A Comparison of Taylor Linearization and Balanced Repeated Replication Methods for Variance Estimation in Medical Expenditure Panel Survey

Sadeq Chowdhury

Agency for Healthcare Research and Quality Working Paper No. 13004

July 2013

Suggested citation: Chowdhury S. A Comparison of Taylor Linearization and Balanced Repeated Replication Methods for Variance Estimation in Medical Expenditure Panel Survey. Agency for Healthcare Research and Quality Working Paper No. 13004, July 2013, http://gold.ahrq.gov.

A Comparison of Taylor Linearization and Balanced Repeated Replication Methods for Variance Estimation in Medical Expenditure Panel Survey

Sadeq R. Chowdhury

ABSTRACT

For computing sampling variances of the Medical Expenditure Panel Survey (MEPS) Household Component estimates, the Taylor linearization method is generally used. The MEPS public use files include variance strata and cluster identifiers to facilitate variance computation using this method. Also a file containing a BRR replication structure (in the form of a set of half sample indicators) is also made available so that the users can form BRR replicate weights from the final MEPS weight to compute variances of MEPS estimates using either BRR or Fay's modified BRR (Fay 1989) methods. These replicate weights are useful to compute variances of complex non-linear estimators for which a Taylor linear form is not easy to derive and not available in commonly used software. However, the BRR replicates derived from the final weight represent a shortcut approach because the replicates are not produced starting with the base weight and all adjustments made in different stages of weighting are not applied independently in each replicate. So the variances computed using the one-step BRR do not capture the effects of all weighting adjustments. The Taylor approach, as implemented in most software, also does not fully capture the effects of different weighting adjustments. Of particular interest here is the effect of adjustments using external control totals which are expected to reduce the variance. To assess the effects of these adjustments on the variances of MEPS estimates, a set of proper Fay's BRR replicates were formed starting with the base weights and independently applying all weighting adjustments to each of these replicates. Variances of selected MEPS estimates were then computed from these properly created replicate weights and compared with those of the linearization method. This report presents the results of this comparison which shows that variance estimates are generally lower under the BRR method indicating that the net impact of various weighting adjustments on variance is generally downward. In most cases, relative standard errors based on BRR are at least 5-10% lower than that of Taylor approach. That means the variance estimates produced by the current recommended method are conservative and can be considered as compensating for the imputation variance and the variability in the control totals, which are unaccounted for under both methods of variance estimation.

Sadeq Chowdhury
Center for Financing, Access, and Cost Trends
Agency for Healthcare Research and Quality
540 Gaither Road
Rockville, MD 20850
Phone: (301)427-1666
Email: Sadeq.chowdhury@ahrq.hhs.gov

A Comparison of Taylor Linearization and Balanced Repeated Replication Methods for Variance Estimation in Medical Expenditure Panel Survey

1.0 Introduction and Background

The Medical Expenditure Panel Survey (MEPS), conducted by the Agency for Healthcare Research and Quality (AHRQ), provides nationally representative estimates of health care use, expenditures, sources of payment, and health insurance coverage for the U.S. civilian noninstitutionalized population. The MEPS Household Component (which will be generally referred to as MEPS hereafter) is a complex national area probability sample survey. The details of the MEPS sample design can be found in Ezzati-Rice et al. (2008). Estimates from the MEPS are produced using the weight calculated based on the base sampling weight and a series of nonresponse and poststratification/raking adjustments. The details of the MEPS weighting procedures can be found in Machlin, Chowdhury, et al. (2010).

For computing variances of MEPS estimates, the Taylor Series Expansion (TSE) linearization approach is generally used. As long as the data file includes identification for variance strata and cluster, the TSE method is very easy to apply. A drawback of the TSE method is that linear forms are not easy to obtain and the method is not easy to apply for many complex estimates. Replication-based variance estimation methods (Wolter, 1975) do not have this drawback. Once the replicate weights are computed, the variances of all forms of estimators can be computed. Therefore, in addition to including MEPS variance strata and cluster to facilitate variance computation using the TSE method, a file containing a BRR replication structure (in the form of a set of half sample indicators) is also produced for MEPS. These half-sample replicate indicators can be used to compute BRR or Fay's BRR replicate weights from the final MEPS weight. However, the BRR weight computed in a single-step from the final weight is a shortcut approach because ideally the replicate weights should be produced in multiple steps starting with the base weight and all adjustments made in different stages of weighting should be applied independently in each replicate. The variances computed using the one-step BRR do not capture the effects of all weighting adjustments. The Taylor approach, as programmed in most commercial software, also does not capture the effects of poststratification and raking adjustments made in different stages of weighting on variance. To assess the variance effects of these adjustments on MEPS estimates, a set of proper Fay's BRR weights are computed starting with the base weight and applying all subsequent adjustments for 2008 MEPS full year weights. This report presents a comparison of MEPS variances estimates using TSE and the Fay's BRR methods.

1.1 Medical Expenditure Panel Survey (MEPS)

The MEPS is a set of large-scale surveys of families and individuals, their medical providers (doctors, hospitals, pharmacies, etc.), and employers across the United States. The MEPS provides estimates of specific health services use by the U.S. civilian noninstitutionalized population, the payments for these services, sources of payment, and the cost and scope of health insurance of U.S. workers. The MEPS has three components: the Household Component, Medical Provider Component and the Insurance Component. The Household Component collects data from individual households and their members in selected communities across the United States, drawn from a nationally representative subsample of households that participated in the prior year's National Health Interview Survey (NHIS) conducted by the National Center for Health Statistics. The data collected from households are supplemented by data from their medical providers collected in the Medical Provider Component. The Insurance Component is a separate survey of employers that provides data on employer-based health insurance.

The MEPS Household Component collects detailed information for each person in the household on demographic characteristics, health conditions, health status, use of medical services, charges and source of payments, access to care, satisfaction with health care, health insurance coverage, income, and employment. The panel design of the survey, which features five rounds of interviewing covering two full calendar years, makes it possible to determine how changes in individuals' health status, income, employment, eligibility for public and private insurance coverage, use of services, and payment for care are related.

A new panel is sampled for MEPS every year from the previous year's responding households of the NHIS and remains in the sample for two years. A MEPS annual file consists of two overlapping sample panels—the first year of a new panel and the second year of the previous panel. The survey can be used to produce national estimates for persons and families as well as subgroups of the population.

1.2 Weighting of MEPS Sample

Each MEPS panel is weighted separately for different rounds of nonresponse and coverage (poststratification/raking) adjustments until the final step when the two panels are combined and a raking adjustment is applied to the combined panels to produce the final full year (FY) weight. Machlin, Chowdhury, et al. (2010) provides details of the weighting and estimation procedures used in the MEPS.

The weighting of the most recent panel starts with computing the dwelling unit (DU) base weight, which is calculated by starting with the nonresponse adjusted NHIS household weight. A poststratified ratio adjustment is then applied to the DU base weight to ensure representativeness of the MEPS sample in terms of the full NHIS sample. The control total for this adjustment is derived from the household reference person's weight in the NHIS sample. A nonresponse adjustment is applied to the poststratified DU weight to compensate for the DU nonresponse to the Round 1 interview. A family-level weight is derived by assigning the DU weight to each family within the DU and then a family-level poststratification adjustment is applied using control totals from the CPS. The Round 1 person weight is then derived by assigning the poststratified family weight to each person in the family and then applying a person-level poststratification adjustment.

