Evaluation of PPS Sampling for the Medical Expenditure Panel Survey

Steven R. Machlin, William Yu, and Robert M. Baskin
Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, MD 20850

KEY WORDS: MEPS, complex multi-stage sampling, PPS sampling, effective sample size

1. Introduction[1]

The Medical Expenditure Panel Survey (MEPS) has been conducted on an annual basis since 1996 by the Agency for Healthcare Research and Quality (AHRQ). MEPS is a complex national probability survey of the civilian noninstitutionalized population. Each new panel of sample households is drawn as a sub-sample of respondents to the prior year's National Health Interview Survey (NHIS) sponsored by the National Center for Health Statistics (Ezzati-Rice et. al.). Due to the combined effects of stratification, clustering and unequal weighting, estimates from complex national surveys like MEPS and NHIS are generally subject to increased variance relative to what would have been obtained under a simple random sample design (Korn and Graubard).

The procedure used for selecting the MEPS samples for 1996-2009 (panels 1-14) was to stratify the NHIS frame eligible for MEPS and select units systematically within strata at varying rates (Ezzati-Rice et. al.). These varying rates were designed to improve sample sizes for targeted population subgroups (i.e., race/ethnicity and/or low income groups) than would be expected from a sample design without stratification. However, this design oversamples some groups that were already oversampled in the NHIS, which has the effect of increasing the overall variation of MEPS weights beyond that of the NHIS weights and increasing the variance of MEPS national estimates.

A strategy to substantially reduce the overall variation of MEPS weights and national estimates is to use probability proportional to size (PPS) sampling (Korn and Graubard) with the NHIS weights as measure of size. While this strategy would have the benefit of reducing the overall variation in sampling weights and thereby improve the precision of national estimates for the total population, PPS sampling in the absence of stratification does not insure specific sample sizes or precision levels for targeted minority subpopulations of interest to policymakers.

This analysis evaluates the potential gains and tradeoffs based on various PPS sampling options for MEPS and presents a strategy for allocating a fixed total sample size across race/ethnicity subgroups. This strategy is designed to balance gains in overall precision from PPS with losses of effective sample sizes in target population subgroups. The results can be used to inform potential revisions in the MEPS sample design for future MEPS panels.

---

[1] 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 are intended or should be inferred. The authors wish to thank Trena Ezzati-Rice, Joel Cohen and Steven Cohen for their helpful reviews of this paper.

2. Methods/Analytic Approach

2.1 Data

We based our evaluation on the MEPS sample selected for 2008 (panel 13), a total of 9,689 households selected from a representative cross-section of approximately 3/8 of all NHIS responding households in 2007. Based on the fixed sample size of 9,689 households in panel 13, we compared effective sample sizes (see definition below) under the stratified systematic sampling approach that was actually used for panel 13 (i.e., current) to simulated results for different sampling approaches using PPS (with and without stratification).

2.2 MEPS Stratified Sampling Approach (Current)

The 12,437 eligible households from the 2007 NHIS were stratified hierarchically into 5 mutually exclusive groups (Ezzati-Rice et. al.) in the following order: 1) households with 1 or more persons identified as *Asian*, 2) households with family income predicted to be below 200% of poverty (*poor*) (Wun et. al.), 3) households with 1 or more persons identified as *Hispanic*, 4) households with 1 or more members classified as *Black*, and 5) all other households (*White/other*). All households in the first 4 strata were selected in order to maximize the sample size in these population subgroups. Survey budget constraints allowed for a sample of about 58% of the remaining stratum (White/other). Also, because a decision was made based on prior research to not maintain stratum 2 as a separate sampling stratum in subsequent MEPS panels (Wun et. al.), we allocated the sample selected for the second stratum to strata 3-5 and eliminated stratum 2 from the analysis (see Table 1).

Table 1: Sample Selection Approach for MEPS Panel 13 (Current)

				Re-categorized for the Evaluation*	
MEPS Sampling Stratum	**No. of NHIS Eligible Households**	**MEPS Sampling Rate**	**No. of MEPS Selected Households**	**No. of NHIS Eligible Households**	**No. of MEPS Selected Households**
Asian	825	100.0%	825	825	825
Poor*	1,800	100.0%	1,800	---	---
Hispanic	1,854	100.0%	1,854	2,512	2,512
Black	1,497	100.0%	1,497	2,053	2,053
White/Other	6,461	57.6%	3,713	7,047	4,299
Total	12,437	77.9%	9,689	12,437	9,689

*Poor households include those below 200% Poverty. These households were distributed to Hispanic, Black or White/other categories for the evaluation.

