

Using Reimputation Methods to Estimate the Variances of Estimates of the American Community Survey Group Quarters Population with the New Group Quarters Imputation

Prepared for 2013 Federal Committee on Statistical Methodology Research Conference

November 5, 2013

Michael Beaghen
U.S. CENSUS BUREAU
U.S. DEPARTMENT OF COMMERCE
Washington, DC 20233



U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU



Problem Statement

- ACS sample is supplemented by a mass imputation of GQ persons
- How do we estimate variances in this context?

Group Quarters Population

- 7,987,323 GQ population (2010 Census)
- Seven major types of GQ facilities
 - Correctional institutions
 - Juvenile facilities
 - Nursing homes
 - Other long-term care facilities
 - College dorms
 - Military facilities
 - Other non-institutional GQs

ACS GQ Sample

- Sampling different for GQ than for household population
- GQ sample designed for state-level estimates; but used for small area estimates of the total population
- Insufficient GQ sample for small area estimates

GQ Imputation

- Starting with ACS data products released last year we implemented GQ imputation
- Impute whole-person records to not-in-sample GQ facilities
- As many imputed persons as sample and interviewed
- Asiala, Beaghen, and Navarro (2011)

Successive Differences Replication

- ACS household population, ACS GQ population through 2010, Current Population Survey, and the Census 2000 long form sample
- Appropriate for systematic sampling
- Replicates amenable to ACS tabulation
- Fay and Train, 1995

SDR with Inflation Factors

- With GQ naively treating the imputed data as sampled would seriously underestimate the variance
- Inflate the naive variance estimate using a set of inflation factors (Asiala & Castro, 2012)
- Factors for each state by seven major types of GQ
- Straightforward incorporation into ACS tabulation

Limitations of SDR with Inflation Factors

- Same inflation factor
 - for all characteristics
 - for the entire state
- Expect some residual underestimation of estimates of variances

Random Groups with Reimputation

- Random groups: divide sample data into groups such that each group has the same sampling distribution as the parent sample
- Shao and Tang (2011) describe how to use the random groups in the context of imputation

Random Groups with Reimputation

- Form groups with the standard random group methodology
- Reimpute again for each group
 - Use only the sample in that group for donors
- Proceed calculating variance estimates as with random groups

Simulations of ACS GQ Sample with Census 2000 Data

- From Census 2000 100% data
 - Age, sex, Hispanic origin, race
- 25 simulations of ACS GQ sample selection
- Used in the research and development of the GQ imputation
- Use it here for comparison purposes
- Erdman & Nagaraja (2010), Weidman (2011)

