

**Assessing the Need for a New Household Panel Study:  
Health Insurance and Health Care**

**This draft: December 16, 2014**

Helen Levy  
Institute for Social Research  
University of Michigan  
426 Thompson St.  
Ann Arbor, MI 48104  
Phone: (734) 936-4506  
Fax: (734) 647-1186  
hlevy@umich.edu

**Abstract:** This paper considers the availability of data for addressing questions related to health insurance and health care and the potential contribution of a new household panel study. The paper begins by outlining some of the major questions related to policy and concludes that survey data on health insurance, access to care, health spending, and overall economic well-being will likely be needed to answer them. The paper considers the strengths and weaknesses of existing sources of survey data for answering these questions. The paper concludes that either a new national panel study, an expansion in the age range of subjects in existing panel studies, or a set of smaller changes to existing panel and cross-sectional surveys, would significantly enhance our understanding of the dynamics of health insurance, access to health care, and economic well-being.

**Keywords:** *Survey data; longitudinal studies; health insurance; access to care; health reform*

## 1. Introduction

The passage of the Affordable Care Act (ACA) marks the most significant domestic policy achievement of a generation. The case for policy action was informed by research using survey data, and the implementation of the law will be tracked using these data as well. At the same time, new policy questions requiring new research will inevitably arise. The pressure that Medicare and Medicaid place on the Federal budget ensures that they will continue to be the subject of policy attention; one can only guess at what changes, if any, these programs may face. There will surely be other new health policy issues that we cannot anticipate. This uncertainty underscores the importance of ongoing, broad-based data collection efforts that do not narrowly target particular interventions or population subgroups, but paint a complete picture of the resources, preferences, and constraints that shape household decisions about health and health care. Comprehensive longitudinal household surveys represent a rich and valuable resource for addressing important policy questions now and in the future.

This paper considers the availability of data for addressing questions related to health insurance and health care and the potential contribution of a new household panel study. I begin by discussing some of the major questions related to current policy and, more speculatively, some of the questions that might arise in the future. Next, I review the existing sources of survey data, including examples of research that illustrate how survey data has contributed to our knowledge of health insurance and health care.

I also discuss, briefly, major contributions to health economics that have relied on other types of data such as administrative data and data from social experiments. I consider what contributions a new household panel study would make in this area and what the attributes of such a study should be in order to make it most useful. I also consider smaller changes to existing surveys that would address some of the same data gaps as a new household panel study.

## **2. What are the most important scientific and policy issues within health insurance and health care now and in the coming years?**

Understanding what data we will need on health care and health insurance requires first thinking about what questions we may want to answer. Broadly speaking, there are two general areas of health and health care that are high priorities for research over the next decade(s), either because of recent policy action or because of the likelihood of significant policy action in the future: evaluating the impact of ACA, which is an urgent priority now, and helping to inform and then evaluate possible changes to Medicare and/or Medicaid in the near future. My discussion of these issues and the research questions they are likely to generate focuses on the impact of public policy on individuals and households rather than firms and suppliers. This focus should not be taken to mean that supply-side impacts are unimportant, but that they are less relevant to the question of what might be needed from a new household panel study.

Evaluating the impact of coverage expansions under the Affordable Care Act (ACA) and their effects on health and economic well-being is perhaps the most pressing item on the current health policy research agenda. The effects of the ACA are likely to be far-reaching, affecting many different domains:

- **Health insurance coverage:** How will coverage from different sources change as a result of the ACA?
  - **Medical care use and spending:** How will changes in insurance coverage affect medical care use and spending?
  - **Other measures of access to care:** How will the ACA affect other measures of access to care such as self-reported difficulty finding providers, cost-related medication non-adherence, foregoing necessary medical care due to cost, and delays in obtaining care due to cost or other reasons?
  - **Health outcomes:** Do increases in coverage translate into improved health for individuals?
  - **Financial well-being:** Does gaining coverage reduce measures of financial stress such as poor credit scores, bankruptcy filing rates, or food insecurity?
- Labor markets:** How will the ACA affect labor demand and supply, including job-to-job mobility?

Thinking beyond the evaluation of the ACA to what other policy questions related to health insurance and health care may arise in the not-too-distant future, it is

instructive to review the Congressional Budget Office's list of health-related options for reducing the deficit (CBO 2013), which are summarized in Table 1. This is not to suggest that all important policies are or should be motivated by deficit reduction; far from it. But the structural deficits associated with Medicare and Medicaid suggest that some action on Medicare and Medicaid seems inevitable in the next decade or so, despite substantial political obstacles. For the reasons outlined above, I focus on options directly affecting beneficiaries. These include: raising the Medicare eligibility age to 67; increasing beneficiary premiums; increasing beneficiary cost-sharing at the point of service; and converting Medicare to a premium support model. The common thread running through all of these proposals is, not surprisingly, the shifting of costs to beneficiaries. The role of research using survey data, if these policies are seriously being debated, is to help predict how they would affect the outcomes already listed above in thinking about the impacts of the Affordable Care Act: access to medical care, health outcomes, financial security, labor supply, consumption, and food insecurity or other material hardship.

