Using Respondent Information and Psychometric Methods to Estimate Response Propensities

William D. Kalsbeek[1], Robert P. Agans[1], Abigail T. Panter[2]
[1]Dept of Biostatistics, University of North Carolina-Chapel Hill, Chapel Hill, NC 27599-2400
[2]Dept of Psychology, University of North Carolina-Chapel Hill, Chapel Hill, NC 27599-3270

Abstract
The nonresponse bias of simple linear estimates can be measured directly from respondent data if the response propensity for each survey respondent is known. In reality these propensities are unknown and must be estimated. Traditional estimation methods for propensities are used to produce weighting class and response propensity modeling adjustments for nonresponse. This research explores the use of psychometric methods to estimate propensities from ancillary information plus respondents' answers to questions about their availability and willingness to participate in surveys more generally. Data from a recent telephone survey conducted by the UNC Survey Research Unit are used to compare estimates of nonresponse bias computed from propensities obtained by this approach versus bias from propensities obtained by the traditional methods. Findings suggest that the use of classical psychometric methods may be more effective than traditional approaches.

Key Words: survey nonresponse; estimating nonresponse bias; nonresponse adjustment

1. Introduction

The precipitous decline in survey response rates in the past 30 years has forced survey practitioners to review and expand their understanding of the origins of survey non-participation and to thereby develop new ways to measure and deal with its effects. For instance, we have moved beyond the earliest notions that populations simply consist of two disjoint subsets of (certain) responders and (certain) nonresponders. It is now common to consider the outcome of sample member recruitment to be the result of a quasi-random process whose outcome depends on the characteristics of the recruiter, the recruited, and the recruitment strategy (Groves and Couper, 1998). This evolution in thinking has moved us from the Hansen and Hurwitz (1946) two-stratum view of nonresponse in sampled populations to the more contemporary stochastic view of the recruitment process of sampled individuals, as first carefully examined by researchers at Statistics Canada (e.g., Platek, 1977; Platek, et al., 1978).

Typical of this change is how one formulates the biasing effect of nonresponse for simple estimates of means, totals, and proportions from survey samples, as noted by Lessler and Kalsbeek (1992). The two-stratum view suggests that the magnitude and direction of nonresponse bias for simple linear estimates depends completely on the proportion of the population in the nonrespondent stratum and differences between the respondents and nonrespondents in what is being estimated. These two quantifiable factors affecting bias have thus contributed to the longstanding interest in minimizing nonresponse rates and to comparing survey respondents and nonrespondents on demographics and other measures that correlate with study outcome measures. By comparison, each (i.e., the *i-th*) member of the sampled population of size *N* under the stochastic perspective of survey participation has a response propensity p_i, (i.e., the probability that if selected under a specific survey design, the member would become a respondent) and the biasing effect of nonresponse depends on the population member mean $(\bar{p})$ of the p_i as well as the covariance (σ_{py}) between p_i and the corresponding member-level survey measurement (y_i) associated with the population parameter to be estimated (e.g., $\bar{y}$, the population mean of the *y*-variable). Specifically, assuming that the realized number of respondents in the population (N_r) is a fixed constant, the bias of a pre-adjustment estimate ($\hat{\bar{y}}_r$) of the population mean $(\bar{y})$, as obtained solely from respondent data using weights for which no nonresponse adjustment has been added, is simply,

$$Bias(\hat{\bar{y}}_r) = \sigma_{py} / \bar{p} \quad . \tag{1}$$

The stochastic viewpoint in general, and this expression in particular, have several somewhat distinctive implications on how to measure and deal with the effects of nonresponse in survey practice. For example, Eq. (1) implies that maximizing response rates (reflected by $\bar{p}$) is not the only way to limit nonresponse bias (Groves and Peytcheva, 2008). One must control the statistical association between what a respondent tells us and his/her likelihood of becoming a respondent. This can be done directly by limiting the between-member variation in p_i in the population since the covariance of any measurement with a constant is zero. Those members with lower response propensities can be increased through targeted additional recruitment efforts (e.g., through conversion of initial refusals, monetary incentives, etc.), and those with higher propensities can be selectively retained (e.g., by randomly rejecting a portion of those agreeing to participate who are typically the most complian**t,** such as middle-aged females) to lower the probability of the final retention probability in the sample. Indeed, regardless of the eventual $\bar{p}$, recruitment strategies that yields a more uniform distribution of response propensities may be more effective in reducing bias expressed this way than global efforts to increase varying propensities.