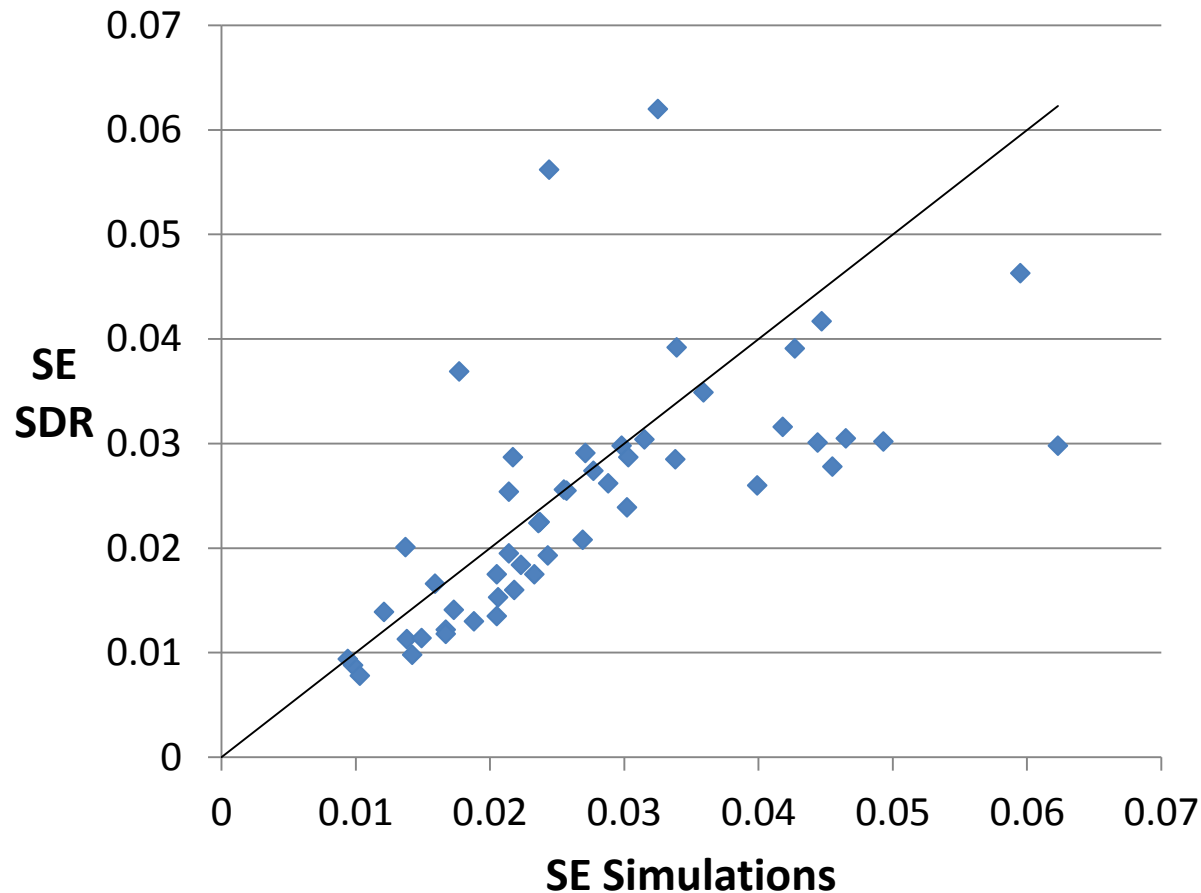
Comparisons

- Assess state- and county-level estimates standard errors of proportions
- Examine the following characteristics
 - Age, sex, Hispanic origin, race (can compare to simulations)
 - Marital status, educational attainment, speaks a language other than English at home, disability status

