Longitudinal Design Options for the Medical Expenditure Panel Survey Insurance Component

Steven B. Cohen, Joel W. Cohen, and Karen Davis
Center for Financing, Access and Cost Trends (CFACT)
Agency for Healthcare Research and Quality
Steve B. Cohen, Director, 540 Gaither Road, Rockville, Maryland 20850
Joel W. Cohen, 540 Gaither Road, Rockville, Maryland 20850
Karen Davis, 540 Gaither Road, Rockville, Maryland 20850

Abstract
The Medical Expenditure Panel Survey Insurance Component (MEPS-IC) is a nationally representative annual survey of over 40,000 business establishments and state/local governments. The survey is designed to produce estimates at the national and State level on the number and types of private health insurance plans offered, benefits associated with these plans, premiums, contributions by employers and employees, eligibility requirements, and employer characteristics. While the current MEPS-IC design provides estimates of employer decisions about health insurance offerings prior and post full implementation of the coverage provisions in the Affordable Care Act, both at the national and State level, the inclusion of a longitudinal arm in the survey would significantly enhance the capacity to interpret direct changes in employer behavior over time. This study provides a summary of the alternative design options under consideration for the MEPS Insurance Component and the most effective and efficient longitudinal design recommended for the MEPS Insurance Component that would permit enhanced analyses of changes in employer behavior associated with the coverage expansions scheduled to occur in 2014.

Key words: Longitudinal design, sample design optimization, establishment surveys, MEPS, ESI, ACA.

Introduction

Most Americans under age 65 obtain their health insurance through their employers. As a result, data on employers' behavior with respect to offering and paying for health care coverage for their employees is critical to understanding the current operation of the health care system in the U.S., and to evaluating how changes in policy are likely to affect that coverage. The Medical Expenditure Panel Survey Insurance Component (MEPS-IC) is a nationally representative annual survey of over 40,000 business establishments and state/local governments sponsored by the Agency for Healthcare Research and Quality (AHRQ). The survey is designed to produce estimates at the national and State level on the number and types of private health insurance plans offered, benefits associated with these plans, premiums, contributions by employers and employees, eligibility requirements, and employer characteristics. The survey is characterized by an integrated design, whereby the sample is selected from the Business Register, a confidential list of nearly all establishments in the United States maintained by the Census Bureau and from the Census of Governments. Use of the register as the MEPS-IC sampling frame permits efficient oversampling of establishments by location, size and industry and also serves as a post-stratification source.

With the implementation of Health Insurance Exchanges in 2014 and other changes made in the U.S. health care system by the Affordable Care Act that have the potential to affect employer decisions about health insurance offerings, it is critical that we ensure the MEPS-Insurance Component design is optimized to permit the necessary evaluations of the effects of those changes. Employers may respond to the new laws in a variety of ways, such as applying for tax credits, instituting vouchers for their employees, offering or discontinuing insurance coverage, instituting wellness programs that affect premiums and varying employee contributions by wage or other characteristics. While the current MEPS-IC design will provide estimates of employer decisions about health insurance offerings both prior and post full implementation of the coverage provisions in the Affordable Care Act, both at the national and State level, the inclusion of a longitudinal arm in the survey could significantly enhance the capacity to interpret direct changes in employer behavior over time. This report provides a summary of analytic objectives, precision requirements, design options and cost constraints for a longitudinal design enhancement to the MEPS-IC.

Cross-sectional and longitudinal survey designs

National health care surveys are generally characterized by cross-sectional or longitudinal designs. The cross-sectional surveys are designed to provide a snapshot of population characteristics that relate to a fixed point or interval in time. Alternatively, longitudinal surveys collect data on more than one occasion from the same sample members of the population of analytical interest in order to measure change and to obtain data for time periods too long to recall accurately in a single interview. Longitudinal observations are essential for characterizing variations in population attributes that are sensitive to changes in time.

Longitudinal survey designs are primarily adopted to provide the necessary information to assess changes in the behavior of the population over a specific time period. Often referred to as panel designs, they have the capacity to permit measurement of seasonal and annual variations in population characteristics and behavior. These longitudinal designs are essential to permit the acquisition of the data necessary to support analyses that measure the impact of changes in the target population's attributes over time. The achievement of well specified sample size requirements for these surveys also permits comparable studies for different economic groups or special populations of interest, such as the poor, the elderly, veterans, the uninsured, or racial/ethnic groups for population based surveys, and for employers, physicians, firms or establishments for business based surveys. While cross-sectional surveys permit analyses of net changes in population parameters at an aggregate level, only a longitudinal survey can discern the extent to which this is attributable to different elements of gross change (Lynn, 2009). For example, under the current MEPS-Insurance Component cross-sectional design, consider a situation in which the annual estimates of health insurance offer rates by employers are estimated to be the same over two consecutive years. Only a longitudinal design could determine whether it was the same set of employers that maintain their offers of coverage or whether there were substantial counter-balancing shifts in employer coverage offerings across individual establishments over time. A longitudinal survey design also allows for the development of economic models designed to produce national, regional and state estimates of the impact of changes in health care financing, health insurance coverage, and reimbursement policy over time, as well as estimates of who benefits and who bears the cost of such changes in policy.

Cross-sectional surveys also can form the baseline cohorts for future active longitudinal follow-ups. These longitudinal designs present extensive opportunities for understanding the relationship between risk factors and disease outcome as well as the natural history of disease for population based surveys and evaluating the impact of new policy initiatives on establishments or businesses. Generation and analysis of such longitudinal data is greatly facilitated in countries (e.g., Sweden) where various governmental information systems and surveys can be linked. Furthermore, the analytical capacity of health care related surveys can be dramatically enhanced through linkage to existing secondary data sources at higher levels of aggregation (both geographic and organizational) as well as through direct matches to additional health and socio-economic measures acquired for the same set of sample units from other sources of survey specific or administrative data. One of the more pervasive uses of existing administrative data bases is to serve as a sampling frame to facilitate a cost efficient identification of an eligible survey population for purposes of sample selection, such as the use of Medicare administrative records as a sampling frame for a survey of Medicare beneficiaries. Health care surveys that are linked to administrative records from their inception also benefit by the capacity for data supplementation that permits more extensive analyses that are beyond the more constrained scope of the core health care survey. Establishing similar connections to existing data sources that will substantially enhance a survey's capacity to address specific research questions are often more difficult to establish after a survey has been administered. This is primarily a consequence of confidentiality restrictions that require respondent permission to link patient records to administrative data sources, in addition to problems with the availability of the necessary identifiers from the survey respondents (Madans and Cohen, 2005).