2.3 MEPS base weights

MEPS annual survey weights are derived through a series of steps (Cohen et. al.). First MEPS household base weights are constructed by multiplying the inverse of the MEPS selection probability by the non-response adjusted NHIS weight.[2] Because the non-response adjusted NHIS weight is not available at the time of MEPS sampling, in this analysis we use an interim annual weight (WTIA_HH) that accounts for the initial sampling and within stratum subsampling components of the NHIS sample design but does not include NHIS nonresponse or poststratification adjustments. It should also be noted that although MEPS base weights are ultimately adjusted for many additional factors (e.g., nonresponse, attrition, raking to population totals) to produce a final annual weight, this analysis is based only on variation in the MEPS base weights across various sampling options.

2.4 Effective Sample Size Definition

Effective sample size (ESS) is the primary measure used for this analysis and is defined as follows:

ESS = Actual # of households in sample / Design effect (DEFF),
where DEFF=1 + (standard deviation of MEPS base weights/mean MEPS base weight)^2

The ESSs are lower than actual sample sizes because they account for the loss in precision associated with the variation in weights (e.g., DEFF>1). Our analysis only assesses the impact of variation in weights on ESSs and does not account for the impact of stratification and clustering relative to different sampling approaches.

Table 2 below shows actual sample sizes, design effects, and effective sample sizes based on the MEPS panel 13 stratified design. In our analysis, we compare ESSs (last column in table 2) under the current design with those obtained by applying various alternative PPS sampling scenarios (see below).

Table 2: Actual vs. Effective Sample Sizes for Panel 13: Current Stratified Design (Evaluation Strata)

MEPS Sampling Stratum	Actual MEPS Sample Size (Households)	DEFF	Effective Sample Size
Asian	825	1.36	605
Hispanic	2,512	1.26	1,994
Black	2,053	1.20	1,708
White/other	4,299	1.24	3,477
Total	9,689	1.66	5,831

[2] This weight is also multiplied by a factor (16/6) that adjusts for the MEPS sample being drawn from only two panels and three quarters of the NHIS.

2.5 PPS Approach

2.5.1 Measure of Size (MOS)

As stated previously, NHIS weights were used as the PPS measure of size (MOS). Table 3 below contains descriptive statistics on NHIS weights by MEPS sampling strata. The average NHIS weight varies among MEPS strata (primarily due to oversampling of minority groups) and these weights also vary within stratum (see coefficients of variation) as a result of various factors used to derive the weights (ref).

Table 3: Variation in NHIS Weights by MEPS Sampling Strata

MEPS Sampling Stratum	Average NHIS Weight	Coefficient of Variation
Asians	2,461	60.3%
Hispanics	2,171	51.0%
Blacks	2,536	45.0%
White/Others	4,661	43.1%
Total	3,662	56.1%

Table 4 below uses a simplified hypothetical example as the basis to describe how NHIS weights are used as MOS and illustrate how PPS has the potential to substantially reduce variation in MEPS weights. The example assumes no variation in NHIS weights within MEPS strata. Column B of the table shows that the MEPS PPS probability of selection ranges from .184 for Hispanics to .394 for the White/other group and increases with the NHIS MOS. The probabilities for MEPS selection were computed by taking the NHIS weight for the group divided by the sum of the NHIS weights for all the groups (i.e. 2,461/11,829=.208 for Asians). The last column of the table shows how this approach (i.e., overall PPS design) would produce MEPS weights with no variation (i.e., 11,829 in all strata).