The research questions and data needs raised by the possibility of changes to the Medicaid program are similar, but even more uncertain because the leading option for significant cuts to Medicaid is capping federal spending (CBO Health Option 1), meaning that each state would decide what changes to make. Changes in policy at the state level would add the additional complexity of different impacts across states, with

richer states more likely to raise additional offsetting revenue and poorer states more likely to cut eligibility or benefits in the face of caps on federal spending. These disparities would be overlaid on existing disparities in public programs and economic well-being across states. This aspect of possible Medicaid changes highlights the need for data that include state identifiers, which I will discuss in more detail in the next section since it is also critical to our ability to evaluate credibly the impact of the Affordable Care Act.

### **3. What data are already available to address these questions?**

The key point of the preceding discussion is that understanding the impact of health policy changes – or forecasting the likely impact of proposed policies – requires ongoing data collection on households’ insurance coverage, health, out-of-pocket-medical spending and access to medical care, in conjunction with a more complete portrait of their economic well-being that includes income, wealth, labor supply, consumption, and measures of material hardship such as food insecurity. These data should also be high-quality in the sense of being carefully measured and representative of the population of interest. Questions about program evaluation (“what was the impact of X on Y”) require in addition some sort of exogenous variation in exposure to the program if they are to support causal inference. In this section of the paper, I consider first what high-quality data are already available at the national level for

answering the types of questions posed in the previous section, and then consider (briefly) the issue of causal inference.

The good news is that high-quality data to evaluate the impact of the ACA or to help inform changes to Medicare and Medicaid are already available from a number of nationally representative surveys and other sources. Table 2 summarizes the major large-scale surveys that include health insurance and at least some health information, grouped according to whether they are cross-sectional, short panel (two to four years), or longitudinal/cohort studies that follow respondents as long as possible. In this section, I assume that readers have basic familiarity with the design of each survey and discuss the strengths and weaknesses of these surveys for addressing the kinds of questions identified in the previous section.

#### *Cross-sectional Census workhorses: CPS and ACS*

The CPS has long been the go-to source for health insurance information at the national level, with official estimates of the uninsured released every fall in the P-60 report “Income, Population, and Health Insurance.” It has also been used in a number of very widely cited papers on health insurance and the labor market and also, in conjunction with other health data, papers on the impact of Medicaid expansions on health [1-4]. The CPS health insurance data have some well-documented limitations. First, the CPS yields unusually high estimates of the uninsured compared to other

studies, perhaps because of how the question is worded [5].<sup>1</sup> Second, Medicaid coverage is underreported in the CPS; this is likely to be true in all surveys that ask respondents about Medicaid [6] but has been particularly well documented in the case of the CPS, thanks the Census Bureau's efforts to analyze and improve the quality of their data products [7-10]. Finally, the CPS includes relatively little health information; a single question on self-reported health has been available since 1995. On the plus side, however, the CPS has a relatively large sample and excellent income and labor supply information. Until 2014, the CPS also had the advantage of having asked relatively consistent questions over time. Starting in 2014, the CPS is introducing new health insurance questions intended to reduce the type of misreporting described above. At the time of this writing, there is some uncertainty about whether these new questions will supplement or replace the questions that the CPS has asked in the past.<sup>2</sup>

The ACS, which grew out of the "long form" of the decennial US Census, has included a question on health insurance since 2008. Strengths of the ACS include its enormous sample size – encompassing 1% of the population, it is approximately 30 times the size of the CPS – which supports estimates for small geographic areas.

---

<sup>1</sup> In a nutshell, the CPS asks respondents in March about whether they had coverage at any time during the prior calendar year, but people seem to answer as if they were asked about their health insurance status at the time of the survey. The result that the CPS should in theory yield an estimate of the number of people uninsured for the entire prior calendar year, but in practice it probably tells us how many people are uninsured at a point in time.

<sup>2</sup> The Census Bureau had originally planned to discontinue asking the original health insurance questions in the March 2015 CPS; however, the spending bill passed by Congress in December 2014 requires that they continue to ask them, in addition to the new questions. This requirement poses a substantial challenge for the Census Bureau, which had re-designed its data collection to reflect the new questions. As a result, it is unclear what will actually be asked about health insurance in the March 2015 CPS.

Weaknesses include a relatively small set of variables; the ACS includes measures of disability but no measures of general health status or access to care. The ACS has not been widely used for analyses of health insurance to date but should be very popular for evaluating changes in coverage as a result of ACA, particularly in light of the changes to the CPS health insurance question discussed above.

*If you want a little more health information but don't need a panel: The NHIS, NHANES, and BRFSS*

Three surveys – the NHIS, the NHANES, and the BRFSS – provide more data on health. Each of these surveys entails different tradeoffs relative to the CPS. All three have much less detail on income or labor supply although they do contain, at a minimum, indicators of whether or not adult respondents are working and categorical measures of family income. Insurance information in BRFSS is limited to a 0/1 indicator of coverage; NHIS and NHANES have extensive health insurance information. BRFSS has a very large sample size (almost half a million) and the capacity to make state estimates, but it also has much lower response rates than the other studies (less than 50 percent) and uses a telephone number listing rather than household listing for sampling. Moreover, CDC recommends against using BRFSS data from before 2011 as a baseline because of changes to weighting in that year.<sup>3</sup> In spite of these limitations,

---

<sup>3</sup> Centers for Disease Control and Prevention. “Behavioral Risk Factor Surveillance System: 2011 Summary Data Quality Report, Version #5.” Revised February 4, 2013. Available at [http://www.cdc.gov/brfss/pdf/2011\\_Summary\\_Data\\_Quality\\_Report.pdf](http://www.cdc.gov/brfss/pdf/2011_Summary_Data_Quality_Report.pdf).

