Using Data from the National Health Interview Survey (NHIS) to Assess the Effectiveness of Nonresponse Adjustment in the Medical Expenditure Panel Survey (MEPS)

Lap-Ming Wun[1], Trena M. Ezzati-Rice[1], Ralph DiGaetano[2], Huseyin Goksel[2], Hongsheng Hao[2]
Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, MD 20850[1]
Westat, 1650 Research Blvd., Rockville, MD 20850[2]

Keywords: weighting; CHAID; response propensity; propensity score; logistic regression

1. Introduction

The Medical Expenditure Panel Survey (MEPS) is a complex national probability sample survey sponsored by the Agency for Healthcare Research and Quality (AHRQ). MEPS, on going since 1996, is designed to provide nationally representative estimates of health care use, expenditures, sources of payment, and insurance coverage for the U.S. civilian noninstitutionalized population. The MEPS consists of three inter-related surveys with the Household Component (HC) as the core survey. The MEPS-HC, like most sample surveys, experiences unit, or total, nonresponse despite intensive efforts to maximize response rates. Survey nonresponse is usually compensated for by some form of weighting adjustment to reduce the potential bias in survey estimates. Nonresponse adjustment methods make use of covariates that are available for both respondents and nonrespondents. Currently, the tree algorithm method, the Chi-squared Automatic Interaction Detector (CHAID), is employed in MEPS to develop models of response probabilities at the household or dwelling unit (DU) level from which nonresponse adjustment cells are created (Cohen, DiGaetano, and Goksel, 1999). An alternative method is the use of logistic models to predict response propensity, i.e., calculate a propensity score, of each sample unit (DU here). Grouping of these propensity scores can be used to form adjustment cells, or the inverse of the propensity score can be used as a weight adjustment to compensate for nonresponse.

Variables from the NHIS were evaluated in 1998, and a set was identified as potentially useful as covariates for response in the CHAID modeling (Cohen and Machlin, 1998). In 2004, additional variables were identified and added to the set (Kashihara, et al (2003)).

The views expressed in this paper are those of the authors and no official endorsement by the Department of Health and Human Services or the Agency for Healthcare Research and Quality is intended or should be inferred.

In this paper we evaluate the effectiveness of several nonresponse adjustment methods by applying the weights adjusted using each of the methods to some analytical variables from the NHIS.

2. Background: MEPS Survey Design and Estimation Strategy

The sample for the MEPS-HC is drawn from respondents to the National Health Interview Survey (NHIS), conducted by the National Center for Health Statistics. The MEPS-HC uses an overlapping panel design in which data are collected through a series of five rounds of interviews over a two and one-half year period. Detailed information on the MEPS sample design has been previously published (Cohen, 1997; Cohen, 2000).

Two separate nonresponse adjustments are performed as part of the process for development of analytic weights in MEPS. The first is an adjustment for DU nonresponse at round 1 to account for nonresponse among those households subsampled from NHIS for the MEPS. The 1996 to 2002 MEPS DU response rates ranged from 80-83 percent (among the NHIS households fielded for MEPS). The second is a person level nonresponse adjustment to account for survey attrition across the various rounds of data collection. This paper deals only with the DU nonresponse adjustment.

The base weight in the MEPS is the reciprocal of an intermediate weight from the NHIS reflecting the disproportionate sampling of minorities in NHIS with a ratio adjustment to NHIS population estimates to account for NHIS nonresponse and undercoverage. This ratio adjusted base weight is then adjusted for nonresponse of MEPS eligible sample DUs at round 1. More specifically, the base weights of MEPS responding DUs are adjusted to compensate for the nonresponding DUs.