The year 1 person weight is derived by first applying a nonresponse adjustment to the Round 1 weight for person-level nonresponse over Round 2 and Round 3 and then applying a raking/poststratification adjustment using the control total for December 31 of the year derived from the subsequent March CPS. This produces the year 1 person weight for responding persons in the most recent panel.

The 2nd year weight for the persons in the preceding panel is derived by starting with the year 1 weight from the previous year and applying a nonresponse adjustment to compensate for nonresponse in year 2. A raking/poststratification adjustment is then applied to the nonresponse adjusted weight. This produces the year 2 weight for responding persons in the preceding panel.

The two panels are then put together to create the FY file for the current year. The panel specific annual weights are scaled down by applying a compositing factor proportional to the sample size in each panel so that the composite weights of both panels jointly add up to the size of the annual target population. Then a raking adjustment is applied to the composite weight in the combined panels using the same set of control totals used for raking of individual panels to produce the preliminary FY weight. When the processing of poverty data is completed later, the final FY weight is produced by applying another round of raking adjustment by adding dimensions involving poverty status to the set of raking dimensions used earlier.

2.0 Variance Estimation

2.1 Variance Estimation Methods for Complex Surveys

A review of commonly used approaches to design-based estimation of the variances of estimates from complex survey data can be found in Wolter (2007) and Shao (1996). The variances of estimates from complex multi-stage surveys are usually obtained by using approximate techniques because the exact forms or properties of such variance estimators are generally complicated. Variances of estimates from complex surveys are generally obtained by either TSE linearization method or by using some replication based methods such as balanced repeated replication (BRR), Fay's BRR, jackknife, and bootstrap methods. The two most commonly used variance estimation methods for complex survey data are TSE and BRR methods.

Under the TSE method, the linear form of a non-linear estimator is derived by using the first-order Taylor series expansion and then the standard formula for variance estimation from complex surveys is used to compute the variance. Approximating a non-linear estimation by a linear function based on the Taylor expansion introduces a bias into the variance estimator but typically such estimators are consistent. The commonly used software for variance estimation from complex surveys are SUDAAN, Stata, and SAS Survey procedures which are used to compute variances using the TSE approach for commonly used non-linear estimators. However, for complicated non-linear estimators the linear form is not easy to derive and software packages do not offer any readily available option for computing the variances of such estimates. When various weighting adjustments are applied to the base sampling weight, the estimator becomes more complicated and the TSE method does not fully capture the variance effects of such adjustments as the linear form is usually derived for the base estimator without the weighting adjustments. So the variance obtained using TSE method does not capture the variance reduction effect of poststratification or raking adjustments.

BRR is a half-sample replication method applied to sample designs where either exactly two sample units or clusters are selected with replacement from each stratum or the sample units within a stratum can be grouped into two clusters. Under this method a series of half sample replicates are formed by taking one unit from each of the strata. Since some replicates include common units they become correlated. However, by using the approach of balancing introduced by McCarthy (1966, 1969), the replicates are formed in a way so that an unbiased estimator of variance can be obtained under the fully orthogonal balanced replicates. Using these balanced replicates, a set of replicate sampling weights are computed by doubling the weights of the units in the replicate and zero-weighting the units not in the replicate. The variance is then computed using these replicate weights and the full sample weight.

The estimated variance, $v_{BRR}(\hat{\theta})$, of an estimate, $\hat{\theta}$, under the BRR method can be expressed as

$$v_{BRR}(\hat{\theta}) = \frac{1}{R}\sum_{r=1}^{R}\left(\hat{\theta}_r - \hat{\theta}\right)^2$$

where R is the number of replicates, $\hat{\theta}$ is the estimate of θ from the full sample, $\hat{\theta}_r$ is the estimate of θ from the r-th replicate.

Fay's method (Fay 1989) is also a BRR method but with a slight variation. As described above, in the BRR method replicate weights are calculated by multiplying the weights of half of the sample units by two and the other half by zero. The problem with this approach is that in computing estimates for a small domain the sample size may become very small because half of the sample units are zero weighted. This problem is avoided under Fay's method by multiplying the weights of half of the sample units by a perturbation factor f $(0 \leq f < 1)$ and the other half by $(2 - f)$. Thus none of the units are excluded from any replicate and hence the sample size in each replicate remains the same as the full sample.

Under the Fay's method, the BRR variance formula presented above is modified by dividing with $(1 - f)^2$ as follows

$$v_{Fay}(\hat{\theta}) = \frac{1}{R(1-f)^2} \sum_{r=1}^{R} \left(\hat{\theta}_r - \hat{\theta}\right)^2$$

Usually, the replicate weights are computed starting with the base weight (inverse of the selection probability) and then all subsequent weighting adjustments are applied to each replicate independently. That is why these methods capture the variances due to all adjustments made to the base sampling weight. In some cases, to save computational steps, a single-step BRR or Fay's BRR weights are derived by using the final full sample weight rather than starting with the base weight. In that case, similar to the TSE approach, the BRR method also does not completely reflect the effects of subsequent weighting adjustments. This approach can produce a substantially biased estimate of variance in some situations (Lemeshow, 1979).

Although it is labor-intensive to compute the replicates, some advantages of the BRR method over TSE are that it can be applied to any complex estimator and is convenient and simple for domain analyses. Moreover, properly formed replicate weights can capture the variance due to various adjustments to the base weight. Once the replicate weights are computed, all sample design and estimation steps are captured in the replicates that can be used to produce variances using the same variance formula for any estimate.

2.2 Variance Estimation in MEPS

The MEPS data are collected using a complex multistage sample design that involves stratification, clustering, and unequal selection probabilities. To obtain accurate estimates of the variances of MEPS estimates for either descriptive statistics or more sophisticated analyses based on multivariate models, the MEPS sample design complexities must be taken into account using special analysis approaches developed for complex surveys. The TSE method is generally used for computing variances of MEPS estimates. The MEPS annual public use files include the two necessary sample design variables for implementing this method. These variables identify the variance estimation strata (VARSTR) and variance estimation clusters (VARPSU). Specifying these variables in conjunction with assuming a 'with replacement' design in software packages that employ the TSE approach (e.g. SUDAAN, STATA, or SAS Proc Survey procedures) will produce acceptable variances of MEPS estimates.

As mentioned before, because it can be extremely difficult to use the TSE method to calculate the variances of complex estimators not readily available in complex survey software packages (e.g., two-part models of health expenditures, the ratio between two medians), a linkage file containing a BRR replication structure (in the form of a set of half sample indicators) is also available for variance estimation. The half sample indicators (1 and 0) can be used to form single-step 'shortcut' BRR replicate weights to compute variances of MEPS estimates using either BRR or Fay's BRR methods.