BRFSS has been used in a number of influential papers documenting health improvements following expansions of public coverage [11, 12]. NHIS has a large sample size (more than 100,000), while NHANES has a sample of 13,000. For both NHIS and NHANES, state identifiers are restricted access and these data must be used at a Census Research Data Center. In the NHANES, a significant limitation for policy evaluation purposes is that it is not necessarily possible to identify the year in which an interview occurred; for example, data from 2013 and 2014 are pooled. NHIS has been used in a number of influential papers in health economics [13, 14].

*Short panels: SIPP, MEPS – household component, and MCBS*

The MEPS, the SIPP, and the MCBS are all panel studies with a panel length of 2 to 4 years. Many studies use these data not for their panel nature, but simply because they provide rich cross-sectional data. The MEPS, sponsored by the Agency for Health Care Research and Quality, is perhaps the most reliable data source for household-level health spending, with extensive imputation and provider follow-backs to verify charges. The MEPS has an overlapping two-year panel so that at any time, half of the sample is in its first year of participation and half is in its second year. The MEPS has a relatively small sample size compared to some of the other surveys, with about 31,000 individuals at a point in time. But the MEPS data are incredibly rich, with multiple measures over the course of a year on health insurance, health, use of medical care, access to care, employment and income.

As a source of information on health and health care, the SIPP is somewhat idiosyncratic but very valuable. As the name suggests, the original goal of the SIPP was to provide high-frequency data on income and program use. The core SIPP questions include health insurance coverage, income, and labor supply. The idiosyncrasy of the SIPP for studying health derives from the fact that valuable information on health, such as access to care or out-of-pocket medical spending, is often found not in the core questions but in the SIPP's "Topical Modules" which are not always predictable in their timing. The SIPP panels themselves also vary in length, and the current design has no overlap in panels; a 2008 panel will end in late 2013, and a new panel will begin in 2014. The timing of the new panel will complicate using the SIPP to evaluate the implementation of the Affordable Care Act. The SIPP also has a reputation for being somewhat hard to use. For example, the technical documentation for the 2008 SIPP begins with the following statement:

The use of the SIPP public use data over the past four years has taught us a number of lessons. Foremost in those lessons is that the relational file structure is too complex for nearly all users. A close second is that the rectangular file, developed to simplify the relational file, is still too complicated for most users.

In my view most of these difficulties are inherent to the use of data from a complex panel study, and the SIPP remains an underused resource for health economists and health services researchers.

Finally, the MCBS contains detailed cost, utilization, health, and access information for Medicare beneficiaries. The MCBS data are made available in two “modules” in each year: access to care and cost & use. Medicare claims data are also available. Despite the richness of these data, the MCBS could also be described as underused. This may be due to the fact that it is necessary to (a) obtain approval from CMS to use the data and (b) pay \$600 per module per year.<sup>4</sup> The other datasets discussed in this review have no such requirements; all are available for download, with the exception of some items that are noted as restricted access (for example, state identifiers in the NHIS or claims data linked to the HRS), and none charges users a fee.

*Ongoing long panel/cohort studies: Panel Study of Income Dynamics (PSID), National Longitudinal Studies (NLS), Health and Retirement Study (HRS) and National Longitudinal Study of Adolescent Health (Add Health)*

Long-running panel and cohort studies represent an exceptionally rich data resource for health policy and health economics research. The tradeoff here, relative to the cross-sectional studies described above, is richer data on a relatively small number of individuals or families. In this discussion, I will focus on the PSID, NLSY, HRS, and Add Health. The PSID and NLSY are longitudinal panels initiated to understand labor market and income dynamics over the life course. PSID includes respondents of all ages, while NLSY focuses on specific cohorts, following them from adolescence onward.

---

<sup>4</sup> <http://www.resdac.org/cms-data/file-family/Medicare-Current-Beneficiary-Survey-MCBS>, visited on 7/29/2014.

Neither study originally included a particular focus on health insurance and health care, although both have collected more data on this over time. In contrast, HRS and Add Health are explicitly intended to study health-related issues, but each samples a particular age group: older Americans in the case of HRS and younger Americans in the case of Add Health.

The National Longitudinal Studies (NLS) comprise a family of surveys following different cohorts of Americans from adolescence into adulthood, with annual or bi-annual interviews. Most NLS interviews include questions about the respondent's health insurance coverage, and some information about health is also obtained: primarily work-limiting disabilities for younger cohorts, but with an increasing emphasis on general health when a cohort of respondents is beyond the age of 40. The NLS has very little information on access to care or use of medical care, however, and has been relatively less used to study health insurance and health care than most of the other datasets considered in this review as a result.