The large majority of the nationally representative population-based health surveys sponsored by the Department of Health and Human Services have benefited by a capacity to link the survey data to county level data on health service resources and health manpower statistics available on the Area Resources File (ARF). More specifically, the ARF is a county-specific health resources information system containing information on health facilities, health professions, measures of resource scarcity, health status, economic activity, health training programs, and socioeconomic and environmental characteristics. Geographic codes and descriptors are provided to enable linkage to health surveys to expand analyses conducted by planners, policymakers, researchers, and other professionals examining the nation's health care delivery system and factors that may impact health status and health care in the U.S. Comparable enhancements to health surveys for supplementation of economic indicators are achievable through linkage of survey data to the socio-economic indicators made available by the Bureau of the Census through the County and City Data Book and public use files from the decennial Census. As noted, comparable enhancements are possible for business, employer and establishment surveys through linkages to the Business Register that is administered by the Bureau of the Census. The information is establishment-based and includes business location, organization type (e.g., subsidiary or parent), industry classification, and operating data (e.g., receipts and employment). As a consequence of confidentiality restrictions, public use files for the MEPS-IC are not produced and linkages to the Business Register are also subject to comparable confidentiality restrictions.

The quality and data content of household specific health surveys are often enhanced through the conduct of follow back surveys to medical providers and facilities that have provided care to household respondents. In terms of data quality, household reported medical conditions can be evaluated for accuracy relative to provider specific records on

medical conditions for the same patient and specific health events. With respect to health care expenditures collected from household respondents for their reported health care events, available linked medical provider level data is a more accurate source of information. The availability of such supplemental data on use and expenditures allows for the conduct of methodological studies to evaluate the accuracy of household reported data and informs adjustment strategies to household data in the absence of provider specific data to reduce bias attributable to response error. Once more, comparable enhancements are possible for business, employer and establishment surveys through linkages to the Business Register (Cohen, S. and J. Rhoades, 2007; Cohen, S. et al. 2006, 2005; Cohen, S. and L. M. Wun 2005; Cohen, S.. and T. Ezzati-Rice 2003).

Current MEPS Insurance Component Sample Design

The total budgeted sample size is approximately 45,000 sample units before non-response and out-of-scope units (AHRQ, 2009). The sample includes private sector establishments which employ at least one person, plus State and local governments. The sampling goal is to produce adequate estimates for 1) the private sector for all 50 States and the District of Columbia; 2) State and local governments by Census Division; and 3) the Nation as a whole. The sample frame is derived from two lists: 1) The Census Bureau's Business Register (BR), a list that contains private sector establishments in the United States which employ at least one person. The list is derived from tax records, and is continually updated to add newly created establishments (births) and remove those establishments that have closed. This list contains over 7,000,000 establishments and is very complete. 2) The Census of Governments, which is collected every five years with data updated in non-Census years using a sample survey. Currently, the most recently available Census of Governments is for the year 2012 and contains over 90,000 units from which the sample of State and Local Governments is selected. Together these two lists cover almost 100 percent of all organizations with at least one employee in the economy, excluding the Federal government.

Allocation to the State and Local Government and Private Sectors: The division of sample between the state and local government and private sectors is based upon past allocations. There are several precision targets for the survey. There are National and State targets for the private sector, and National and Census Division targets for government.

The national relative standard error (RSE) targets for the private sector for the survey are the following:

- a .005 RSE for national estimates of single and family premiums
- a .0150 RSE for national estimates of single and family employee contributions
- a .0075 RSE for national estimates of important proportions, such as the percent of employees enrolled in health insurance

The national RSE targets for the state and local government sector for the survey are the following:

- a .0075 RSE for national estimates of single and family premiums
- a .020 RSE for national estimates of single and family contributions
- a .010 RSE for national estimates of important proportions

State estimate targets for the private sector are that the RSE for state estimates have errors less than 6 times the similar national private sector targets. Census Division targets for the state and local government sector are that the RSE be less than 5 times the national state and local government targets. A MEPS-IC longitudinal design

enhancement will need to continue to meet these existing precision specifications for annual survey estimates.

Given these goals and the budget limitation for the sample of about 45,000 units, approximately 42,000 sample units are allocated to the private sector and 3,100 to governments. Within the private sector, the allocation includes a small number of approximately 200 large certainty units and the remaining sample is allocated to individual states (Sommers, 2007, 2004, 1999; Davis, 2013)

State Allocation for the Private Sector: From experience with past MEPS-IC surveys, it has been determined that a sample of approximately 500 responding establishments per state yields estimates that meet most state estimation goals using state stratification and allocation processes. Consequently, the sample initially allocates 17,000 responding sample establishments proportionally by size among the states. The allocation is then supplemented for the 42 smallest states so that each of the 11 smallest states receive about 495 sample establishments and each of the next 31 largest states receive 535 sample units. The 9 largest states receive their entire sample allocation from the proportional allocation of the 17,000 units. Additional adjustments are incorporated to allow for expected nonresponse and out of scope establishments in order to arrive at the final sample size per state.

Private Sector Sample Selection: Once the allocations are complete (excluding the certainty units) for each state, samples are selected for the private sector within each state. Before final sample allocation and selections, the universe in each state is stratified into 14 stratification cells. In addition, there is a certainty stratum for each State which contains establishments with projected enrollments of above 5,000 employees. Once these cells are created, the frame within each state is classified into the 14 strata. The Neyman optimal allocation formula (Cochran, 1977) is then used to obtain the State-level non-certainty allocation for the i^{th} stratum within each State:

$$n_{si} = \frac{N_{si} S_{1si}}{\sum_{i=1}^{14} N_{si} S_{1si}} n_s$$

where

N_{si} is the number of establishments in the i^{th} stratum in the s^{th} State,

n_s is the State sample size,

S_{1si} is the average standard deviation for the s^{th} State and the i^{th} stratum calculated based on the percent of all establishments that offer health insurance and

n_{si} is the allocation to the i^{th} stratum in the s^{th} State based on establishments that offer health insurance.

