

Overview of Feasibility Assessment of Using Matrix Sampling and Other Methods to Reduce Respondent Burden

Workshop On Respondent Burden in the
American Community Survey

March 9, 2016

Mark E. Asiala

Chief, American Community Survey Statistical Design

Overview

- Background
- Feasibility Report of Reducing Respondent Burden
 - Options considered
 - Assessment criteria
 - Highlights assessment
 - Literature review and empirical work
 - Recommendation

Background

- 2014 Content Review provided a thorough look at the federal uses of the ACS by question
- It documented the needs for each topic by Federal agency by
 - Frequency
 - Level of geography
 - Legal justification
- From that work, recommendations were made whether certain questions should remain on the ACS questionnaire

Background

- Review provided a rich database for investigating other means of shortening the ACS
- For example:
 - Could any questions be asked less frequently?
 - Could any questions be asked of fewer respondents if small-area estimates are not needed?
 - Could administrative data be used to create estimates rather than collecting from respondents?

Feasibility Report of Reducing Respondent Burden

The current design of the ACS asks all of the survey questions from all sampled households every year. Alternative survey designs might reduce that burden.

- Formed an internal Census Bureau group to identify ways to reduce respondent burden by tailoring the ACS design to satisfy the federal needs established in the Content Review
- Developed four options while acknowledging that combinations of options are also possible
- Posted high-level feasibility report on September 30, 2015 on the ACS internet site (link at end of slides)

Options Considered in Report

- Option #1: Periodic inclusion of questions not required every year
- Option #2: Targeted matrix sampling for a small set of questions
- Option #3: More comprehensive matrix sampling for a broad set of questions
- Option #4: Hybrid direct substitution of admin records augmented with respondent data

Criteria Used to Select Topics for Options #1 and #2

- Periodic inclusion (three topics)
 - No required or mandatory uses at tract level
 - All required or mandatory uses are needed less frequently than every year (or is unclear)
- Targeted matrix sampling (nine topics)
 - No required or mandatory uses at tract level
- Nine topics
 - Preliminary and subject to verification
 - Provided good examples for exploration

Illustration of Option 1: Periodic Inclusion

In this example, Topic C is needed only every three years, while Topic E is needed every other year. One version of the questionnaire is used for all households in sample each year.