The PSID started in 1968 with a sample of about 4,800 families; these families, and their "descendants" – the families formed by the children and grandchildren of the original PSID cohort – have been interviewed biannually since then. Health and health insurance were not a focus of PSID in most early waves of the study, but much more is available starting in 1999. Health insurance estimates in the PSID benchmark well to those in the MEPS (Levy 2007). The comparative advantage of a very long-running

panel is the ability to identify the long-run impacts of early life circumstances, a particularly important area of inquiry for health. PSID has been the basis of several important studies in this area [15, 16].<sup>5</sup>

The HRS began in 1992 with a sample of about 7,000 older couples (at least one member born between 1931 and 1941). Over time, the sample has expanded so that it is nationally representative of individuals over the age of 50; the current cohort of respondents includes about 20,000 individuals, who are interviewed biannually. The strengths of the HRS include detailed income, wealth, and self-reported health data collected in every wave; physical measures and biomarkers (e.g. grip strength and blood pressure) collected every other wave; and a linkage to Medicare claims data. The weaknesses, for purposes of the questions considered here, include the absence of data on individuals under age 50 and the fact that state identifiers are not publicly available. Nonetheless, the HRS has been used in a series of important papers documenting the impact of Medicare on health at age 65 [17-23] as well as several papers exploring the relationship between health and wealth [24-27].

The Add Health study was first conducted during the 1994-1995 school year. An in-school questionnaire was administered to more than 90,000 students. Approximately 20,000 of these students subsequently completed an in-home interview, including questionnaires for both parents and children, with extensive information on health and

---

<sup>5</sup> There is a lot on disability as well; I am assuming the labor market chapter covers this, but can add here if necessary.

health behaviors. The parent interview asked about health insurance status.

Approximately 15,000 respondents were re-interviewed in 1996, 2001-2002, and 2006-2007. Since the second wave (1996) of the survey did not include a parent interview, no health insurance information was collected. Therefore, Add Health has health insurance data on adolescents and young adults at three points in time, spaced 5 to 7 years apart. Given the volatility in health insurance coverage among young adults and the numerous significant transitions that occur from year to year (aging out of Medicaid coverage, gaining one's own coverage through a job, gaining dependent spousal coverage through marriage) this window is too long to answer many questions about the dynamics of health insurance for young adults. As a result, while Add Health represents a tremendous resource for studying the health trajectories of young adults, its usefulness for studying health insurance is more limited.

*Other data: vital statistics, health insurance claims, and social experiments*

Survey data are not the whole story. To give some context for the relative utility of survey data, Table 3 lists the recipients of the Arrow Award for best paper in health economics from the International Health Economics Association since the award began in 1993. Eight of the twenty papers rely on survey data, and nearly all of the surveys discussed in this review are represented at least once. Thinking beyond just this list of papers, many important papers have used vital statistics data, sometimes in conjunction with survey data [2, 28-30], Medicare claims data [31, 32], or hospital discharge data

[14]. Data from sources like these will surely play an important role in the evaluation of the Affordable Care Act. New administrative data – for example, data on enrollment in health insurance exchanges – may play an important role as well, although the extent to which such data will be developed and made available for research remains unclear.

Social experiments have also played a very important role in health policy. The RAND Health Insurance Experiment [33] and, more recently, the Oregon Health Insurance Experiment [34-36] provide the best information we have on the impact of health insurance on health and economic well-being. There is a longstanding debate about the adequacy of econometric methods to address selection in non-experimental data (including survey data) and the relative merits of social experiments modeled on randomized controlled trials [37, 38].<sup>6</sup> The relevance of this debate for population-based surveys is that these surveys should be designed with the explicit goal of supporting credible identification. I will return to this point below in discussing the covariates that a new national household survey should include in order to be most useful.

#### **4. Does the nation need a new national household panel study to address important questions about health and health care?**

Will we be able to answer the questions outlined in Section 2 using the data described above? I believe that the answer to this question is a qualified yes. Given the current arsenal of data sets reviewed above, the most significant gap may be the

---

<sup>6</sup> A recent debate (2014) between two eminent scholars presenting both sides of this issue is here: <http://www.nyudri.org/initiatives/deaton-v-banerjee/>

absence of an ongoing panel study enrolling new cohorts of children or adolescents and following them through young adulthood and into middle age. This is a particularly significant omission for studying health insurance since the risk of uninsurance has historically been highest among young adults [39]. High rates of uninsurance among this age group led to a provision in the Affordable Care Act requiring insurers who provide dependent coverage to give policyholders the option of keeping their children on their plan up to age 26, a change that is estimated to have increased coverage for several million young adults [40, 41].

A new national household panel study, or an expansion in the scope of one of our existing panel studies, would greatly improve our understanding of the dynamics of health, health insurance, and economic well-being during the transition to adulthood and during the prime working years of an individual's thirties and forties. The HRS does an excellent job (in my admittedly biased view) of covering these issues for individuals ages over the age of 50; however, 50 may be too late to start collecting data in the sense that many of the underlying processes that will govern the dynamics of health and well-being in older age are already in place. For example, more than one-quarter of individuals in their early fifties already have already experienced the onset of at least major chronic disease (cancer, diabetes, heart disease, lung disease, or stroke); if we add hypertension to that list, the fraction rises to one-half (author analysis of 2010 HRS data). Indeed, substantial evidence exists that these processes have their roots in

childhood and infancy [42, 43], emphasizing the need for intergenerational panel data so that we can understand parental contributions to children's adult outcomes. A long panel study would allow us to understand life-course trajectories of health behaviors and health outcomes. It would also help to illuminate the long-run impacts of different policies. In the case of health care, it seems likely that many behaviors that are the targets of current policy, such as the ACA's focus on increasing the use of preventive services, will affect health outcomes significantly in the long run but not necessarily enough in the short run to be measurable in a short-panel (2-4 year) survey. Indeed, one caveat for both the RAND and the Oregon studies mentioned above is that their failure to find an effect, on average, of health insurance on health outcomes may be due to relatively short follow-up (3 to 5 years in RAND and 2 years, so far, in Oregon). A long-running panel study also means that we are ready to assess the impact of changes in policy that we cannot anticipate; the marginal value of ongoing panel data is increased by a rich stock of data on the history of the individuals and households in the study.

