

UNIVERSITY OF MINNESOTA

Twin Cities Campus

*Minnesota Population Center
Office of the Director*

*50 Willey Hall
225 19th Avenue South
Minneapolis, MN 55455
612-624-5818
Fax: 612-626-8375
<http://www.pop.umn.edu>*

May 9, 2015

Sent via email: OIRA_Submission@omb.eop.gov

RE: Proposed Information Collection: Comment Request: The American Community Survey
Content Review Results 80 FR 23501/April 28, 2015/pp. 23501-23503

To whom it may concern:

I am writing to respond to the request for comments about the content of the American Community Survey (ACS). I applaud the Census Bureau's recommendation to retain key questions about marital history and college field of degree in the 2016 ACS. I am nevertheless deeply troubled by the suggestion that the marital history question in the future may not be available on an annual basis. An even graver concern is that there is no indication that the Census Bureau intends to fix its procedure for evaluating the benefit of ACS questions. The current procedure is so deeply flawed that it identifies many of the most important ACS questions as low benefit, and classifies some of the most expendable and high-cost questions as high benefit.

To understand why the marital history questions must be asked on an annual basis, it is important to understand why they were introduced and what they are used for. As a member of the Census Scientific Advisory Committee (CSAC), I presented a detailed comment describing this history at the April 17 meeting of CSAC at Census Bureau headquarters. I attach the text of that report as an appendix to this memo, and the video of the meeting may be viewed [here](#) (Day 2, Part 5).

Very briefly, the marital history questions were introduced at the request of the Department of Health and Human Services to address a pressing need for reliable data on trends in divorce, marriage, and widowhood. The measurement strategy was carefully designed and thoroughly tested over four years before being introduced in 2008, and it yields highly reliable data on age-specific marriage rates and duration and age-specific divorce and widowhood rates. The ACS is the only reliable source available for annual U.S. statistics on marriage, divorce, and widowhood. The data have many Federal and non-Federal uses. Especially notable are the population projections of the Social Security Administration and the Congressional Budget Office, which are essential for estimating future entitlements. Any effort to project future incomes for retirees—and the costs of providing those incomes—must take marriage, divorce, widowhood, and remarriage into account. Without up-to-date rates broken down by age, marriage duration, and marriage order, such population projections will be inaccurate. This in turn will lead to unreliable data on the solvency of the Social Security and Health Care systems.

The ingenious measurement strategy used to obtain those data depends on measuring marital events—marriage, divorce, or widowhood—within the past year. Limiting the questions to events within one year of the survey reduces recall bias and respondent burden. Accuracy is high, even when respondents are reporting on the marriage or divorce of another household

member. But there is one disadvantage of this measurement strategy: it requires a very large number of cases, since a relatively small proportion of the population experiences a marital event within the previous year. The only survey large enough to accommodate this measurement strategy was the ACS, which is why that survey was chosen for these questions.

Because a very large number of cases is needed for analysis of the marital history questions, they were not designed for sub-state analysis. In many cases, sub-state analysis would yield unacceptably high standard errors. For example, the 2013 ACS includes just 45 marriages in Vermont, and just 21 divorces in South Dakota. With numbers like these, sub-state analysis is impractical. Indeed, even to get reliable age-specific rates at the state level, when looking at smaller states it is necessary to combine multiple ACS years.

There are two major reasons why asking the marital history questions intermittently is a very bad idea. First, we would lose our only source of annual estimates of marriage, divorce, and widowhood rates, making it impossible to quickly adjust our projections to changing trends in behavior. Second, we would lose the ability to make reliable state-level estimates of marriage in the smaller states, or to do sub-state analysis in the larger states. If the intermittent approach is adopted, the United States will become the only industrialized country in the world that lacks annual statistics on these core demographic indicators.

Why did the Census Bureau target the marital history questions for removal in the first place, and why are they still under threat as annual questions? The answer is simple: the ACS Content review was deeply flawed. In particular, the Census Bureau adopted a deeply flawed method for assessing the benefits of each ACS question. The approach focused on whether each question had official federal uses for small-area analysis. There were separate benefit categories for tract level or county level, and for three types of use, but within each category a question got benefit points if there was one or more federal uses. There was no consideration of the *number* of federal uses of the data, the *importance* of those uses, federal uses for any *small population subgroups* not defined by geography, or *non-federal* uses. The lack of alternative data sources was a factor, in the benefit computation, but it only counted for 7.7% of the score, and the analysis did not assess the cost of developing alternative data sources.

