes both survey and Internal Revenue Services (IRS) data. The corresponding person level data allowed computation of relative differences for the standard method, the alternative combination methods and the mode–specific estimates relative to a benchmark. For wage and salary income data in this specific dataset, the mode effects seem to be ignorable and variances of the mode-specific estimates were equal. Ignorable mode effects and equal variance properties yielded a special case of the combination weights in alternative (CM3) above that minimizes the mean square error of combined estimator. As a result, performance differences were not significant between the alternative combination methods. On the other hand, they all outperformed the standard method. Between the ignorable and the nonignorable imputation model simulations, the nonignorable imputation model eliminated the differences in the relative differences between the telephone and the in-person means.

One of the open research questions in the first study was the direction of the relative differences between the estimates and the benchmark statistics. Despite the widely documented underreporting of wage and salary income (Moore, Stinson, & Welniak, Jr., 2000), the relative differences were reported to be positive in Suzer-Gurtekin et al. (2012). Further investigation revealed the variables for survey and IRS variables were switched in the imputations and the computations of the relative differences. This paper reports the corrected relative differences and absolute relative differences for the ignorable mode choice imputation models, in the Appendix Figures 2-3. Other plots are available upon request.

The next section presents the measurement error model as extended to a mixed-mode survey context. In the following section, the bias properties of the standard and the alternative methods under the measurement error model are shown. The Study Description section details the dataset, the method to create replicate samples and the computations. The paper concludes with a discussion of the empirical results and suggestions for future research.

---

[1]Current Population Survey, 1973, and Social Security Records: Exact Match Data [ICPSR 7616]. ICPSR version. Washington, DC: U.S. Dept. of Commerce, Bureau of the Census and Social Security Administration, Long-Range Research Branch [producer], 197?. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2001. doi:10.3886/ICPSR07616

1.1 Measurement Error Model

A simple measurement error model can be formulated for a mixed-mode survey context as shown in equation in (1). More generally, the model can be extended to as many response modes as included in the mixed-mode design. For simplicity of presentation, only telephone and in-person modes are considered here:

$$y_j = \mu_j + R_{Tj}B_{Tj} + R_{Ij}B_{Ij} + \varepsilon_j \text{ , where:} \qquad (1)$$

$j = 1, 2, 3, ...N$ indexes individual persons in the survey population,

μ_j is the mean true value for person j which can depend on X_j, a vector of covariates,

Subscripts T and I correspond to telephone and in-person modes,

$B_{Tj} =$ reporting error if person j responds by telephone,

$B_{Ij} =$ reporting error if person j responds in-person,

$$R_{Tj} = \begin{cases} 1 & \text{if population unit } j \text{ responds in telephone mode} \\ 0 & \text{if otherwise} \end{cases},$$

$$R_{Ij} = \begin{cases} 1 & \text{if population unit } j \text{ responds in in-person mode} \\ 0 & \text{if otherwise} \end{cases},$$

$$\varepsilon_j \overset{iid}{\sim} N(0, \sigma^2).$$

Errors associated with coverage, unit and item nonresponse are not accounted for in this analysis. The simple response model in (1) assumes independence of residuals among all population members, i.e. $\text{cov}(\varepsilon_j, \varepsilon_{j*}) = 0$.

The systematic reporting errors, B_{Tj} and B_{Ij}, enter the measurement error model through response mode indicators, R_{Tj} and R_{Ij}. Under multiple theoretical realizations of response, R_{Tj} and R_{Ij} can be considered as random variables in (1) in which case a more elaborate formulation can be considered. Suppose $R_{Tj} = (1 - R_{Ij})$, R_{Ij} is a random variable with $E_R\left(R_{Tj}\right) = g(X_j, \psi) \equiv g_j$, where $g(\bullet)$ is logistic, probit, or some other binary regression equation. Note that X_j is a vector of covariates for person j that can contain dummies for social-demographic group and ψ is a vector of regression model parameters. We label the systematic reporting errors, B_{Tj} and B_{Ij}, as mode effects and $g(\bullet)$ as mode choice. When sample units chose the mode by which they respond, B_{Tj} and B_{Ij} are not separately estimable unless one of them is known to be 0.

1.2 Bias Properties of Alternative Methods

In this section, we derive the bias of the standard method of estimation that ignores mode effects and the alternative imputation methods under the same measurement error in (1). To simplify the presentation, we do not consider sampling, although that feature could be added.

1.2.1 Bias of Standard Method

It can be shown that for the finite population mean, $\bar{Y}_0 = \frac{1}{N}\sum_{j \in U} y_j$, that ignores mode effects, the expectation of the estimation error with respect to the measurement error is the average combined effect of random mode choice, g_j and mode effects (B_{Tj}, B_{Ij}):

$$E_{MEAS}\left[\bar{Y}_0 - \bar{Y}\right] = \left(\overline{gB}\right)_U \text{ , where:} \tag{2}$$

$\bar{Y}$ is the finite population mean with no measurement error, E_{MEAS} denotes expectation with respect to model (1), and $\left(\overline{gB}\right)_U = \frac{1}{N}\sum_{j \in U}\left(g_j B_{Tj} + (1 - g_j) B_{Ij}\right)$.