After this allocation is completed, a second allocation is performed where a different key MEPS-IC estimate (total enrollees) is used to calculate the average standard deviation.

$$m_{si} = \frac{N_{si}S_{2si}}{\sum_{i=1}^{14} N_{si}S_{2si}} n_s$$

N_{si} is the number of establishments in the i^{th} stratum in the s^{th} State,

n_s is the State sample size,

S_{2si} is the average standard deviation for the s^{th} State and the i^{th} stratum calculated based on total enrollees and

m_{si} is the allocation to the i^{th} stratum in the s^{th} State based on total enrollees.

The final allocation, r_{si} , is the weighted allocation obtained by taking the weighted value of the optimal allocations for the two variables as follows:

$r_{si} = .44\ n_{si} + .56\ m_{si}$

The weighting factors for the final allocation (.44 and .56) were determined based on an evaluation of the best overall balance in precision of estimates for the 2 variables. Once these allocations are completed, each establishment in a stratification cell is given the same chance of selection equal to $p_{si} = r_{si}/N_{si}$ where r_{si} is the final allocation within the State. In order to reduce the reporting burden on large firms, the probabilities are further adjusted. Once these probabilities of selection are finalized, the allocated samples are selected using systematic sampling. To perform this selection, the file is sorted by State, strata, industry and number of employees.

Analytical Objectives of a Longitudinal Design for the MEPS Insurance Component

Given its large sample size and very high response rates (on average 78%), the MEPS-IC is the leading source of data on employment-related health insurance coverage in the U.S. The MEPS-IC has been used extensively by analysts and policymakers to examine access to, enrollment in, and the cost of private employer-sponsored health insurance both nationally and among the states. Currently, however, the survey is cross sectional, which limits its capacity to evaluate changes in behavior within businesses over time. Given changes in the market for employer-sponsored insurance in recent years and the implementation of the Affordable Care Act (ACA), the ability to conduct longitudinal analyses of employer behavior with respect to health insurance will be vital to efforts to evaluate the impacts of the ACA and provide policymakers with the information they need to monitor and respond to trends in employment-related coverage.

The Affordable Care Act (ACA) was signed into law in 2010 to initiate a comprehensive set of changes to the U.S. health care system designed to expand insurance coverage and improve the efficiency and quality of care provided under that system. Prominent among these changes were requirements for individuals to have health insurance or face a tax penalty, and for large employers (those with 50 or more employees) to offer health insurance to their employees or face a penalty. The legislation also created state-based American Health Benefit Exchanges through which individuals will be able to purchase coverage, with subsidized premium credits available, on a sliding scale, to those with family incomes between 139-400% of the Federal poverty line. Individuals are eligible for exchange subsidies, however, only if they do not have access to “affordable”

coverage through their employer. Insurance is deemed to be affordable if the out-of-pocket employee contribution for single coverage is less than 9.5 percent of the worker's modified adjusted gross family income (MAGI) and the offered plan has an actuarial value of at least 60%.

Employers with 50 or more full-time workers that do not offer coverage and have at least one full-time employee who receives a premium tax credit will be assessed a fee of $2,000 per full-time employee, excluding the first 30 employees. Employers with 50 or more full-time employees that do offer coverage but have at least one full-time employee receiving a premium tax credit, will pay the lesser of $3,000 for each employee receiving a premium credit or $2,000 for each full-time employee, excluding the first 30 employees. Employers with less than 50 full-time workers are exempt from these penalties.

The ACA also provides tax credits for small employers that purchase health insurance for their employees. To receive these credits, the employer must contribute at least 50% of the total premium cost. The first phase, which was implemented in 2010, provided a tax credit of up to 35% of the employer's contribution toward the employee's health insurance premium. The full credit is available to employers with 10 or fewer employees and average annual wages of less than $25,000 and phases-out as firm size increases (to a limit of 25) and average wage increases (to a limit of $50,000). When fully implemented in 2014 the tax credit will pay for up to 50% of employers' contributions toward employees' health insurance premiums in the State Exchanges.

Finally, the legislation establishes an excise tax on the insurers of employer sponsored plans for coverage with an aggregate value that exceeds a specified threshold, starting in 2018 at $10,200 for individual coverage and $27,500 for family coverage. The thresholds will increase after 2020 at the rate of increase in the CPI-U. The tax will be set at 40% of the amount a plan exceeds the threshold and is assessed on the issuer of the insurance policy, which in the case of self-insured plans may be the employer. It also eliminates, in 2013, the tax deduction for employers who receive Medicare Part D subsidy payments for coverage of prescription drugs for retirees.

The current MEPS-IC has already served to inform several components of the Affordable Care Act (Cohen, 2011). In collaboration with the Office of the Secretary, DHHS and the Department of Treasury, AHRQ staff have provided MEPS-IC national and State level estimates of average premiums that were utilized to determine the small business tax credits for 2010 and subsequent years. More specifically, data from the 2012 MEPS Insurance Component were used to provide estimates of health insurance premiums by state for employer sponsored coverage provided by small employers of size 50 or less. The small employer health insurance tax credit was then determined based on the MEPS derived estimates of the average premium for the small group market in each State for the 2012 taxable year. On a related topic, the 40% "Cadillac plan" excise tax is supposed to take effect in 2018 and initially apply to health benefits packages that cost more than $10,200 for single coverage and more than $27,500 for family coverage. The MEPS data on the distribution of employer-sponsored health insurance premiums will continue to be utilized this decade to improve estimates of the number of plans that will likely be subject to this excise tax as we move closer to 2018. In addition, other characteristics of these plans, including employer and employee contributions, plan co-pay levels, and deductibles will be evaluated to assess trends in benefit structures over time.

There are a number of ways employers may alter their behavior with respect to the provision of health insurance after the implementation of the ACA. At a most basic level, one potential outcome is that employers will drop their insurance coverage in response to ACA, because for some large employers, the cost of the penalty for not offerring insurance may be less than the cost of providing it. Employer behavior, however, is not determined simply by the nominal cost of insurance, therefore, how employers will actually respond is an open question. Insurance coverage has become a part of employees' overall compensation package, both because of tax preferences associated with paying compensation in the form of insurance rather than salary, and constraints in the individual insurance market that make it difficult for some people to obtain insurance outside of employer-sponsored coverage. Consequently, the decision to offer insurance depends on a host of factors associated with the labor market, including labor supply, equilibrium compensation levels, and tax preferences that alter the cost of providing compensation in one form versus another.