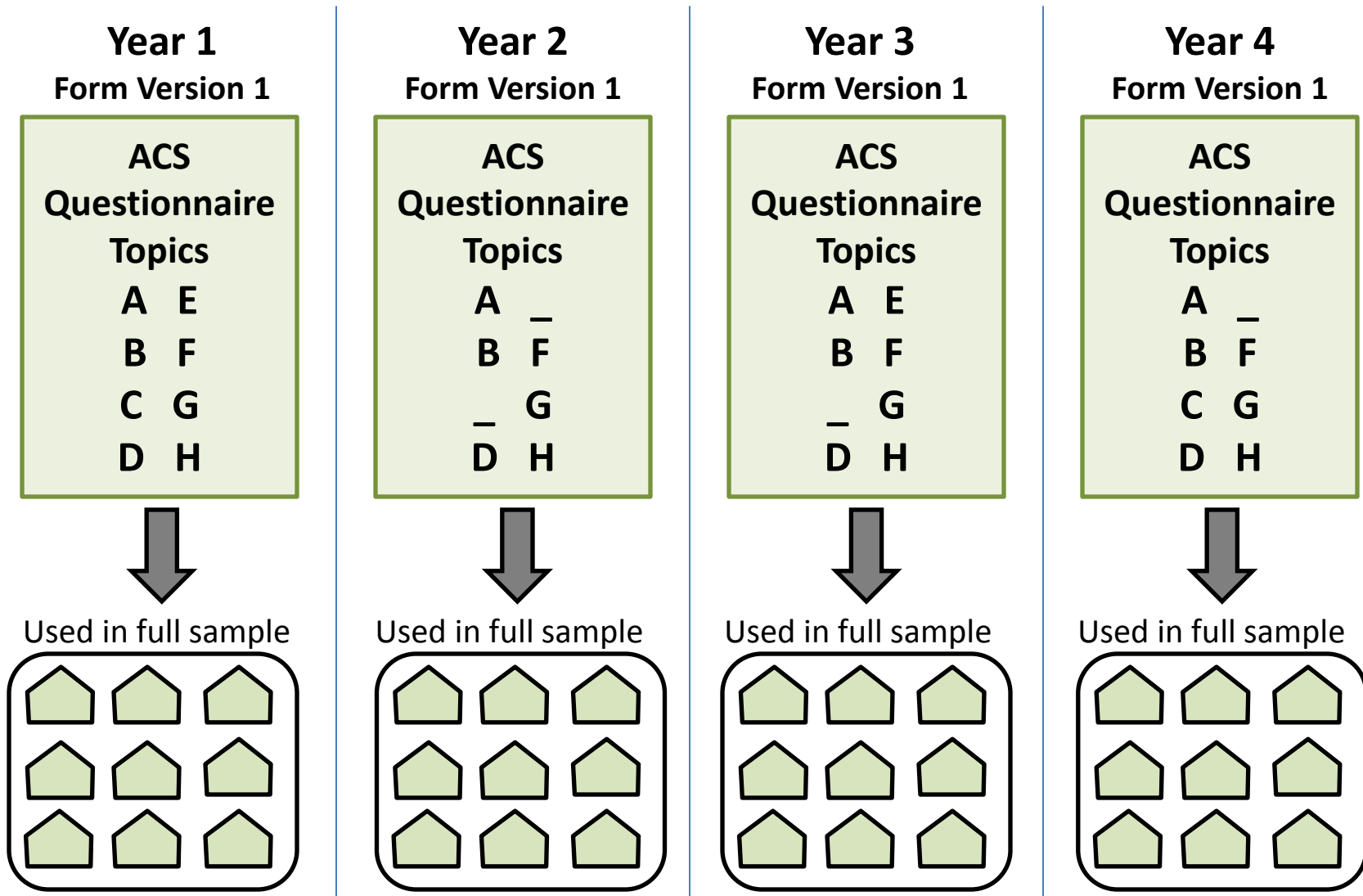
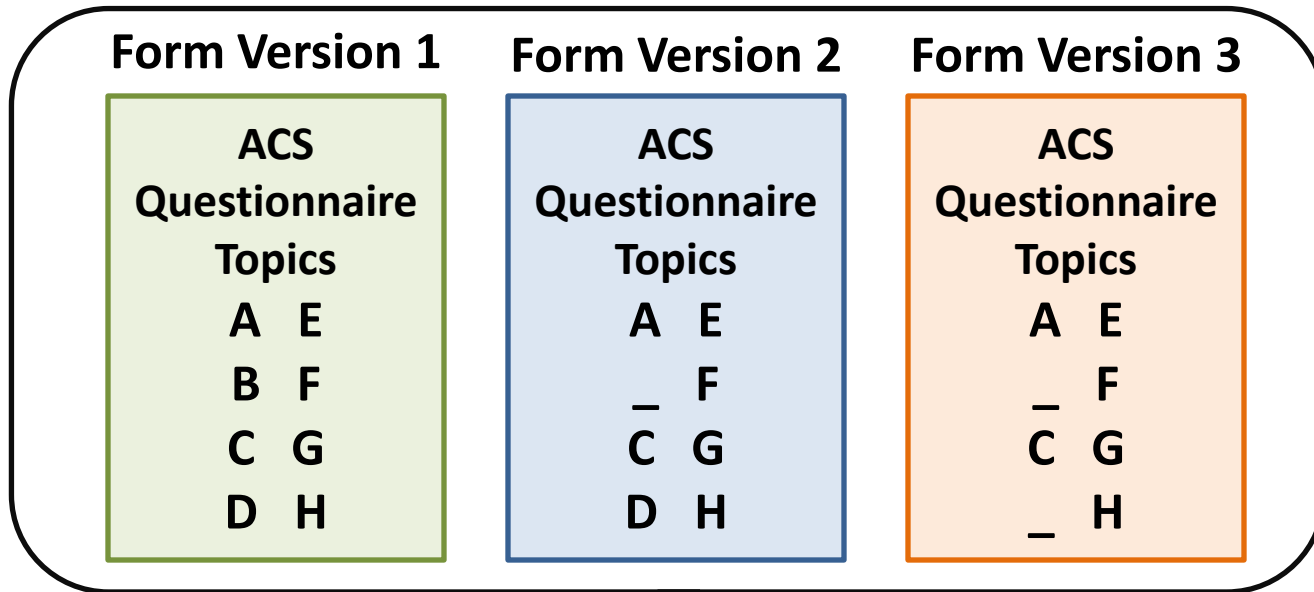


Illustration of Option 2: Targeted Matrix Sampling

In this example, Topic B is needed only at the state level, while Topic D is needed at the county level.