3. Nonresponse Weighting Adjustment

The use of classifying or auxiliary variables, i.e., covariates, to form nonresponse adjustment cells is a commonly used method for nonresponse adjustment. It has been shown by Cochran (1968) that it is effective in removing nonresponse bias in observational studies. Rosenbaum and Rubin (1984) have indicated that as the number of covariates increases, the number of classes grows exponentially and suggest using predicted response probabilities or propensity scores from a logistic regression model based on the covariates to form the weighting classes or cells. Another use of propensity scores from logistic regression is to use the inverse of each respondent's predicted propensity score as an adjustment factor (see Kalton and Flores-Cerantes (2003)). In our study, we call this latter method the "direct use" of propensity scores. A propensity score of response in surveys is essentially the conditional probability that a person or household responds given the covariates. More elaboration of the propensity score and its application in nonresponse adjustments can be found in Little (1986) and Little and Rubin (2002) among others. A previous comparison of the use of covariates versus the use of response propensities to form classes for nonresponse adjustment for a complex sample survey, the third National Health and Nutrition Examination Survey (NHANES III), was reported by Ezzati-Rice and Khare (1994). Comparisons of the various methods of adjustment using MEPS' data for earlier years were reported by Wun et al (2004), and Wun et al (2005a).

The current method implemented by Westat to compensate for nonresponse in the MEPS at the DU level uses CHAID's "tree algorithm" response propensity approach (see Breiman, Friedman, Olshen, and Stone, 1993) to form nonresponse adjustment cells. In this research study, we investigate the CHAID method with the newly enhanced set of potential covariates as well as two alternative ways of using response propensities from logistic regression modeling to adjust weights to compensate for nonresponse.

As noted in Rizzo, Kalton, and Brick (1996), to the extent that sizable interactions exist, one might expect models that reflect only main effects to perform somewhat differently from those where interactions have been incorporated into the modeling. One logistic model and the CHAID model reflect components of interaction.

4. Methods

In this study, we assess the differences among the various methods of DU nonresponse adjustment at round 1 of the 2002 MEPS (panel 9). In the method currently used for MEPS, Westat uses a tree diagram generated by the computer package CHAID to form nonresponse adjustment cells based on response propensity using a set of classifying variables. Cells are collapsed, if necessary to ensure that the number of respondents in a cell are no less than 20 (Göksel, Alvarez-Rojas, and Hao, (2001)). Adjustment factors are not permitted to exceed two in value in order to limit the impact of such factors on the variability of sample estimates. It should be noted that because of the unique sample linkage of MEPS and the NHIS, a sizeable number of variables are available from the NHIS for responding and non-responding eligible MEPS DUs. The following is the list variables used by Westat as potential predictors of response propensity to construct subclasses for the DU nonresponse adjustment in MEPS-HC through 2001. These classifying variables were determined based on analysis of 1996 MEPS-HC data (Cohen and Machlin, 1998).

1. Age of the reference person
2. *Race/ethnicity of the reference person
3. *Marital status of the reference person
4. *Gender of the reference person
5. *Number of persons in the DU
6. *Education of the reference person
7. *Family income of the reference person
8. Employment status of the reference person
9. *Phone number refused in NHIS
10. *Major work status – working or reason for not working
11. DU level health status
12. *If anyone in the DU needs help with daily activities
13. *Census region
14. *Metropolitan Statistical Area (MSA) size
15. *MSA/Non MSA residence
16. *Urban/Rural residence
17. *Type of primary sampling unit (PSU)
18. *Predicted poverty status of the household
19. Any Asian in the household
20. *Any Black in the household.

The following additional covariates were identified (Kashihara, et al (2003)) and added to the list in 2004 for the 2002 MEPS:

1. Interview language
2. US citizenship of the reference person
3. Born in US - reference person
4. Type of home, e.g., house, apartment etc.
5. Time period without phone
6. Family medical expenses category
7. *Homeowner status of the reference person
8. *Number of nights in the hospital last year
9. Healthcare coverage.

In the two lists above, the variables that actually entered the CHAID models for the 2002 MEPS are identified by an *.

An alternative to the current CHAID propensity nonresponse adjustment method is to develop a logistic regression model to predict response status based on the selected set of covariates. A propensity score of response in surveys is essentially the conditional probability of response given the covariates. It was calculated through the following steps:

1. Identify an appropriate logistic model from the set of potential covariates with

response/nonresponse indicator as the dependent variable
2. Calculate logit for each sample unit using the model established in 1 above.
3. Convert the estimated logit value obtained from the logistic model established in step 2 into the predicted probability of response, i.e., the propensity score, through the following equation:

PROB=EXP(LOGIT)/(1+EXP(LOGIT)).