To facilitate analysis of subpopulations and/or low prevalence events, it may be desirable to pool together more than one year of MEPS-HC data to yield sample sizes large enough to generate reliable estimates. MEPS-HC samples from year to year are not completely independent because households are drawn from the same sample geographic areas and many persons are sample respondents for two consecutive years (see MEPS-HC Methodology Reports for more details at http://www.meps.ahrq.gov). Despite this lack of independence, it is valid to pool multiple years of MEPS-HC data and keep all observations in the analysis because each year of MEPS-HC is designed to be nationally representative. However, to obtain appropriate standard errors when pooling years of MEPS-HC data, it is necessary to ensure a variance structure that consistently specifies MEPS geographic sampling units across years.

Starting in 2002, the annual MEPS public use files were released with a common variance structure that allows users to seamlessly pool annual files from 2002 onward. Prior to 2002, however, each annual MEPS public use file was released with a variance structure unique to the particular MEPS sample in that year. Therefore, when one or more years of data being pooled precede 2002, it is necessary to obtain a common variance structure. A pooled estimation linkage file is produced to enable pooling with years prior to 2002 and estimation using the Taylor Series method. This file provides a common variance structure (i.e., consistent specification of MEPS geographic strata and primary sampling units) across all years since the inception of MEPS in 1996. In addition, the BRR replicates file provides standardized replicates across all panels to facilitate appropriate BRR variance estimation from pooled data.

2.3 Effect of Raking Adjustment on Variances

In addition to mitigating bias for coverage error, a poststratification/raking[1] adjustment reduces variances of survey estimates to the extent the relevant target variable is correlated with raking variables. A raking adjustment is essentially a model-based adjustment where a survey variable can be viewed as a dependent variable and configurations of raking dimensions with appropriate main effects and interactions of raking variables as independent variables. Since a single weight is derived using a raking adjustment for a survey sample, effectively the same adjustment model is used for all survey variables. The effectiveness of a raking adjustment for a particular target variable depends on the explanatory power of the raking model for that variable. Hence, a raking adjustment can be effective to varying degrees for different survey variables. When a variable is used in a raking adjustment with known control totals, variances of the estimates of that variable at the level of raking dimensions become zero because the estimates are adjusted by benchmarking to known population values. For example, if population totals in cells of poverty status x census region are known and used as a raking dimension then the survey variances of the estimates of population totals in these cells will be zero. The variances of the estimates of other target variables which are related to the poverty status will also decrease depending on the correlation of a target variable with poverty status. Generally, weighting adjustments increase the variances of estimates through increased variation in weights. However, model-based adjustments such as poststratification and raking can reduce the variance even after increasing the variation in weights. This reduction in variance is not captured by the TSE method of variance estimation. The TSE method captures the increase in variance due to the increased variation in weights but fails to capture the reduction in variance due to the use of control totals. In contrast, the BRR method captures the reduction in variance due to adjustments with control totals.

[1] Since poststratification is a particular case of raking, both poststratification and raking will be referred to as raking generally.

3.0 Comparison of Variance Estimates

3.1 Methods

For comparison, variance estimates of selected MEPS estimates are produced using TSE and BRR methods from the MEPS 2008 FY file. To produce variance estimates using the TSE method, the usual variance strata (VARSTR) and cluster (VARPSU) included in the MEPS PUF are used. For the BRR method, a set of Fay's BRR replicate weights were developed for Panels 12 and 13. The development of replicate weights started by creating replicates using Round 1 MEPS initial DU weight i.e., the MEPS base weight. A total of 128 replicates were formed using the 125 variance strata and two PSUs in each stratum provided in 'HC-036BRR: 1996-2008 Replicates for Variance Estimation File' for computing replicate weights using the single-step BRR approach. Similar to the formation of HC-036BRR replicates, a Fay's factor of $f=0.5$ is used to form the replicates. Then all subsequent weighting adjustments applied to the full sample weight were also applied to each replicate for both panels separately until the panels were combined to create the final FY weight. The replicate weights computed at different weighting steps were stored to enable the comparisons at different stages of weighting.

Using these replicate weights and the full sample weight, estimates of variances for selected MEPS estimates are produced under both TSE and BRR methods using SAS. Variance estimates under both methods were produced and compared at several important stages of MEPS weighting using the weight from each stage to examine the impact of various adjustments up to that stage on the variances. The stages of weighting at which comparisons are made are:

DU Base Weight (DUWT1) – Round 1 DU initial weight before any MEPS nonresponse or benchmarking adjustment is applied. However, this weight includes a simple trimming of extreme weights that is done to reduce the impact of large weights on MEPS estimates. At this stage, both methods are expected to produce very similar estimates of variances.

DU Final Weight (DUWTF) – Round 1 DU final weight computed by applying a poststratification/raking adjustment using DU-level control totals. For Panel 12, control totals by MSA, Race/ethnicity, income, employment status, health status and health insurance coverage status for the DU or the reference person of the DU obtained from the NHIS were used. For

Panel 13 control totals on raking dimensions involving age, region, MSA, race/ethnicity, marital status and education obtained from the CPS were used.

Final Round 1 Person Weight (PNFWT) – Round 1 person weight derived from the DU weight with a poststratification with control totals on dimensions involving region, race/ethnicity, sex and age obtained from the CPS.

FY Panel Specific Weight (FYWT) – FY weight after nonresponse and raking adjustment for each panel separately, the nonresponse adjustments include adjustment for year 1 nonresponse for Panel 13 (the new panel) and adjustments for both Years 1 and 2 nonresponse for Panel 12 (the older Panel) and raking adjustment to CPS control totals on dimensions involving region, race/ethnicity, sex and age.

Final FY Weight (FYWTF) – FY weight after combining the two panels and applying a final raking adjustment to the composite weight of both panels using control totals on dimensions involving poverty status, region, MSA, race/ethnicity, sex and age from the CPS.

For comparison of variance estimates, the cases that are in-scope on 12/31/08 with positive final FY weight and also positive initial Round 1 DU weight are included. The nonresponding cases may have positive weight for earlier round(s) but had to be excluded from the analysis as they are not in the FY file and do not have utilization and expenses data.

The variances in terms of relative standard error expressed as percentages (RSE%) are compared for selected estimates of proportions using categorical variables and also for estimates of means of selected expense variables. Variances are compared using weights at different weighting stages to see impacts of nonresponse and raking adjustments on TSE and BRR estimates.