The average, $\left(\overline{gB}\right)_U$, is affected by the random mix of the types of units that respond by each mode and by the sizes of the mode effects themselves. This raises to two important points: (1) The size and the direction of the mode effects should be investigated, and (2) In case of substantial mode effects, mean estimator should be adjusted for mode choice and mode effects.

1.2.2 Bias of Alternative Methods

Given the same mixed-mode survey design in which telephone and in-person modes are used for persons who responded by I, we impute values as if they had responded by T and the reverse for persons who responded by T :

y^*_{Tj} $(j \in U_I)$ is the imputed telephone value for persons who responded by I

y^*_{Ij} $(j \in U_T)$ is the imputed in-person value for persons who responded by T , where: U_T : set of persons with R_{Tj} =1, and U_I : set of persons with R_{Ij} =1.

Assume $U = U_T \cup U_I$ is the full population. To do the imputations, we will use the telephone reports to create the y^*_{Tj} imputations for the cases that respond in-person. Similarly, the in-person reports will be used to create imputations for the cases that respond by telephone. In those circumstances, it is reasonable to suppose that, on average, the imputations for a set of cases (U_T or U_I) are contaminated by the reporting errors associated with the cases used to create the imputations. In particular, suppose that the expectations with respect to the imputation mechanism are

$$E_{IMP}\left[y^*_{Tj}\right] = \mu_j + B_{Tj} \text{ and}$$
$$E_{IMP}\left[y^*_{Ij}\right] = \mu_j + B_{Ij}. \tag{3}$$

Define the population means that use imputed data as

$$\bar{Y}^*_T = \frac{1}{N}\left[\sum_{j \in U_T} y_j + \sum_{j \in U_I} y^*_{Tj}\right] \text{ and } \bar{Y}^*_I = \frac{1}{N}\left[\sum_{j \in U_I} y_j + \sum_{j \in U_T} y^*_{Ij}\right] \tag{4}$$

In the next section, we discuss ways of combining $\bar{Y}^*_T$ and $\bar{Y}^*_I$ to estimate the population mean.

E_M denotes expectation with respect to the response model (Y-model), E_R is the expectation over the mode choice model (R-model), and E_{IMP} denotes expectation with respect to the imputation model. Suppose that $\tilde{y}_j$ is the true value for unit *j* which obeys the model:

$\tilde{y}_j = \mu_j + \delta_j$ where $\delta_j \overset{iid}{\sim} \left(0, \sigma^2\right)$ are independent error terms.

The estimation error for the mean computed as if all cases responded by telephone is $\bar{Y}_T^* - \bar{Y}$, which can be written as

$$\begin{aligned}\bar{Y}_T^* - \bar{Y} &= \frac{1}{N}[\sum_{j\in U_T} y_j + \sum_{j\in U_I} y_{Tj}^*] - \frac{1}{N}[\sum_{j\in U_T} \tilde{y}_j + \sum_{j\in U_I} \tilde{y}_j] \\ &= \frac{1}{N}[\sum_{j\in U_T}(y_j - \tilde{y}_j) + \sum_{j\in U_I}(y_{Tj}^* - \tilde{y}_j)]\end{aligned} \tag{5}$$

The expectation with respect to the Y-model and the imputation model, conditional on the sets of units that responded by *T* or *I* is

$$\begin{aligned}E_M E_{IMP}[\bar{Y}_T^* - \bar{Y} \mid U_T, U_I] &= P_T \frac{1}{N_T}\left\{\sum_{j\in U_T} B_{Tj}\right\} + P_I \frac{1}{N_I}\left\{\sum_{j\in U_I} B_{Tj}\right\} \\ &= \frac{1}{N}\sum_{j\in U} B_{Tj} \equiv \bar{B}_{UT}\end{aligned} \tag{6}$$

where $P_T = N_T / N$, $P_I = N_I / N$.

Thus, the imputations for U_I are contaminated by the telephone mode effects for individuals, and conditional on the realized modes selected by respondents, the imputed estimate $\bar{Y}_T^*$ inherits the average reporting error associated with the telephone mode. Similar, to (6), the expectation of the mean as if all cases had responded in person is

$$\begin{aligned}E_M E_{IMP}[\bar{Y}_I^* - \bar{Y} \mid U_T, U_I] &= P_I \frac{1}{N_I}\left\{\sum_{j\in U_I} B_{Ij}\right\} + P_T \frac{1}{N_T}\left\{\sum_{j\in U_T} B_{Ij}\right\} \\ &= \frac{1}{N}\sum_{j\in U} B_{Ij} \equiv \bar{B}_{UI}\end{aligned} \tag{7}$$

$\bar{B}_{UI}$ in (7) is the average mode effect in the population if all cases responded by I. Unlike (2), the results in (6) and (7) are independent of the mode choice. In principle, B_{Tj} and B_{Ij} are not estimable unless a study is conducted using an experimental design as described in Tourangeau, Rips, & Rasinski (2000b).

The sizes of the mode effects is usually unknown, i.e. it may not be known whether $\bar{B}_{UT} < \overline{(gB)}_U$ or $\bar{B}_{UI} < \overline{(gB)}_U$. We will have $\bar{B}_{UI} \le \overline{(gB)}_U \le \bar{B}_{UT}$ or the reverse depending on which of $\bar{B}_{UI}$ or $\bar{B}_{UT}$ is smaller since $g_j B_{Tj} + (1-g_j)B_{Ij}$ is a convex combination.