In this study, with a propensity score calculated for each sample unit, the propensity score from the logistic regression is used in two different ways:

1. Direct:

The estimated propensity score of each respondent is used directly as the adjustment factor, i.e., each individual respondent's base weight is multiplied by the inverse of their propensity score. No constraint was imposed on the size adjustment values could take. The sum of the resulting weights across respondents using this mode of nonresponse adjustment is not generally the same as the sum of the weights prior to nonresponse adjustment across respondents and nonrespondents.

2. Grouping scores to form adjustment cells:

Using the propensity scores, the sample is grouped into classification cells. In this study, we used five groups, based on the discussion in Cochran (1968) and extended to propensity scores in observational studies by Rosenbaum and Rubin (1984). Comparisons between the use of five and 100 cells as part of previous MEPS nonresponse research indicated little gain with the use of 100 cells (Wun et al (2004), and Wun et al (2005a)).

5. Models Evaluated

A total of five models were evaluated (one based on CHAID and four based on Logistic regression). The CHAID model is:

1. The current CHAID approach as used for the 2002 MEPS with 29 potential covariates to form adjustment cells. This method is coded CC (CHAID, Current) in the tables of results.

Two logistic models were selected based on a separate modeling study reported in Wun et al (2005b). Each set of propensity scores from each of the two models were then used directly. In addition, two additional models were employed, based on five adjustment cells formed by grouping direct propensity scores by size. These four models are:

2. Model with main effects only - model with all 29 potential covariates plus the MEPS base weight as a covariate following the rationale of Little and Vartivarian (2003); resulting propensity scores were used to form 5 adjustment groups. This method is coded M5 (Main effect, 5 groups) in the tables of results.

3. The same main effects only model as in 2 above but resulting propensity scores were used directly for adjustment. This method is coded Md (Main effect, direct) in the tables of results.

4. Model with main effects and interactions; resulting propensity scores were used to form 5 adjustment groups. This method is coded F5 (Full model, 5 groups) in the tables of results.

5. The same model with main effects and interactions as in 4 above but the propensity scores were used directly for adjustment. This method is coded Fd (Full model, direct) in the tables of results.

6. Analytical Variables used for Evaluation

The following 6 variables/conditions from NHIS were selected for the evaluation of the adjustment methods:

- Head start - children in Head Start program.
- Medical test - received any medical advice or test results in the past two weeks.
- Honorably discharged - ever been honorably discharged from active duty in the US armed forces.
- Doctor's visits - seen a doctor or other health care professionals in the past two weeks.
- Limitations - limitation in daily activities.
- Barrier to health care - Medical care was delayed due to cost concerns some time during the past 12 months.

For each of these six items, a dichotomous variable was constructed. A record for an NHIS DU received the value 1 if any DU members had a given condition, and was assigned the value 2 otherwise.

7. Evaluation Criteria and Results

For each of the six selected variables, the estimated proportion of DUs satisfying a given condition (i.e., coded 1) was computed using each of the nonresponse adjusted weights as well as the MEPS base weight (i.e., the pre-nonresponse adjustment weight but applied only to respondents).

Estimates calculated for the total MEPS sample (i.e. all households selected from the 2001 NHIS for the 2002 MEPS) using the base weight represent the full sample or "target" values to which the estimates obtained from weights after adjustment for nonresponse can be compared for evaluation purposes. The difference (in absolute value) between the estimates calculated using the nonresponse adjusted weights and the "target values" is used as the measure of bias. In addition, the Mean Square Error (MSE) has been calculated to provide a measure reflecting both sample variance and bias (MSE = SE^2 + $bias^2$, where SE represents the standard error of the estimate).

The results are given in tables 1 and 2. In table 1, column (1) is the full sample or target value for the estimated proportion of DUs in category 1 of each of the six variables. Column (2) is the absolute bias of the estimates using only the respondents with their MEPS base weight, while columns (3) to (7) provide the absolute biases associated with each of the five models ordered by magnitude, e.g., for Head start, Md has the lowest bias and CC has the highest bias etc. Table 2 provides a similar arrangement but for MSE results.

8. Observations

Estimates using weights based on the direct propensity score adjustment method from the main effects only model (i.e., model Md) had the smallest bias and MSE for five of the six analytical variables. Estimates using weights based on the CHAID approach generally had higher bias and MSE compared to the logistic regression models.