Whether the aim is to directly estimate nonresponse bias or to minimize bias by controlling the distribution of propensities, Eq. (1) implies the need for a plausible measure of p_i for each survey respondent. Estimates ($\hat{p}_i$) of the p_i are also needed to adjust (by a factor of $\hat{p}_i^{-1}$) the base weights that are used to produce weighted estimates in survey data analysis (Kalsbeek and Agans, 2007). [1]

Either of two approaches, the weighting class adjustment and response propensity modeling, has traditionally been used to adjust sample weights for nonresponse (Kalton and Flores-Cervantes, 2003). The weighting class approach uses an estimate of the response rate for members of the subgroup (i.e., "class") of population members with similar characteristics and response tendencies as the respondent. The choice of characteristics to use in defining these groupings (i.e., the "weighting classes") is strategically important, since bias reduction associated with the weighting class adjustment directly depends on the correlation between the class-specific response rate and the parameter of interest for these classes (Kalton, 1983). The estimated response propensity for any respondent is often computed as a weighted variation of the RR4 response rate (AAPOR, 2008). To avoid seriously imprecise propensity estimates, minimum size requirement are typically set for the number of sample members whose weights are summed to produce the denominator of each subgroup's response rate.

Estimates of p_i following the response propensity modeling approach for nonresponse adjustment are obtained from a logistic regression model that has been fitted to a binary (0/1) response outcome using ancillary data that are available for all eligible members of the selected sample. Sources of ancillary data may include the sampling frame and/or measurements taken during the recruitment process (see Iannacchione, et al., 1991, and Lepkowski, et al., 1989, for examples). The best-fit model is used to predict the response propensity for all respondents. In principle this approach has the benefit of utilizing the prediction of all important main effects and interactions in estimating propensity, although empirical comparisons of response propensity modeling and weighting class adjustments has show little difference in bias reduction for the two approaches (Kalsbeek, et al, 2002). While both approaches have the advantage of reducing nonresponse bias, any resulting reduction in the mean squared error of estimates is at least partially offset by the increased variability of the adjusted weights which, in turn, increases the variance of weighted estimates.

Traditional methods of estimating response propensities apply common methods of basic quantitative analysis to general information about those in the sample who were recruited for participation in the study. To date, directly obtained information regarding the member's views on or history of survey participation has only begun to be explored as a source of ancillary information (Peytcheva and Olsen, 2007). Nor has the vast array of psychometric methodology for measuring latent constructs been examined as a tool for turning the response experience of and ancillary data for respondents into estimates of response propensities. Both of these possibilities are investigated in this research. We begin by developing relevant psychometric methodology. We then describe our evaluation criterion for comparing estimates from these and traditional methods, and end by suggesting some future directions to our work.

---

[1] A sample weight that has been adjusted for nonresponse is therefore the product of the base weight, computed as the inverse of the sampling probability, and an adjustment for nonresponse.

2. Methods

2.1 Psychometric Approaches to Estimate Response Propensities

Psychometric tools are used in the social sciences to measure latent variables (i.e., hidden variables, model parameters, and hypothetical constructs) that are not directly observed but are rather inferred through a mathematical model from other variables that are observed and directly measured (Nunnally & Bernstein, 1994). Measurement in this form typically involves scaling or the assignment of numerals to objects or events according to some rule (Stevens, 1946). We treat response propensity as a latent variable among individuals that can be measured at any point-in-time by psychometric techniques. Much like attitudes guide behavior, response propensities, as defined here, are a set of internal attributes that can be used to discriminate among survey respondents. .

To estimate response propensities, we needed to develop a set of questions that were likely to be good measures of this construct. To guide us, we turned to the nonresponse literature to identify predictors of respondent recruitment outcomes (see Groves & Couper, 1998) and developed questionnaire items within three basic groupings: i) respondent data, ii) interviewer data, and iii) call history data (see Table 1).