As an alternative estimator, we propose

$$\bar{Y}^* = \alpha \bar{Y}_T^* + (1-\alpha)\bar{Y}_I^*,\ 0 \le \alpha \le 1 \tag{8}$$

The bias of this estimator is

$$E_R E_M E_{IMP}\left[\bar{Y}^* - \bar{Y}\right] = \alpha \bar{B}_{UT} + (1-\alpha)\bar{B}_{UI} \tag{9}$$

implying that the bias of $\bar{Y}^*$ depends on α and sizes of average mode effects. It is not straightforward to compare (2) and (9), but the empirical results show that smaller bias is achievable by the alternative methods.

The smallest bias would be obtained by choosing the mode with the smaller bias (if this were known) and using only $\bar{Y}_T^*$ or $\bar{Y}_I^*$ but this would waste the data collected via the other mode. Note that $E_R E_M E_{IMP}\left[\bar{Y}_T^* - \bar{Y}_I^*\right] = \bar{B}_{UT} - \bar{B}_{UI}$. So $\bar{Y}_T^* - \bar{Y}_I^*$ can be used to estimate the difference in the average mode effects in the population. This difference incorporates the possibility that the effect of mode can differ among persons.

1.2.3 Combination Methods

The optimal value of α in (8) can be derived as

$$\alpha_{opt} = \frac{v_I - \rho_{TI}\sqrt{v_T v_I}}{v_I + v_T + 2\rho_{TI}\sqrt{v_T v_I}}, \tag{10}$$

where $\rho_{TI} = corr\left(\bar{Y}_T^*, \bar{Y}_I^*\right)$, $v_T = \mathrm{var}_M(Y_T^*)$, and $v_I = \mathrm{var}_M(Y_I^*)$.

The combination methods CM1 and CM2 are special cases of α_{opt}. Simple average combination method which uses $\alpha = 1/2$ is a special case when $\bar{B}_{UT} = \bar{B}_{UI} = 0$, $v_T = v_I$ and $\rho_{TI} = 0$. When $\bar{B}_{UT} \neq \bar{B}_{UI} \neq 0$, the combination method that combines mode-specific means by inverse of variances is the special case. In this study only these two methods have been evaluated. Future research will test the α_{opt} empirically.

2. Study Description

2.1 2012 CPS March Data

2012 CPS March respondent data are used to perform an empirical comparison of the proposed estimation methods in a condition where no benchmark values are available. Estimates for two population statistics are studied: (1) mean personal income and (2) proportion of persons covered by health insurance coverage. CPS March Supplement measures of income and health insurance coverage are merged with the CPS March data to determine the response mode. The nonrespondents to the CPS March 2012 are excluded from the analysis. Values that CPS imputed for item nonresponses are also excluded from the analyses. The unit of analysis is householders. Tables 1 and 2 present the household and householder covariate percentages. The imputation model covariates are selected from among these covariates.

To fit the imputation models, two sets of analysis were conducted for personal income and health insurance coverage data: (1) Mode choice logistic regression models (dependent variable: In-person vs. telephone), and (2) Regression analysis of personal income and health insurance coverage which tested for the mode interactions and significance of covariate parameter estimates. The exploratory analysis of predicted probabilities for health insurance coverage showed that it is more informative to stratify the sample in four groups by age and work status of householders: 65+ vs. <65, and worker vs. nonworker (Table 3). The mode choice logistic regression models were fit separately for these four groups. For both sets of analyses, the model structures were finalized based on overall likelihood ratio tests. Table 4 shows the likelihood ratio tests for the mode choice logistic regressions. Table 5 shows the covariates selected for the imputation models.

2.2 Study Description

Following the model selection exercise described in the previous section, the following 5 steps were applied:

1- Bootstrap replications: Since the base weights were not available separately, the sampling weights were recomputed at the state and month in sample (MIS) level to reflect the unequal probabilities of selection for the 2012 CPS observations. In the bootstrapping computations, units were defined as the Primary Sampling Unit (PSU) and state x MIS were considered to be the strata. Although the replicate weights were computed and applied using the bootstrap function in R survey package, this method should incorporate a more comprehensive approach in the future work as was done in (Kennickell, 1991). The current method does not re-estimate the parameters of the mode choice and response regression models in each replicate, as should be done in a comprehensive approach

2- Parameter estimation: Models were estimated to compute the parameters of beta coefficients for the response models. The ignorable mode choice models include only the response regression models (Y-model). For the personal income, the log transformation was used in the imputation models. In the prediction computations the bias correction for the log transformation was applied (Newman, 1993).

3- Imputation: Using the parameter estimates from Step 2, telephone and in-person completed data vectors were created for a given bootstrap replicate sample. These completed data vectors include both observed and imputed data values conditioned on the response mode--telephone or in-person. Five completed data vectors were computed.

4- Estimation: Using the survey weights and the completed data vectors, mode-specific means for personal income and health insurance coverage were computed. These mode-specific means were compared against the means generated by the standard method using a repeated measurement ANOVA model to detect significant differences for possible mode effects.