3.2 Results

Explanatory Power of the Final Raking Model

In this section, a comparison of variance estimates of selected MEPS estimates at different stages of weighting computed using TSE and BRR methods is presented. However before presenting the comparison, an idea of explanatory powers of the final raking model for different survey variables is provided. As discussed in Section 2.3, the effectiveness of a raking adjustment and its impact on variances of estimates depends on correlations between survey variables and variables used in the raking model. In MEPS weighting, raking/poststratification adjustments are done in different weighting stages but we will concentrate on the explanatory power of the final raking adjustment as the final adjustment is likely to have a major impact. Table 1 presents explanatory powers of the raking dimensions used in the final raking adjustment (that produces the weight FYWTF as discussed above) for different target variables used in the analysis. The R^2 values are obtained from fitting regression models with the target variable as the dependent variable and configurations of raking dimensions as independent variables. Generally, the raking model appears to be more effective for categorical variables than continuous variables. This may be because most of the continuous variables are expense variables which are harder to model. Annual expenses for a particular type of service (e.g., prescribed medicines, office-based visits) can vary widely within an adjustment cell which reduces the effectiveness of the model. The raking model is also less effective for those categorical variables which are not applicable to majority cases such as 'ER visits', 'poor mental health', and 'unable to get health care'. This may be because the raking adjustment is done at the overall level and not conditional on nonzero expenses. The presence of many zero values in a target variable can reduce the correlation with the raking variables. On the other hand, the raking model is more effective for some continuous variables which are directly related to any raking variable. For example, R^2 is high for 'wage income' as it is directly related to poverty status which was used in the raking model. Among other continuous variables, R^2 for prescription expense and out-of-pocket expense are relatively high. These variables are not rare and may also be somewhat related to poverty status. The decrease in RSE of an estimate attributable to the final raking adjustment is expected to depend on the R^2 of the raking model.

Table 1. Explanatory power (R^2) of the final raking model for different target variables

Categorical Target Variables		Continuous Target Variables	
Variable	R^2	**Variable**	R^2
Uninsurance Status	12.8%	Total Expense	8.3%
5+ Office-based Provider visits	15.2%	Office-based Expense	6.4%
1+ Outpatient Visits	8.6%	Prescription Expense	10.8%
1+ Inpatient Stays	4.01%	Out-of- pocket Expense	9.2%
1+ ER Visits	2.3%	Outpatient Expense	1.9%
5+ RX - Prescription	25.3%	Inpatient Expense	2.4%
Daily Activity Limitation	9.3%	Emergency Room Expense	2.3%
Any Limitation	20.3%	Wage Income	50.8%
Poor Health	8.7%	Dental care expense	3.4%
Poor Mental Health	3.8%		
Unable to Get Healthcare	4.5%		

Comparison of RSEs for Estimates of Categorical Variables

Table 2 presents the comparison of RSEs under TSE and BRR methods for several estimates of proportions often produced from the MEPS. It shows that the estimates under both methods are very close at the initial stage of weighting i.e., when using DU initial weight before making any nonresponse or poststratification/raking adjustments. The ratios of RSEs of BRR over TSE are mostly less than 103%. However, as different weighting adjustments are made at subsequent stages, RSEs under the TSE method increase or remain the same while RSEs generally decrease rapidly under the BRR method. With the final FY weight, the ratios of BRR over TSE estimates of RSEs are considerably less than 100% for most estimates. For example, for the estimate of proportion uninsured, the RSEs are 3.27 under TSE and 3.28 under BRR with the initial DU weight, while the RSEs are 3.16 under TSE and 2.21 under BRR with the final FY weight (Figure 1). The ratio of RSEs is about 100% with the DU initial weight but decreases to 70% with the final FY weight. The RSE under BRR is considerably lower for the final estimate.

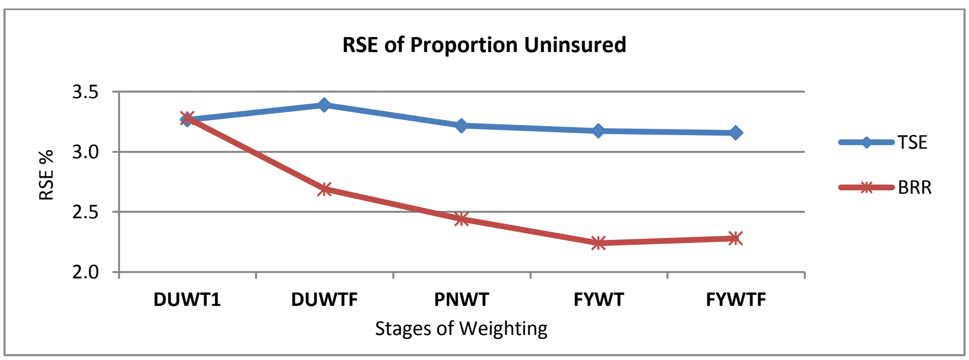

Figure 1. Comparison of RSEs under TSE and BRR for estimates of proportion uninsured

Just to illustrate the impact of raking, the comparison of RSEs for estimates of poverty status, which is used as a raking variable, is presented in Figure 2. The RSEs under TSE remain almost the same over different stages of weighting but decrease under BRR with a sharp drop at the final stage when the weights are raked for poverty status. The ratios of RSEs of BRR over TSE at the final stage are very close to zero and range only from 14% to 25%. Ideally, these RSEs under BRR would be zero as the raking adjustment is done at that level but it deviated slightly from zero as collapsing of some raking cells had to be done in the raking procedure for insufficient cell sizes. This shows how the impact of raking on variance is captured by the replication method but ignored by the TSE method. RSEs under the BRR method decreased with poststratification/raking adjustments at various stages but increased slightly when TSE is used.

If a variable itself is used in one of the raking dimensions then the raking adjustment for that variable will be the most effective and the RSE of the estimate of that variable at the level of raking dimension should be zero. Since the poverty status is a variable in the raking adjustment, the RSEs under BRR for poverty status estimates are close to zero. However, the health insurance status is not used in raking adjustment but related to other variables used in the raking model such as poverty status, race-ethnicity, etc., the RSEs under BRR decreases but not as much as in the case of poverty status. For the TSE method, the variance remained stable or increased because while the benefit of poststratification/raking in terms of variance reduction is not captured, the negative impact of increased variation in weights due to different adjustments on variance is captured.