2.2 Criterion to Evaluate Results from Alternative Approaches

The product of each alternative approach considered here is a set of estimated propensities $(\hat{p}_i^{(approach)})$ for survey respondents. The basis for the criterion we used to evaluate the quality of these estimates relative to the actual (unknown) propensities (p_i) is the well-known fact that random error of estimation $(\varepsilon_i = \hat{p}_i^{(approach)} - p_i)$ would reduce the magnitude of measures of association between $\hat{p}_i$ and the study measures used to define what is being estimated (Biemer and Trewin, 1997, p. 620). For example, when estimating the population mean $(\bar{y})$ of some member measurement (y_i), the covariance $(\sigma_{\hat{p}y})$ between $\hat{p}_i^{(approach)}$ and y_i in the sampled population will be directly related to the amount of random error in estimating the propensities. Thus, if the combination of sampling error and/or modeling error associated with the process leading to the propensity estimates can be viewed as effectively random, then the estimation approach yielding the set of propensity estimates with the largest estimated value of $\sigma_{\hat{p}y}$ among the alternative approaches for any member measurement can be considered the "best" estimation approach. The usual weighted estimates ($\hat{\sigma}_{\hat{p}y}$) of $\sigma_{\hat{p}y}$ can be obtained from respondent data only by using sample weights that have been adjusted for nonresponse.[2] Finally, to enhance comparability among both estimation approaches and member measurements for various population means we will consider, we report the estimated relative bias *(ERB)* of the pre-adjustment estimated mean ($\hat{\bar{y}}_r$) from any approach as,

$$ERB(\hat{\bar{y}}_r; approach) = \frac{\hat{\sigma}_{r:\hat{p}y}^{(approach)} / RR4_w}{\hat{\bar{y}}_r} \quad , \tag{2}$$

where $\hat{\sigma}_{r:\hat{p}y}^{(approach)}$ is the estimated covariance between an approach's set of estimated response propensities and the corresponding member measurements for the survey estimate, and $RR4_w$ is the survey's weighted response rate. Notice that differences in *ERB* among approaches for the same survey estimate will only be due to differences in the estimated propensities, since $\hat{\bar{y}}_r$ and $RR4_w$ will be the same for each approach in these comparisons. Only values of $\hat{\sigma}_{r:\hat{p}y}^{(approach)}$ will differ among approaches.

2.3 Application to the 2007 North Carolina Parenting Survey

The three psychometric approaches and two traditional approaches were applied to recruitment data from a 2007 statewide telephone survey of 2,884 parents of small children conducted by the Survey Research Unit at UNC-CH. The overarching goal of the North Carolina Parenting Survey (NCPS) was to study child discipline in families with children who had been born in the state between October 1, 2005 and July 31, 2007. Respondents were identified from a disproportionately allocated stratified simple random sample of 38,334 live births occurring in North Carolina during this time interval. Three of the variables used to form an initial set of 24 sampling strata were the mother's

[2] Note that we used the same set of estimated propensities in the basic formulation for $\hat{\sigma}_{\hat{p}y}$.

education ($\leq$ high school / > high school), mother's age (<25 / $\geq$25), and mother's tobacco or alcohol use during pregnancy (yes/no). Since disproportionately larger sample sizes were sought for two specially targeted counties (Durham and Guilford), stratification was also by county (Durham County/Guilford County/all other NC counties). Relatively small population counts in the two specially targeted counties required that the number of initial strata be reduced to 18 by collapsing each county's four initial strata with births to mothers who had smoked or drank during pregnancy.

Sample recruitment for the NCPS telephone interview required two steps. Addresses for all 38,334 selected birth certificates were first matched against files of residential telephone numbers maintained by MSG/GENESYS (matching rate of 49%). A random subset of 12,828 of the birth certificates with matched telephone numbers were then assigned to trained SRU interviewers who called these numbers to invite eligible parents to complete an interview that averaged about 20 minutes in length (interview agreement rate of 54% based on a weighted RR4 conditioned on matching success and as defined by AAPOR, 2008).