5- Combination of mode-specific means: Mode-specific means were combined using two methods: (1) simple average estimator (CM1), and (2) inverse variance weighted estimator (CM2). These are comparable to the ones used in the previous empirical/simulation study but the MSE weighted estimator cannot be used as there are no benchmarks available. These combined estimates were compared using a repeated measurement ANOVA model to detect significant differences for possible mode effects on the estimates.

3. Results and Discussion

The unadjusted means for personal income are 33,162 and 41,704, respectively for the in-person and telephone modes. When values are imputed as if all persons responded by a single mode as described in section 1, there are still differences in means. Figure 4 shows the averages of mode-specific means over bootstrap replicates. The F test for equality of means is significant at 95% confidence level for both the mean personal income and health insurance coverage percent. The direction of the difference is same for health insurance coverage where the unadjusted means for in-person and telephone are 0.83 and 0.89, respectively.

The F test for equality of means is also significant for the means between the combination methods, although the difference in means is numerically small. While the personal income estimates based on the combination methods are higher than the standard method mean, the health insurance coverage percent is lower. Thus, both personal income and health insurance coverage as measured in the CPS 2012 March are sensitive to the methods applied. Although these results cannot address the sources for differences, they may be considered as motivation for further investigation of mode effects.

4. Future Research

The empirical comparison study used a subset of public CPS March 2012 data. This data set allowed us to implement the imputation method for a continuous variable, personal income, and a binary variable, health insurance coverage. Empirical analyses were conducted to detect possible differences as a result of mode effects. The differences under ignorable and nonignorable mode choice imputation models (not shown) emphasized the importance of modeling assumptions. Although bootstrapping is a way of reflecting imputation variance, the approach used here needs to be extended to fully account for the steps used. The application of the bootstrapping method reported here does not re-estimate the model coefficients for each bootstrap replicate.

Also, the imputation models did not incorporate the likelihood that an in-person report would be correlated with a telephone report for most persons. Future research should explore multivariate distribution modeling techniques to incorporate possible correlations between the responses in different modes in addition to studying mode effects in explicit experimental designs as described in Tourangeau et al. (2000b). Furthermore, these methods can be extended to panel surveys that switch from one mode to another. Panel surveys should also be a case in which correlations between mode responses can be estimated.

The results are shown to be sensitive to the modeling assumptions. Although a general measurement model is used in this dissertation, social and cognitive theories may be helpful when formulating models and assumptions. In addition, item and unit nonresponse adjustments should be incorporated in these methods. Finally, the usefulness of the optimal combination parameter, α_{opt}, will be tested empirically.

Acknowledgements

Partial funding for this research was also provided by the DDIG - Doctoral Dissertation Improvement Grants (NSF) (Award Number SES 1238612).

References

Aquilino, W. S. (1994). Interview Mode Effects in Surveys of Drug and Alcohol use: A Field Experiment. Public Opinion Quarterly, 58(2), 210–240.

Beland, Y., & St-Pierre, M. (2008). Mode effects in the Canadian Community Health Survey: A Comparison of CATI and CAPI. In J. M. Lepkowski, C. Tucker, J. M. Brick, E. D. De Leeuw, L. Japec, P. J. Lavrakas, M.W. Link, & R. L. Sangster (Eds.), Advances in Telephone Survey Methodology (pp. 297–314). Wiley.