There also may be incentives for some employers to shift the way they offer health insurance as a result of the altered regulatory landscape. For example, one possible outcome is that the ACA may provide an incentive for small employers to begin self insuring to avoid the Act's requirements with respect to what types of plans they must offer. Under this scenario and in concert with the new mandates, small firms with healthier than average employee populations may have an incentive to self insure because, under the Employee Retirement Income Security Act (ERISA), they may be able to avoid the requirements with respect to mandated benefits. The risk of incurring unanticipated medical costs may be mitigated by purchasing stop loss insurance with a low attachment point. The subsequent removal of individuals with low risks of high levels of medical expenditures from the small employer market could result in higher costs for the small employers who remain in the ACA regulated insurance market. At present, very few small employers self insure, about 12% of employers in establishments with less than 50 employees according to 2011 MEPS-IC data, but this indicates there is a great deal of potential for change following ACA implementation. Again, how small employers will actually respond to the change in regulatory environment is an open question. There may also be incentives under the ACA for employers to alter their numbers of employees to fit within the definition of a small employer.

Although trends can be monitored with cross sectional data, longitudinal panel type data are much more powerful for determining how change is occurring over time. For example, cross sectional trend data might show that the percent of employers offering health insurance remains constant over time, but those data will not show whether the stability is the result of new entrants into the market or existing establishments that have not altered their behavior. In addition, panel data is necessary to sort out exactly which characteristics of employers are associated with changes in behavior, with each establishment serving as its own control for determining whether differences are a result of behavioral change rather than unmeasured differences in the sample being examined. This is particularly important for examining issues such as how the ACA has affected employers overall compensation and workforce characteristics. A longitudinal design also permits cohort analyses of employers with specific characteristics over time to identify transitions in their approach to employer sponsored health insurance coverage and analyze the factors that are associated with these changes.

Given that the insurance markets for small and large employers are fundamentally different, and small and large employers are treated differently in the ACA, modifying

the design to include a longitudinal component would significantly enhance the capacity of the survey to address the following questions:
For Establishments in small firms (< 50 FTE workers): Separate analyses for small, low-wage employers and for other small employers to assess what changes have occurred in employer decisions on ESI:
Have small employers moved from: not offering health insurance to offering health insurance - either on their own or through a SHOP exchange; offering health insurance to not offering coverage?

What proportion of small employers have continuously offered health insurance?

What changes have occurred in plan offerings, including: changes in the required employee/employer premium contributions; changes in covered benefits?

How has the rate of self insurance changed?

Do the self insured employers have stop loss coverage?

How have workforce characteristics changed, including the number of FTEs and the mix of full- and part-time employees to analyze whether the change was related to the ACA definition of small/large firms?

How have wage levels changed in relation to whether or not health insurance is offered?

For Establishments in large firms (>= 50 FTE workers)

What changes have occurred in employer decisions on offering insurance: moving between the states of offering and not offering coverage?

How many and which types of employers have shifted their workers from coverage provided by the employer to coverage provided through an exchange?

How have employer/employee contribution levels changed, particularly with respect to the definition of affordability of coverage in the ACA.

What percent of workers obtain coverage through the individual exchange?*

How have benefits changed for workers who have shifted to an exchange?*

How have premiums, benefits and numbers and types of offered plans changed?

What changes have occurred in the number of FTE employees and the mix of full- and part-time employees?
*Note: – the ability to answer these and related questions will be dependent on being able to link to data from the exchanges.

Developing a longitudinal capacity for the MEPS Insurance Component
The objective of this study is to identify a longitudinal design option for the MEPS Insurance Component that would permit planned analyses of changes in employer behavior pre and post the 2014 Affordable Care Act planned coverage expansions. The design prototype would then serve as a model to allow for a redesign of the MEPS

Insurance Component sample selection strategy to support longitudinal analyses, with the 2013 MEPS-IC serving as the baseline. One of the primary design requirements for a MEPS-IC longitudinal design enhancement is to continue to meet existing precision specifications (noted above) for annual survey estimates. Another design constraint requires the cost of the longitudinal design to not exceed the current budget for the survey. As the alternative design options for a longitudinal design enhancement are specified, the sample sizes necessary to support each of the alternative design under consideration will also be specified.

In addition to the analytical attractions of enhancing the capacity of the MEPS-IC to conduct longitudinal analyses that assess changes in employer sponsored coverage over time, the use of each establishment as its own control in analyses of time trends has additional benefits in terms of gains in precision using paired comparisons. To illustrate this expected gain in precision for analyzing changes in employer behaviors through a design modification to allow for longitudinal analyses, the following analysis was conducted based on the MEPS Household Component, which has a longitudinal design. The following estimates derived from the survey for calendar years 2009 and 2010 were identified: annual healthcare expenditures, annual out of pocket healthcare expenditures, annual number of hospital stays, annual number of Dr. visits, the percent with fair/poor health status, and the percent of the population uninsured throughout the entire year. The sample was further restricted to those individuals who were classified as respondents for both years under study. The standard errors of the mean differences in survey estimates over the two years were then analyzed under two alternative survey design assumptions: 1) the samples for each of the survey years were independently selected, and 2) the sample observations were obtained from a longitudinal survey design.

Table 1 provides a summary of the respective estimates of the standard errors of the mean differences in survey estimates for the specified health care measures under the two design options. The results clearly indicate that the standard errors obtained from a design with two independent sample selections for the two year period are consistently higher than those obtained from a longitudinal design, ranging from 1.17 to 2.24 times as large. However, it is important to note that a longitudinal design is often characterized by lower survey response rates for subsequent years post the initial contact relative to cross-sectional design as a consequence of survey attrition. Survey estimates under longitudinal designs are also subject to potential bias due to conditioning effects over time. Consequently, a decision regarding the optimal design for a given survey is often based upon weighing the competing benefits and limitations of the alternative designs under consideration.