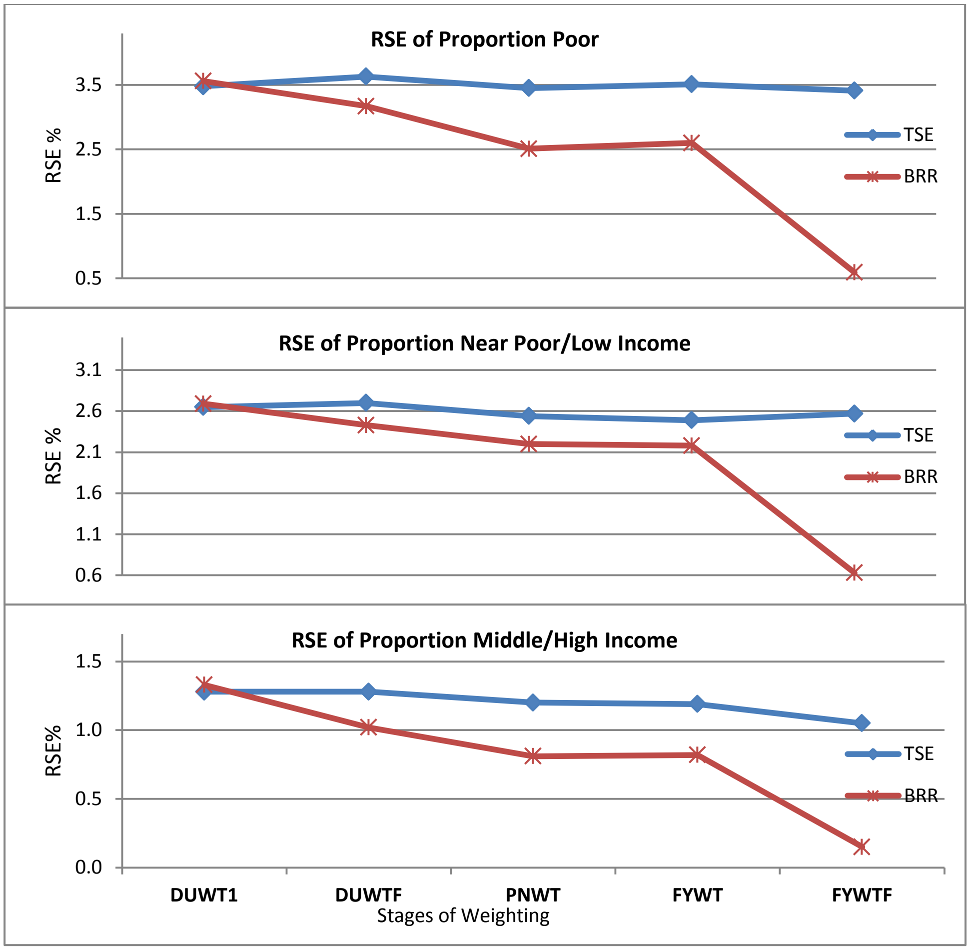

Figure 2. Comparison of RSEs under TSE and BRR for estimates of proportions in different poverty categories

Except for poverty status, the pattern of differences in RSEs between TSE and BRR are very similar for the estimates presented in Table 2. For estimates of proportions with Daily Activity Limitation, 1+ Inpatient Discharge, 1+ Outpatient Visit, 1+ ER Visit, Any Limitation, 5+ Office Visit and 5+ RX prescription, Figure 3 shows all RSEs under TSE tend to remain flat or go up while the RSEs under BRR go down. Ratios of RSEs range from 72% to 96% at the last stage when the final FY weight is used. Again, the decreases in RSEs under BRR are roughly proportional to explanatory powers of the raking model for different variables.

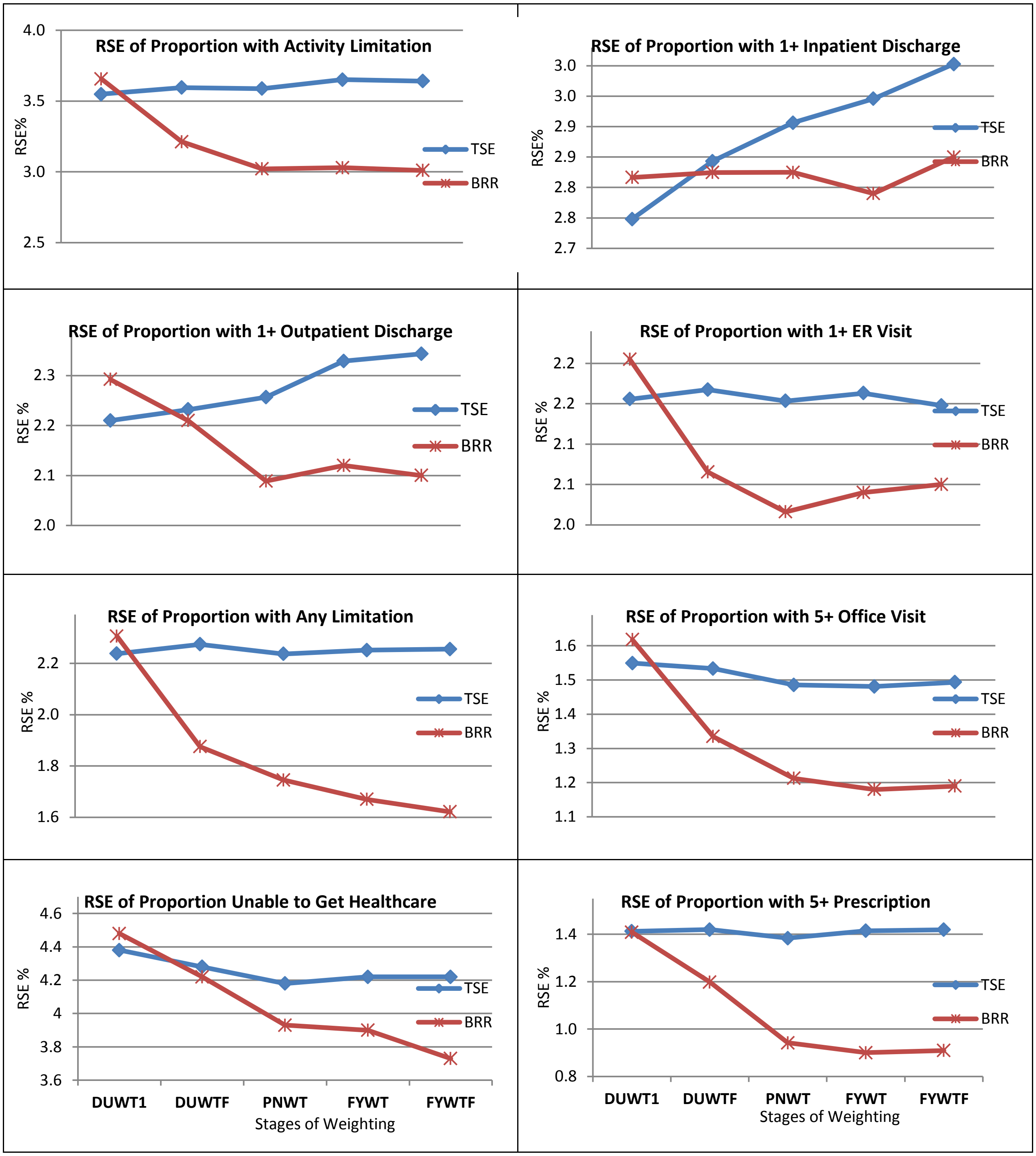

Figure 3. Comparison of RSEs under TSE and BRR for various estimates

Comparison of RSEs for Estimates of Continuous Variables

Table 3 presents the comparison of RSEs under TSE and BRR methods for estimates of means of selected healthcare expenditure variables. For this comparison, the pattern of difference between TSE and BRR are mixed and not as pronounced as in the case of categorical variables. This is because expense variables are harder to model and the raking model used has less explanatory power for these variables (see Table 1). Figure 4 shows the comparison of RSEs between TSE and BRR for estimates of mean expenses for prescription medicines, office-based visits, total expenditures, and out of pocket expenditures, which are more correlated with the raking variables and the raking adjustment is more effective for these variables then other variables. For these estimates, although less pronounced, the pattern is somewhat similar to that of categorical variables. The RSEs under TSE broadly remain at the same level while the RSEs under BRR consistently decrease for most estimates over different stages of weighting (Figure 4). The final ratio of RSEs ranges from 87% for prescription expenses to 94% for office-based expenses.