Because of the two-step recruitment process in the NCPS, member-level values of $\hat{p}_i^{(approach)}$ used to compute $\hat{\sigma}_{r:\hat{p}y}^{(approach)}$ and thus *ERB* for each estimation approach were the product of an estimated matching propensity $(\hat{p}_{match,i})$ for the member and each member's approach-specific estimated propensity to agree to be interviewed $(\hat{p}_{agreement,i}^{(approach)})$. In other words,

$$\hat{p}_i^{(approach)} = \left[\hat{p}_{match,i}\right]\left[\hat{p}_{agreement,i}^{(approach)}\right]. \tag{3}$$

The same set of matching propensities were used for each set of $\hat{p}_i^{(approach)}$, since the additional ancillary data we applied to psychometric methods were intended to improve estimates of agreement propensities rather than matching propensities.

Estimated member propensities for matching were obtained as predicted values of the dependent variable from a forward-elimination stepwise logistic regression model-fitting process in SAS (v.9.1.3), with all main effects and first-order interactions included. This stepwise procedure was run on those 31,537 members of the original sample of birth certificates that were sent for telephone number matching but that had no item nonresponse on either the 0/1 matching dependent variable or any of the 10 independent variables that were used in the model (see Table 1). The final stepwise model was run again using SUDAAN, v.9.0.3 (RTI, 2007) to identify the final set of significant predictors and to produce the estimated matching propensities.

2.4 Weighting Class (WC) Approach

Propensities based on the two traditional approaches were computed as follows. Adjustment cells for the weighting class approach were formed by the 18 sampling strata, and participation propensities were estimated for any respondent simply as the unweighted AAPOR-RR4 conditional response rate for sample members in the respondent's sampling stratum.[3]

2.5 Response Propensity Modeling (RPM) Approach

Model-fitting leading to the estimated participation propensities under the response propensity modeling approach followed a comparable stepwise process to the one used for matching propensities, except that here the binary dependent variable for the logistic regression model was an indicator of participation (1) or not (0) among known survey eligibles, and a somewhat different list of independent variables was considered (see Table 2).

2.6 Reliability Scaling (RS) Approach

Three psychometric approaches were used to measure response propensity. The first approach, what we are calling *reliability scaling (RS)*, simply retained items with the highest reliability as measured by Cronbach's alpha ($\alpha > .70$). These items included i) respondent's interest in the research topic; ii) interviewer perception of respondents' interest, engagement, patience, and forthcomingness throughout the interview process as well as the respondents' level of understanding in terms of the questions read to them; and iii) respondent race and income status. An overall scale,

[3] Unweighted response rates could be used here since base weights for all members of each sampling stratum in this design are equal.

equaling the sum of the ordinal responses to the retained items, was standardized so that: (i) scores ranged between 0 and 1 and (ii) the mean equaled the NCPS response rate.

2.7 Fitted Model Scaling (FMS) Approach

The second psychometric approach, labeled as *fitted model scaling (FMS)*, involved fitting a regression model to identify significant predictors of response propensity among the items seen in Table 1. The dependent variable was the respondents' answer to the question "Would you say, in general, you are the type of person who does surveys?" The significant predictors prove to be the respondents' interest in the topic, the interviewers' expectation of completing an interview, the number of interruptions during the interview, the number of call attempts made to the household, and the respondents' household income. Again, items were scaled and standardized as done above.

2.8 Factor Scaling (FS) Approach

The third psychometric approach, referred to as *factor scaling (FS)*, employed factor analysis to determine the items which best measure response propensity. To accomplish this, the respondent dataset with the items in Table 1 was randomly split in half. The first half was used to perform exploratory factor analysis (EFA). The EFA model produced four factors (with eigen values greater than 1), but only two factor made intuitive sense and were retained for the confirmatory factor analysis (CFA) model which was ran on the second half of the dataset. Two factors emerged and included an *interest factor* (as measured by the respondents' interest in the topic and the interviewers' perceptions of respondent engagement, interest, forthcomingness & patience), and a *demographic factor* (as measured by the respondents' age, income, race & smoking status during pregnancy). A factor score was produced in MPlus (Muthén & Muthén, 2007) for each respondent based on the CFA model and was standardized as described above.