Table 1: Comparison of Precision in Estimates Under Alternative Design Assumptions

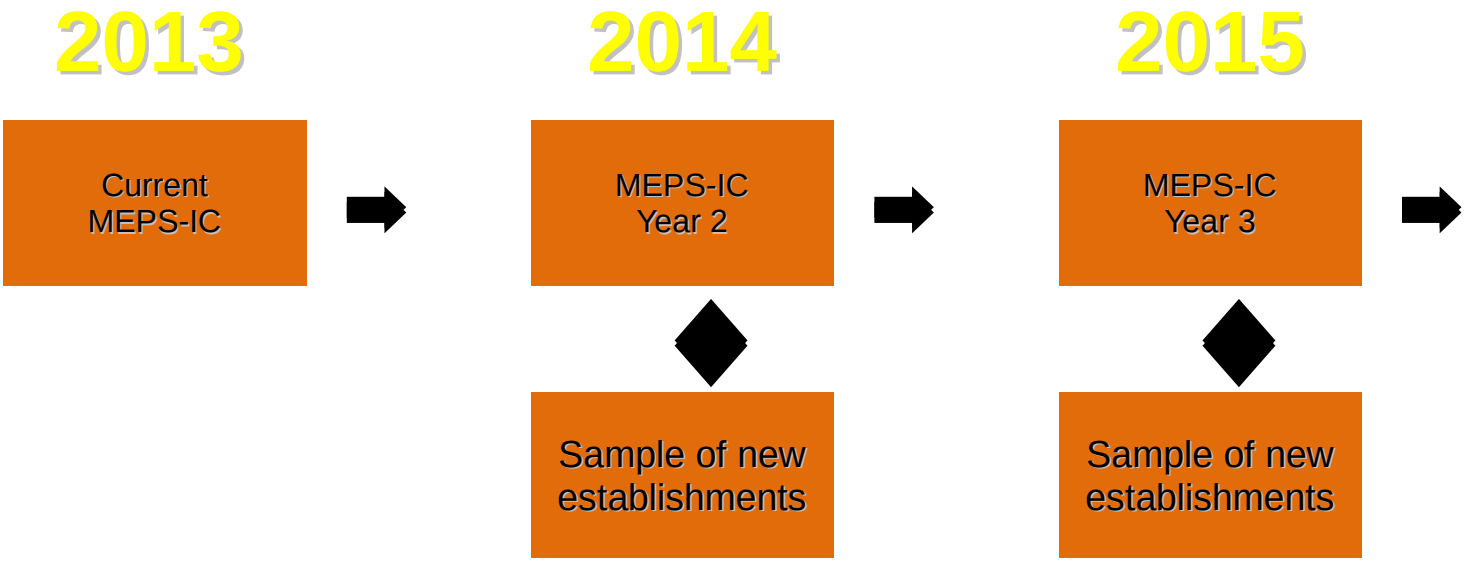
Measure	Mean Difference over time (2010-2009)	Standard error – Independent Design	Standard error – Longitudinal Design	Ratio of S.E.s Independent Design/Longitudinal Design
annual healthcare expenditures	**69.4815**	**115.79543**	**91.23015**	**1.26927**
annual out of pocket healthcare expenditures	**56.9348**	**15.70015**	**13.30513**	**1.18001**
annual number of hospital stays	**-0.0021**	**0.00437**	**0.00375**	**1.16521**
annual number of Dr. visits	**0.1437**	**0.05914**	**0.03694**	**1.60095**
percent with fair/poor health status	**0.5431**	**0.36504**	**0.25401**	**1.43713**
percent of the population uninsured	**-0.5116**	**0.41347**	**0.18420**	**2.24472**

Source: Medical Expenditure Panel Survey 2009-2010, Agency for Healthcare Research and Quality

After a series of internal deliberations on alternative longitudinal design prototypes that were feasible to achieve the specified analytic objectives of the survey, the following three distinct models were specified for further consideration. The three design options varied in both scale (sample size) and duration (length of longitudinal period) to permit longitudinal studies using the MEPS-IC. More specifically, the three design options under consideration included 1) a fully longitudinal design; 2) an overlapping panel design; and 3) a composite design consisting of cross-sectional and longitudinal arms. All the design options assume only minor if any changes to the MEPS-IC questionnaire. The more detailed descriptions of the designs are provided in the following section together with a summary of their advantages and limitations:

Design Option A: Fully Longitudinal Design. Under this design option, the sample of MEPS-IC establishments that defined the 2013 sample would be re-interviewed annually for a total of K years. For planning purposes K was set to a maximum of 4 years, to balance the analytical capacity for the conduct of longitudinal analyses against potential bias in survey estimates that would be attributable to significant levels of survey attrition as the number of survey contacts increased over time. In addition, each year the design would include a nationally representative sample of newly formed establishments (new births to the frame that were not in existence at the time the sample for time t-1 was selected) to supplement the longitudinal cohort in the derivation of nationally representative annual estimates from the MEPS-IC for a given year. The sample size for the refreshment sample of new businesses would be specified to be comparable to the sum of the number establishments that go out of business in a given year in addition to the number expected to attrite. Consequently, the overall sample size for this longitudinal survey would mirror the overall sample size of the current MEPS-IC. At the end of K years, a new MEPS-IC sample would be selected that was comparable to the current MEPS-IC cross sectional design, with an option to continue the older longitudinal survey for one additional year based on a representative subsample to allow for longitudinal estimates (Figure 1).

Figure 1. Design Option A: Fully Longitudinal Design



Advantages:

- Purest longitudinal design-largest sample allocation to support longitudinal studies.
- Provides the greatest precision for the conduct of longitudinal analyses over time. This can be demonstrated by comparing the power obtained by conducting a test for trends over two points in time using a paired t-test for a sample of size n interviewed at two distinct points in time in contrast to a comparison of two independent sample of size n at the distinct time points.
- Estimation strategy relatively straight-forward.
- Variance estimation and analysis relatively straight-forward