Brick, J., & Lepkowski, J. M. (2008). Multiple Mode and Frame Telephone Surveys. In J. M. Lepkowski, C. Tucker, J. M. Brick, E. D. de Leeuw, L. Japec, P. Lavrakas, M.W. Link, R. L. Sangster (Eds.), Advances in Telephone Survey Methodology (pp. 149–170). Wiley.
Buelens, B., & Van den Brakel, J. (2011). Inference in Surveys with Sequential Mixed-Mode Data Collection. Discussion Paper. Statistics Netherlands. Retrieved from http://www.cbs.nl/NR/rdonlyres/C70ED9D5-6199-4E27-B37D-792A161D614D/0/2011x1021.pdf
Camillo, F., & D'Attoma, I. (2011). Tackling the Problem of Self Selection in the Integration of Different Data Collection Techniques. Retrieved from http://amsacta.unibo.it/3063/1/1.pdf
Couper, M. P. (2011). The Future of Modes of Data Collection. Public Opinion Quarterly, 75(5), 889–908.
De Leeuw, E. D. (1992). Data Quality in Mail, Telephone and Face to Face Surveys. Amsterdam: T. T. Publikaties. Retrieved from http://edithl.home.xs4all.nl/pubs/disseddl.pdf
De Leeuw, E. D. (2005). To Mix or not to Mix Data Collection Modes in Surveys. Journal of Official Statistics, 21(2), 233–255.
Groves, R. M., Fowler, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2009). Survey Methodology (2nd ed.). Hoboken, New Jersey: Wiley Series in Survey Methodology.
Hansen, M., & Hurwitz, W. (1946). The Problem of Nonresponse in Sample Surveys. The American Statistician, 41(236), 517–529.
Hochstim, J. R. (1967). A Critical Comparison of Three Strategies of Collecting Data from Households. Journal of the American Statistical Association, 62(319), 976–989.
Jäckle, A., Roberts, C., & Lynn, P. (2010). Assessing the Effect of Data Collection Mode on Measurement. International Statistical Review, 78(1), 3–20.
Kennickell, A. B. (1991). Imputation of the 1989 Survey of Consumer Finances: Stochastic Relaxation and Multiple Imputation. In Joint Statistical Meetings - Section on Survey Research Methods (pp. 1–10).
Lugtig, P., Lensvelt-Mulders, G. J. L. M., Frerichs, R., & Greven, A. (2011). Estimating Nonresponse Bias and Mode Effects in a Mixed-mode Survey. International Journal of Market Research, 53(5), 669–686.
Moore, J. C., Stinson, L. L., & Welniak, Jr., E. J. (2000). Income Measurement Error in Surveys: A Review. Journal of Official Statistics, 16(4), 331–361.
Newman, M. C. (1993). Regression Analyses of Log-transformed Data: Statistical Bias and Its Correction. Environmental Toxicology and Chemistry, 12, 1129–1133.
Schwarz, N., Strack, F., Hippler, H. J., & Bishop, G. (1991). The Impact of Administration Mode on Response Effects in Survey Measurement. Applied Cognitive Psychology, 5(3), 193–212.
Soulakova, J., Hartman, A., Gibson, J., & Davis, W. (2009). The Impact of Survey and Response Modes on Current Smoking Prevalence Estimates Using TUS-CPS: 1992-2003. Survey Research Methods, 3(3), 123–137.
Suzer-Gurtekin, Z. T., Heeringa, S. G., & Valliant, R. (2012). Investigating the Bias of Alternative Statistical Inference Methods in Sequential Mixed-Mode Surveys. In JSM Proceedings, Survey Research Methods Section (pp. 4711–4725).
Tourangeau, R., Rips, L. J., & Rasinski, K. (2000a). Mode of Data Collection. In The Psychology of Survey Response (pp. 289–312). Cambridge, England: Cambridge University Press.

Tourangeau, R., Rips, L. J., & Rasinski, K. (2000b). Designs for Assessing the Accuracy of Survey Reports. In The Psychology of Survey Response (pp. 267–269). Cambridge University Press.
Vannieuwenhuyze, J., Loosveldt, G., & Molenberghs, G. (2010). A Method for Evaluating Mode Effects in Mixed-mode Surveys. Public Opinion Quarterly, 74(5), 1027–1045.
Vannieuwenhuyze, J., Loosveldt, G., & Molenberghs, G. (2012). A Method to Evaluate Mode Effects on the Mean and Variance of a Continuous Variable in Mixed-Mode Surveys. International Statistical Review, 80(2), 306–322.