In general, as expected, the bias and MSE were greatest when no adjustment was done. However, there were some exceptions. For four of the six analytical variables (Head start, Doctor's visit, Limitation, and Barrier to health care), the estimates using respondents only with unadjusted weights had higher bias and MSE than any of the adjusted weights. For the other two analytical variables (Medical tests, and Honorably discharged), the bias and MSE of the respondents with unadjusted weights were lower than the measures for the F5, and CC weights .

9. Summary and Discussion

The models developed using logistic regression, in general, had somewhat lower measures of bias and MSE for the selected analytical variables. The main effects only models and the direct propensity score for the full model with interaction components had the lowest values.

There were some limitations associated with our investigation. First, this current investigation was limited to dichotomous categorical variables created at the DU (household) level, derived from person or family level variables. A major focus for MEPS is on expenditure information. The range of values for adjustment factors for the propensity models (as high as 5.5) could potentially result in unstable estimates of expenditures. Thus, additional research should be carried out to investigate the impact of constraining the inverse propensity scores as is done in CHAID where adjustment factors are capped at 2.0. Also, future research should include an evaluation of estimates based on continuous variables, including those with skewed distributions. In addition, MEPS oversamples a number of subgroups of policy and analytic interest including Hispanics, blacks, Asians, and households containing persons predicted to be poor. The effectiveness of the models in regards to these analytic domains remains to be investigated. Finally, implementation advantages and disadvantages of the CHAID and logistic regression response propensity modeling approaches merit consideration. The effectiveness of the CHAID and logistic regression methods in reducing nonresponse bias will continue to be investigated.

References

Brieman, L., Friedman, J.H., Olshen, R.A., and Stone, C.J. (1993). Classification and Regression Trees. New York: Chapman and Hall.

Cohen, S.B., and Machlin, S.R., (1998). "Nonresponse Adjustment Strategy in the Household Component of the 1996 Medical Expenditure Panel Survey," Journal of Economic and Social Measurement, 25, 15-33.

Cohen S.B. (1997). "Sample Design of the Medical Expenditure Panel Survey Household Component". Agency for Health Care Policy and Research, MEPS Methodology Report, No. 2, Rockville, MD., AHCPR Pub. No. 97-0027.

Cohen, S.B., DiGaetano, R., and Goksel, H., (1999). "Estimation Procedures in the 1996 Medical Expenditure Panel Survey Household Component," Agency for Health Care Policy and Research, MEPS Methodology Report No. 5, AHCPR Pub. No. 99-0027.

Cohen S.B. (2000). "Sample Design of the 1997 Medical Expenditure Panel Survey Household Component". Agency for Healthcare Research and Quality, MEPS Methodology Report, No. 11, Rockville, MD., AHRQ Pub. No.01-0001.

Cochran, W.G. (1968). Removing Bias in Observational Studies," Biometrics, 24, 295-313.

Ezzati-Rice, T.M., and Khare, M. (1994). "Modeling of Response Propensity in the Third National Health and Nutrition Examination Survey," 1994 Proceedings of Survey Research Methods Section of the American Statistical Association, 955-959.

Göksel, H.A., Alvarez-Rojas, L., Hao, H., (2001). "MEPS Panel V Round 1 – DU level weights" Unpublished internal memo of Westat, WGTS # 469.01, P5R1 # 1.01.

Kashihara, D., Ezzati-Rice, T.M., Wun, L.-M., and Baskin, (2003). "An Evaluation of Nonresponse Adjustment Cells for the Household Component of the Medical Expenditure Panel Survey (MEPS)". 2003 ASA Proceedings of the section on Survey Methodology, pp.3758-3765, (CD-ROM).

Kalton, G., and Flores-Cervantes, I., (2003) "Weighting Methods," Journal of Official Statistics, Vol. 19, No. 2, 81-97.

Little, R.J.A., (1986). "Survey Nonresponse Adjustments for Estimates of Means," International Statistical Review, 54, 139-157.

Little, R.J.A. and Rubin, D.B. (2002). Statistical Analysis With Missing Data (2nd Ed.). New York: Wiley.