For all of these reasons, either a new panel study or modifications to existing studies would be very valuable in terms of our ability to monitor the impact of policy changes. In a nutshell, there are three options, ordered from most to least ambitious: (1) begin a new panel study; (2) significantly expand the scope of existing panel studies (e.g. lower the eligibility age for inclusion in the HRS sample; introduce bi-annual Add

Health cohorts); (3) make smaller fixes to existing surveys, including cross-sectional and short panel surveys, to allow us to address some though not all of these questions.

If we were to undertake a new panel study – or to significantly expand existing panels to achieve the same result – what information should such a study collect? Table 4 summarizes the key information that should be collected in different domains. In addition to essential information on health insurance coverage, access to and use of services, and physical health, we would need data on income, assets, consumption, employment, and material hardship that would paint a more complete picture of the individual or family's economic situation. Moreover, in many cases, whether or not research using survey data can credibly address the identification probably relies on the availability of specific covariates. For example, studies using regression discontinuity to identify the impact of Medicare on health outcomes require exact age [14]. In terms of the Affordable Care Act, variation across states will provide much of the basis for evaluation. The Supreme Court decision in June 2012 to allow states to decide whether or not to expand their Medicaid programs for childless adults represents a gift to social scientists (at the expense of approximately 6 million very low-income adults) – but this evaluation approach can only be realized with data that make state identifiers available, a limitation of many current studies.

Finally, linkages to other data can dramatically increase the usefulness of survey data. Even if a linkage is not in the works when a survey is conducted, the data

collection and policies governing access to them should be developed with the potential for linkages in mind. Ideally, the data should be linked to Medicare claims; Medicaid claims; and Social Security records, including earnings, benefit receipt, and DI/SSI receipt. Survey data could be linked to administrative data on coverage and subsidy receipt through health insurance exchanges. Administrative records of SNAP (Food Stamp) and TANF receipt would also be useful.

What about smaller fixes to existing surveys? If a new panel study is not feasible, are there smaller changes to existing surveys that could help achieve some of the same goals? The short answer is yes. Table 6 summarizes some relatively minor changes (for example, asking more detailed health insurance questions in the BRFSS, or following splitoff households in the SIPP) that would help fill some of the same knowledge gaps that a new long panel study would, presumably at lower cost.

## **5. Conclusion**

The Affordable Care Act ushers in a period of change in health insurance and health care on a scale not seen since the implementation of Medicare and Medicaid in the mid-1960s. More changes are likely to come in relatively short order, particularly to Medicare and Medicaid. Whether the social science community will be able to inform the design of policies before they are enacted and to evaluate their impact once they have been implemented depends in large part on investments we make now in our data infrastructure.

**Acknowledgements:** The author gratefully acknowledges helpful comments from Charlie Brown, Sherry Glied, Bob Schoeni, an anonymous reviewer, and participants in the workshop “Assessing the Need for a New Nationally Representative Household Panel Survey in the United States,” held June 5-6, 2014 in Washington DC. Financial support from the National Institute on Aging (grant number NIA K01AG034232) is also gratefully acknowledged.