In all, 62% of the benefit score was based on the existence of federal uses of each question below the state level. The marital history questions, of course, were not designed for sub-state analysis and cannot generally be used for sub-state analysis, so they were targeted for removal, even though they were judged to be low cost.

This was not the only counter-intuitive result. Indeed, *most* of the questions identified as low benefit belong to the core set of questions that have been asked on the census since 1940 and which form the backbone of the nation's statistical infrastructure. These include:

- Migration (residence 1 year ago)
- School enrollment
- Educational attainment
- Food stamp benefits
- Full or part-time employment (hours worked last week)
- Weeks worked in past year

These are among the most intensively-used questions in the ACS, and it would be absurd to cut them from the survey. How were they spared, even though they fell into the same benefit category as the marital history questions? To preserve these questions, the Census Bureau applied a *post hoc* rule after all the cost and benefit scores were calculated: they decided that any low-benefit questions with required sub-state uses would be spared. This saved all the low-benefit questions except for marital history and college field of degree.

The benefit analyses used by the Census Bureau is a travesty. Marriage, education, and work are considered low-benefit topics, whereas dial-up internet subscriptions, amount of water bill, and value of agricultural sales were all deemed to be high benefit. This benefit analysis is broken, and should never be used to affect the content of the ACS or any other survey.

As described in the Appendix, there are ways to simplify the ACS and eliminate questions that are both high cost and low benefit. The ACS content review process failed to identify those questions, however, and it should be scrapped.

Sincerely yours,

A handwritten signature in black ink, appearing to read "Steven Ruggles". The signature is fluid and cursive, with the first name "Steven" and last name "Ruggles" clearly distinguishable.

Steven Ruggles

President, Population Association of America
Member, Census Scientific Advisory Committee
Member, National Science Foundation Advisory Committee for the Social, Economic, and
Behavioral Sciences
Principal Investigator, Integrated Public Use Microdata Series
Regents Professor of History and Population Studies
Director, Minnesota Population Center, University of Minnesota

Appendix

Census Scientific Advisory Committee Comment on ACS Content Review Steven Ruggles April 17, 2015

Background

Since the nineteenth century, statistics on marriage and divorce were derived from the vital registration system. These data were originally compiled by the Census Bureau, which published detailed annual statistics about marriages and divorces. After World War II the responsibility for marriage and divorce statistics was shifted to a new agency, now known as the National Center for Health Statistics (NCHS). The main job of NCHS is to collect and analyze data on health. Marriage and divorce statistics were never central to the NCHS mission, and after 1990 the statistics on these topics began to deteriorate.

In 1995, the agency concluded that the detailed statistics were no longer reliable, and decided to discontinue collection of detailed data on these topics. This created a huge problem.

Without reliable age-specific and duration-specific data on marriages and divorces, we can't do population projections that include marital status. This wasn't just a problem for a few academic demographers. The projected expenses for the largest programs in the federal government—Social Security and Medicare—depend on marital status.

NCHS continued to collect a raw count of the number of marriages and divorces from each state. By 2005, six states representing 20% of the population had dropped out of the program, and with little funding the quality of reporting continued to decline. NCHS has reported no data whatsoever based on marriage and divorce records since 2012.

We now know what we had long suspected—even the raw counts of marriages and divorces published by NCHS were drastically undercounted. Because of the ACS, we can estimate that the number of divorces published by NCHS is undercounted by about 20%.

From 1995 to 2008, demographers relied on two federal surveys for information about marriage and divorce: NSFG and SIPP. Both of these surveys ask respondents to provide complete marital histories, going back many years, and people often neglect to report previous marriages and divorces. This problem is especially severe in SIPP, probably because it is really a survey about program participation. The long-run marital histories for all household members in SIPP are reported by a single respondent, and they often don't know the details. Almost half of the SIPP divorce dates are imputed! And the divorce rate in SIPP is 20% below the vital statistics, and probably at least a third lower than reality.

