Precision of Survey Estimates Derived from the Medical Expenditure Panel Survey (MEPS)

William Yu, Agency for Health Care Research and Quality
2101 E. Jefferson St., Suite 5W29, Rockville, MD 20852

Key Words: design effect, health care expenditure, utilization

Introduction

The sample design of the 1997 Medical Expenditure Panel Survey(MEPS) is an overlapping panel design characterized by a multistage, complex area probability design that includes disproportionate sampling of specified policy relevant population groups(e.g.,blacks, Hispanics, the functionally impaired, children with activity limitations, individuals predicted to have high medical expenditures, and persons predicted to have family income less than 200 percent of poverty level). Standard methods of variance estimation which assume simple random sampling generally result in an under-estimation of variance, when used with data from a complex survey design (Cohen S., 1982). The extent of this departure from simple random sampling assumptions and its impact on the variances of survey estimates may be measured by the design effect. The design effect is defined as the ratio of the true variance of a statistic to the variance derived under simple random sampling assumptions. Based on data from the Household Component of the 1996 MEPS, design effect variations on estimates of health care expenditure and sources of payment was reported (Yu, W., 2000). Using data from the 1997 MEPS-HC, this paper will evaluate and contrast the design effects achieved for national estimates of health care utilization and expenditures; the level of design effect variation in related survey estimates; and design effect variation by alternative population subgroups and by different geographic regions of the nation.

Design of the 1997 MEPS Household Component

The 1997 MEPS Household Component (HC) was specified as a continuous survey with an overlapping panel design. Health care data are collected for each new MEPS sample (panel) to cover a two-year period, with the first two MEPS panels spanning 1996-97 and 1997-98, respectively. To produce health care estimates for calendar year 1997, the data are pooled across the two distinct nationally representative MEPS samples.

The views expressed in this paper are those of the author and no official endorsement by the Department of Health and Human Service or the Agency for health Care Research and Quality is intended or should be inferred. The author wishes to thank Dr. Steven Cohen and Dr. Doris Lefkowitz for their helpful reviews of the paper.

The National Health Interview Survey(NHIS) serves as the sampling frame for the MEPS. The NHIS is an ongoing annual household survey of approximately 42,000 households (109,000 individuals) conducted by the National Center for Health Statistics (NCHS) to obtain national estimates on health care use, health conditions, health status, insurance coverage, and access for the U.S. civilian noninstitutionalized population (Cohen S., 2000).

The 1996 MEPS HC sample was selected from households that responded to the 1995 National Health Interview Survey (NHIS). This selection consists of 195 Primary Sampling Units (PSUs), 1,675 sample segments (second-stage sampling units) and 10,597 responding households. It is designed to produce unbiased estimates for the four Census regions, with over-sampling of households with Hispanics and blacks at a ratio of approximately 2.0:1 for Hispanics and 1.5:1 for blacks. This 1996 MEPS sample constitutes a panel that was surveyed to collect annual data for 2 consecutive years (Cohen S., 1997).

A new 1997 MEPS sample was selected from the 1996 NHIS. This subsample was concentrated within the same 195 PSUs selected for the 1996 MEPS household sample and consisted of 14,706 responding NHIS dwelling units. A nationally representative subsample of 6,300 NHIS responding dwelling units (consisting of 6,480 reporting units) was selected to serve as the new 1997 MEPS sample. In addition to retaining the over sample of minorities that characterized the NHIS sample design, the 1997 MEPS was designed to over sample of the following policy-relevant subgroups: functionally impaired adults, children limited in activities, adults predicted to have high medical expenditures, and persons predicted to have family income less than 200 percent of the poverty level. The new 1997 MEPS panel was also designed to collect annual data for two consecutive years. Consequently, the full 1997 MEPS-HC sample consists of the first year of the 1997 MEPS panel pooled with the second year of the 1996 MEPS sample.

Source and Definition of Data

This study is based on the 1997 full year consolidated data file (MEPS HC-020). Expenditures on this file refer to what is paid for health care services. More specifically, expenditures in MEPS are defined as the sum of direct payments for care provided during the year, including out-of-pocket payments and payments by private insurance, Medicare, Medicaid, and other sources. Payments for over the counter drugs and for alternative care services are not included in MEPS total expenditures. Indirect payments not related to specific medical events such as medicaid Disproportionate Share and Medicare Direct Medical Education subsidies, are also not included.