## References

1. Cutler, D.M. and J. Gruber, *Does public insurance crowd out private insurance?* The Quarterly Journal of Economics, 1996. **111**(2): p. 391-430.
2. Currie, J. and J. Gruber, *Health insurance eligibility, utilization of medical care, and child health.* The Quarterly Journal of Economics, 1996. **111**(2): p. 431-466.
3. Currie, J. and J. Gruber, *Saving babies: The efficacy and cost of recent changes in the Medicaid eligibility of pregnant women.* Journal of Political Economy, 1996. **104**(6): p. 1263-1296.
4. Baicker, K. and A. Chandra, *The Labor Market Effects of Rising Health Insurance Premiums.* Journal of Labor Economics, 2006. **24**(3).
5. Swartz, K. and T.D. McBride, *Spells without health insurance: distributions of durations and their link to point-in-time estimates of the uninsured.* Inquiry: a journal of medical care organization, provision and financing, 1989. **27**(3): p. 281-288.
6. Meyer, B.D., W.K. Mok, and J.X. Sullivan, *The under-reporting of transfers in household surveys: its nature and consequences.* 2009, National Bureau of Economic Research.
7. Davern, M., et al., *An examination of the Medicaid undercount in the Current Population Survey: preliminary results from record linking.* Health services research, 2009. **44**(3): p. 965-987.
8. Klerman, J.A., et al., *Understanding The Current Population Survey's Insurance Estimates And The Medicaid 'Undercount'.* Health Affairs, 2009. **28**(6): p. w991-w1001.
9. Call, K.T., et al., *Medicaid undercount and bias to estimates of uninsurance: new estimates and existing evidence.* Health Services Research, 2008. **43**(3): p. 901-914.
10. Pascale, J., M.I. Roemer, and D.M. Resnick, *Medicaid Underreporting in the CPS Results from a Record Check Study.* Public opinion quarterly, 2009. **73**(3): p. 497-520.
11. Sommers, B.D., K. Baicker, and A.M. Epstein, *Mortality and access to care among adults after state Medicaid expansions.* New England Journal of Medicine, 2012. **367**(11): p. 1025-1034.
12. WEES, P.J., A.M. ZASLAVSKY, and J.Z. AYANIAN, *Improvements in Health Status after Massachusetts Health Care Reform.* Milbank Quarterly, 2013. **91**(4): p. 663-689.
13. Card, D., C. Dobkin, and N. Maestas, *The impact of nearly universal insurance coverage on health care utilization and health: evidence from Medicare.* American Economic Review, 2008. **98**(5): p. 2242-2258.
14. Card, D., C. Dobkin, and N. Maestas, *Does Medicare save lives?* The quarterly journal of economics, 2009. **124**(2): p. 597-636.
15. Johnson, R.C. and R.F. Schoeni, *Early-life origins of adult disease: national longitudinal population-based study of the United States.* American journal of public health, 2011. **101**(12): p. 2317.
16. Johnson, R.C. and R.F. Schoeni, *The influence of early-life events on human capital, health status, and labor market outcomes over the life course.* The BE journal of economic analysis & policy, 2011. **11**(3).
17. McWilliams, J.M., et al., *Impact of Medicare coverage on basic clinical services for previously uninsured adults.* JAMA, 2003. **290**(6): p. 757-764.
18. McWilliams, J.M., et al., *Health of previously uninsured adults after acquiring Medicare coverage.* JAMA, 2007. **298**(24): p. 2886-2894.
19. McWilliams, J.M., et al., *Health insurance coverage and mortality among the near-elderly.* Health Affairs, 2004. **23**(4): p. 223-233.
20. McWilliams, J.M., et al., *Medicare spending for previously uninsured adults.* Annals of internal medicine, 2009. **151**(11): p. 757-766.

21. McWilliams, J.M., et al., *Commentary: assessing the health effects of Medicare coverage for previously uninsured adults: a matter of life and death?* Health services research, 2010. **45**(5p1): p. 1407-1422.
22. Polsky, D., et al., *The Health Effects of Medicare for the Near-Elderly Uninsured*. Health services research, 2009. **44**(3): p. 926-945.
23. Polsky, D., et al., *Response to McWilliams Commentary: "Assessing the Health Effects of Medicare Coverage for Previously Uninsured Adults: A Matter of Life and Death?"*. Health services research, 2010. **45**(5p1): p. 1423-1429.
24. Dranove, D. and A. Sfekas, *Start spreading the news: a structural estimate of the effects of New York hospital report cards*. Journal of Health Economics, 2008. **27**(5): p. 1201-1207.
25. Adams, P., et al., *Healthy, wealthy, and wise? Tests for direct causal paths between health and socioeconomic status*. Journal of Econometrics, 2003. **112**(1): p. 3-56.
26. French, E. and J.B. Jones, *The Effects of Health Insurance and Self-Insurance on Retirement Behavior*. Econometrica, 2011. **79**(3): p. 693-732.
27. French, E., *The effects of health, wealth, and wages on labour supply and retirement behaviour*. The Review of Economic Studies, 2005. **72**(2): p. 395-427.
28. Currie, J. and J. Gruber, *Saving Babies: The Efficacy and Cost of Recent Changes in the Medicaid Eligibility of Pregnant Women*. Journal of Political Economy, 1996: p. 1263-1296.
29. Meyer, B.D. and L.R. Wherry, *Saving teens: Using a policy discontinuity to estimate the effects of medicaid eligibility*. 2012, National Bureau of Economic Research.
30. Finkelstein, A. and R. McKnight, *What did Medicare do? The initial impact of Medicare on mortality and out of pocket medical spending*. Journal of public economics, 2008. **92**(7): p. 1644-1668.
31. Kessler, D. and M. McClellan, *Do doctors practice defensive medicine?* The Quarterly Journal of Economics, 1996. **111**(2): p. 353-390.
32. Kessler, D.P. and M.B. McClellan, *Is hospital competition socially wasteful?* The Quarterly Journal of Economics, 2000. **115**(2): p. 577-615.
33. Newhouse, J.P. and R.C.I.E. Group, *Free for all?: lessons from the RAND health insurance experiment*. 1993: Harvard University Press.
34. Baicker, K. and A. Finkelstein, *The effects of Medicaid coverage—learning from the Oregon experiment*. New England Journal of Medicine, 2011. **365**(8): p. 683-685.
35. Baicker, K., et al., *The Oregon experiment—effects of Medicaid on clinical outcomes*. New England Journal of Medicine, 2013. **368**(18): p. 1713-1722.
36. Finkelstein, A., et al., *The Oregon Health Insurance Experiment: Evidence from the First Year\**. The Quarterly Journal of Economics, 2012. **127**(3): p. 1057-1106.
37. LaLonde, R.J., *Evaluating the econometric evaluations of training programs with experimental data*. The American Economic Review, 1986: p. 604-620.
38. A Smith, J. and P. E Todd, *Does matching overcome LaLonde's critique of nonexperimental estimators?* Journal of econometrics, 2005. **125**(1): p. 305-353.
39. Levy, H., *Health insurance and the transition to adulthood*. The price of independence: the economics of early adulthood, 2007: p. 84-106.
40. Sommers, B.D., et al., *The Affordable Care Act has led to significant gains in health insurance and access to care for young adults*. Health affairs, 2013. **32**(1): p. 165-174.
41. Antwi, Y.A., A.S. Moriya, and K. Simon, *Effects of Federal Policy to Insure Young Adults: Evidence from the 2010 Affordable Care Act's Dependent-Coverage Mandate*. American Economic Journal: Economic Policy, 2013. **5**(4): p. 1-28.
42. Currie, J., *Healthy, wealthy, and wise: Is there a causal relationship between child health and human capital development?* Journal of Economic Literature, 2009. **47**(1): p. 87-122.