Table 4: MEPS PPS Sampling Probabilities and Base Weights: Hypothetical Illustration (MOS varies between but not within stratum)

MEPS Stratum	NHIS Weight (MOS) (A)	MEPS PPS Probability of Selection (B)	MEPS Base Weight A*(1/B)
Asians	2,461	0.208	11,829
Hispanics	2,171	0.184	11,829
Blacks	2,536	0.214	11,829
White/Other	4,661	0.394	11,829
Sum	11,829	--	--

2.5.2 Sampling scenario alternatives

In order to examine tradeoffs among different approaches, we evaluated the following 3 alternative PPS scenarios relative to the current method:

PPS1: Overall PPS without stratification,
PPS2: PPS within White/other stratum and sampling rates of 1.0 in all other strata, and
PPS3 Sampling rate of 1.0 for Asians, PPS sampling within all other strata with sample size of 8,869 (after excluding the 825 Asian households selected with certainty) allocated to maximize sum of ESSs across strata. ESSs were computed for all possible combinations of sample sizes across strata to identify the optimal allocation.

3. Results

Table 5 below compares stratum-specific sampling rates and ESSs for the current method and the 3 PPS sampling alternatives evaluated. Compared to the current stratified approach, the overall PPS approach with no stratification (PPS1) resulted in substantially larger ESSs for the total population (8,692 vs. 5,831) and the White/other stratum (5,977 vs. 3,477), but ESSs for the minority strata were substantially lower—400 vs. 605 for Asians, 1,210 vs. 1,994 for Hispanics, and 1,205 vs. 1,708 for Blacks. The PPS2 approach preserves the higher ESSs for the minority stratum (i.e. sampling rate=100%) and increases the ESS for the White/other group and for the total population relative to the current method. However, the increases for the White/other group and total population are much smaller than under the PPS1 method. PPS3 achieves the largest ESS for the total population, with ESSs for the population subgroups falling between those for PPS1 and PPS3.

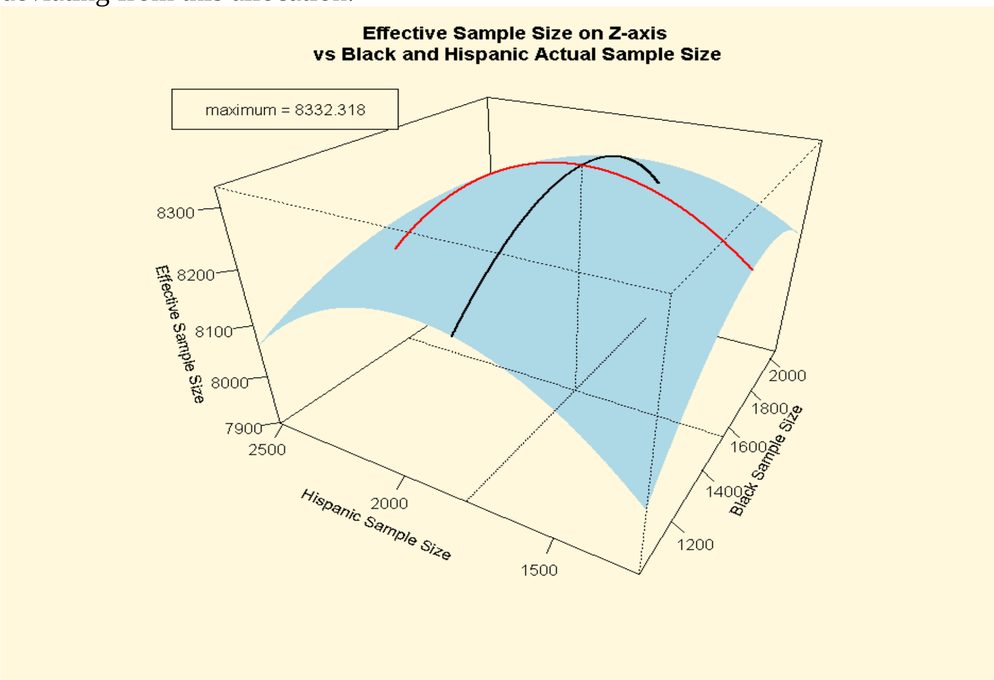
Table 5: Comparison of Sampling Rates and Effective Sample Sizes for Alternative Sampling Approaches