3. Findings

The unweighted frequency distribution of the estimated agreement propensities ($\hat{p}^{(approach)}_{agreement,i}$) under each approach was determined (but not presented) for NCPS respondents. These results indicated that the mean propensity for all but the RPM approach were the same, since the psychometric approach results were rescaled so that their means equaled the agreement rate from the NCPS. The higher mean for the RPM approach can be explained by the fact that those sample members of unknown eligibility were excluded in fitting the logistic regression model. Although all distributions were negatively skewed, this is especially true for the traditional approaches (skewness = -1.66 and -0.66 for the WC and RPM approaches, respectively) and for the reliability scaling approach (skewness = -0.82). Finally, estimated propensities obtained via the three psychometric approaches were found to be more variable (0.115<sd<0.122) than those obtained from the traditional approaches (0.036<sd<0.049).

The intercorrelation between pairs of the five approaches is shown in Table 3. Here we note that the correlation between the three psychometric approaches is relatively high (0.65 to 0.85), while the traditional RPM approach is only modestly correlated with all others (0.40 to 0.60) with the correlation between the two traditional approaches being the lowest in this group . The correlation between the weighting class approach and the psychometric approaches is quite low (0.20 to 0.25).

The main comparison of the quality of estimated propensities from the traditional and psychometric approaches is found in Figure 1. Presented here are the magnitudes of the (all negative) values of $ERB(\hat{\bar{y}}_r)$ obtained from Eq. (2) using values of estimated response propensities $(\hat{p}_i^{(approach)})$ based on the five approaches. These findings are presented for five of the key survey outcome measures from the NCPS as defined explicitly in Table 3. If it is reasonable to consider the quality of the response propensity estimates from an approach as being directly related to the magnitude of *ERB* from that approach, then the results in this figure generally suggest that the propensities obtained from the psychometric approaches are generally superior to those from the traditional approaches. The only exception to this trend occurs when values of *ERB* are smallest, particularly for the frequency of participation in religious activities. Among the psychometric approaches, we found for the three measurements with the largest *ERB* that the most sophisticated of the three (factor scaling) produced the best propensity estimates, although the simplest approach (reliability scaling) is nearly as good for two of these survey outcome measurements ("general health" and "neglect"). The weighting class approach consistently produced the worst propensity estimates among all approaches here.

Demonstrating the validity of any derived scale by showing that it performs consistently with other external evidence is a key part of establishing its utility (Nunnally & Bernstein, 1994). Our strategy to provisionally establish the validity of the propensity estimates obtained from each of the five psychometric and traditional approaches was to compares values of $ERB(\hat{\bar{y}}_r)$ from Eq. (2 for estimated rates and means associated the five NCPS measures in Table 4 between dichotomous subgroup defined by the following three demographic characteristics of the mother that were suspected to be statistically associated with survey response success: race-ethnicity, education, and urbanization of residence. Seventy-five (75) subgroup comparisons were therefore made for each combination of demographic variable, measurement, and approach. For instance, we compared *ERB* between white and non-white mothers for estimates of the percentage of parents who shouted, yelled, or screamed at their child in the past year, based on the survey measurement, "Yell," in Table 4. The results of this comparison are seen in Figure 2, where it is apparent that for all approaches that *ERB* is greater for non-white mothers than white mothers, especially for the psychometric approaches. The relative sizes of these subgroup values by race-ethnicity are consistent with other evidence on the role of race-ethnicity and survey recruitment outcomes (Kalsbeek, et al., 2002). Unfortunately, less than a third of the *ERB* findings were as expected (24 of 75 overall; 16 of 45 for the psychometric approaches only). A partial explanation for these findings may lie in the low rate of matching birth certificate addresses to phone numbers and the significant percentage (19.1%) of assigned telephone numbers found to not be in service, thus raising uncertainty the quality linkages between selected birth certificates and households where telephone interviews were attempted. The lack of clear validation in this way suggest the need for further improvement in the psychometric approaches

References

American Association for Public Opinion Research (2008). *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Survey (5th edition)*. Lenexa, Kansas: AAPOR.