The expenditure data included on this file were derived from the MEPS HC and Medical Provider Components(M.C.). Only HC data were collected for nonphysical visits, dental and vision services, other medical equipment and services, and home health care not provided by an agency while data on expenditures for care provided by home health agencies were only collected in the M.C. In addition to HC data, M.C. data were collected for some office-based visits to physicians(or medical providers supervised by physicians), hospital-based events (e.g. inpatient stays, emergency room visits, and outpatient department visits), and prescribed medicines. For these types of events, M.C. data were used if complete; otherwise HC data were used if complete. Missing data for events where HC data were not complete and M.C. data were not collected or complete were derived through an imputation process(Cohen S. and Carlson B., 1994).

Design Effect in the 1997 MEPS HC

Given the complex nature of the 1997 MEPS HC survey design, the assumptions of independence and equal selection probabilities are not satisfied. Its impact on variance estimation is best described as follows:

$$\sigma^2_{complex} = \sigma^2_{SRS} [1 + \rho \, (\tilde{n} - 1)]$$

where

$\sigma^2_{complex}$ is the true variance of a statistic given the complex survey design,

σ^2_{SRS} is the variance estimate obtained for the statistic under sample random sampling assumptions,

ρ is the intra cluster correlation coefficient, and

$\tilde{n}$ is the average cluster size.

The design effect is consequently expressed as:

$$\text{Design Effect} = (\sigma^2_{complex} / \sigma^2_{SRS}) = [1 + \rho \, (\tilde{n} - 1)]$$

The design effect deviates from unity when the effects of clustering are dominant in a survey design and the average cluster size is moderate to large. Variances of all estimated parameters presented in this paper were derived using the Taylor series linearization method to account for survey design complexities (shah, 1996).

Evaluation of Design Effect Variation

Based on the 1997 MEPS HC data, design effects are determined for a representative set of 30 survey statistics which estimate health care use and expenditures of the U.S. civilian noninstitutionalized population. For the nation, the design effects ranged from 1.14 for the estimate of total ambulatory (office based + outpatient) nurse/practitioner expenditure to 3.59 for the estimated number of dental care visits with an overall average of 2.10. Figure 1 is a bar chart comparing the level of design effects achieved for a subset of national estimates of health care use and expenditure.

Variables used to form population subgroups in this analysis include gender (male, female), age (<=17, 18-44, 45-64, 65+), race/ethnicity (Hispanic, black/non-Hispanic, others), and Census region (Northeast, Midwest, South, West), poverty level (0-199%, 200%+), children (age < 18 years) with limitations (yes, no), adults(age 18+ years) with functional impairments (yes, no), adults(age 18+ years) with other impairments (yes, no).

Figure 2 presents a comparison of average design effects from the selected health care use and expenditure measures across all the alternative population groups and by different geographic regions of the nation. Overall, age group 45-64 has the lowest average design effect at 1.47 while the South census region has the highest average design effect of 2.27. The average design effect is higher for males than females at 1.94 and 1.69 respectively. There is a notable downward trend for the value of average design effect by ascending age groups 0-17, 18-44, and 45-64. For race/ethnicity, the black/non-Hispanic group has the lowest average design effect at 1.47. For the census regions, persons living in the Midwest had the lowest average design effect at 1.81 and those in the South had the highest at 2.27. There is no notable difference in average design effect for people classified below or above 200 percent of poverty level. The specified policy relevant population groups who are functionally impaired (adults with 1 or more ADL/IADL) and children with activity limitations, have lower average design effects at 1.76, 1.69, and 1.54 respectively compare to their respective counter parts in the population.

The following subset of representative 1997 medical use and expenditure measures were selected for a more detailed study of design effect variation:

- Total health care expenditures,
- Total office-based expenditures,
- Total Rx-expenditures,
- Total inpatient expenditures,
- Total # of office-based provider visits,
- Total # of Rx medicine including refills,
- Total # of hospital discharges.