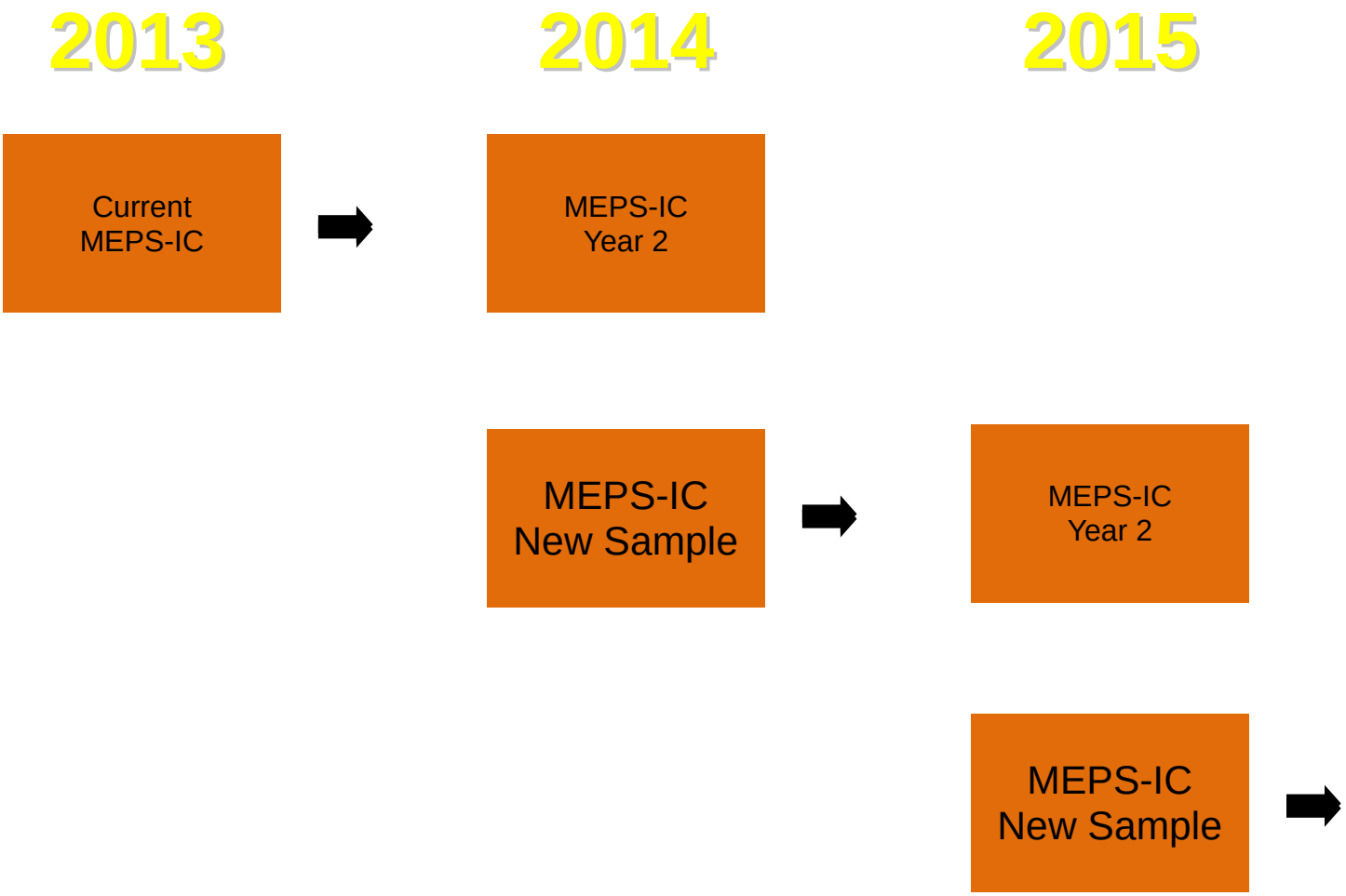
Limitations:

- Lower response rate attributable to attrition
- Potential impact of nonresponse bias attributable to attrition.
- Incremental level of respondent burden attributable to multiple contacts
- Potential bias in estimates associated with conditioning effects.

Design Option B: Overlapping Panel Design. Under this design option, a subsample (X %) of MEPS-IC establishments that defined the 2013 sample would be re-interviewed annually for a total of L years. In addition, a new nationally representative sample of establishments (100% -X %) would be selected each year. Estimates would be derived from the survey by pooling the estimates from the respective panels to meet the core MEPS-IC precision requirements. For planning purposes L was set to 2 years, to balance the analytical capacity for the conduct of longitudinal analyses against the added complexity of implementing composite estimation that pooled the panel specific estimates and potential bias in survey estimates that would be attributable to significant levels of survey attrition as the number of survey contacts increased over time. The pooled estimates for national and state characteristics would be based on composite estimation for the establishments that were in existence in the prior year t-1 and were on the Business Register sampling frame to support year t-1 estimates (pooling the longitudinal sample with the new sample cases that existed in the prior year t-1). They would be supplemented with estimates of new establishments that were newly formed

(new births to the frame that were not in existence at the time the sample for time t-1 was selected) solely based on the new sample. It should be noted that the MEPS Household Component has adopted an overlapping panel sample design in which each panel is surveyed to acquire calendar year estimates for a consecutive two year period (Figure 2).

Figure 2. Design Option B: Overlapping Panel Design.



Advantages:

- When sample design or methodological changes are made in the implementation of an ongoing survey, a common strategy is to conduct a bridging survey, using new design features in one random part of the sample and the previous older design for the remainder. The bridging survey, sometimes referred to as a parallel or overlap survey provides a means for assessing the impact of the changes on the survey estimates, and
  it can also serve to link past trends based on the old design with future trends based on
  the new design.
- No special provisions needed for a sample of new businesses each year, as this is taken care of in the annual selection of a fresh sample of establishments to represent the nation.
- The survey response rate is higher than a full longitudinal design.