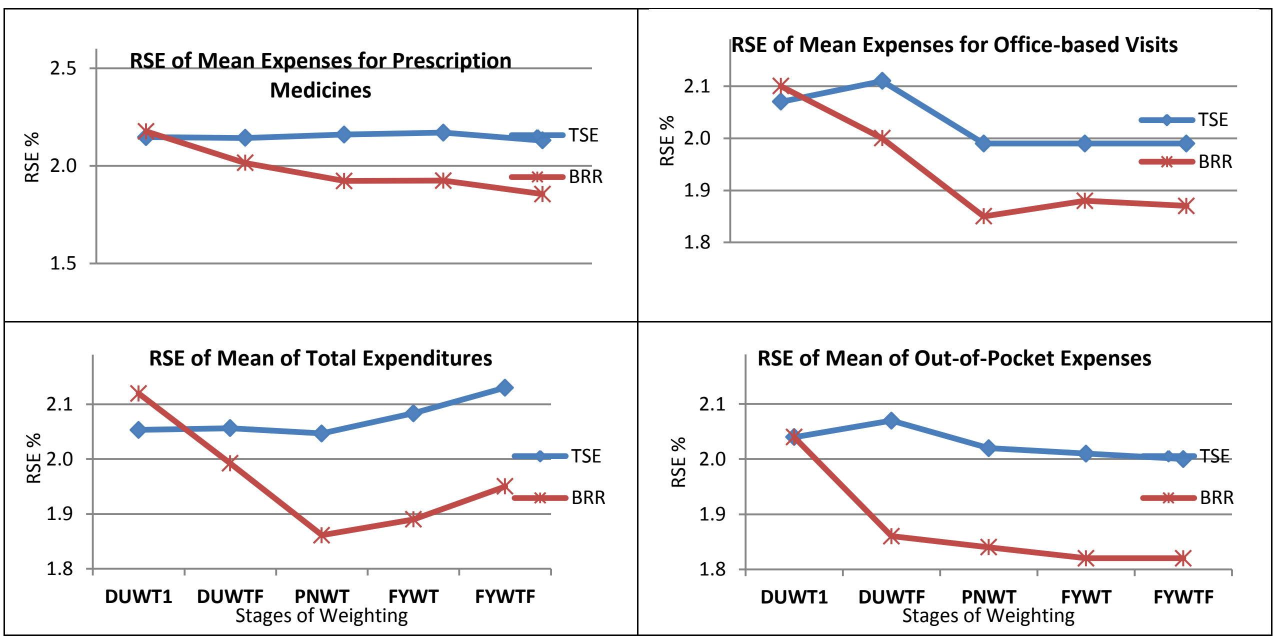

Figure 4. Comparison of RSEs under TSE and BRR for estimates of mean expenses those are more correlated with raking variables

Figure 5 shows a comparison of RSEs under TSE and BRR for estimates of means of some expense types (e.g., outpatient, inpatient, and emergency room visits) for which the raking model is less effective (see Table 1). Here, RSEs appear to go up under both TSE and BRR with the progress of weighting adjustments. The BRR estimates do not show any impact of raking on variance reduction. The expenses for the events like ER visit or outpatient or inpatient are not very correlated with the variables used in raking adjustments (R^2 less than 3% in all cases as shown in Table 1). Consequently, the adjustments increased variation in weights without any gain from raking adjustments.

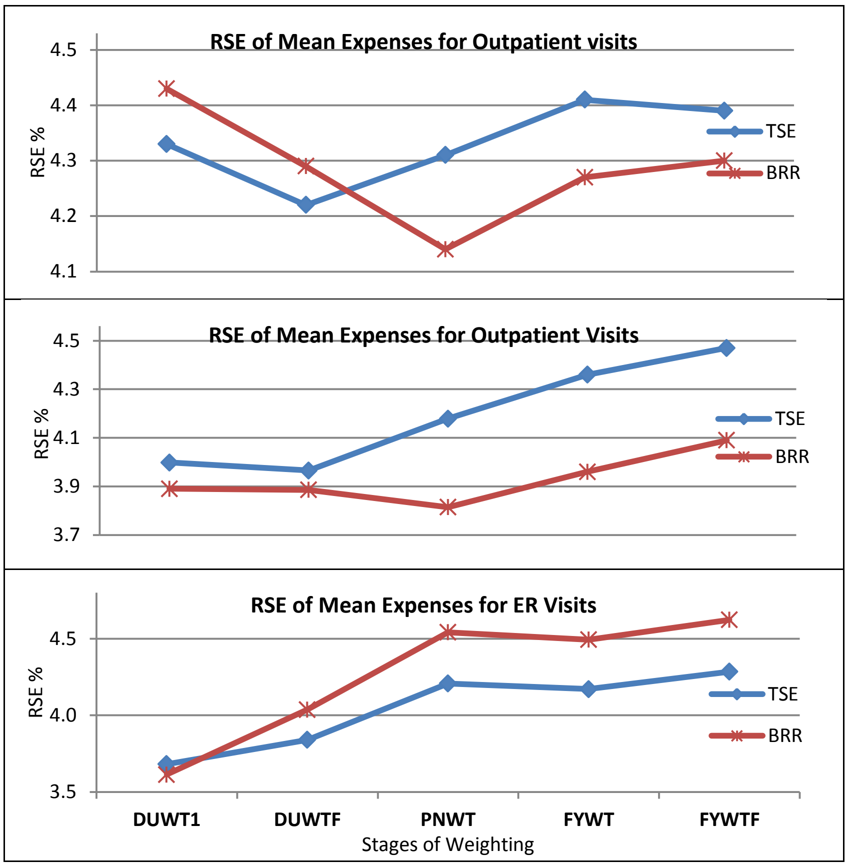

Figure 5. Comparison of RSEs under TSE and BRR for estimates of mean expenses those are less correlated with raking variables

Impact of Outliers on TSE and BRR

In the above comparison of the RSEs of mean expense of outpatient visit, an influential observation (large weight with extreme total expenditure) has been excluded. Table 4 and Figure 6 show the impact of this observation on TSE and BRR methods. With the outlier included (_O), RSEs with DU initial weight is 7.19 under TSE while the same is 9.42 under BRR. For all the estimates we discussed above, generally there are agreements between TSE and BRR estimates of RSEs at least at the beginning but in presence of this outlier the difference in variance estimates is wide even at the initial stage. At subsequent stages of weighting, with outlier included, RSEs under both TSE and BRR decreased and with the final FY weight, the RSE is 6.86 under TSE and 8.24 under BRR; the final ratio of RSEs of BRR over TSE is 120%, while this is less than 100% for almost all estimates compared. With the outlier observation excluded (_T), the RSEs with the initial DU weight under both methods are fairly close (4.33 under TSE and 4.43 under BRR) and with the final FY weight, the RSEs under both methods are again very close (4.39 under TSE and 4.30 under BRR). The final ratio RSEs of BRR over TSE is 98%, the usual pattern observed for most other variables. This comparison shows that RSE estimates under TSE and BRR methods can be very different in presence of an outlier. Which method is computing the correct variance in presence of the outlier is unknown. This deserves further investigation which we plan to conduct in a separate study.