NSFG does better. It is a fertility survey. The big problem is it can't be used even to estimate a crude marriage rate or divorce rate, because it doesn't cover the whole population. Much of the recent change has been occurring in the population 45 and older—a population of special interest for health and retirement projections—but NCHS does not cover that population. And the sample sizes are very small--too small for most applications.

A new approach

In 2004, the Department of Health and Human Services (DHHS) decided to pursue an alternative strategy to provide reliable, timely, and cost-effective data that would meet the needs of Federal agencies. DHHS requested that the Census Bureau investigate the feasibility of adding a set of questions to the American Community Survey (ACS) to address the problem.

The ACS is large, covering about 1.5% of the population each year. This large size offered the opportunity for a new strategy for the measurement of marriage and divorce. Instead of asking for complete marital histories like the NSFG and the SIPP, the ACS could be limited to asking about marriages or divorces within the past year.

Asking about events in the past year instead of entire marital histories had huge advantages. First, it reduced the impact of recall bias and virtually eliminated underreporting. Second, it yielded information about current demographic behavior, providing a direct substitute for the vital statistics data. With the addition of the questions on year of marriage and number of times married, the suite of questions allowed calculation of age-specific and duration-specific rates of marriage, divorce, and widowhood, which are the raw materials for population projections. The questions on year of marriage and times married have a long pedigree, and were first asked in 1900 and 1910, respectively. They give us everything we need to build marital status life tables, which provide a detailed look at marriage patterns over the life course, and enables us to estimate the percentage of marriages ending in divorce.

The result was much more powerful than the vital statistics ever were. Not only did the ACS provide age-specific marriage rates and duration-specific divorce rates, it also allowed correlation of marriage, divorce, and widowhood with other variables such as hours worked, fertility, and other social, demographic and economic characteristics. These correlations were incorporated into the projections of future benefits for both Social Security and Medicare, greatly increasing their accuracy.

It was also a gold mine for family demographers. Even though the data have been available for just five years, they have already led to a series of major publications that have reshaped our understanding of marriage and divorce.

But there is one cost to asking about events in the previous year instead of complete marital histories, as is done in SIPP or NSFG. This measurement strategy requires 10 to 20 times as many cases to obtain comparable precision. The ACS is the only survey we have that is large enough to support this measurement strategy.

A flawed cost-benefit analysis

So if these questions were so badly needed and so successful, why are they being targeted for removal from the survey? The answer is a badly flawed cost benefit analysis. The analysis of costs seems pretty good. There were 6 criteria based on a survey of interviews and some testing, and the results seem quite plausible. All the marital history questions were determined to be low cost.

The big problem of the cost-benefit analysis was on the benefit side. The Census Bureau identified 13 benefits, all of which focused on the uses of the data by federal agencies. The critical flaw in the analysis is that 8 of the 13 benefits they identified focused on uses of the data to describe small geographic areas, such as census tracts or counties.

Whoever designed the cost-benefit analysis apparently assumed that the only reason a question would need to be on the ACS would be to generate small-area statistics, because the ACS is the only survey large enough to be used for small-area statistics. The implicit assumption of the ACS benefits analysis is that if a question is not used to analyze small geographic areas, it should be moved from the ACS to a smaller survey.

If the questions on marital events within the past year were moved to a smaller survey, however, they would be useless: only the ACS has sufficient cases to enable this powerful measurement strategy. Because the number of times married and marriage duration questions are necessary to interpret the questions on incidence of marriage, divorce, and widowhood, there is an inherent need for large sample sizes for all five marital history questions. It is an interdependent package of questions.

The benefit analysis determined that no agency was using the marital history questions for small-area analysis. This should have been self-evident: most of the marital history questions cannot be used for small-area analysis!

