Using Imputation Methods to Improve the American Community Survey Estimates of the Group Quarters Population for Small Geographies[1]

Mark Asiala, Michael Beaghen, Alfredo Navarro
U.S. Census Bureau, 4600 Silver Hill Road, Washington, DC 20233

Abstract

This paper describes the background and methodology of a Census Bureau program under development to improve the American Community Survey (ACS) estimates of the group quarters (GQ) population for small areas. What motivates this work is that while the ACS GQ sample was designed to produce estimates at the state-level, the estimates of the GQ population contribute to ACS estimates of the total resident population for substate areas such as counties and tracts. Consequently, there are small geographies which either do not have GQ sample or have GQ sample that is not representative of the area, which can lead to distorted estimates of characteristics and/or total population for these geographies. The approach taken is to impute whole person records (and weight them appropriately) to GQ facilities which appear on the sampling frame but were not selected into sample.

Key Words: sample design, small area estimation

1. Introduction

The Census Bureau has undertaken a research program aimed at improving the American Community Survey (ACS) estimates of characteristics of the group quarters (GQ) population for substate geographies such as counties and tracts. The plan is to have a new estimation methodology for GQ population in place for the ACS estimates produced in 2012, which necessitates deciding on a methodology in the fall of 2011. The research is spurred both by limitations in the usability and questions in the accuracy of the ACS data pointed out by ACS data users, and by long-term concerns from within the Census Bureau about the design of the ACS GQ sampling and weighting. At the heart of the matter is that the current ACS sample design and weighting were designed to produce state level estimates of characteristics of the GQ population, whereas estimates of the GQ population contribute to substate estimates of the characteristics of the total resident population. ACS estimates of characteristics of the GQ population are published for states and larger geographies, but not for substate geographies. However ACS data products which include GQ population are released for substate areas as small as block groups.

We have focused our resources on developing a new estimation methodology because we do not have other good alternatives at this time. No changes in the GQ sample design can be made quickly enough to remedy the problem for the ACS estimates produced in 2012. Further, the sampling rate for GQ persons is fixed by budget constraints, and any changes to the sampling plan that increase the number of GQs in sample would increase the cost

[1] Any views expressed on statistical, methodological, technical, or operational issues are those of the authors and not necessarily those of the U.S. Census Bureau.

of the survey. Also, while publishing estimates for only the household population for substate areas is an option, it is not appealing, as data users expect estimates for the total resident population, such as had been provided by the Census 2000 sample (long form) data.

The new methods we are investigating also seek to use as auxiliary data the 2010 decennial census GQ Universe file. For our research we have used the current ACS sampling frame, which is based on the 2000 decennial census, and the 2010 decennial census GQ Enumeration file, a listing of GQ facilities compiled in 2009 and used as the basis for the 2010 decennial census GQ enumeration. The approach we are investigating involves imputing GQ person records into facilities that are either on the ACS sampling frame or on the 2010 GQ enumeration listing.

First we describe some aspects of the ACS GQ sampling and estimation processes in Section II. Then we review evidence illustrating the concerns that motivate this research. In Section III we show gaps in representation of the ACS GQ sample across tracts and counties. The remainder of the paper describes the research carried out to develop more useful ACS estimates for areas with smaller GQ populations. Section IV describes our general approach and some of the issues that were addressed. Section V, Section VI, and Section VII, and Section VIII present the details of the imputation and weighting methods. Section IX provides a brief overview of the evaluation study with simulated data, and Section X provides an overview of the second evaluation which uses ACS data. Section XI discusses the next steps in the evaluation and implementation of the new methodology.

2. ACS Sampling and Estimation for the Group Quarters Population

For a better understanding of the issues in this paper, some description of the ACS GQ sample design and estimation is needed. Of salience is that the sampling and estimation methodologies for the GQ population are designed to produce optimal state level estimates, as it is only for states or larger geographies that estimates of the characteristics of the GQ population are produced. Only the estimates of the total GQ population are published for geographies smaller than the state. While the sample stratification includes type of GQ and geography, the sampling rates are such that many counties and tracts do not have sample for particular major types of GQ which nevertheless exist within them. Further, the GQ population estimates are controlled at the state level, whereas the ACS estimates of the total resident population are controlled at the level of county-based weighting areas.