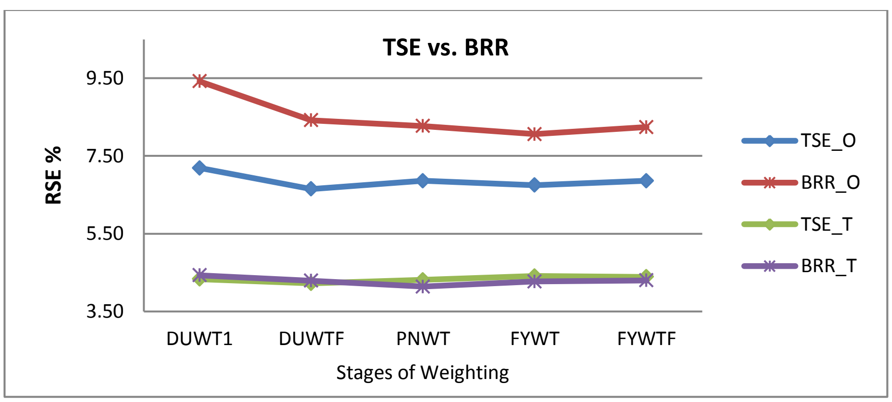

Figure 6. Comparison of RSEs under TSE and BRR with and without an outlier

Table 2. Comparison of RSEs computed using TSE and BRR methods for estimates of proportions at different stages of weighting

			RSE% of Percentage Estimates														
			Taylor Series					Fay's BRR method					Ratio of RSEs (BRR/TSE)				
Characteristic for which proportion was estimated		Category counts from total sample size of 30,287	DU initial weight	Round 1 DU final weight	Round 1 final person weight	Panel specific FY weight	Final FY weight	DU initial weight	Round 1 DU final weight	Round 1 final person weight	Panel specific FY weight	Final FY weight	DU initial weight	Round 1 DU final weight	Round 1 final person weight	Panel specific FY weight	Final FY weight
Insurance status	Private	17,271	1.23	1.24	1.19	1.21	1.11	1.26	0.93	0.74	0.72	0.58	103%	75%	62%	60%	52%
	Public	8,117	2.68	2.68	2.68	2.73	2.72	2.79	2.30	1.89	1.97	1.60	104%	86%	71%	72%	59%
	Uninsured	4,899	3.27	3.39	3.22	3.17	3.16	3.28	2.69	2.44	2.30	2.21	100%	79%	76%	72%	70%
Poverty status	Poor/neg	6125	3.48	3.63	3.45	3.51	3.41	3.56	3.17	2.51	2.60	0.59	102%	87%	73%	74%	17%
	Near poor/ Middle Inc	7148	2.65	2.70	2.54	2.49	2.57	2.69	2.43	2.20	2.18	0.63	102%	90%	87%	88%	25%
	High Inc	17014	1.28	1.28	1.20	1.19	1.05	1.33	1.02	0.81	0.82	0.15	104%	80%	68%	69%	14%
5+ Office-based Visits		7,684	1.50	1.48	1.44	1.43	1.44	1.57	1.28	1.16	1.12	1.15	105%	87%	81%	78%	80%
1+ Outpatient Visits		3,879	2.21	2.23	2.26	2.33	2.34	2.29	2.21	2.09	2.15	2.13	104%	99%	93%	92%	91%
1+ Inpatient Stays		1,931	2.75	2.84	2.91	2.95	3.00	2.82	2.82	2.82	2.77	2.82	102%	99%	97%	94%	94%
1+ ER Visits		3,988	2.16	2.17	2.15	2.16	2.15	2.21	2.07	2.02	2.06	2.06	102%	95%	94%	95%	96%
5+ RX - Prescription Medicines		10,321	1.41	1.42	1.38	1.41	1.42	1.41	1.20	0.94	0.89	0.91	100%	84%	68%	63%	64%
Daily Activity limitation		1,937	3.55	3.60	3.59	3.65	3.64	3.66	3.21	3.02	3.09	3.06	103%	89%	84%	85%	84%
Any limitation		5,604	2.24	2.27	2.24	2.25	2.26	2.31	1.88	1.75	1.68	1.62	103%	82%	78%	74%	72%
Poor health		3,516	2.44	2.47	2.44	2.46	2.52	2.36	2.29	2.32	2.30	2.27	97%	92%	95%	93%	90%
Poor mental health		1,885	3.20	3.30	3.35	3.43	3.45	3.24	3.26	3.30	3.33	3.33	101%	99%	98%	97%	97%
Delay in getting health care		1,942	3.92	3.85	3.75	3.77	3.74	4.02	3.82	3.55	3.59	3.58	102%	99%	95%	95%	96%
Unable to get health care		1,875	4.38	4.28	4.18	4.22	4.22	4.48	4.22	3.93	3.90	3.73	102%	99%	94%	92%	88%

Table 3. Comparison of RSEs computed using Taylor Series and BRR methods for expense estimates at different stages of weighting

		RSE% of Percentage Estimates														
		Taylor Series					Fay's BRR method					Ratio of RSEs (BRR/TSE)				
Mean Annual Expenses	Number of cases with expense	DU initial weight	Round 1 DU final weight	Round 1 final person weight	Panel specific FY weight	Final FY weight	DU initial weight	Round 1 DU final weight	Round 1 final person weight	Panel specific FY weight	Final FY weight	DU initial weight	Round 1 DU final weight	Round 1 final person weight	Panel specific FY weight	Final FY weight
Prescription Medicines	17560	2.15	2.14	2.16	2.17	2.13	2.18	2.02	1.92	1.92	1.86	101%	94%	89%	89%	87%
Office-based Visits	20658	2.07	2.11	1.99	1.99	1.99	2.10	2.00	1.85	1.88	1.87	101%	95%	93%	94%	94%
Total Expenditures	22703	2.05	2.06	2.05	2.08	2.13	2.12	1.99	1.86	1.89	1.95	103%	97%	91%	91%	92%
Out of Pocket	22703	1.99	2.02	1.97	1.96	1.95	1.99	1.81	1.79	1.77	1.77	100%	90%	91%	90%	91%
Dental Expenses	10727	2.47	2.47	2.48	2.43	2.39	2.65	2.57	2.44	2.40	2.38	107%	104%	98%	99%	100%
Wage Income	15980	1.39	1.38	1.38	1.41	1.40	1.51	1.19	1.17	1.13	0.96	109%	86%	85%	80%	69%
Outpatient* Visits	3878	4.33	4.22	4.31	4.41	4.39	4.43	4.29	4.14	4.27	4.30	102%	102%	96%	97%	98%
Inpatient Stays	1931	4.00	3.97	4.18	4.36	4.47	3.89	3.89	3.81	3.96	4.09	97%	98%	91%	91%	91%
Emergency Room	3988	3.68	3.84	4.21	4.17	4.28	3.61	4.04	4.54	4.49	4.62	98%	105%	108%	108%	108%

*excludes 1 outlier

Table 4. Impact of an outlier on Taylor series and BRR methods of variance estimation at different stages of weighting

