Preliminary Evaluation of Impact of PPS on Precision of MEPS Estimates

Robert Baskin and William Yu and Sadeq Chowdhury

Agency for Healthcare Research and Quality Working Paper No. 12002

February 2012

Suggested citation: Baskin R, Yu W, Chowdhury S. Preliminary Evaluation of Impact of PPS on Precision of MEPS Estimates. Agency for Healthcare Research and Quality Working Paper No. 12002, February 2012, http://gold.ahrq.gov.

Preliminary Evaluation of Impact of PPS on Precision of MEPS Estimates

By Robert Baskin, William Yu, and Sadeq Chowdhury

ABSTRACT

This report presents a preliminary evaluation of the impact of probability proportional to size (pps) sampling within stratum in the Medical Expenditure Panel Survey (MEPS) Household Component. Based on the limited data from a MEPS pps design that are currently available, we examine insurance estimates as well as coefficients of variation of the sampling weights. While the sampling weights show a marked reduction in variability, the impact of pps on the standard errors of the insurance estimates is ambiguous. Expenditure data for the first year of panel 15 (2010) are scheduled for release in the fall of 2012 which will enable further evaluations of the effect of PPS on MEPS estimates.

Robert Baskin
Center for Financing, Access, and Cost Trends (CFACT)
Agency for Healthcare Research and Quality
540 Gaither Road
Rockville, MD 20850
Phone: (301)427-1669
Email: Robert.Baskin@ahrq.hhs.gov

Preliminary Evaluation of Impact of PPS on Precision of MEPS Estimates

By Robert Baskin, William Yu, and Sadeq Chowdhury

Introduction

The Medical Expenditure Panel Survey (MEPS) is an annual survey that has been conducted since 1996 by the Agency for Healthcare Research and Quality (AHRQ). The National Health Interview Survey (NHIS) which is a complex probability survey conducted by the National Center for Health Statistics is used as the frame for MEPS. From 1996 through 2009 the MEPS subsample of NHIS was obtained by stratifying the NHIS responding units by target populations such as race-ethnicity or low-income and using a systematic sampling procedure within the MEPS sample strata. In 2010 AHRQ implemented a probability proportional to size (pps) method of subsampling within the target strata with the measure of size based on the sampling weights from NHIS. The purpose of this pps sampling approach within strata was to reduce the sampling variance of the MEPS estimates by reducing the variation in sampling weights while retaining any benefit from the stratification methodology. The 2010 point-in-time (PIT) file is now available which contains early part of the year data from the first panel in which pps sampling was implemented. This project is an empirical evaluation of the weights and variance of early insurance estimates from the first year of pps sampling in comparison to the corresponding weights and variance estimates from the previous two years of systematic sampling without pps.

Data

The data used for this evaluation come from the first round of data collection from MEPS years 2008, 2009, and 2010. These MEPS years are also referred to as panels 13, 14, and 15. The panel 13 and 14 units were selected under systematic sampling within MEPS strata whereas panel 15 units were selected pps within MEPS strata. MEPS is selected as a subsample of the previous year's eligible responding households from NHIS. However, NCHS retains two of the four NHIS panels for their use and only NHIS quarters one, two, and three are available in time for MEPS selection so MEPS is only selected from about three eighths (2/4 panels times 3/4 quarters) of the eligible responding households in NHIS. In panels 13, 14 and 15 the available NHIS households were stratified hierarchically into households categorized in order as Asian[1], Black, Hispanic, or Other. If a household contained any member classified as Asian the entire household was classified as an Asian household for sampling purposes. If a household contained any Black but no Asian it was classified as a Black household. If a household contained no member in the previous two categories but did contain a Hispanic member the household was classified as a Hispanic household. And finally if a household was not any of the three previous strata it was classified as an Other household. The number of eligible and selected units in the strata and overall are for panels 13, 14 and 15 are given in Table 1.

Note that for panels 13 and 14 the total number of households selected was close to 9,700 but for panel 15 the total number selected was almost 1,000 smaller at 8,750. Thus, the evaluation of the pps sampling must take into consideration that the pps sample has a smaller sample size than the previous two panels.

Method of Evaluation

Two types of information will be compared across the panels. The first comparison is based on the Coefficient of Variation (CV) of the weights where the CV is the standard deviation of the weights divided by the mean of the weights. The CV of the weights contributes to the Design Effect (DEFF) of a survey estimate, where DEFF is the ratio of the variance of an estimate under a complex design to the variance under a simple random sample design. The relationship between the portion of DEFF due to unequal weights and CV of the weights is given by: $DEFF=(1+CV^2)$. In Machlin et. al. (2009), *effective*

[1] In panel 13 Asian households were combined with households with income below the poverty level as a certainty domain for sampling purposes.

sample size, which is the raw sample size divided by the DEFF, was used for comparison based on different sampling approaches with the same raw sample size but here we are comparing a raw sample of size 8,750 in panel 15 to samples of size approximately 9,700 panels 13 and 14 so the effective sample size will not be considered.