The GQ sampling selects groups of GQ residents, not the GQ facilities themselves, in contrast to the HU address sample. The GQ frame is divided into two sampling strata within each state, a small GQ stratum and a large GQ stratum, each with different sampling methods. The small strata consist of GQs with expected populations of 15 or fewer and GQs closed on April 1, 2000. Small stratum GQs are sampled systematically within each state, sorted by small versus closed on census day, new GQ facility versus previously existing, GQ type, and geographical order (county, tract, block, street name, and GQ identifier). The sampling rate is 1-in-40 for most states but is higher for states with the smallest GQ populations. If there are 15 or fewer people found in a small stratum GQ, then everyone in the GQ is in sample. If there are 16 or more people found in a small stratum GQ, then ten people are systematically selected from the GQ. The large strata include GQs with expected populations of 16 or more. The primary sampling

unit for large stratum GQ facilities is a group of ten people, not the facility itself. For each large stratum GQ selected to be in sample, one or more systematic samples of groups of ten people are taken to achieve the state sampling rate. All large GQ facilities in a state are sorted by GQ type and geographical order in the large GQ frame. On the 2007 GQ sampling frame, there were approximately 105,000 small stratum GQ facilities, 77,000 large stratum GQ facilities, and 3,000 facilities with an unknown population (U.S. Census Bureau, 2009).

3. Representation of ACS Group Quarters Sample Across Tracts and Counties

The distribution of ACS sample GQ facilities across counties and tracts illustrates the limitations of the sample design with respect to producing small area estimates. Table 1 and Table 2 show the representation of the ACS sample across tracts in the years 2006-2009. (The 2005-2009 ACS 5-year estimates will be based on only four years of sample, from 2006-2009, as the ACS did not collect from a GQ sample in 2005.) The limitations in tract representation are ameliorated slightly by the fifth year of ACS GQ sample starting with the 2006-2010 estimates released in 2011. In Table 1 we see that about half of the tracts with GQ facilities had at least one GQ facility in the ACS sample from 2006-2009, that is, 21,596 of 43,367. The number of tracts and counties with GQs is determined from the ACS sampling frame, which is based on the 2000 decennial census. For perspective, note that the Census 2000 long form was given to a sample of 1-in-6 persons residing in GQs, so there was better coverage of GQ facilities.

Table 1: ACS GQ Sample in Tracts in 2006-2009

	Frequency
Tracts with GQs	43,367
Tracts with ACS GQ sample	21,596
Tracts without ACS GQ sample	21,771

Table 2 shows the representation of ACS sample in tracts by seven major types of GQ facilities. The categorization by major type of GQ is more relevant because people in different types of GQ facilities differ from each other in consistent, predictable ways. The categorization by seven major types shown in the tables is used in assigning the weights and is a convenient categorization here. Table 2 shows that large numbers of tracts with GQs do not have ACS sample of the same major type of GQ. For example, of the 4,994 tracts with an adult correctional facility, 2,113 did not have any facilities in the ACS sample from 2006 to 2009.

Table 2: ACS GQ Sample in Tracts by Major Type of GQ in 2006-2009

Major Type of Group Quarters	Tracts with ACS Sample	Tracts without ACS Sample	Total Tracts with Type of GQ
Adult correctional facilities	2,881	2,113	4,994
Juvenile facilities	1,133	1,685	2,818
Nursing/Skilled nursing facilities	9,847	6,736	16,583
Other health care facilities	985	2,648	3,633
College/university student housing	2,428	923	3,351
Military group quarters	287	289	576
Other noninstitutional facilities	10,037	24,934	34,971

Table 3 shows the analogous counts as Table 2 but for counties instead of tracts. Even at the county level there is a significant proportion of counties which lack representation by major type of GQ.

Table 3: ACS GQ Sample in Counties by Major Type of GQ in 2006-2009

Major Type of Group Quarters	Counties with ACS Sample	Counties without ACS Sample	Total Counties with Type of GQ
Adult correctional facilities	1,792	953	2,745
Juvenile facilities	660	522	1,182
Nursing/Skilled nursing facilities	2,601	354	2,955
Other health care facilities	551	781	1,332
College/university student housing	987	168	1,155
Military group quarters	216	180	396
Other noninstitutional facilities	1,888	935	2,823