Only the portion of the full sample that is needed to produce estimates at the necessary geographic level receives the corresponding form version. In this example, only a small subset of housing units get Topic B while most, but not all, get Topic D.

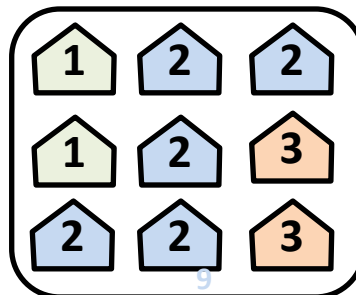
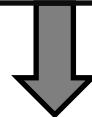
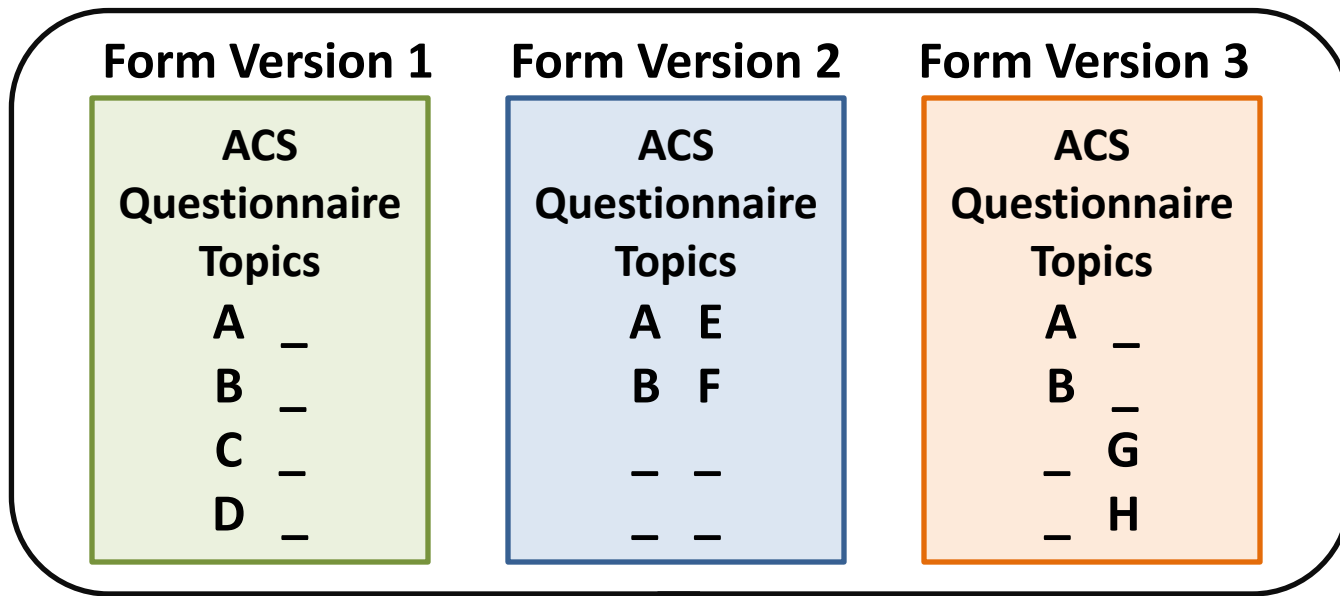
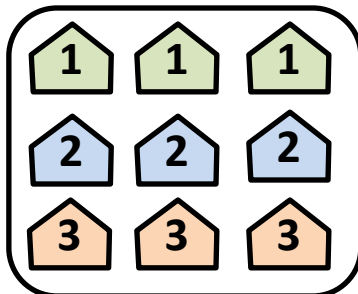


Illustration of Option 3: Comprehensive Matrix Sampling

In this example, topics are assigned to form versions in a partially overlapping manner.



Use either statistical tools or an increase in total sample to help mitigate the impact of the missing data.