Figure 1 - Design Effects for Mean Use and Expenditure Estimates (1997 MEPS-HC)

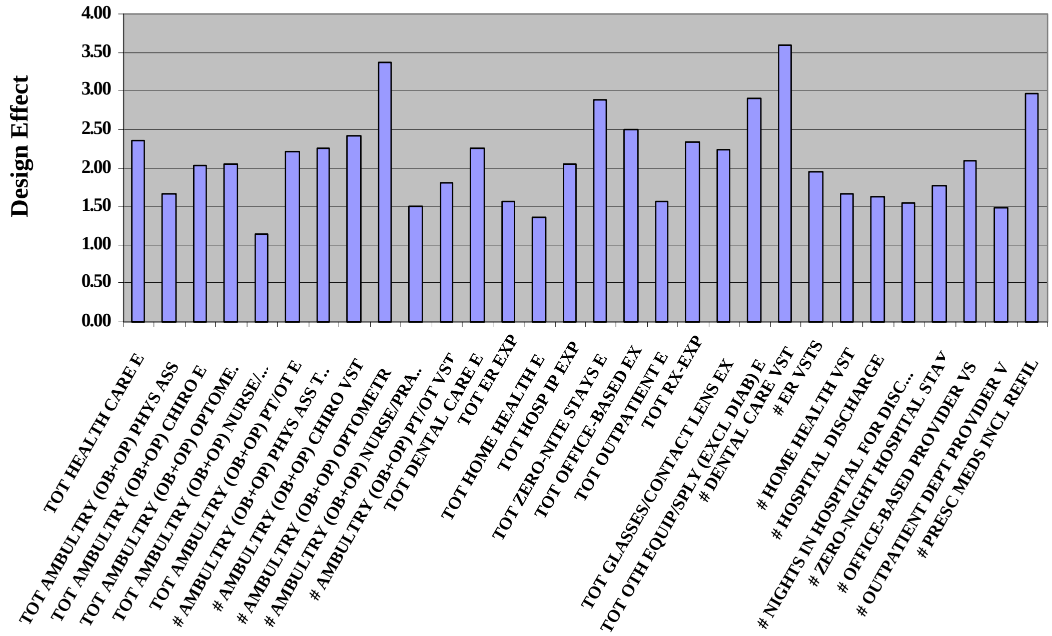
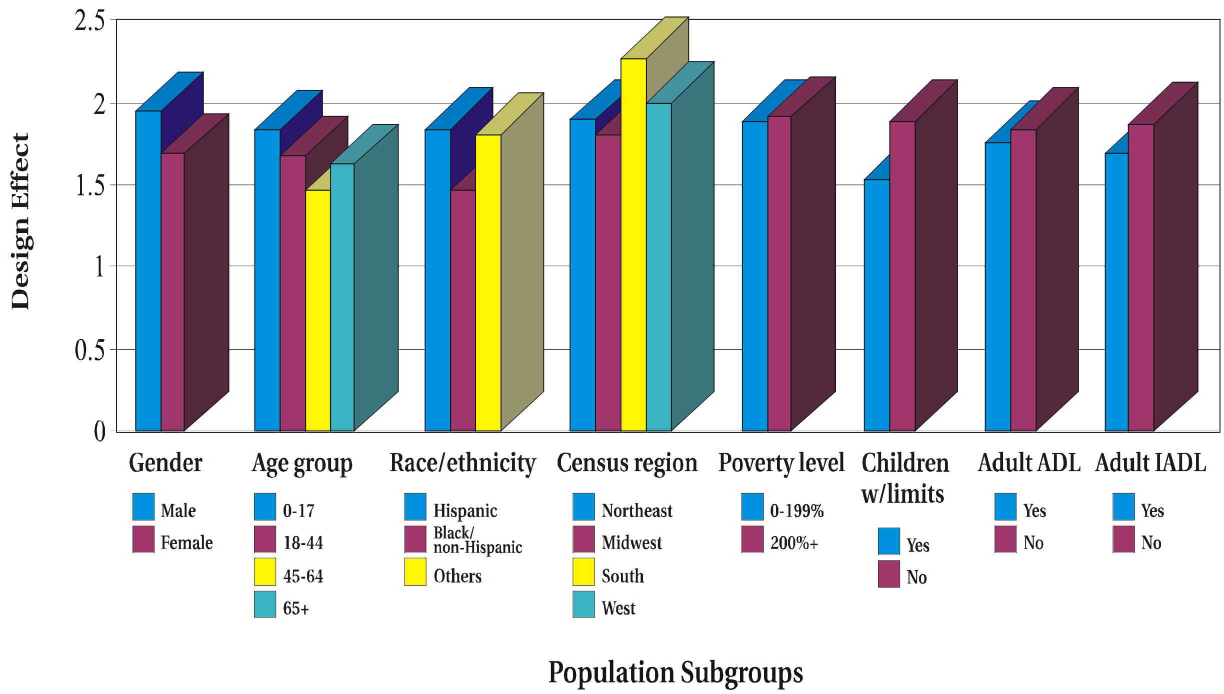