Because the GQ sampling is conducted at the state level, some counties can have disproportionately small or large numbers of GQ persons in sample for a given year, contributing to high sampling variability for county level estimates of the total GQ population. As a result, data users see large year-to-year fluctuations of county estimates of GQ population, and unrealistic estimates of persons per household, among other unexpected estimates. The median difference between the 2007 and 2008 ACS GQ population estimates over the 790 counties for which one-year estimates are published is 1,090 (the average number of people residing in GQs per county in 2007 was only 7,903 for these 790 counties).[2] For an example, consider the estimates of GQ population for Harford County, Maryland, for 2006, 2007, and 2008, which were 2,897, 6,138, and 1,463. Another relatively extreme example is Benton County, Oregon, which had 6,129 and 2,709 according to the 2006 and 2007 ACS (the Population Estimates Program (PEP) estimate of GQ population for these years was 4,280).

These year-to-year fluctuations highlight the point that which GQ facilities fall into sample has a disproportionate effect on the estimates of substate geographies. In the 5-year estimates this phenomena will be weaker at the county level because there will be five times as much sample. However, in 2010 the ACS published its first 5-year estimates down to the tract and block group, where we expect to see analogous problems, as suggested by the distributions of GQ sample across tract shown earlier. We also expect to see effects on the estimates of characteristics of the total resident population. Though the actual number of people residing in GQs is small, their effect on estimates of the characteristics of the total resident population can be disproportionately large for characteristics that are strongly related with GQ residence. Such characteristics include disability status, income, and variables derived from income, such as poverty. Beaghen and Stern (2009) document such concerns with estimates of poverty rates.

[2] The Census Bureau releases 1-year ACS estimates of total GQ population for counties with total resident populations of 65,000 or more.

4. Overview of the New Methodology

The objective of the new methodology is to improve the estimates of the GQ population for counties and tracts, thereby also improving estimates of the total resident population for counties and tracts. Note that the limitations in the sample design can be viewed both in terms of high variances of estimates of the GQ population for substate geographies, as well as in a lack of representation of ACS sample in counties and tracts which are known to have GQ facilities. Though we approach the problem from the point of view of trying to have GQ person records, sampled or imputed, in the smallest geographies, a successful methodology should shrink the variances of small area estimates.

4.1 The Basic Approach to the Problem

The approach to the problem is to populate GQ facilities without ACS sample with person records copied from in-sample GQ facilities, with appropriate weighting adjustments. This imputation is a whole person imputation and not an item-level imputation. The whole set of person characteristics of the donor is copied to the recipient record (with a few exceptions discussed later). The recipient record maintains the recipient GQ type characteristics and current residence geography. Imputing to not-in-sample facilities has appealing advantages conceptually and for data processing. The imputed person records function as pseudo-sample and are transparent to the data processing and production of estimates. For this approach we identified the following key challenges.

- How do we construct the frame which we later populate with GQ person records?
- Which not-in-sample GQ facilities do we impute to?
- How many person records should be imputed to each GQ facility?
- How do we select GQ person records to serve as donors?
- How do we assign weights to imputed and sample GQ person records?

4.2 Establishing the Enhanced Frame

We will refer to the listing of GQ facilities to which we will potentially impute for as the ‘enhanced frame’. The enhanced frame could have been constructed in several ways and the following were considered.

- Use only the ACS GQ sampling frame.
- Use only the 2010 decennial census listing of GQs.
- Use a merged listing of the ACS GQ sampling frame and the 2010 decennial census listing.

Using the ACS GQ sampling frame would be consistent with the time period over which we collected the data and which the estimates represent. However, the 2010 decennial census listing is more up-to-date than the 2000-based sampling frame. Using a merged listing of the current ACS sampling frame and the decennial census listing is potentially more comprehensive than either alone. This is the method we chose. We note, however, it could be limited by duplication, as some ACS frame GQs will be deleted by decennial census operations and added again, possibly with a new address, a new name, or a new geocoding.

4.3 Which Group Quarters Facilities to Impute to?

Once the enhanced frame is determined, the next question was to determine which subset of not-in-sample GQ facilities on the enhanced frame to impute to. Imputing to all of them has the appeal that there would be GQ person representation in every geography for every detailed type of GQ (there are more than 60 detailed types of GQ facilities) where GQ facilities exist. However, it would require imputing a prohibitively large number of records. Thus we have imputed to only a subset of GQ facilities, prioritizing which GQs to impute to as follows.