The second comparison is based on the standard errors of insurance estimates from panel 15 to those from panels 13 and 14 using data from Point in Time (PIT) files. More specifically estimates of the Percent Insured at any time during the early part of the year were made using the variable INSRD13X. Percents, standard errors, CVs in percents of the percents, and estimated design effects were calculated using SAS.

Evaluation of the CV of the Weights

Table 2 shows the CVs, in percents, of the Point in Time person weight, WGTSP13, and of two household weights starting with the NHIS final weight denoted WTFA and the MEPS household base weight denoted MEPS WT. The NHIS final weight represents the inverse of the probability of selection of the household with an adjustment for non-response and a final disposition of response. The MEPS base weight uses the available NHIS weight at the time of MEPS weight construction as the starting point and adjusts that weight for the probability of selection in MEPS. The table has the values for panels 13 through 15 broken down by sampling strata as well as for all weights together. First, in the Asian strata all units were selected in all 3 panels so pps sampling had no effect. Now, in the other three strata for panels 13 and 14 the systematic sampling from NHIS was used and the CVs of the MEPS WT are almost identical to the CVs of the NHIS final household weight, WTFA. However in panel 15, with the use of pps sampling, the CVs of MEPS WT are approximately two thirds of the CVs of the NHIS final weight for the Hispanic and Other sampling strata. In the Black sampling stratum, the CV of MEPS WT is approximately four fifths of the CV of the NHIS final weight. This is a very good confirmation that the pps sampling is accomplishing the goal of reducing the variation in the MEPS sample weights.

Under two assumptions it is possible to produce confidence intervals for the CVs to assess whether differences between panels are statistically significant. McKay (1931) developed a method for constructing confidence intervals for CVs assuming the measurements are ratio type measures that are approximately normal and that the underlying population CV is less than 0.33, or 33%. Since no pair of the CVs except panels 14 and 15 in the Other stratum have confidence intervals with upper bounds less than 0.33 only those two CVs will be compared using McKay type intervals. A Bonferroni technique for non-overlapping intervals will be used to compare the intervals. The CV for the Other stratum in panel 15 is 0.2095, or 20.95% as listed in the table. A 97.5% confidence interval for this CV is (0.2043, 0.2150). The CV for the Other stratum in panel 14 is 0.2936 from Table 2 and a 97.5% confidence interval for the CV is (0.2866, 0.3010). Since the two confidence intervals at the 97.5% are non-overlapping then the Bonferroni inequality implies that the two CVs are significantly different at the 95% level. Thus there is clear evidence in the MEPS household base weights that the pps sampling is reducing the within strata variance in the household weights as intended. Despite declines in CVs for the MEPS WT in panel 15 relative to the other panels within sampling domains, the overall CV of the weights is similar in panel 15 to panel 14 (51.37 versus 52.78). This lack of an overall difference is likely attributable to variations in the mean of the weights between the strata.

The final column in Table 2 shows the CVs for PIT person-level weights. Since the CVs for these weights are not consistently lower for panel 15 than the other two panels, it is not clear if the reduction in variance in the DU weights described above will be reflected in person-level estimates.

Evaluation of PIT Insurance Estimates

The estimate of Percent Insured from the variable INSRD13X was calculated for the PIT files for 2007, 2008 and 2009. This was done for each sampling strata as well as for the overall sample at the person level. Since all households in the Asian domain were selected with certainty, the pps sampling had no effect (it is in the table for completeness but will not be further mentioned).

In the Hispanic domain the CVs of the estimates are 1.85%, 2.42%, and 1.94% for panels 13, 14, and 15 respectively. The panel 15 CV is lower than the panel 14 CV but slightly higher than the panel 13 CV. The design effects of the estimates are 3.68, 5.88, and 3.43. In the Hispanic domain the panel 14 CV and design effect were both much larger than either panel 13 or 15 but the differences between panels 13 and 14 appears to be negligible.

In the Black domain the largest CV is 1.37% in panel 15 compared to 1.16% in the other two panels. The design effects for panels 13 and 14 are 2.18 and 2.01 compared to 2.42 in panel 15. Although the pps sampling appears to have been least effective in the Black domain, a large proportion of the stratum was sampled (85%) and therefore the pps sampling could only have a limited impact.

