1. INTRODUCTION

1.1 Overview of the Medicare Current Beneficiary Study Redesign Evaluation Task

The sample design for the Medicare Current Beneficiary Study (MCBS), conducted for the Health Care Financing Administration (HCFA), was developed in 1990. The first stage units (primary sampling units, or PSUs) were selected using, for the most part, 1980 Census data, and the sample was based largely on the 1980 "master" sample, with some modifications to meet the special needs of the MCBS. Since 1990, of course, newer data have become available that could be used to redesign the MCBS sample. For example, 1990 Census data became available in 1992, although now these data are nearly 10 years old. Data from the 2000 Census will become available in 2001. Another source of data would be counts of Medicare beneficiaries taken directly from the HCFA records.

The primary purpose of this report is to review the advantages and disadvantages of redesigning and reselecting the MCBS sample using current counts of Medicare beneficiaries. Note that there are two separate issues here: (1) updating and changing the measure of size (used to select PSUs) from population counts to Medicare beneficiary counts and (2) reselecting the sample. While using actual counts of Medicare beneficiaries would seem to be a superior measure of size (MOS) for designing the PSU sample, it is not clear whether the improvements in precision and other benefits would outweigh the costs of changing to a new PSU sample for MCBS.

Westat's evaluation attempted to answer the following questions:

- How do the 1980 measures of size compare with 1990 data and current HCFA data?
- How would a change in the MOS affect the number of certainties?
- How would a new MOS affect cluster sample sizes and workloads?
- What costs would be associated with selecting and fielding a new sample?
- Could HCFA expect substantial gains in precision with a new MOS?
- What degree of overlap with the old PSUs could be achieved in new sample selection?
- How would the transition to the new sample be achieved?

- How would the redesign and reselection of the sample affect ongoing reporting of MCBS data?
- How would the redesign affect such ongoing activities as imputation, sample weighting, and variance estimation?

These issues are discussed in Chapters 3 through 7. Chapter 2 gives an overview of the MCBS sample design. The remainder of this chapter summarizes Westat's findings.

1.2 Summary of the Evaluation Results

1.2.1 Alternative Measure of Size

The three measures of size considered here included 1980 and 1990 Census counts and 1999 Medicare beneficiary counts. All these measures are highly correlated, though, in some specific PSUs, the number of Medicare beneficiaries gave a substantially different measure of size.

The number of certainty selections with an alternative measure of size depends heavily on the choice of the certainty cut-off. If 50 percent is used (i.e., all PSUs with at least 50% chance of selection), then the number of certainties with Medicare beneficiary counts would be substantially larger than in the current MCBS design. If 100 percent is used, the number of certainties would be substantially reduced.

1.2.2 Improved Precision through a New Measure of Size

The improvements in precision to be gained by using Medicare beneficiaries as a MOS are modest. For most survey variables, we would expect improvements in the range of 7 percent to 10 percent in the design effects, meaning that the sample size could be reduced by that amount and maintain the same level of precision. This improvement would result from balanced sample sizes rather than from reduced between-PSU variance components. There is evidence that design effects have been moving upward as the current MCBS panel ages; however, this increase could be due to variability in sampling weights to adjust for nonresponse.

1.2.3 Costs of Selecting and Fielding a New MCBS Sample

The cost of selecting and fielding a new MCBS sample depends heavily on the degree of overlap with the original PSU sample. These costs would be expected to be approximately $900,000 for an overlap of 30 percent, $750,000 for an overlap of 50 percent, and $500,000 for an overlap of 70 percent. There are other, less tangible, costs of working with an aging panel design, primarily because of uneven workloads that evolve over time. However, these costs are difficult to quantify and can often be minimized by operational solutions.

1.2.4 Evaluation of Overlap of Current and New Samples

Using the 1986 Ernst algorithm to maximize the overlap between two samples, we expect that the overlap between the new and old PSU samples would be approximately 50 percent for noncertainties and about 60 percent overall. This algorithm worked acceptably for us in strata that have less than 55 or so PSUs; with larger strata, we encountered difficulties, but these problems could probably be surmounted with further effort.

While the 1986 Ernst algorithm is well-established as a method for controlling sample overlap between surveys, we have some concern about the selection of new samples conditional on the original sample. Like other methods that operate by conditioning on earlier selections, the Ernst method is unbiased when *all* earlier samples are considered, but it is strongly influenced by the specific sample selected, regardless of possible changes in the measure of size.

Another method, developed by Ohlsson (1999), does not have this property. This method has good overlap properties for repeated future surveys, though the overlap with the current PSUs would be much smaller than with the 1986 Ernst method.

1.2.5 Impact on Ongoing Operations

We expect that there would be little impact of a redesign of the PSU sample on ongoing survey operations. In particular, the current operations of imputation, sample weighting, and standard error calculations would be relatively unaffected. The primary reason for this is that all these procedures

operate on specific panels. Since the new PSUs will be phased in by panels, the current procedures can be used with little or no modifications.

There is some possibility that time series trends may be affected by the change to a new PSU sample. However, we expect that this is not likely.

1.3 Conclusions

We recommend moving to the use of Medicare beneficiaries for several reasons. First, the gains in precision, while modest, are not negligible. Second, there is evidence that design effects have increased with the aging panel structure; if the same PSUs remain in place for another 10 year cycle, further deterioration of survey estimates can be expected. Third, while it is hard to quantify savings of balanced PSU workloads, balanced workloads do facilitate survey operations. Finally, the use of a 20-year-old measure of size, which will become 30 years old by the end of the cycle, seems inappropriate for a major national survey. No matter how unwarranted, the aging sample and size measure give an appearance of statistical unsoundness and an opportunity for questions regarding the accuracy of survey results.