43. Almond, D. and J. Currie, *Killing me softly: The fetal origins hypothesis*. The Journal of Economic Perspectives, 2011: p. 153-172.

Table 1  
Summary of large-scale surveys that collect information on health insurance  
What other data do they include?

	Frequency & sample restrictions	Current sample size <sup>1</sup>	Health care use	Health status	Physical measures/ biomarkers	State IDs
<b>Repeated cross-section surveys</b>						
ACS	Annual	5,600,000	No	No	No	Yes
CPS	Annual	205,000	No	Yes	No	Yes
NHIS	Annual	108,000	Yes	Yes	No	Restricted
NHANES	Annual	13,000	Yes	Yes	Yes	Restricted
BRFSS	Annual; adults only	477,000	Yes	Yes	No	Yes
<b>Short panel studies (2-4 years)</b>						
MCBS	4 year panel; only Medicare enrollees	18,000	Yes	Yes	No	Restricted
MEPS	2 year panel	31,000	Yes	Yes	No	Restricted
SIPP	2.5-4 year panel	100,00	Yes <sup>2</sup>	Yes <sup>2</sup>	No	Yes
<b>Long panel/cohort studies</b>						
PSID	Every 2 years since '68 Original cohort + their descendants	10,000	Yes	Yes	No	Yes
NLS	Several different birth cohorts; interviewed annually or bi-annually starting in adolescence	10,000	Yes	Yes	No	Restricted
Add Health	Adolescents (grades 7-12) in 1994/5, re-interviewed in 1996, 2002, and 2007	15,000	Yes	Yes	Yes	Restricted
HRS	Interviewed bi-annually since '92 Ages 51+ only	20,000	Yes	Yes	Yes	Restricted

Notes:

All datasets in this table include at least some information on household income.

<sup>1</sup>Approximate, especially for cohort studies since it depends on cohort & wave

<sup>2</sup>Some waves only

Table 2  
 Deficit Reduction Options Related to Health  
 Adapted from “Options for Reducing the Deficit, 2013 – 2014” (CBO, November 2013)

Option Number	Title	Savings, 2014–2023 (Billions of Dollars)
	<i>Mandatory spending</i>	
1	<a href="#">Impose Caps on Federal Spending for Medicaid</a>	105 to 606
2	<a href="#">Add a “Public Plan” to the Health Insurance Exchanges</a>	158
3	<a href="#">Eliminate Exchange Subsidies for People With Income Over 300 Percent of the Federal Poverty Guidelines</a>	109
4	<a href="#">Limit Medical Malpractice Torts</a>	64
5	<a href="#">Introduce Minimum Out-of-Pocket Requirements Under TRICARE for Life</a>	31
6	<a href="#">Convert Medicare to a Premium Support System</a>	22 to 275
7	<a href="#">Change the Cost-Sharing Rules for Medicare and Restrict Medigap Insurance</a>	114
8	<a href="#">Raise the Age of Eligibility for Medicare to 67</a>	19
9	<a href="#">Increase Premiums for Parts B and D of Medicare</a>	287
10	<a href="#">Bundle Medicare’s Payments to Health Care Providers</a>	17 to 47
11	<a href="#">Require Manufacturers to Pay a Minimum Rebate on Drugs Covered Under Part D of Medicare for Low-Income Beneficiaries</a>	123
	<i>Discretionary spending</i>	
12	<a href="#">Modify TRICARE Enrollment Fees and Cost Sharing for Working-Age Military Retirees</a>	20 to 71
13	<a href="#">Reduce or Constrain Funding for the National Institutes of Health</a>	13 to 28
14	<a href="#">End Enrollment in VA Medical Care for Veterans in Priority Groups 7 and 8</a>	48
	<i>Revenues</i>	
15	<a href="#">Reduce Tax Preferences for Employment-Based Health Insurance</a>	240 to 537
16	<a href="#">Increase the Excise Tax on Cigarettes by 50 Cents per Pack</a>	37

Table 3  
What data did they use?