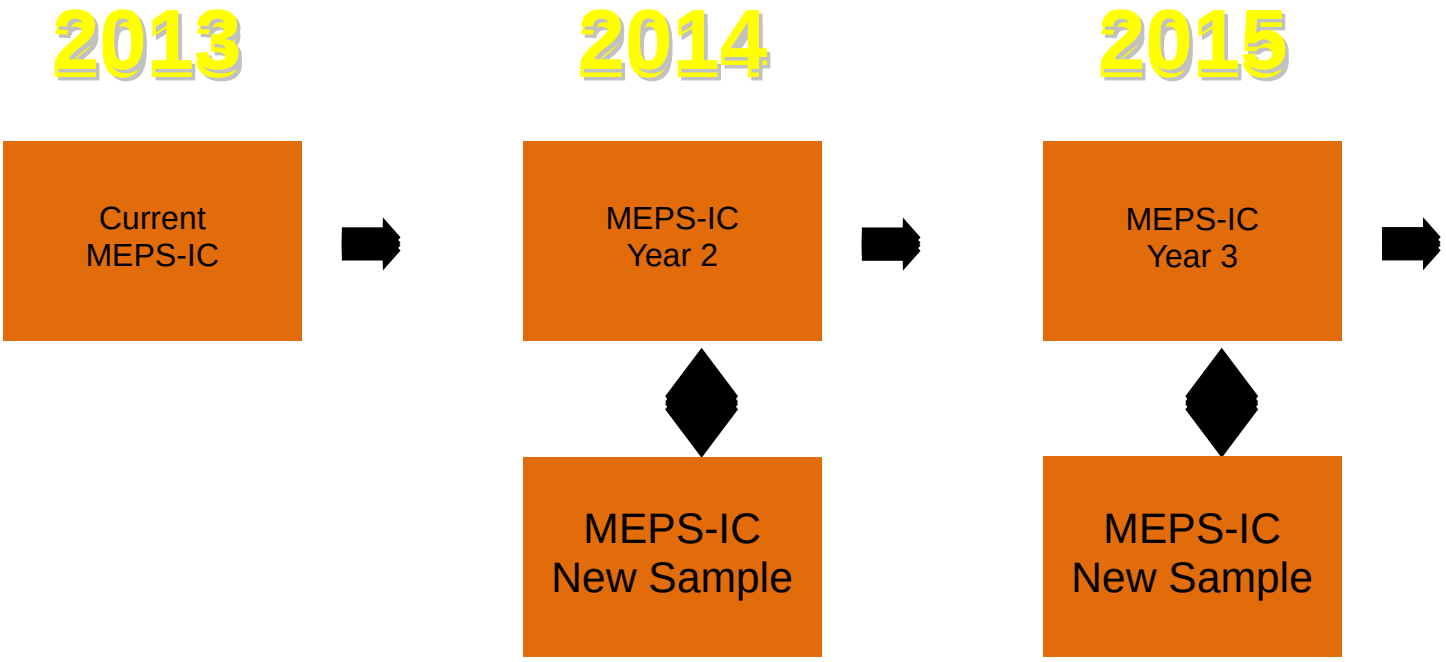
- If the longitudinal period for the overlapping panel design is specified as a shorter interval than the full longitudinal design, the impact of bias due to survey attrition and conditioning effects is lower.

Limitations:

- Greater complexities in estimation as a consequence of composite estimation pooling the panel specific estimates.
- Greater complexities in variance estimation and analysis as a consequence of composite estimation pooling the panel specific estimates

Design Option C: Composite design consisting of cross-sectional and longitudinal arms. Under this design option, X percent of the sample of MEPS-IC establishments that defined the 2013 sample would be re-interviewed annually for a total of K years. For planning purposes K was set to a maximum of 4 years, to balance the analytical capacity for the conduct of longitudinal analyses against the added complexity of implementing composite estimation that pooled the panel specific estimates and potential bias in survey estimates that would be attributable to significant levels of survey attrition as the number of survey contacts increased over time. Each year, the remainder of the MEPS-IC sample would consist of a new nationally representative cross-sectional sample of establishments. Estimates for national and state characteristics would be based on composite estimation which pooled the estimates from the respective survey arms. The pooled estimates for national and state characteristics would be based on composite estimation for the establishments that were in existence in the prior year (pooling the longitudinal sample with the new sample cases that existed in the prior year) supplemented with estimates of new establishments in the current year solely based on the new cross-sectional sample (Figure 3).

Figure 3. Design Option C: Composite design



Advantages:

- Attributes of bridging survey as noted above. No special provisions needed for a sample of new businesses each year, as this is taken care of in the annual selection of a fresh sample of establishments to represent the nation.
- The survey response rate is higher than a full longitudinal design.

Limitations:

- Greater complexities in estimation as a consequence of composite estimation pooling the panel specific estimates.
- Greater complexities in variance estimation and analysis as a consequence of composite estimation pooling the panel specific estimates
- Incremental level of respondent burden attributable to multiple contacts
- Potential bias in estimates associated with conditioning effects.
- Lower precision in estimates for the longitudinal component relative to Option A as a consequence of the sub-sampling of establishments (X percent of year t-1 selected to continue in year t).

Discussion

When survey redesigns are implemented, there is always the risk of being affected by unintended consequences of the modifications. With respect to an ongoing survey such as the MEPS-IC, a decision regarding the best approach to transform the design to include a longitudinal capacity will need to carefully balance the anticipated gains in analytic capacity and precision against the attendant additional sources of bias attributable to burden, survey attrition and conditioning effects. Furthermore, all the longitudinal designs under consideration will be affected by nonresponse due to survey attrition and loss of sample due to employers going out of business. To minimize the impact of these deleterious factors and allow time to evaluate their effects on survey estimates, survey operations and costs, an incremental approach to enhancing the capacity of the MEPS-IC to permit longitudinal studies is proposed. From that perspective, consideration of Design Option B with 1) a restricted sample size for the longitudinal arm not to exceed 2500 responding employers, 2) a two year constraint on the longitudinal period and 3) the inclusion of a fresh sample to evaluate the impact of a second year of data collection on survey estimates and operations, would best achieve that criterion.

In terms of determining the sample allocation for the longitudinal component, the precision requirements for the underlying analyses would need to be specified and the gains in precision anticipated by the longitudinal design would also need to be determined. The sample size requirements to detect differences in survey estimates of a given magnitude with an alpha level of .05 and a power of 0.80 under independent sampling assumptions is provided in Table 2. The results are based on an assumption of simple random sampling. While the MEPS-IC is characterized by a stratified sample design with a probability proportional to sample selection scheme, average survey design effects are close to 1. Consequently, the results in Table 2 should also be applicable to the MEPS-IC.

For paired comparisons over time that would occur with a longitudinal sample design, sample size reductions as large as 50 percent are possible relative to those required from independent designs. For example, if 90 percent of large firms offer coverage to their employees at baseline and the underlying design must have the capacity to detect a 2

percent change in the next year at an alpha level of .05 and a power of 0.80, a sample size of 3528 is necessary under independent sampling and as low as 1764 under a longitudinal design (Table 2).

