

Recommendations and Observations of the Census Scientific Advisory Committee Working Group on Group Quarters in the American Community Survey

Barbara A. Anderson, University of Michigan
Robert A. Hummer, University of Texas
Irma T. Elo, University of Pennsylvania

April 18, 2013

This working group was formed in response to recognized problems with the Group Quarters data in the American Community Survey. A National Academy of Sciences issued a report in 2012 on the same topics. The main issues with the GQ data in the ACS are the quality of the data and the high cost of collecting data for GQ respondents as compared to non-GQ respondents. The purpose of the working group was to discuss and make recommendations about the collection and quality of ACS GQ data. Any cost-saving suggestions in collecting ACS GQ data were also welcome.

The working group has had numerous internal discussions. It also has had bi-weekly conference calls with the ACS Headquarters staff since the Fall 2012 CSAC meeting. These calls have been constructive and productive. The working group has received materials that have been very helpful in the discussion of how to improve GQ data in the ACS. The following observations and recommendations are based on reading and discussion within the Working Group, as well as with the ACS Headquarters staff and with Paul Voss.

1. **Need for Continued Collection of GQ Data in the ACS:** The ACS is an extremely important source of demographic and social data for policy, business and scientific purposes. It is essential that, as a replacement for the long form of the census and as a source of data about the entire United States population, the ACS continue to collect data about the GQ population.
2. **Additional Support from the Upper Levels of the Census Bureau:** Several of the issues addressed below require increased cooperation between the ACS and other units of the Census Bureau and between the ACS and government units outside the Census Bureau. In some areas, the ACS GQ staff has been able to obtain improved cooperation, but often progress appears to have been very slow.

Although there have been many indicators of commitment by the Census Bureau to the quality of the ACS GQ data, additional higher-level Census Bureau support could speed up this process. For example, it could be made clearer that cooperation among Census Bureau units to improve the GQ data in the ACS is a Census Bureau priority. This would especially be important in implementing the suggested changes in the Charter of the GQ Forum. This could lead to standardization of GQ definitions and improvement of the GQ sampling frame. It also would be desirable if the Census Bureau made clear that it is justifiable for non-ACS Census personnel to devote some time to these cooperative efforts and if some funds were provided to non-ACS units to compensate for whatever additional time is needed to help the ACS in this area.

3. **GQ Forum:** The GQ Forum involves various bodies that collect data from the GQ population. This group is a very good idea and has facilitated communication among various groups concerned with GQ data. However, it would be desirable if it moved to function more as a working group to take action on issues and initiatives such as

improvement of the sampling frame and agreeing on consistent definitions relevant to GQ facilities. For this to happen, the Charter of the GQ Forum would need to be revised. Support from the upper levels of the Census Bureau, mentioned above, would probably be needed to effect a substantial change. It would also be useful to review whether the Inter-Agency Task Force has been helpful in improving GQ data in the ACS.

4. **Sampling Frame:** Partially due to the uneven distribution of GQ's across space, the accuracy of the sampling frame for GQ's is at least as important, if not more important, than for the non-GQ population. The updating of the GQ sampling frame should be a high priority. Several later points, such as on sampling and weights, are highly dependent on the quality of the sampling frame. Thus, the sampling frame should be updated frequently between censuses. To date, it has only been updated after the decennial census except for some minimal specific annual updates.

The Population Estimates Program and the Survey of Business Owners and Self-Employed Persons include almost all types of GQs. Nursing home surveys also include old age homes. There is not a consistent well-worked out procedure to update the ACS GQ sampling frame between censuses. This is also a topic that could be addressed by the GQ Forum.

We strongly endorse continued collaboration with the Justice Department regarding data on federal prisons and encourage the staff to work closely with state agencies that administer the state prison system in an effort to improve data collection from state prisons.

We also strongly encourage the ACS GQ staff to work with state agencies that monitor data on GQs and take advantage of such data, even if such collaboration varies by state, with appropriate attention to avoiding of duplicates in the sampling frame.

5. **Sampling:** The ACS has already changed to data collection from persons in college dormitories so that now it only occurs during the academic year. This change was a very good idea. But other sampling issues remain. For example, if a sampled GQ has moved, the ACS tries to find the sampled GQ in the new location, but then no interviews are conducted at the old location of the GQ. The information on the change of GQ type or shift out of being a GQ is not used to update the sampling frame. It seems that this type of information should be used to update the sampling frame. The ACS could gain useful information by conducting frequent analyses of GQ's that change GQ type and of shifts of residential units between GQ and non-GQ status.
6. **Imputation:** In the GQ data, often imputation of specific characteristics or of entire responses is necessary. One concern is whether the same person's characteristics are used for imputation numerous times. The ACS has decided to only impute the appropriate sex for single-sex GQs. It would be desirable in the PUMS data for there to be a flag if an entire case or specific characteristics were imputed. Currently there is no flag on the PUMS data that a GQ person record was imputed. If this were flagged, then users could deal with imputed values however they wished.
7. **Weighting:** One problem with accuracy of GQ estimates for sub-state divisions is the omission of GQ's. We discussed a new plan developed by the ACS staff for imputing data when it is known a GQ exists, but it was not in the sample. The ACS staff has conducted research showing that the results of this new procedure are better than the

current procedure, based on comparison with census data shortly after the census. Much of the value of this new procedure depends on the quality of the sampling frame.

8. **Concerns of Different User Constituencies:** Different user constituencies are concerned with the ACS data for small geographic areas than those concerned with the PUMS data. It would be useful to review the concerns of the different user constituencies and what aspects of GQ data in the ACS are used and would be desirable for each constituency. For users of data for small geographic areas, a question of what information besides the age, sex, racial and ethnic composition is important. For users of the PUMS data, the main concerns are cross-tabulations and statistical analyses concerning the interrelations of different characteristics. It will be very important to include discussions of GQ data with the newly established ACS Data Users Group. It may also be important to make sure that at least one person on the new ACS Data Users Group is a knowledgeable and experienced user of GQ data in the ACS.
9. **GQ Questionnaire:** A main concern about the GQ population concerns the institutional population, the largest part of which is prisons, although some of the same concerns also pertain to old age homes and treatment facilities. Some questions are not asked of respondents in institutions in the automated GQ questionnaire, such as about journey to work. However most of the questionnaire is identical for the institutional and the non-institutional population, and the paper questionnaire is identical for the institutional and the non-institutional population. We had a discussion about getting a shorter set of information from records for the institutional contexts. We suggested that person's age, race and ethnicity might be enough. ACS agreed that this might be possible for Federal prisons. It was left as a point of investigation and discussion whether it would be possible to get forms with this limited set of information filled out in exchange for not bothering non-Federal correctional institutions with in-person visits. This change could save money in data collection and could allow collection of a narrower set of data *from many more institutions*. Such a change could be of tremendous benefit to addressing some of the sampling and imputation issues discussed above. To begin it would be a good idea for ACS to look at the correspondence between data from administrative records and ACS survey data for Federal prisoners.

Questionnaires that are administered using CAPI include appropriate automatic skip patterns. This is not true where paper questionnaires are used. Changing to CAPI data collection wherever possible would be desirable.

10. **Research:** Several observations and recommendations have referred to research undertaken by ACS staff, such as the need for analyses of changes of GQ type and between GQ and non-GQ status. Also the recent ACS research on imputation methods is quite important. In order to further improve the quality of the ACS GQ data, continued research in these and other areas will be essential.