Little, R.J. and Vartivarian, S., (2003) "On Weighting the Rates in Non-response Weights," Statistics in Medicine, 22, 1589-1599.

Rizzo, L., Kalton, G., and Brick, J.M., (1996) "A Comparison of Some Weighting Adjustment Methods for Panel Nonresponse," Survey Methodology, Vol.22, No. 1, pp. 43-53.

Rosenbaum, P.R., and Rubin, D.B., (1984). "Reducing Bias in Observational Studies Using Subclassification on Propensity Score," Journal of the American Statistical Association, 79, 516-524.

Wun, L.-M., Ezzati-Rice, T.M., Baskin, R., Greenblatt, J., Zodet, M., Potter, F., Diaz-Tena, N., and Touzani M., (2004), "Using Propensity Scores to Adjust Weights to Compensate for Dwelling Unit Level Nonresponse in the Medical Expenditure Panel Survey," 2004 ASA Proceedings of the section on Survey Methodology, pp. 4625-4631, (CD-ROM).

Wun, L.-M., Ezzati-Rice, T.M., Diaz-Tena, N., and Greenblatt, J., (2005a),"On Modeling Response Propensity for Dwelling Unit (DU) Level Nonresponse Adjustment in the Medical Expenditure Panel Survey (MEPS)," presented in the Tenth Biennial CDC and ATSDR Symposium on Statistical Methods, 2005, Washington DC.

Wun, L.-M., Ezzati-Rice, T.M., Diaz-Tena, N., and Greenblatt, J., (2005b),"Evaluation Alternative Propensity Models for Adjusting Weights to Compensate for Dwelling Unit Nonresposne in the Medical Expenditure Panel Survey (MEPS)," to appear in 2005 ASA Proceedings of the section on Survey Methodology, (CD-ROM).

Definitions of abbreviations used in tables 1 and 2:

Resp. = Respondents using base weight. CC = CHAID (current)
M5 = Main effect, 5-group; Md = Main effect, direct.
F5 = Full model (w/ interactions), 5-groups. Fd = Full model (w/ interactions), direct.

Table 1

Percentage absolute bias
(five models (columns (3) to (7)) are in increasing rank order according to size of bias)

Variable Name	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Head start	Target 7.03%	Resp. 0.78	Md 0.16	M5 0.23	Fd 0.29	F5 0.34	CC 0.36
Medical tests	Target 10.20%	Resp. 0.21	Md 0.10	M5 0.12	Fd 0.14	F5 0.22	CC 0.26
Honorably discharged	Target 19.36%	Resp. 0.99	Md 0.76	Fd 0.81	M5 0.91	F5 1.02	CC 1.09
Doctor's visit	Target 32.60%	Resp. 1.80	M5 1.40	Fd 1.47	Md 1.48	F5 1.50	CC 1.67
Limitation	Target 25.36%	Resp. 2.05	Md 0.69	Fd 0.75	M5 0.79	F5 0.91	CC 1.13
Barrier to health care	Target 14.72%	Resp. 1.15	Md 0.59	Fd 0.59	M5 0.68	F5 0.70	CC 0.70

Table 2

MSE (in percentage points)
(five models (columns (2) to (6)) are in increasing rank order according to size of MSE)

Variable Name	(1)	(2)	(3)	(4)	(5)	(6)
Head start	Resp. (0.74)	Md (0.14)	M5 (0.17)	Fd (0.21)	F5 (0.24)	CC (0.25)
Medical tests	Resp. (0.25)	Md (0.21)	Fd (0.22)	M5 (0.23)	F5 (0.26)	CC (0.28)
Honorably discharged	Resp. (1.34)	Md (0.94)	Fd (1.02)	M5 (1.19)	F5 (1.41)	CC (1.56)
Doctor's visit	Resp. (3.56)	M5 (2.38)	Md (2.64)	Fd (2.62)	F5 (2.69)	CC (3.22)
Limitation	Resp. (4.83)	Md (1.05)	Fd (1.14)	M5 (1.22)	F5 (1.41)	CC (1.85)
Barrier to health care	Resp. (1.61)	Md (0.64)	Fd (0.64)	M5 (0.76)	F5 (0.79)	CC (0.79)