OR

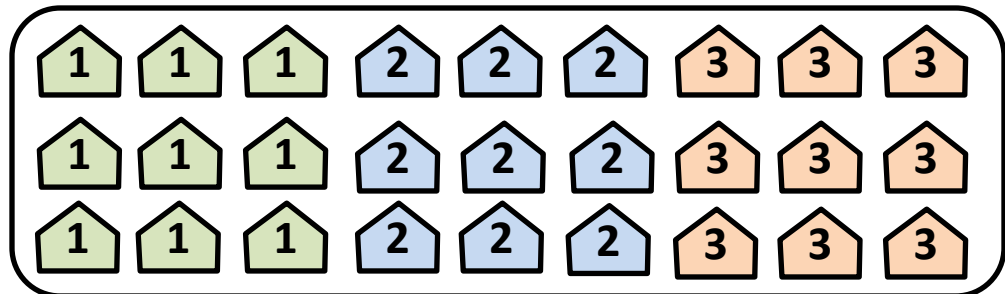
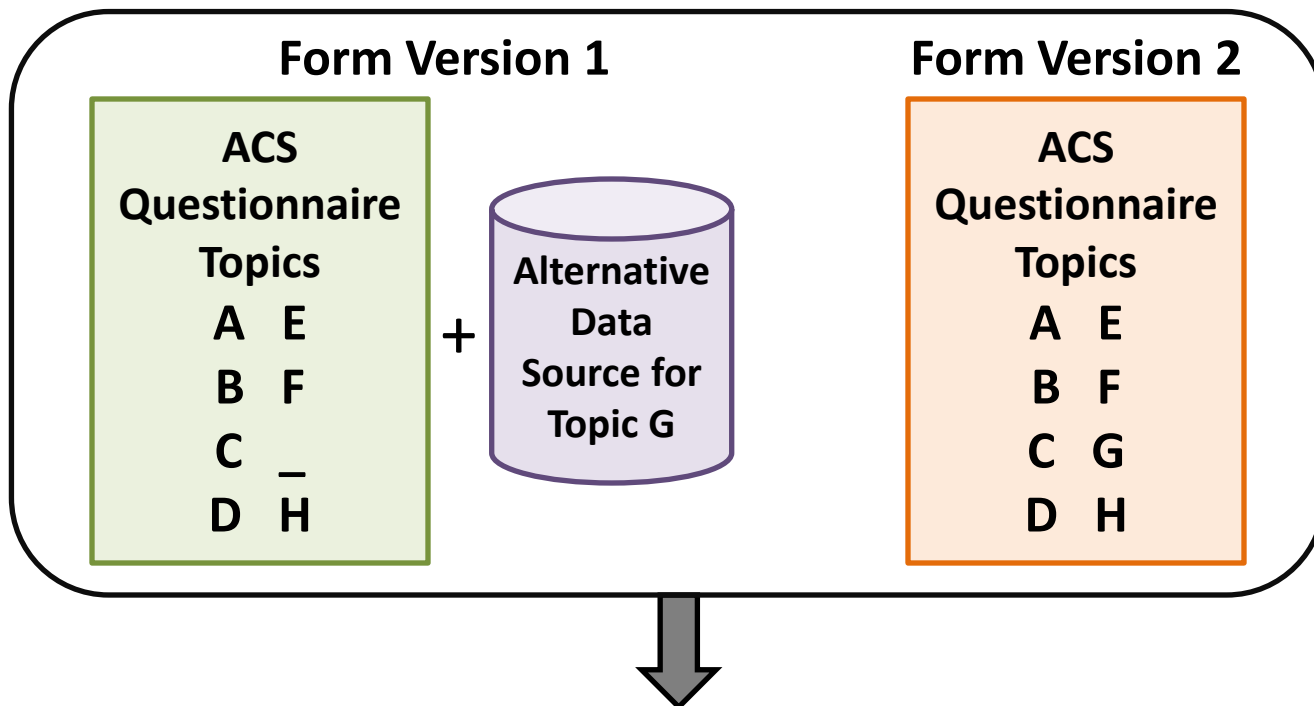
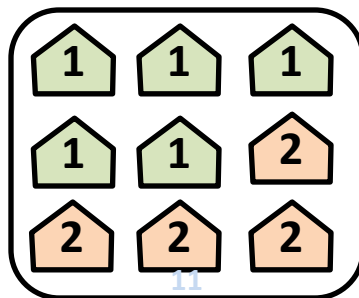


Illustration of Option 4: Administrative Records Hybrid

In this example, Topic G has an alternative data source with good quality and coverage for most geographic areas that can be used directly in place of collecting the data on the questionnaire in those areas.