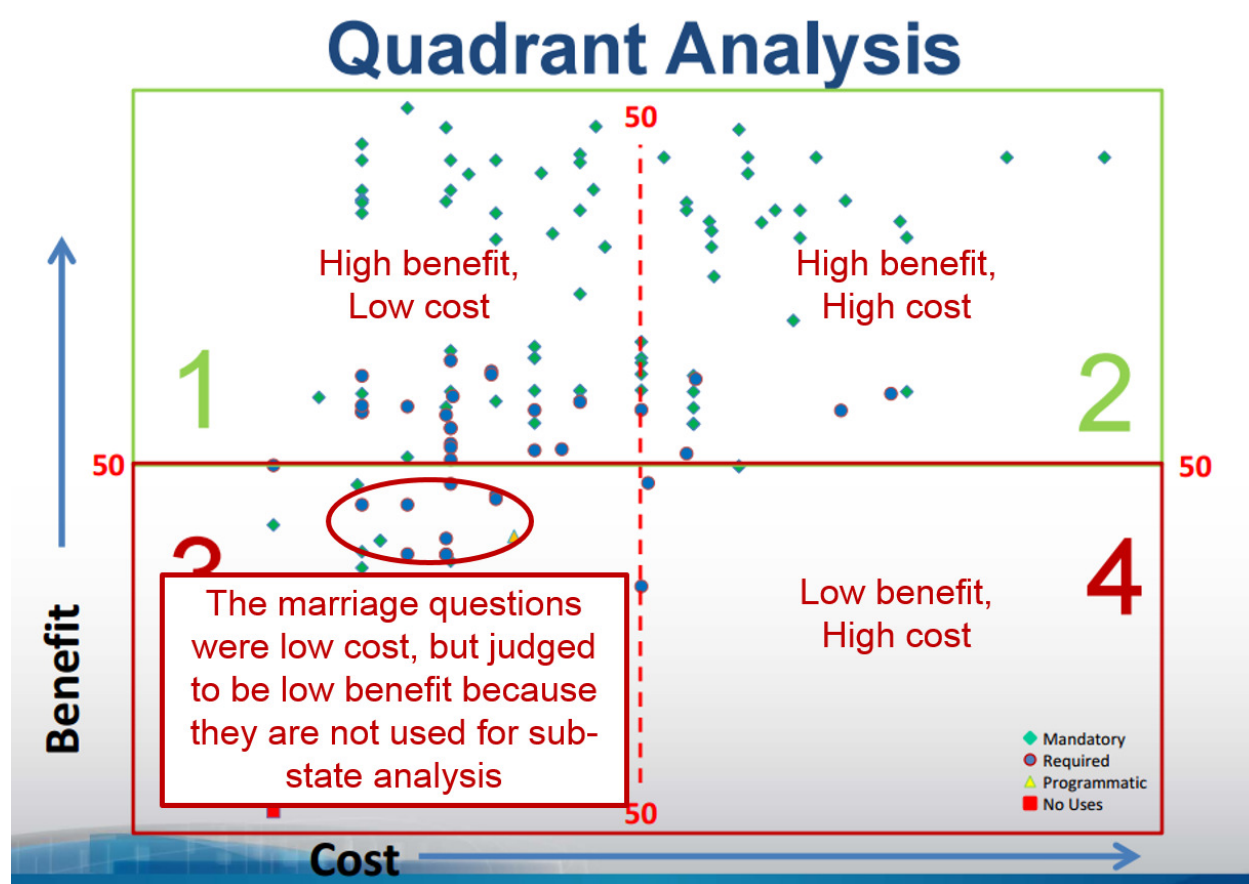
The measurement strategy focusing on marital events within the past year requires a large number of observations, and there simply are not enough cases in the ACS to conduct small-area analysis using these questions. For example, in 2013 the ACS collected data on marriages of just 45 women who resided in Vermont, and divorce information for just 32. Here are the number of cases in nine small states.

	Marriages	Divorces
Vermont	45	32
Alaska	51	27
Wyoming	66	29
North Dakota	74	23
South Dakota	81	21
Delaware	83	38
District of Columbia	87	38
Rhode Island	96	44
Montana	98	50

Some sub-state analysis is probably feasible for the largest states, and could be done for some smaller states by aggregating multiple survey years. But that was never the goal of these questions. They are designed to provide the detailed state and national statistics that are no longer available from NCHS, and they are currently the only available source for those statistics.

If the questions are dropped, the United States will be the only developed country in the world without a reliable annual count of the number of marriages and divorces.