Figure 2 - Average Design Effect for Alternate Population Subgroups (1997 MEPS)



For each of the selected variables, domainestimates of design effect were generated in terms of population means. The domain estimates are defined by marginal or cross-classified distributional categories of the selected variables shown in Figure 2. Children with activity limitations and adults who are functionally impaired are combined into one category as persons with activity limitations or functionally impaired.

The quartile boundaries on sample size for the set of domain estimates under investigation were cross-classified by the quartile boundaries on the resultant mean estimates of the respective health care utilization and expenditure measures, yielding sixteen strata. Within each of these strata and their marginal classes, the average design effect and the standard error of the design effect were derived.

The most notable pattern in design effect variability was the positive incremental effect of sample size on the value of average design effect. As shown in Figure 3, the pattern was most obvious for the estimates of design effect of the mean on total number of prescription medicines including refills. The average design effects ranged from 1.01 (SE = .015) on sample size less than or equal to 45, to 1.71 (SE = .017) for sample size greater than 739. Similar patterns were observed for the other selected health care utilization and expenditure measures. This pattern of positive incremental effect was also reported in an earlier study of design effect variation on health care expenditure and sources of payment measures (Yu, 2000). No distinct relationship was observed (Figure 4) between the average design effect and the respective quartile boundaries which characterized the distribution of criterion variable domain estimates.

Summary

An overall precision requirement for the 1997 MEPS survey was the achievement of an average design effect of 1.7 for the survey estimates of policy relevant population subgroups(e.g., households with Hispanics and blacks, persons with family incomes less than 200% of poverty line, persons 65 years or older, adult and/or children with functional impairments). The study findings revealed that this requirement was generally satisfied (figure 2) with respect to measures of health care utilization and expenditures.

Figure 3 - Positive Incremental Effect of Sample Size on the Values of Average Design Effect (1997 MEPS)