Biemer, P.P. and D. Trewin (1997) "A Review of Measurement Error Effects on the Analysis of Survey Data," in Lyberg, L. et al. (Eds.), *Survey Measurement and Process Quality,* New York: John Wiley and Sons, p. 620.

Groves, R. and E. Peytcheva (2008). "The Impact of Nonresponse Rates on Nonresponse Bias: A Meta-Analysis." *Public Opinion Quarterly* **72**(2): 167-189

Groves, R. and M.P. Couper (1998). *Nonresponse in Household Interview Surveys*. New York, John Wiley & Sons, Inc.

Hansen, M. and W. Hurwitz (1946). "The problem of nonresponse in sample surveys." *Journal of the American Statistical Association* **41**(236): 517-529

Iannacchione VG, Milne JG, and Folsom RE (1991). Response Probability Weight Adjustment Using Logistic Regression. American Statistical Association Proceedings.

Kalsbeek, W. D., and Agans, R. P. (2007)"Sampling and weighting in household telephone surveys," in J. M. Lepkowski, C. Tucker, J. M. Brick, E.D. deLeeuw, L. Japec, P.J. Lavrakas, M.W. Link, and R.L. Sangster (Eds.), Advances In Telephone Survey Methodology, New York: J.W. Wiley and Sons, Inc.,

Kalsbeek, W.D., Yang, J., and Agans, R. P. (2002). "Predictors of Nonresponse in a Longitudinal Survey of Adolescents," Proceedings of the Section on Survey Research Methods, American Statistical Association. 1740-1745

Kalton, G. and Flores-Cervantes, I. (2003). Weighting Methods. *Journal of Official Statistics, 19*(2), 81-97.

Kalton G (1983). *Compensating for Missing Survey Data.* Research Report Series. Ann Arbor, MI: Institute for Social Research, University of Michigan.

Lepkowski, J., G. Kalton, and D. Kasprzyk (1989). "Weighting Adjustments for Partial Nonresponse in the 1984 SIPP Panel" *Proceedings of the Section on Survey Research Methods, American Statistical Association* 296-301:

Lessler JT, and Kalsbeek, WD (1992). Nonsampling Errors in Surveys. New York: Wiley and Sons.

Muthén & Muthén (2007). MPlus (Version 5.1) [computer software]. Los Angeles, CA

Nunnally, J. C., Bernstein, I. H. (1994). Psychometric Theory (3rd ed.). New York: McGraw-Hill, Inc.

Peytcheva, A. and K. Olson (2007). "Using Interviewer Observations to Improve Nonresponse Adjustments: NEW 2004." *Proceedings of Section on Survey Research Methods, American Statistical Association*: 3362-3371

Platek, R. (1977). "Some Factors Affecting Non-response." *Survey Methodology,* **3**, 191-214.

Platek, R., M. P. Singh, and V. Tremblay (1978) "Adjustments for nonpresonse in Surveys". *Survey Sampling and Measurement,* N.K. Namboodiri (ed). New York, Academic Press: 157-174

SAS (Version 9.1.3). SAS Institute, Cary, North Carolina.

Stevens, S.S. (1946). "On the Theory of Scales and Measurements", Science 103, 667-680

SUDAAN (Version 9.0.3). RTI, International, Research Triangle Park, NC

Table 1: Questions Used to Measure Response Propensity

TYPES OF ITEMS USED		
RESPONDENT DATA	**INTERVIEWER DATA**	**CALL HISTORY DATA**
- Interest in the topic - If interest increased or decreased during the survey - Part of the *Do Not Call Registry* - Type of person to do surveys - Reasons for not participating in surveys - Number of biological children - If mother smoked or drank during pregnancy - Household income - Age - Race	- Respondent engagement - Respondent interest - Respondent patience - Respondent forthcomingness - Respondent understanding of the research questions - Interviewer expectation of completing the interview - Differences in speech accent with regard to respondent - Differences in flow of speech with regard to respondent - Number of interruptions during the interview	- Number of call attempts made to household - Refusal status