In the Other domain the panel 15 CV was 0.71% as compared to 0.75% for panel 13 and 0.73% for panel 14. The design effects for panels 13 and 14 are 2.61 and 3.07 compared to a design effect of 2.27 for the Other domain in panel 15. Given that panel 15 had a smaller sample than the other two panels and slightly smaller CV as well as smaller design effect it might be considered that the pps sampling could have had an effect in this domain. It is not clear theoretically if the McKay interval applies to the CV of a percent from a complex sample, but if it does then the panel 15 CV would be conservatively significantly different from the panel 13 CV at the 0.05 level and conservatively significantly different from the panel 14 CV at the 0.10 level.

Looking at the overall estimates, the CV for panel 15 is 0.64% compared to 0.70% for the other two panels. Again this is based on a smaller sample size of almost two to three thousand persons. If the McKay interval does apply to this situation then the overall panel 15 CV is significantly different from both of the other panels at a conservative 0.05 level. The overall design effect for panels 13 and 14 are 4.12 and 4.26 compared to a design effect of 3.01 for panel 15.

Conclusion

The pps sampling was introduced to reduce the variability of the weights and there is evidence from the panel 15 data that the reduction did occur within the sampling strata. However, there are other factors besides the variability of the weights that affect the variability of the MEPS estimates. It was hoped that reducing the variability of the weights would play some part in reducing the variability of the MEPS estimates. At this point there is at best mixed evidence from early data from the Point in Time file that the pps sampling had any substantial beneficial impact on the overall variability of the insurance estimates. However, the overall CV% for panel 15 was significantly smaller than the other two panels and the design effect for panel 15 was less than 75% of the other two design effects. Note that the estimates from the PIT on insurance have relatively small variability compared to expenditure estimates so it remains to be seen if the pps sampling can be demonstrated to have an impact on variability of other estimates. Expenditure data for the first year of panel 15 (2010) are scheduled for release in the fall of 2012 which will enable further evaluations of the effect of PPS on MEPS estimates.

References

Machlin, S.R., Yu W., and Baskin, R.M. (2009) Evaluation of PPS sampling for the Medical Expenditure Panel Survey *ASA Proceedings of the Joint Statistical Meetings*, 2389-2395

McKay, A.T. (1931) Distribution of the Estimated Coefficient of Variation, *Journal of the Royal Statistical Society*, 94, #4, 564-567

Table 1
MEPS Households selected from NHIS for panels 13, 14, and 15

	Panel 13			Panel 14			Panel 15		
	NHIS 2007	MEPS sample size	Percent Selected	NHIS 2008	MEPS sample size	Percent Selected	NHIS 2009	MEPS sample size	Percent Selected
Asian/Poor[1.]	2,632	2,632	100%	834	834	100%	860	860	100%
Hispanic	1,855	1,855	100%	2,386	2,066	87%	2,579	1,961	76%
Black	1,498	1,498	100%	2,019	1,816	90%	1,994	1,705	86%
Other	6,467	3,718	57%	6,942	4,984	72%	6,957	4,224	61%
Total	12,452	9,703		12,181	9,700		12,390	8,750	

1. Poor only applies to panel 13.

Table 2
CV% of Weights Associated with NHIS and MEPS

		Household Weight		Person Weight
	Panel	WTFA	MEPS WT	WGTSP13
Asian HHs				
	13	50.68	50.40	65.50
	14	46.74	46.56	59.62
	15	45.47	44.84	56.81
Hispanic HHs				
	13	42.82	42.59	61.84
	14	47.96	47.90	60.19
	15	48.46	33.40	57.00
Black HHs				
	13	37.58	37.35	52.98
	14	44.82	44.32	47.70
	15	36.08	29.26	50.41
Other HHs				
	13	30.66	34.04	48.95
	14	29.97	29.36	48.16
	15	31.25	20.95	46.92
All HHs				
	13	48.83	68.11	78.37
	14	46.56	52.78	67.49
	15	46.66	51.37	69.88

Table 3
MEPS, PIT person - Percent Insured
(INSRD13X)

Panel		N	Mean	Stderr	CV%	Design Effect
Asian						
	13	1406	0.82	0.0144	1.76	3.15
	14	1386	0.80	0.0148	1.85	3.38
	15	1358	0.81	0.0186	2.29	3.66
Hispanic						
	13	6327	0.63	0.0116	1.85	3.68
	14	5124	0.66	0.0159	2.42	5.88
	15	4731	0.65	0.0126	1.94	3.43
Black						
	13	4273	0.78	0.0090	1.16	2.18
	14	3605	0.79	0.0092	1.16	2.01
	15	3454	0.78	0.0107	1.37	2.42
Other						
	13	8603	0.84	0.0063	0.75	2.61
	14	9076	0.86	0.0063	0.73	3.04
	15	7703	0.85	0.0061	0.71	2.27
All						
	13	20609	0.80	0.0056	0.70	4.12
	14	19191	0.82	0.0058	0.70	4.26
	15	17246	0.81	0.0052	0.64	3.06