- The primary objective is to establish representation of county by major type of GQ in the tabulations for each combination that exists on the frame.
- A secondary objective is to establish representation of tract by major type of GQ for each combination that exists on the frame, as is reasonably feasible.

These priorities lead to a scheme where all large stratum GQs are imputed to, but only a sample of small stratum GQs are imputed to so that the second objective is met.

4.3 How Many Group Quarters Person Records to Impute to Each Group Quarters Facility?

How many imputed person records each not-in-sample GQ facility receives will be a function of its population, which is either modeled or observed. For this determination we make a distinction between small and large stratum GQ facilities. For large facilities we will impute so that there is about 2.5 percent of the expected population, which is to be similar to the overall one in 40 overall sampling rate in most states. The allocation to small GQ facilities is 20 percent of the expected population.

5. Identify GQ Facilities that Require Imputations and Determine How Many GQ Persons to Impute

The GQ selection procedure gives priority to obtaining representation for each major GQ type group in each county. Then facilities are selected to establish representation for each major type group at the tract level. A detailed outline of the procedure is given next.

1. For each year and each large GQ not in sample, impute the number of records equal to 2.5% of the expected GQ population.

2. For each year and for each combination of county and major GQ type on the year's frame that is not in the year's sample nor in the year's imputes, select a small GQ facility at random with probability equal to the reciprocal of the number of small GQ facilities in the county and of the same major GQ type.

3. For each GQ selected in Step 2, impute the number of records equal to 20% of the expected GQ population or 1, whichever is larger.

4. Select all combinations of tract and major GQ type that exist on any year's sampling frame but are not in any year's sample, nor in any year's imputed records.

5. For each combination in Step 4 and for each year that the combination exists on the sampling frame, select a small GQ facility at random with probability equal to the reciprocal of the number of small GQ facilities in the tract and of the same major GQ type.

6. For each GQ selected in Step 5, impute the number of records equal to 20% of the expected GQ population or 1, whichever is larger.

6. Select Donors

The donor selection is referred to as the expanding universe approach. Note that for each year, donors are selected only from that same year.

The donor selection procedure chooses from within specific type when the donor to imputation ratio within the specific type is feasible, and gives preference to donors from facilities that are geographically close. Once GQ facilities have been selected for imputation, the donor pool for each facility is set to be the first combination of geography and GQ type in the following list in which there is at least one donor per five imputed records needed. Donors are recruited first in the lower ranking step starting with step 1. If a suitable donor is not found in a given step, then proceed to the next step.

1. County and specific type
2. County and major type
3. State and specific type
4. State and major type
5. Division and specific type
6. Division and major type
7. Region and specific type
8. Region and major type
9. Specific type
10. Major type

For example, suppose that in a particular county we wish to impute one hundred records into college dormitories. If at least twenty dormitory residents in the county have been interviewed, we sample these interviews for imputation, with replacement, one hundred times. If fewer than twenty dormitory residents in the county have been interviewed, we expand the geography of the donor pool (to the state, division, region, or nation) as necessary so that there are at least twenty records from which to sample.

7. Geography-Dependent Characteristics

Certain geography-dependent characteristics such as migration (residence one year ago) and journey to work present particular challenges for the imputation based method, as assigning the geography of the donor can lead to unreasonable results for certain geography-dependent characteristics. For example, say we assign a donor person from Smith County to a recipient GQ listing in Springfield County. Further, say the donor indicates he lived in Smith County the previous year. If we assign Smith County as the residence a year ago, it will seem as if the respondent had moved between counties over the year – although the respondent had indicated they had not moved. An analogous quandary exists with journey to work; do we want the imputed respondent to indicate that he commuted to the neighboring county?

Koerber (2011) found that for characteristics which are geography-dependent, such as Place of Work and Migration, always taking the donor values leads to a poor distribution of estimates of the characteristics. He found, however, replacing the donor geographies for these characteristics with the current residence geography of the recipient for certain cases leads to markedly improved distributions of these characteristics in the sense they are more like the distribution based on sample. These replacements or special edits amount to an additional processing step to be conducted after the GQ small area imputation, which itself takes place after the traditional production edits.

The rules of the special edits are summarized in these two points.

- Place of work - use the geography of the recipient's current residence for place of work if the donor's current residence and place of work are in the same tract or if they are in the same county.
- Migration (residence one year ago) - use the geography of the recipient's current residence for previous residence if the donor's current residence and previous residence a year ago are in the same tract or if they are in the same county.