Appendix

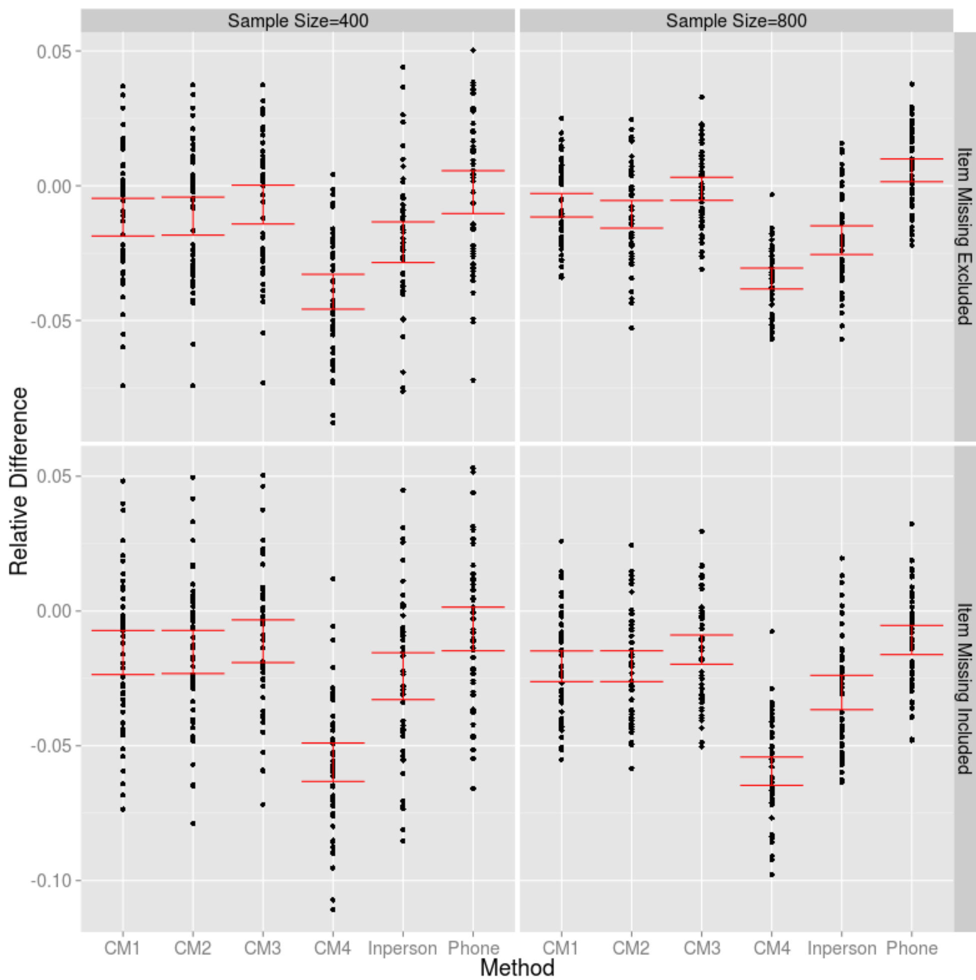

Figure 2: CPS/IRS Exact Match Study. Relative differences , $\text{RelDiff}_{CM} = (\bar{y}_{CM_i} - \bar{y}_{IRS})/\bar{y}_{IRS}$, in 50 samples[2] of estimates of mean wage and salary income. CM1, CM2, CM3, and CM4=Standard Combination Method denote four alternative combination methods; Inperson = mean as if all persons had responded in-person; Phone = mean as if all persons had responded by telephone. Sample sizes are 400 and 800 each for the samples; five imputations were performed for each sample, the red error bars represent the 95% confidence intervals for the mean relative difference.

[2] The model parameters are not estimated in one replicate in sample size=400 simulations due to zero sample size cells

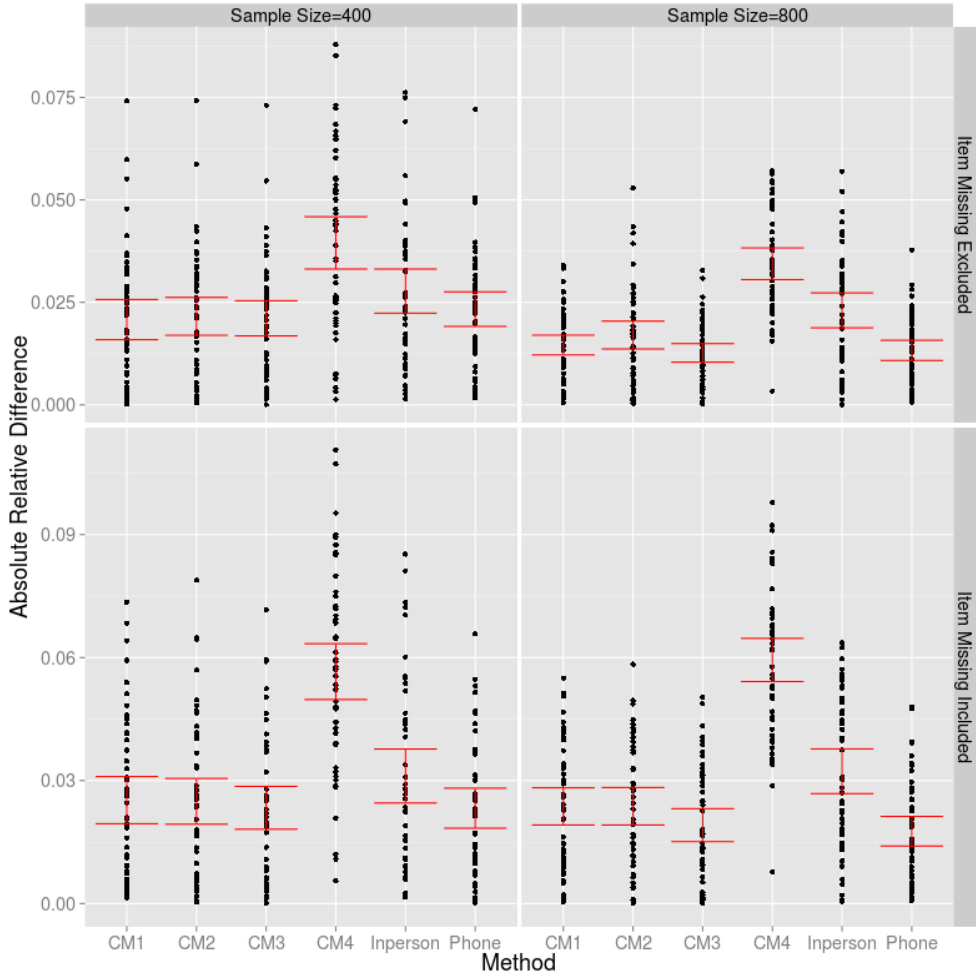

Figure 3: CPS/IRS Exact Match Study. Absolute relative differences, $\text{AbsRelDiff}_{CM} = \left|\bar{y}_{CM_i} - \bar{y}_{IRS}\right| / \bar{y}_{IRS}$, in 50 samples[3] of estimates of mean wage and salary income. CM1, CM2, CM3, and CM4=Standard Combination Method denote four alternative combination methods; Inperson = mean as if all persons had responded in-person; Phone = mean as if all persons had responded by telephone. Sample sizes are 400 and 800 each for the samples; five imputations were performed for each sample, the red error bars represent the 95% confidence intervals for the mean relative difference.