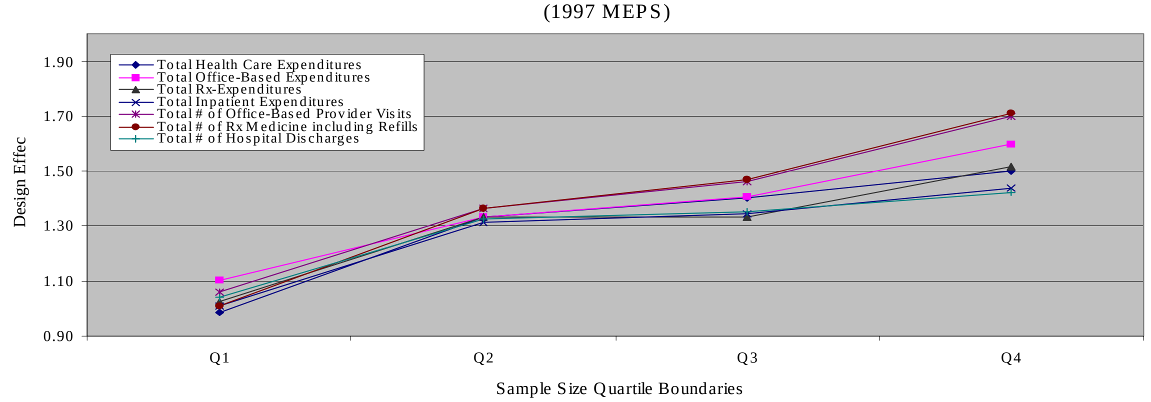
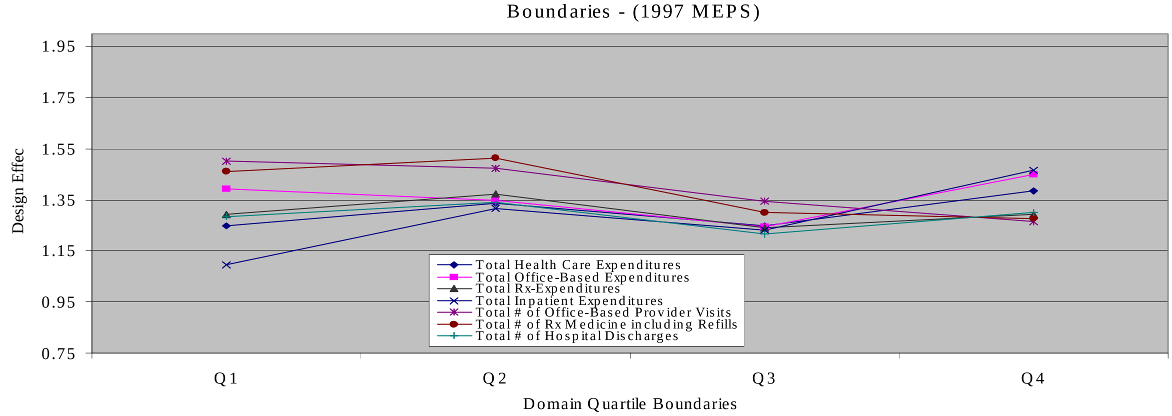


Figure 4 - No Distinct Relationshp Between the Values of Average Design Effect and Domain Quartile Boundaries - (1997 MEPS)



Overall, for the 30 selected measures of health care utilization and expenditures, the average design effects are approximately the same between persons with family incomes less than 200% or greater than 200% of poverty line but varied appreciably between gender, race/ethnicity groups, age categories, and among Census regions.

Positive incremental effects on the average design effect were observed in relation to sample size for all the selected variables. It is also observed that average design effect for uncommon events(e.g., 1.62 for # of hospital discharges) is closer to unity than other planned events(e.g., 3.59 for # dental care visits). One possible explanation is that the ultimate cluster units in the 1997 MEPS HC sample design are the household or family. It is to be expected that a strong positive correlation exists between individuals in the same household with respect to the total number of dental care visits compare to the number of times discharged by hospitals.

References

Cohen, S. B. "Comparison of Design Effect and Relative Variance Curve Strategy for Variance Estimation from Complex Survey Data" presented to the annual meetings of the American Public Health Association, 1982

Yu, W.W. "Design effects of survey estimates derived from the Medical Expenditure Panel Survey(MEPS)," Proceedings of the section on Survey Research Methods, American Statistical Association, 2000

Cohen S. B. "Sample design of the 1997 Medical Expenditure Panel Survey Household Component". Rockville (MD): Agency for Health Research and Quality; 2000. MEPS methodology Report No.11. AHCPR Pub. No. 01-0001.

Cohen S. B. "Sample design of the 1996 Medical Expenditure Panel Survey Household Component". Rockville (MD): Agency for Health Care Policy and Research; 1997. MEPS methodology Report No. 2. AHCPR Pub. No. 97-0027.

Cohen S. and Carlson B. "A comparison of household and medical provider reported expenditures in the 1987 NMES". Journal of Official Statistics 1994; 10(1):3-29.

Shah, B.V., B.G. Barnwell, G.S. Bieler, K.E. Boyle, R.E. Folsom, L. Lavange, S.C. Wheeless, and R.Williams Technical Manual: Statistical Methods and Algorithms Used in SUDAAN Release 7.0, Research Triangle Park, NC: Research Triangle Institute, 1996.