Areas with good coverage for the alternative data source receive form Version 1, while areas without good coverage for the alternative data source receive form Version 2.



Assessing Each Option

A Census Bureau team assessed each option according to a set of factors that demonstrate the feasibility and impact of the method. These assessments were based on the professional judgment of the team members, and not on empirical criteria. These factors are:

1. Operational and processing complexity
2. Impact on the accuracy of the data
3. Impact on data availability for small geographies and groups
4. Estimated reduction in respondent burden
5. Impact on richness of the data products
6. Assessment of additional costs and resources required

Assessment of Option #1 and Option #4

- Option #1 (Periodic Inclusion):
 - Most challenges are operational
 - Challenges are driven by the adding and removing items from the form
- Option #4 (Administrative Records):
 - Central challenge is assessing quality of each potential administrative record source
 - Operational challenges of adjusting data processing

Identified Challenges to Matrix Sampling (Options #2/3)

- Operational
 - An incomplete microdata file
 - Ability to create cross-tabulations
 - User-friendly public use files
- Statistical
 - Imputation of missing data (potential biases)
 - Loss in precision of estimates
 - Estimating variances

Literature Review

- Goal of literature review was to find
 - examples of large surveys using matrix sampling
 - statistical methods used in estimation
- Generally, most papers involved simulation studies that had varied success
- Papers explored impacts on bias and variances
- Unable to find examples of implemented solutions for surveys similar to ACS

Literature Review

- Estimation approaches include:
 - Imputation techniques to fills in holes
 - Multiple imputation (Raghunathan and Grizzle, 1995, Thomas, et al., 2006)
 - Hot deck (examined at IRS, described in Gonzalez and Eltinge, 2007)
 - Best Linear Unbiased Estimators (Chipperfield et al., 2013)
 - Generalized regression (Merkouris, 2015)
- Papers stressed importance of optimizing how topics are grouped by form

Empirical Explorations

- Estimate how much we could subsample certain topics that do not need tract-level data
- Nine topics: 5 at state level, 4 at county level
- Estimated potential reduction in sample for
 - Target geography
 - Target coefficient of variation (CV)
- How much a topic could be subsampled depends on its prevalence and the geography

State Level – Reductions

- Target CV of 10% on average for each topic
- Reduction in sample by prevalence:
 - Higher prevalence estimates: 95-97% reduction
 - Middle prevalence estimates: 75% reduction
 - Low prevalence: 25% reduction

County Level – Reductions

- Target CV of 20% on average for each topic
- Some counties have CVs greater than 20%
– these would not be subsampled
- Large counties can have more substantial reductions in sample for a given topic
- Reduction in sample by prevalence:
 - Higher prevalence estimates: 90% reduction
 - Middle prevalence estimates: 75% reduction
 - Low prevalence: 60% reduction

Team's Recommendations

- Periodic Inclusion (Option #1)
 - Recommend implementing this wherever possible given relatively low operational and statistical complexity
- Administrative Records (Option #4)
 - Recommend assessing administrative record data sources that could be used to partially remove questions from the form
 - Recognized that this option could translate to significant reductions in burden with relatively few undesirable impacts
- Matrix Sampling (Options #2 and #3)
 - Recognized that there are potentially large impacts to cost and also negative impacts on accuracy and richness of survey estimates
 - Recommend that the ACS seek input to help develop research into efficient and effective designs for use of matrix sampling

References

- 2014 Content Review

[www.census.gov/programs-surveys/acs/
operations-and-administration/
2014-content-review.html](http://www.census.gov/programs-surveys/acs/operations-and-administration/2014-content-review.html)

- Agility in Action Report

Reducing Respondent Burden Report

[www.census.gov/programs-surveys/acs/
operations-and-administration/
2015-16-survey-enhancements.html](http://www.census.gov/programs-surveys/acs/operations-and-administration/2015-16-survey-enhancements.html)