Table 2: Sample Size Requirements to Detect Differences of Size d Over Two Years for Independent Samples

Measure	Proportion at baseline	Difference to Detect	Alpha level	Power	Sample Size
Change in Proportion of Firms with Coverage Offers	0.9	0.01	0.05	0.8	14112
	0.9	0.02	0.05	0.8	3528
	0.9	0.03	0.05	0.8	1568
	0.9	0.04	0.05	0.8	882
	0.5	0.01	0.05	0.8	39200
	0.5	0.02	0.05	0.8	9800
	0.5	0.03	0.05	0.8	4356
	0.5	0.04	0.05	0.8	2450

Summary

This report has focused on methods to expand the analytic capacity of the MEPS Insurance Component to include a longitudinal design component. While the current MEPS-IC design provides estimates of employer decisions about health insurance offerings prior and post full implementation of the coverage provisions in the Affordable Care Act, both at the national and State level, the inclusion of a longitudinal arm in the survey would significantly enhance the capacity to interpret direct changes in employer behavior over time. Consequently, particular attention has been given to distinguishing the benefits of both cross-sectional and longitudinal surveys as well as their constraints.

A description of the current MEPS Insurance Component sample design and precision requirements is provided to frame the underlying design parameters for the survey. The report also provides a detailed summary of the analytical objectives of a longitudinal design enhancement to the survey. Emphasis is then given to the health policy questions that could best be addressed with time dependent data on employer sponsored coverage characteristics.

To achieve the specified analytic objectives of the survey, three distinct longitudinal design options were specified for further consideration. The three design options varied

in both scale (sample size) and duration (length of longitudinal period) to permit longitudinal studies using the MEPS-IC. More specifically, the three design options under consideration included 1) a fully longitudinal design; 2) an overlapping panel design; and 3) a composite design consisting of cross-sectional and longitudinal arms. All the design options assumed only minor if any changes to the MEPS-IC questionnaire and required the design to also continue to satisfy current analytical objectives and precision requirements.

The implementation of design modifications to an on-going large national healthcare related survey such as the MEPS Insurance Component is not "risk-free". To mitigate the potential impact of unanticipated adverse effects from the design enhancement under consideration, we considered a conservative approach was viewed as the most attractive approach to adopt. From that perspective, we conclude that a longitudinal design enhancement characterized by an overlapping panel design would be the best option to achieve that objective. To further advance this strategy, we propose that the sample size for the longitudinal arm not exceed 2500 responding employers and the longitudinal period be restricted to a two year interval. The recommendation is based on 1) the capacity to detect changes of ~ 2% at the national level and ~3-4% for 25% subsample, 2) limitations in available budget and 3) recognizing the need to conduct a pretest to assess feasibility. This incremental approach to enhancing the analytic capacity of the MEPS-IC with a longitudinal component serves to permit careful evaluations of the effects of those changes on survey operations, data quality, accuracy and cost. In essence, this model aligns with implementing a design modification through a "pretest" mechanism with an evaluation component to guide design "fine-tuning", prior to initiating a full scale survey design enhancement.

Note: 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 is intended or should be inferred.

References

AHRQ (2009). MEPS OMB Clearance Supporting Statement - Part B. Medical Expenditure Panel Survey –Insurance Component 2012-2013, Version September 22, 2011, Agency for Healthcare Research and Quality (AHRQ).

Cohen S. B. (2011) "The Capacity of MEPS to Inform Provisions of the Patient Protection and Affordable Care Act." Agency for Healthcare Research and Quality Working Paper No. 11003, February 2011, http://gold.ahrq.gov.

Cohen, S. and J. Rhoades (2007) "Estimation of the long term uninsured in the Medical Expenditure Panel Survey". Journal of Economic and Social Measurement. Volume 32: 235-249.

Cohen, S., T. Ezzati-Rice, and W. Yu (2005) "Integrated Survey Designs: A Framework for Nonresponse Bias Reduction." Journal of Economic and Social Measurement Volume 30, No.2-3, 101-114.

Cohen, S. and L. M. Wun (2005) "A Comparison of Household and Medical Provider Reported Health Care Utilization and Estimation Strategies to Correct for Response Error". Journal of Economic and Social Measurement. Volume 30, No.2-3, 115-126.

Cohen, S., T. Ezzati-Rice, and W. Yu (2006) "The Impact of Survey Attrition on Health Insurance Coverage Estimates in a National Longitudinal Health Care Survey". The Journal of Health Services and Outcomes Research Methodology. Volume 6: 111-125.

Cohen, Steven B. and T. Ezzati-Rice (2003) "Design and Estimation Strategies and Innovations in the Medical Expenditure Panel Survey for the Measurement of Trends in Health Care Expenditures". Proceedings of the American Statistical Association, Section on Health Policy Statistics.

Davis, K. (2013). *Sample Design of the 2011 Medical Expenditure Panel Survey Insurance Component***.** Methodology Report #27. March 2013. Agency for Healthcare Research and Quality, Rockville, MD. http://www.meps.ahrq.gov/mepsweb/data files/publications/mr27/mr27.pdf

Lynn, Peter (2009). *Methodology of Longitudinal Surveys*. John Wiley & Sons Ltd. Chichester, West Sussex, United Kingdom.

Madans, Jennifer H, and S. B. Cohen (2005). "Health Surveys: A Resource to Inform Health Policy and Practice" In Health Statistics in the 21st Century: Implications for Health Policy and Practice.

Sommers, J. P. (2007). Additional Imputations of Employer Information for the Insurance Component of the Medical Expenditure Panel Survey since 1996. Methodology Report No. 17. January 2007. Agency for Healthcare Research and Quality, Rockville, Md. http://www.meps.ahrq.gov/mepsweb/data_files/ publications/mr17/mr17.pdf

Sommers, J. P. (2004) Updates to the Medical Expenditure Panel Survey Insurance Component List Sample Design, 2004. Methodology Report No. 18. January 2007. Agency for Healthcare Research and Quality, Rockville, Md. http://www.meps.ahrq.gov/mepsweb/data_files/publications/mr18/mr18.pdf

Sommers JP. (1999). Construction of weights for the 1996 Medical Expenditure Panel Survey Insurance Component list sample. Rockville (MD): Agency for Health Care Policy and Research; 1999. MEPS Methodology Report No. 8. AHCPR Pub. No. 00-0005