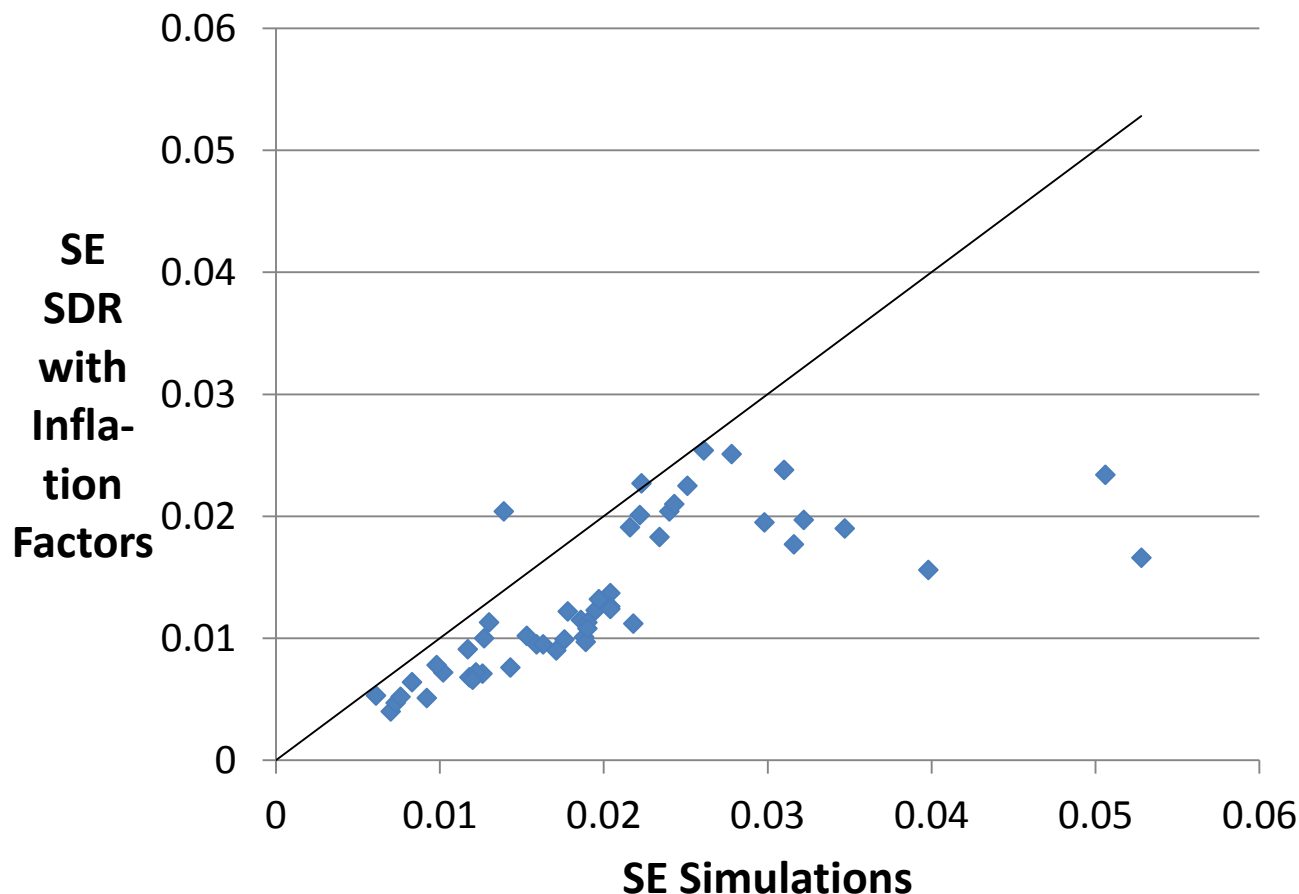
Comparisons

- No GQ imputation
 - SDR against simulations
- With GQ imputation
 - SDR with inflation factors against simulations
 - Ten random groups with reimputation against simulations
 - SDR with inflation factors against ten random groups with reimputation

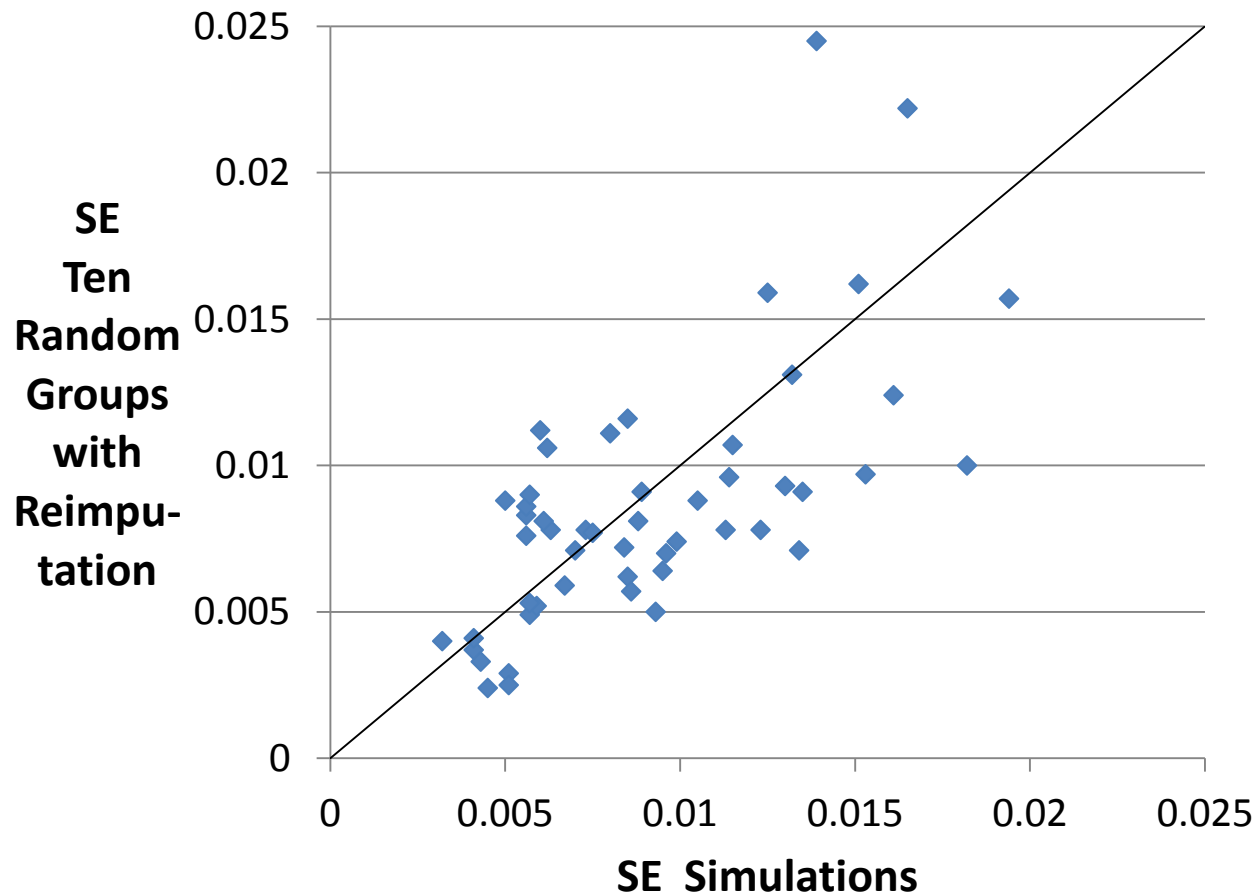
Standard Errors of Proportion Male for States: SEDR against the Simulations (no GQ Imputation)