		RSE% of Estimates of Mean Expense														
		Taylor Series					Fay's BRR method					Ratio of RSEs (BRR/TSE)				
Mean Expense for outpatient visit	**Number of cases with expense**	**DU initial weight**	**Round 1 DU final weight**	**Round 1 final person weight**	**Panel specific FY weight**	**Final FY weight**	**DU initial weight**	**Round 1 DU final weight**	**Round 1 final person weight**	**Panel specific FY weight**	**Final FY weight**	**DU initial weight**	**Round 1 DU final weight**	**Round 1 final person weight**	**Panel specific FY weight**	**Final FY weight**
With Outlier	3879	7.19	6.65	6.86	6.75	6.86	9.42	8.42	8.27	8.06	8.24	131%	127%	121%	119%	120%
Without 1 Outlier	3878	4.33	4.22	4.31	4.41	4.39	4.43	4.29	4.14	4.27	4.30	102%	102%	96%	97%	98%

4.0 Conclusions

This report presents a comparison of Taylor linearization and BRR methods for sampling variance estimation in MEPS. Variances are computed and compared at different stages of weighting to see the impact of different weighting adjustments on two methods of variance estimation. The comparison shows that the variance estimates from both methods are almost the same at the initial stage of weighting. But with various subsequent adjustments for nonresponse and poststratification/raking adjustments using available control totals, generally the variances computed using TSE increase and the variances computed using BRR decrease. The pattern is more pronounced for estimates which are more correlated with variables used in raking adjustments. For the MEPS FY estimates produced with the final FY weight, the variances under BRR are 5-10% lower than the variances under TSE for most estimates and variances are even more than 20-25% lower under BRR for some estimates. This reduction in variance is due to the fact that TSE does not capture the variance reduction due to poststratification/raking adjustments using known or highly accurate control totals while it inflates the variance because of the increased variation in weights due to these additional adjustments. On the other hand, BRR does capture both the increased variation in weights for additional adjustments and the reduction in variance for adjustments with external control totals.

The variables such as age, sex, race-ethnicity, marital status, census region, MSA status, and poverty status for which control total adjustments are made are correlated to different degrees with most MEPS estimates as shown by R^2 of the model used for the final raking adjustment. The extent of lower variance estimates under BRR depends on the strength of correlation measured by R^2. The variables which are more correlated with the variables used in different raking adjustments have a lower estimate of variance under BRR. For example, the differences in variances are higher or more pronounced for estimates of proportions of persons with healthcare events and less pronounced for estimates of expenses from different events because expenses vary widely within a poststratification or raking adjustment cell. Variance estimates of expenses from event categories which are less frequent and have little correlation with control totals (such as emergency room or outpatient or inpatient visits) increase at a similar rate at different adjustment stages for both BRR and TSE. In general, the incidences of having expenses of a particular type are more correlated with control total categories than the mean expenses from these events.

In the presence of outliers, the variance estimates under the two methods disagree even more. Even with the base weight i.e., before any adjustment, the estimates under the two methods differ considerably and with different subsequent stages of adjustments the variances go down under both methods but does not follow the pattern observed for most other estimates. However, when an outlier is removed both methods agree at the initial stage and then follow the usual pattern for all subsequent stages of weighting and at the final stage the variance under BRR is lower than TSE consistent with the general pattern. This issue deserves further investigation to see which method provides the more reliable estimate of variance in the presence of outliers.

The analysis presented in this report shows that the most commonly used TSE method overestimates the sampling variances of MEPS estimates as it does not capture the variance reduction affect of poststratification/raking adjustments. In other words, TSE produces conservative estimates of the accuracy of MEPS estimates. However, this approach has some unintended positive implications. The control totals that we use in MEPS are not all known values and are estimated from NHIS and CPS. Both these sources of estimates are subject to sampling errors. The NHIS estimates are subject to higher sampling errors and the CPS estimates are highly accurate but still subject to a small sampling error. In computing variances, we did not incorporate the sampling variation of control totals and variance estimates are obtained assuming control totals are known without any sampling error. Consequently, some additional variation in estimates is ignored. However, the variability in MEPS estimates due to the variance in control totals is likely to be relatively small. Thus even after compensating for that we can still say that estimates under TSE are conservative and overestimate actual sampling variances.

Another source of variation that we ignore in computing the variances of MEPS estimates is the imputation variance. A significant portion of the MEPS expense data are imputed because the expense information is difficult for respondents to report and there are budgets constraints to the MEPS Medical Provider Component (Stagnitti et al. 2008). Hence the estimates of these variables are also subject to imputation variances which are not captured by either the TSE or the BRR method. The extent of such variance relative to the variance captured under the TSE or

BRR methods is not completely known. Further research is required to assess the relative impact of imputation on the overall variance and determine whether TSE method still overestimates the MEPS variance by more than the increased variance due to imputation. However, for all other non-expense type estimates from MEPS which are not subject to substantial imputation, the TSE method is likely to overestimate the variance of the MEPS estimates. So for most of the MEPS estimates, particularly the healthcare utilization estimates, the recommended methods such as TSE or single-step BRR replicates that we provide are conservative. While the TSE method generally overestimates the sampling and weighting components of the variances of MEPS expenditure estimates, it is probably a safer approach when considering that imputation and variation in control totals have inflationary effects on variances that are not accounted for by either method.

The BRR method does capture the reduction in variance due to poststratification/raking adjustments, an important source of variance reduction. This can be considered a more accurate method of variance estimation for MEPS healthcare utilization estimates which are not subject to any major imputation. However, the BRR method can underestimate the overall variances of MEPS expense estimates which are subject to significant imputation variance.

5.0 References

Botman S.L., Moore T.F., Moriarity C.L., Parsons V.L. Design and Estimation for the National Health Interview Survey, 1995–2004. National Center for Health Statistics. Vital Health Stat 2(130). 2000.

Ezzati-Rice, T.M., Rohde, F., Greenblatt, J., (2008). *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.shtml

Kass. G. V. (1980). "An Exploratory Technique for Investigating Large Quantities of Categorical Data." *Journal of Applied Statistics*, Vol. 29, No. 2 (1980), pp. 119–127.

Korn E. and Graubard B. (1999). *Analysis of Health Surveys*. John Wiley and Sons Inc. New York.

Machlin S.R., Chowdhury S.R., Ezzati-Rice T., DiGaetano R., Goksel H., Wun L.-M., Yu W., Kashihara D. (2010). *Estimation Procedures for the Medical Expenditure Panel Survey Household Component.* Methodology Report #24. September 2010. Agency for Healthcare Research and Quality, Rockville, MD. http://www.meps.ahrq.gov/mepsweb/data_files/publications/mr24/mr24.pdf

Shao, J. (1996). "Resampling Methods in Sample Surveys (with discussion)." *Statistics*, 27, 203–254.

Stagnitti, M. N., Beauregard, K., and Solis, A. (2008). *Design, Methods, and Field Results of the Medical Expenditure Panel Survey Medical Provider Component (MEPS MPC)—2006 Calendar Year Data*. Methodology Report No. 23, November 2008. Agency for Healthcare Research and Quality, Rockville, MD. http://www.meps.ahrq.gov/mepsweb/data_files/publications/mr23/mr23.pdf

Wolter, K.M. (2007). *Introduction to Variance Estimation*. 2nd Edition. New York: Springer-Verlag.