The ACS Content review relied on what they called a “quadrant analysis.” Using some method of aggregating the 6 costs and the 13 benefits, they got a score for each and plotted them on a graph. The marriage questions fell into quadrant 3, which was low benefit and low cost.



This was not the only counter-intuitive result. Indeed, most of the questions identified as low benefit belong to the core set of questions that have been asked on the census since 1940 and which form the backbone of the nation’s statistical infrastructure. These include:

- Migration (residence 1 year ago)
- School enrollment
- Educational attainment
- Food stamp benefits

- Hours worked last week
- Weeks worked in past year

These are among the most intensively-used questions in the ACS, and it would be absurd to cut them from the survey. How were they spared, even though they fell into the same benefit category as the marital history questions? To preserve these questions, the Census Bureau applied an ad hoc rule after all the cost and benefit scores were calculated: they decided that any low-benefit questions with required sub-state uses would be spared. This saved all the low-benefit questions except for marital history and college field of degree.

So the geographic factors effectively counted double. The marriage questions ended up being classified as “low benefit” because they had no sub-state uses. But all the other low-benefit questions survived anyway—only because they had sub-state uses. In effect, this is the sole criterion being used—even though these questions were never intended for sub-state use.

When the marriage questions were added in 2008, the Census Bureau provided a detailed justification, including about 25 statutory uses of the new questions. Among the most important needs are determining state eligibility for performance related funding for TANF. And projections of future entitlements, which depend on marital status. Both SSA and CBO rely on these data to develop and calibrate their projections.

There is no other source for these data. Marriage is now changing faster than at any time in our history. Age at first marriage is soaring. An unprecedented percentage of young people are forgoing marriage altogether. After slowing briefly in the 1980s, divorce is once again accelerating. And 40% of children are born outside marriage. This is not a good moment to stop collecting data on marriage and divorce.

Undergraduate field of degree

I want to briefly comment on one other question proposed for elimination. The question on undergraduate field of degree has been wildly successful, and has been used in hundreds of publications on the impact of educations and the STEM workforce.

At IPUMS, we have received 22,000 requests for this variable. It was proposed by the National Academy of Sciences, and one major rationale was to allow a redesign of the Scientists and Engineers Statistical Data System. The large size of the survey enabled NSF to redesign SESTAT, their statistical system on scientists and engineers. By using the ACS as a sampling frame, they were able to eliminate the separate survey of recent college graduates. NSF spent seven years working on the redesign, which was finally launched in 2013.

I am a member of the National Science Foundation Advisory Committee for the Social, Behavioral, and Economic Sciences. The fall meeting of the Committee was occurring just as the Federal Register notice of cuts was published. It immediately became apparent that the content review team did not consult with the subject experts at NSF before conducting the review; just like the members of CSAC, NSF learned of the

outcome of the review only after it was complete, and never had an opportunity to weigh in on measurement of the benefits of the question.

It was the same kind of quadrant analysis. As in the case of the marriage questions, the field of degree question was eliminated solely because it has no sub-state uses, even though the large number of cases in the ACS is needed to obtain a sufficient sample of recent stem graduates.

Alternative cuts to the ACS

So what should be eliminated from the survey? I have a few obvious candidates. I think the worst is the set of four questions on the cost of gas, electricity, water, and heating fuel. People hate these questions—they are very high cost. They are also low benefit. These questions appear to be unreliable—they differ dramatically from what utilities report. It looks like people really don't know what their bills are, and they just guess. But the Census Bureau could actually guess better. The farm questions could also be cut. The data could probably be obtained from private sources. There are other questions that are fairly redundant. An example is number of rooms. There are definitional issues of what constitutes a room, so the number of bedrooms is generally more reliable. It is not clear that number of rooms provides any useful information beyond the number of bedrooms.

Summary

Because the content review used an inappropriate standard for benefit, it recommended dropping five of the most important and widely used questions from the survey.

If this goes into effect, it will do serious damage to the Nation's statistical infrastructure.

Small-area analysis is not the sole purpose of the ACS. We need large samples to study any sort of small population. Among those are small geographic areas. But they also include people with a demographic event in the previous year, and recent college graduates in STEM fields.