Winners of the Arrow award for best paper in health economics from the International Health Economics Association, 1993-2013

<b>Year</b>	<b>Author &amp; title</b>	<b>Data source</b>
1993	Richard Hirth. Nursing Home Quality: Roles of Information and Ownership	Theory; no data
1994*	Phillip Cook and Michael Moore. Drinking and schooling	NLSY
1995*	Jonathan Gruber. The Incidence of Mandated Maternity Benefits	1977 NMCES (precursor to MEPS) and CPS
1996*	Martin Gaynor and Paul Gertler. Moral Hazard and Risk Spreading in Partnerships	1978 national survey of medical groups, plus AHA & AMA data
1997	Daniel Kessler and Mark McClellan. Do Doctors Practice Defensive Medicine?	Medicare claims data
1998	Ching-To Albert Ma and Thomas G. McGuire. Optimal Health Insurance and Provider Payment	Theory; no data
1999*	Donna B. Gilleskie. A Dynamic Stochastic Model of Medical Care Use and Work Absence	1987 NMES (precursor to MEPS)
2000*	Will Dow, Tomas J. Philipson and Xavier Sala-i-Martin. Longevity Complementarities Under Competing Risks	Health surveys in several African countries
2001	David M. Cutler, Mark McClellan and Joseph P. Newhouse. How Does Managed Care Do It?	Claims data from a single large firm plus all hospital discharge data from Mass. In 1994 and 1995
2002	Willard G. Manning and John Mullahy. Estimating Log Models: To Transform or Not to Transform?	Simulated data
2003*	Anne Case, Darren Lubotsky and Christina Paxson. Economic Status and Health in Childhood: The Origins of the Gradient	NHIS, PSID, and NHANES
2004	Kenneth Chay and Michael Greenstone. The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession.	EPA data; mortality & natality data from NCHS
2005	Edward Miguel and Michael Kremer. Worms:	Evaluation of Kenyan

	Identifying impacts on education and health in the presence of treatment externalities	program with randomized design
2006	Gary S. Becker, Tomas J. Philipson, and Rodrigo R. Soares. The Quantity and Quality of Life and the Evolution of World Inequality	Country-level data from Penn World Tables, World Bank, WHO
2007*	Kevin M. Murphy and Robert H. Topel. The Value of Health and Longevity	Theory; CDC vital statistics; 1977 NMCES and 1987 NMES
2008	Amitabh Chandra and Doug Staiger. Productivity Spillovers in Health Care: Evidence from the Treatment of Heart Attacks	Medicare claims data
2009*	Hanming Fang, Michael P. Keane, and Dan Silverman. Sources of Advantageous Selection: Evidence from the Medigap Insurance Market	MCBS and HRS
2010	Kate Ho. Insurer-Provider Networks in the Medical Care Market	Characteristics of health insurance plans from commercial sources, HEDIS, and CAHPS
2011	Carol Propper and John Van Reenen. Can Pay Regulation Kill?	Panel data on hospitals in England from administrative sources
2012	Amy Finkelstein et al. The Oregon Health Insurance Experiment: Evidence from the First Year	Administrative and custom survey data from randomized Medicaid expansion
2013	Jonathan Kolstad. Information and Quality when Motivation Is Intrinsic: Evidence from Surgeon Report Cards	Data on bypass surgery from Pennsylvania Health Care Cost Containment Council

\* Paper used survey data.

Table 4

What data should be collected in a new panel study?

Domain			
Health insurance	Health care	Health	Other
Point-in-time coverage of all household members	Use of acute care services (doctor visit, dentist visit, ER visit, hospital stay in past year)	Self-reported health and the presence of major chronic conditions, including a short screen for depression	Income
Who is policyholder			Assets
Source of coverage (employer, insurance company, Medicare, Medicaid, etc.)	Use of preventive services such as mammograms, colonoscopies, prostate-specific antigen (PSA) testing, and flu shots	Health behaviors, including diet, exercise, tobacco and alcohol use	Consumption/spending
If uninsured, when did respondent last have insurance?			Labor supply, including hours and sector (self-employed versus wage & salary)
If insured, have there been any spells without insurance since last wave?	Self-reported access barriers (financial/non-financial), including medication non-adherence	Work/activity limitations for everyone; for the elderly, functional status (ADL/IADL or Nagi limitations)	Material hardship, including food insecurity
	Does respondent have a usual source of care	Biomarkers and physical measures (blood pressure, BMI)	Exact date of birth
		Cognitive ability	Exact age as of survey date
		Subjective life expectancy	State of residence

Table 5

Some smaller fixes that would improve the usefulness of national surveys for questions related to health economics and policy

<b>Survey</b>	<b>Proposed change</b>
ACS, CPS, SIPP, MEPS, NHIS, NHANES, BRFSS	Link to administrative data: <ul style="list-style-type: none"> <li>• Medicare and Medicaid claims</li> <li>• Health insurance exchange data</li> <li>• Social Security records</li> <li>• Food Stamps</li> <li>• TANF benefits</li> </ul>
HRS, PSID	Link to health insurance data, Food Stamps, TANF
All but PSID	Add measures of consumption/expenditures
All but HRS and PSID	Add measures of wealth
ACS	Add self-reported health status
SIPP, MEPS	Follow splitoff households
SIPP	Adopt overlapping panel design
NHIS, NHANES, HRS	Include state on public use file; expand sample if necessary
BRFSS	Ask about source of health insurance (e.g. ACS question)

Note: Surveys considered in this table include ACS, CPS, SIPP, MEPS, NHIS, NHANES, BRFSS, PSID, NLS, HRS, and Add Health.