The distributions of imputation-based place of work and migration estimates show marked improvement using the special edits in the sense that they are more like the distributions based on sampled data. However, the special edits tend to overcorrect the estimates of tract-oriented interpretations of these characteristics. We noted also that the special edits could lead to anomalous data for small geographies. On the whole however, the distributions of these characteristics are greatly improved.

8. Weighting

The new imputation methodology implies a new weighting scheme (Asiala, 2011a), which makes a clean break from the old weighting design that was used for 2006-2009. The weighting procedure is applied to the augmented data: the data set containing both the sampled and imputed records. It makes no distinction between sampled and imputed GQ person records. Further, the weighting scheme ensures that when computing estimates for small areas, the weighted data for GQ persons only represent persons within that tract or county, depending on the length of the estimation period.

For the purposes of weighting we define large and small GQ facilities differently than the large and small GQ facility strata used in sampling. This is because when a GQ is in sample and interviews are conducted there, we may find that the number of persons is much smaller or larger than expected, in particular that the average number of persons across all hits in a year is less than 15 when it was expected to be more or vice versa.

> **Large (self-representing) GQ**: a sampled GQ with an average number of persons across all hits in a year or a GQ which is not sampled with an expected population of more than 15 persons.
> **Small (non self-representing) GQ**: a sampled GQ with an average number of persons across all hits in a year or a GQ which is not sampled with an expected population of 15 or fewer persons.

8.1 General Outline of the Weighting Methodology

The following describes the weighting for the 2010 estimates. For multiyear estimates, the base weights are simply divided by the period length (3 or 5). The state-level controls are those defined for the periods in the standard methodology and are described in U.S. Census Bureau (2009).

The weighting has three basic steps.

1. The first step is to separately define the base weights for the persons in large and small GQs.
2. Apply tract- and county-level constraints based solely on the modeled populations on the frame.
3. Apply the independent population estimates as controls at the state by major type level.

We start with the following notation:

np_{Tctg} = Number of sample person records (real or imputed) in GQ *g* in Tract *t* of County *c* of Type Group *T*

NP_{Tctg} = Population of GQ *g* in Tract *t* of County *c* of Type Group *T*

NP_{Tct} = Population of all GQs in Tract *t* of County *c* of Type Group *T*

NP_{Tc} = Population of all GQs in County *c* of Type Group *T*

NP_{T} = Population of all GQs in a given state of Type Group *T*

PEP_{T} = Population Estimate Program Estimate (PEP) of GQ Population for Type Group *T* in a given state

ngq_{Tct} = Number of sample GQs in Tract *t* of County *c* of Type Group *T*

NGQ_{Tct} = Number of frame GQs in Tract *t* of County *c* of Type Group *T*

8.2 Assign the Base Weights to the Sampled and Imputed Records

Large GQs
In the assignment of base weights, no distinction is made between sampled and imputed records for the large GQ population. The sum of the base weights for the sampled and imputed person records will be constrained to the modeled current population of the large GQ since all large GQs on the enhanced frame will have either sampled or imputed person records in them.

Define BW_{Tctgi} to be the base weight for person *i* in GQ *g*, tract *t*, county *c* and type group *T*. The base weights would be defined to be equal across all persons in the GQ so that the GQ would be self-representing. The base weight for all persons in that large GQ is thus defined to be:

$$BW_{Tctgi}(\text{all large}) = NP_{Tctg} / np_{Tctg}\ ,\ i = 1,\ldots, np_{Tctg} \qquad (1)$$

Note that as defined, the weight, BW_{Tctgi} accounts for both the within-GQ subsampling and the nonresponse (for the sampled GQs) in one step.

Small GQs
For small GQs, it will not be the case that all small GQs on the enhanced frame will have sampled or imputed records in them. Therefore, the sum of the base weights within a small GQ will still need to represent the population in non-sampled, non-imputed small GQs. In order to reflect the selection procedure for the imputed GQs and to preserve the subcounty distribution of the GQ weights, the base weights for the small GQs will be defined so that any sampled or imputed GQ person records represent that tract by type group combination.