Table 2: Independent Variables Used to Estimate Matching and Agreement Propensities

Variable	Categories	Used for Matching Propensities	Used for Agreement Propensities
Alcohol usage during pregnancy	Yes/No	X	X
Tobacco usage during pregnancy	Yes/No	X	X
County of mailing address	Durham or Guilford County/ Rest of NC	X	X
Rural / urban	1 Urban 2 Rural	X	X
Mother's race	Hispanic or Black/Other	X	X
Marital status	Married/Not Married	X	X
Mother's education	≤12 years of school/ >12 years of school	X	X
Father's education	≤12 years of school/ >12 years of school	X	
Mother's age	Continuous	X	
Child's age	Continuous	X	
Number of living children	Continuous		X
Number of prenatal care visits	Continuous		X

Table 3: Intercorrelation of Estimated Agreement Propensities Among Approaches

Approach:	WC	RPM	RS	FMS	FS
WC	1.000	0.376	0.237	0.218	0.228
RPM	---	1.000	0.512	0.443	0.579
RS	---	---	1.000	0.845	0.673
FMS	---	---	---	1.000	0.714
FS	---	---	---	---	1.000

Table 4: Y-variables from the NCPS

Variable	Type	Attribute/Description
General Health	0/1	Health is "excellent" or "very good"
Neglect	0/1	Respondent failed at least once to get child to doctor when needed in the past year
Yell	0/1	Either parent Shouted, yelled, or screamed at the child in the past year
Hit	0/1	Either parent hit the child with an object in the past year
Religious Activities	count	Number of times respondent participated in religious services in the past month

Figure 1: Estimated Relative Bias (Negative) of NCPS Estimates

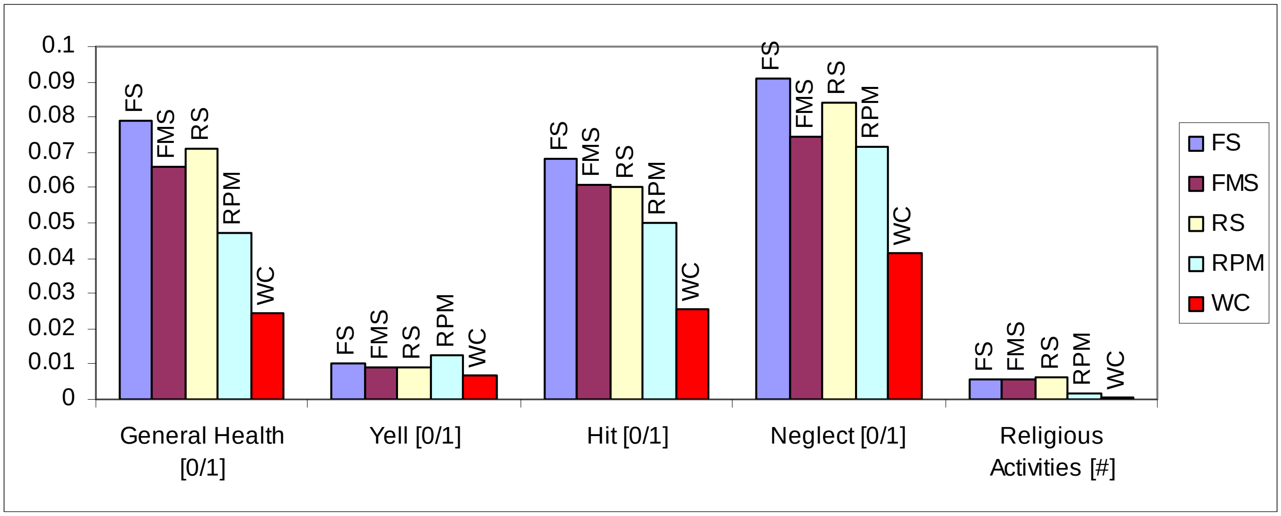

Figure 2: Example of a Subgroup Comparison of Estimated Relative Biases (ERBs): Prevalence of "Yell" by Mother's Race-Ethnicity `