	Current		**PPS1**		**PPS2**		**PPS3**	
Sampling Stratum	**Sample Size (Rate)**	**ESS**	**Sample Size (Rate)**	**ESS**	**Sample Size (Rate)**	**ESS**	**Sample Size (Rate)**	**ESS**
Asian	825 (100%)	605	438 (--)	400	825 (100%)	605	825 (100%)	605
Hispanic	2,512 (100%)	1,994	1,240 (--)	1,210	2,512 (100%)	1,994	1,777 (71%)	1,661
Black	2,053 (100%)	1,708	1,235 (--)	1,205	2,053 (100%)	1,708	1,551 (76%)	1,469
White/other	4,299 (61%)	3,477	6,776 (--)	5,977	4,299 (61%)	4,186	5,536 (79%)	5,201
Total	9,689 (--)	5,831	-- (78%)	8,692	9,689 (--)	7,021	9,689 (--)	8,936

Summary/Discussion

Given a fixed sample size and assuming that all households with Asian members are sampled with certainty due to the relatively small number of Asian households in the NHIS frame eligible for MEPS, the results of this evaluation suggest that incorporating PPS (with NHIS weights as MOS) in the MEPS sample design has the potential to substantially reduce the variance of MEPS national estimates without severe loss of precision for race/ethnic subgroup estimates. The current stratified MEPS design without PPS maximizes sample sizes for race/ethnic subgroups by selecting 100% of households in those groups with the remaining budgeted sample selected from the White/other group. Our evaluation indicated that it is possible to increase the overall ESS for national estimates by about 20% (from 5,831 to 7,021) without any loss of precision to race/ethnic subgroups by simply modifying the current design to incorporate PPS sampling within the White/other group (i.e., PPS2 option). Moreover, notable further improvements in precision in national estimates may be achievable with modest sacrifices in precision for subgroup estimates. For example, we estimated that using an approach that maximizes the sum of ESSs across subgroups (PPS3 option) increased the ESS for national estimates by an additional 27 percent (from 7,021 to 8,936) with the tradeoff of relatively smaller reductions in ESSs for Hispanics (decrease of 17%) and Blacks (decrease of 14%).

Our analysis provides a basis for evaluating the pros and cons of different PPS sample design options for the MEPS. While the PPS3 option was developed to identify an "optimal" sample allocation across strata, there is not one "optimal" design since there are tradeoffs across different designs. For example, an approach that would increase ESSs for minority groups but sacrifice some precision in overall national estimates would be an allocation between the PPS2 and PPS3 options. The figure below illustrates the optimal allocation based on PPS3 (i.e., intersection of the two curves where Hispanic sample size=1,777 and Black sample size=1,551--see Table 5) and can be used to assess tradeoffs in ESSs for subgroups and the overall population (excluding Asians) when deviating from this allocation.



While this paper presents an empirical evaluation illustrating the potential benefits of PPS for MEPS sampling, areas for further research regarding the impact and feasibility of adopting a PPS approach include: 1) comparing the effect of different sampling strategies on standard errors for primary MEPS analytic variables such as healthcare expenditures (which are especially subject to high variability due to skewness), 2) development of methods to estimate NHIS weights for observations where the appropriate NHIS weight is not available at the time of MEPS sampling and 3) an assessment of the robustness of these findings by conducting multiple simulations per sampling approach rather than just one sample per approach.

References

Cohen SB, DiGaetano R, Goksel H. Estimation procedures in the 1996 Medical Expenditure Panel Survey Household Component. Rockville (MD):Agency for Health Care Policy and Research; 1999. MEPS Methodology Report No. 5. AHCPR Pub.No. 99-0027.

Ezzati-Rice, TM, Rohde, F, Greenblatt, J, *Sample Design of the Medical Expenditure Panel Survey Household Component*, 1998–2007. Methodology Report No. 22. March 2008. Agency for Healthcare Research and Quality, Rockville, MD. http://www.meps.ahrq.gov/mepsweb/data_files/publications/mr22/mr22.pdf

Korn E. and Graubard B., Analysis of Health Surveys, Wiley and Sons, 1999.

Wun, LM, Ezzati-Rice, TM, Baskin, R. "Evaluation of the Prediction Model and Oversample of Low Income Persons in the Medical Expenditure Panel Survey", 2008 Proceedings of the American Statistical Association, Survey Research Method Section, Alexandra, VA American Statistical Association.