Thus under the assumption that those sampled or imputed small GQs within a type group-tract combination represent all small GQs within that same combination and that we select a subsample of all persons at the sampled GQs, the base weight for the i^{th} person record is:

$$BW_{Tctgi(all\ small)} = \left(\frac{NGQ_{Tct}}{ngq_{Tct}}\right) \times \left(\frac{NP_{Tctg}}{np_{Tctg}}\right), i = 1 \ldots np_{Tctg} \quad (2)$$

8.3 Calculate the County- and Tract-Level Constraints for the Weighting

In order to preserve the sub-state geographic distribution of total GQ population by type group from the enhanced frame, we implement county- and tract-level constraints. These constraints make use of the modified state-level population controls by type group and the modeled current population created for each GQ on the enhanced frame rolled up to the county- and tract-level.

Using our previous notation established for the GQ-level modeled populations, the tract-level modeled population, NP_{Tct}, would equal:

$$NP_{Tct} = \sum_{g=1}^{NGQ_{Tct}} NP_{Tctg} \quad (3)$$

We also define two county-by-type-group totals. The first population total is the sum of the tract-level populations across all tracts in a given county *c* of a type group *T*.

$$NP_{Tc} = \sum_{t \in c} NP_{Tct} \quad (4)$$

The second population total is like the previous but is restricted to all tracts within a county that contain either sampled or imputed GQs in them. This total would be situations including where a GQ is in sample in a tract in an early year but is not in sample for 2010. In that case, we may not have any imputed records in the tract where the sample GQ is located but we would still need to include its population in our county totals.

$$NP_{s,Tc} = \sum_{t \in c \mid ngq_{Tct} \neq 0} NP_{Tct} \quad (5)$$

Finally, we define the modeled state total by type group *T*.

$$NP_T = \sum_c NP_{Tc} \tag{6}$$

8.4 Apply the Tract- and County-Level Constraints and State-Level Controls

Once the BW_{Tctgi} weight has been established separately for persons in small and large GQs, the remaining steps will apply to all GQ persons combined without distinction to small or large (or sampled / imputed). The tract- and county-level constraints will ensure that the sum of the weights will agree with the frame with its modeled populations at the tract and county level. We now define $GQWT_{Tctgi}$ as the final unrounded weight as follows:

$$GQWT_{Tctgi} = BW_{Tctgi} \times \frac{NP_{Tct}}{\sum_{g,i} BW_{Tctgi}} \times \frac{NP_{Tc}}{NP_{s,Tc}} \times \frac{PEP_T}{NP_T}, i = 1, \ldots, np_{Tctg} \tag{7}$$

The first ratio adjustment is so that the sum of $GQWT_{Tctgi}$ within a tract would equal the tract-level constraint. The second ratio adjustment is so that sum of $GQWT_{Tctgi}$within a county would equal the county-level constraint in the event that not every tract contains GQs that have either sampled or imputed records within them. Finally, the third ratio adjustment is so that the sum of $GQWT_{Tctgi}$ within a state will equal the PEP total for that type group within a given state.

The final weight $GQWR_{Tctgi}$ would be created by controlled rounding the $GQWT_{Tctgi}$ weight.

9. Evaluation with Simulated Data

Two evaluation studies of the imputation procedures have been undertaken. We present only an overview of them as their details are beyond the scope of this paper. The first evaluation, which has been completed, was a simulation study with Census 2000 100 percent GQ data which used an inventory of all residents of GQs on April 1, 2000 and their basic demographic information – sex, age, race, and Hispanic origin. Estimates of these characteristics were compared and evaluated on simulated samples and four imputation procedures. The purpose of this study was three-fold:

a. To operationalize the imputation software and ensure that it worked correctly
b. To select between four different imputation procedures (we only discuss in this paper the chosen imputation procedure).
c. To analyze the differences between the imputation and sample only (design-based) estimates to determine if there are any major problems with the imputation procedures that need to be addressed or would cause us to discontinue this research and continue with the design-based estimates for the 2010 ACS production.

The simulations found that nearly half of the augmented data are comprised of imputed records. In addition, the number of imputed records can far exceed the number of sampled records for some major GQ types. Most donors were found with the specific GQ

type. It generally found donors within the state, or many times within the county of the GQ to be imputed.

We found that the imputation-based method was systematically biased even at the state level. The variances were smaller than the design-based estimate variances. Comparisons of the mean squared error, mean absolute deviation, and percent absolute deviation gave mixed results. Some of the shortcomings of the imputation methods identified in the simulation study were mitigated by minor changes to the imputation methodology in the evaluation with ACS data (see the next section). In particular, we limited the number of times a donor can serve, generally to five times. And we have identified a list of single-sex GQ facilities to which we restricted donors to be of the correct sex.