[3] The model parameters are not estimated in one replicate in sample size=400 simulations due to zero sample size cells

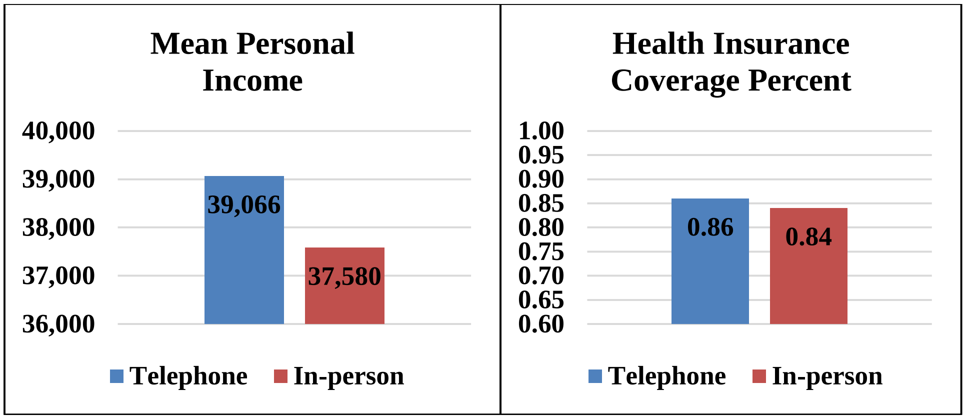

Figure 4: Means of Mode-Specific Mean Personal Income and Health Insurance Coverage Percent over Bootstrap Replicates (n_b=200)

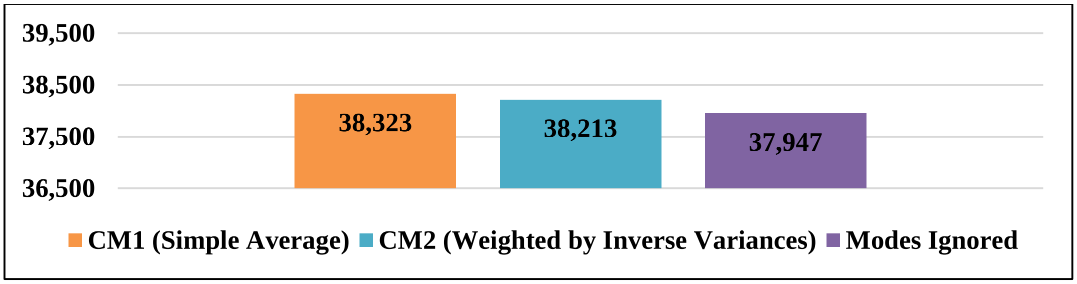

Figure 5: Means of Combined Mean Personal Income over Bootstrap Replicates (n_b=200)

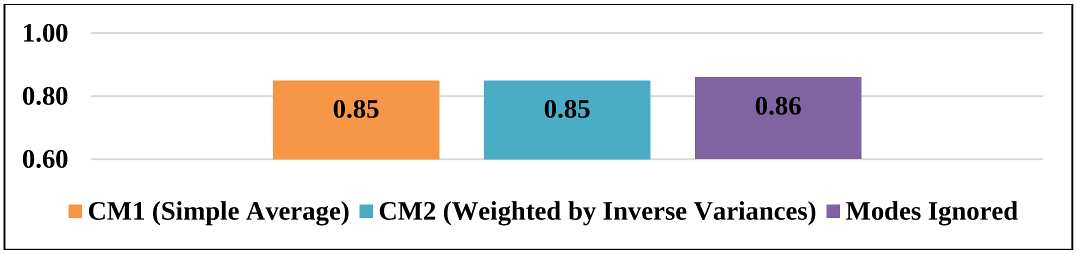

Figure 6: Means of Combined Health Insurance Coverage Percent over Bootstrap Replicates (n_b=200)

Table 1: Household Covariate Percentages in the 2012 CPS March Data (n=42,470)

Covariates	Category	2012 Weighted %
Presence of children	No kids under 14	71.74
	Kids under 14	28.26
Principal city/Balance status	Principal city	28.34
	Balance of CBSA	41.69
	Non CBSA	15.42
	Not identified	14.55
Occupied Unit Tenure	Owned or being bought	64.41
	Rented for cash	34.15
	Occupied without payment or cash rent	1.45
Living Quarters	Other	95.45
	Trailer-Permanent	4.55
Census Region and Division of Residence	Northeast	17.65
	North Central (Midwest)	22.55
	South	38.24
	West	21.56
Welfare Receipt Status	1	1.11

Table 2: Householder Covariate Percentages in the 2012 CPS March Data (n=42,470)

Covariate	Category	2012 Weighted %
Marital Status	Married	49.51
	Single	50.49
Sex	Male	50.67
	Female	49.33
Age	15-24	5.26
	25-29	7.77
	30-34	8.99
	35-39	8.70
	40-44	9.40
	45-49	9.80
	50-54	10.10
	55-59	9.80
	60-64	8.70
	65-69	6.70
	70-74	4.90
	75+	10.00
Education Attainment	None	9.50
	Elementary School	12.00
	High School	28.50
	College	50.00
Race-Ethnicity-White	1	68.20
CPS Income	Nonworker	34.00
	Wage and salary	61.30
	Self-employment	4.70
Work Class	Management	7.75
	Business and financial operations	3.34
	Computer and mathematical sciences	1.93
	Architecture and engineering	1.50
	Life, physical, and social sciences	0.66
	Community and social service	1.15
	Legal	0.85
	Education, training, and library	3.93
	Arts, design, entertainment, sports	1.29
	Healthcare practitioner and technician	3.64
	Healthcare support	1.70
	Protective service	1.62
	Food preparation and serving related	2.95
	Building and grounds cleaning and maintenance	2.50
	Personal care and service	2.16
	Sales and related	6.72
	Office and administrative support	7.89
	Farming, fishing, and forestry	0.44
	Construction and extraction	3.44
	Installation, maintenance, and repair	2.17
	Production	3.99
	Transportation and material moving	4.04
	Armed Forces	0.34
	Nonworker	34.00
Industry	Agriculture, forestry, fishing and hunting	1.00
	Mining	0.50
	Construction	4.30
	Manufacturing	7.10
	Wholesale and retail trade	8.30
	Transportation and utilities	3.40
	Information	1.50
	Financial activities	4.70
	Professional and business	8.10
	Educational and health services	14.90
	Leisure and hospitality	5.00
	Other services	3.20
	Public administration	3.70
	Armed Forces	0.30
	Nonworker	34.00
Employment Status of Spouse	Not working	15.60
	Full-time	26.50
	Part-time	5.70
	Single	52.10

Table 3: Response Mode % by Age x Work Status (n=42,323[4])

Age x Work status	n	In-person %	Telephone %
65+, Worker	2,040	35.74	64.26
65+, Nonworker	7,531	42.7	57.3
<65, Worker	25,991	40.44	59.56
<65, Nonworker	6,761	48.41	51.59

Table 4: Mode Choice Logistic Model Type III Tests, CPS March 2012 Respondents

		65+, Worker (n=2,040)		65+, Nonworker (n=7,531)		<65, Worker (n=25,991)		<65, Nonworker (n=6,761)	
Covariates	Df	LR Chisq	Pr(> Chisq)	LR Chisq	Pr(> Chisq)	LR Chisq	Pr(> Chisq)	LR Chisq	Pr(> Chisq)
Month in sample (MIS)	7	473.67	0	1750.46	0	4935.96	0	1069.24	0
State	50	74.55	0.01	205.3	0	285.53	0	126.16	0
Living quarters	1	-	-	11.96	0	-	-		
Tenure	2	6.26	0.04	30.12	0	77.24	0	33.93	0
Telephone in household	1	-	-	-	-	43.52	0	29.74	0
Telephone available (Universe=No telephone in household)	2	-	-	29.24	0	-	-	-	-
Telephone interview acceptable (Universe=Telephone available)	1	112.58	0	455.84	0	904.01	0	391.58	0
Principal city/Balance status	3	-	-	-	-	10.72	0.01	-	-
Metropolitan area (CBSA) size	6	-	-	-	-	18.42	0.01	-	-
Sex	1	-	-	12.12	0	-	-	7.44	0.01
Age (Categorical)	8	-	-	-	-	34.16	0	15.68	0.05
Level of school completed/degree received	3	-	-	18.32	0	53.24	0	19.97	0
Race-ethnicity	3	-	-	27.79	0	98.06	0	12.85	0.01
Occupation of longest job	4	-	-	-	-	22.75	0	-	-
Employment status	3	-	-	-	-	9.58	0.02	-	-
Spouse's employment status and presence of children	9	-	-	-	-	43.98	0	34.36	0
Householder March respondent	1	4.9	0.03	2.7	0.1	126.15	0	0.83	0.36
MIS x State	350	-	-	413.55	0.01	405.33	0.02	-	-
MIS x Telephone in household	7	-	-	-	-	17.42	0.01	-	-
MIS x Telephone interview acceptable	7	22.53	0	15.71	0.03	30.92	0	27.13	0
MIS x Metropolitan (CBSA) size	42	-	-	-	-	69.46	0	-	-
MIS x Householder March Respondent	7	-	-	-	-	-	-	20.03	0.01

Table 5: Covariates included in the Response Imputation Models, CPS March 2012 Respondents

	Personal Income	Health Insurance Coverage			
Covariate		65+, Worker (n=2,040)	65+, Nonworker (n=7,531)	<65, Worker (n=25,991)	<65, Nonworker (n=6,761)
Month in sample (MIS)	x	x	x	x	x
State	x				
Living quarters	x			x	
Tenure	x	x	x	x	x
Telephone available (Universe=No telephone in household)					x
Telephone interview acceptable (Universe=Telephone available)		x			x
Principal city/Balance status	x		x	x	
Metropolitan area (CBSA) size	x			x	
Age (Categorical)	x				x
Sex	x			x	x
Level of school completed/degree received	x			x	x
Race-ethnicity	x		x	x	x
Spanish speaking households	x		x	x	
Occupation of longest job	x			x	
Industry of longest job	x	x		x	
Part-time/Full-time Status	x	x		x	
Sources of earnings	x			x	
Spouse's employment status and presence of children	x	x	x	x	x
Householder March Respondent	x	x	x	x	x
Employment status	x				

[4] Excludes Armed Forces (n=147)