10. Evaluation with ACS Data

The second evaluation, which has been started, uses ACS data so that estimates of the diverse characteristics produced by ACS can be analyzed, not just the demographic characteristics analyzed in the simulation. The purpose of this evaluation is to establish the reasonableness and to confirm the usability of the GQSAE imputation methodology. In particular, the analyses in this evaluation are based on the following points.

- It is impossible to establish that the imputation estimates are superior to the design-based with traditional statistical measures such as bias, variance, MSE, or other loss functions. This is because the true population characteristics are unknown and thus we can't compute these statistics, nor do we have a sound variance estimator for the imputation methodology in place.
- We believe the design-based estimates of the GQ population have tolerably small bias at the state level. The ACS state-level estimates of the GQ population have been generally accepted to be sound. Thus any large differences between the design-based and imputation estimates at the state level suggest problems with the imputation methodology. This is one way of establishing that the imputation methods are unreasonable.
- In contrast to the state-level estimates, we don't assume the design-based estimates of the GQ population for substate geographies are sound. It is for estimates of the substate geographies where we hope the imputation methods will show advantages.
- We expect the imputation methods to enhance the usability of the ACS data, in particular, that they have a better distribution of GQ population by major type of GQ across substate geographies such as counties and tracts. We will attempt to confirm that the imputation methodology enhances data usability for small areas, for example, that data users have nonzero estimates of the GQ population in small areas where they expect them, and have estimates of GQ population of about the right size.

The specific goals of this study are described below.

- Evaluate changes made to the imputation procedure as a consequence of the results from the simulation study, to determine if they are accomplishing their goals.
- Compare the full range of characteristics produced for the GQ population and the total population from the design-based and imputation procedures.
- Compare county estimates of demographic characteristics by the two methodologies to their PEP estimates.
- Inform a final decision on applying the GQ imputation methodology to the ACS 2011 production cycle.

This evaluation requires five years of data since 1-, 3-, and 5-year estimates are to be investigated. Four years, 2006-09, of ACS GQ data that have been through production are available for use. A fifth year of data, representing 2010, was created for the GQs in the 2010 ACS sample. Preliminary results of this evaluation for state-level comparisons (Asiala, 2011b) showed no biases of an order of magnitude that would cause concern or preclude the use of the imputation-based methodology.

11. Next Steps

The second evaluation is underway. Depending on the results of this evaluation, the GQSAE methodology will be implemented in the ACS in 2012 for the 2011, 2009-2011, and 2007-2011 ACS estimates. We have planned a thorough analysis for 2009 estimates of counties, places with population of 65,000 or more, and 2006-2010 estimates of smaller geographies down to tracts. Analysts from the Census Bureau's Social, Economic, and Household Statistics Division and Population Division will study the estimates for their reasonableness. If the 2009 imputation-based estimates are deemed sound, then there will be an analysis of the 2006-2010 small area estimates will be examined, leading to a recommendation whether or not to implement the imputation-based methodology in 2012.

Acknowledgements

The development of the GQ imputation for small areas is a large project with contributions from staff throughout the Census Bureau. Among the contributors to the design of the methodology are: Edward Castro, Chandra Erdman, Patrick Joyce, and Lynn Weidman.

References

Asiala, M. (2011a). "Description of the 1-, 3-, and 5-year GQ Weighting for 2010". U.S. Census Bureau Memorandum.

Asiala, M. (2011b). "Comparison of State-Level GQ Estimates using the Design-Based and the Imputation-Based Methodologies". U.S. Census Bureau Memorandum.

Beaghen, M., and Stern, S. (2009). "Usability of the American Community Survey Estimates of the Group Quarters Population for Substate Geographies". *2009 Joint Statistical Meetings: Proceedings of the Survey Research Methods Section.* American Statistical Association.

Erdman, C. (2010). "Development of Imputation Procedures for American Community Survey Group Quarters Small Area Estimation". U.S. Census Bureau memorandum.

Koerber, W. (2011). "Issues Paper – Place of Work, Migration, and Place of Birth Geography for GQ Imputed Records". Internal U.S. Census Bureau Memorandum

U.S. Census Bureau (2009). "Design and Methodology: American Community Survey". Issued April 2009. http://www.census.gov/acs/www/methodology/methodology_main/