Standard Errors of Proportion Male for States: SDR with Inflation Factors against the Simulations (with GQ Imputation)



SEs of Proportion Age 65 + for States: Ten Random Groups with Reimputation against the Simulations (with GQ Imputation)



Root Mean Square Differences from Simulations of SE of Proportions (with GQ imputation)

Characteristic	Ten Random Groups with Reimputation	SDR with Inflation Factors
Male	0.001449	0.001346
Hispanic Origin	0.000555	0.000287
Age 65+	0.000476	0.000352
Age under 18	0.000572	0.000623
Age 18 to 34	0.000630	0.000506

Conclusions

- SDR used for ACS GQ (before imputation) sound
- The current methodology, SDR with inflation factors, seems adequate
 - appeared to moderately underestimate SEs
- The random groups with reimputation yielded apparently sound variance estimates in the context of a mass imputation
- Ten random groups less reliable than SDR with inflation factors

Future Research

Research is ongoing

- Investigate county results
- Test successive differences replication with higher values of inflation factors
- Attempt 20 random groups with reimputation (some complications)

Reference List

- Asiala, M., Beaghen, M., and Navarro, A. (2011). "Using Imputation Methods to Improve the American Community Survey Estimates of the Group Quarters Population for Small Geographies". *2011 Joint Statistical Meetings: Proceedings of the Survey Research Methods Section*. American Statistical Association.
- Asiala, M., and Castro, E., (2012). "Developing Replicate Weight-Based Methods to Account for Imputation Variance in a Mass Imputation Application". *2011 Joint Statistical Meetings: Proceedings of the Survey Research Methods Section*. American Statistical Association.
- Erdman, C. and Nagaraja, C. (2010). "Imputation Procedures for the American Community Survey Group Quarters Small Area Estimation". Research Report (Statistics #2010-09), Statistical Research Division, U.S. Census Bureau, Washington, DC.
- Fay, R. and Train, G. (1995). "Aspects of Survey and Model-Based Postcensal Estimation of Income and Poverty Characteristics for States and counties". *Proceedings of Government Statistics Section, American Statistical Association*, 154-159.
- Shao, J. and Tang, Q., (2011). "Random Group Variance Estimators for Survey data with Random Hot Deck Imputation". *Journal of Official Statistics*, Vol. 27, No. 3, 2011.
- Weidman, L. (2011). "Research to Improve American Community Survey Group Quarters Estimates for Small Areas". Paper presented at the February 16, 2011 Committee on National Statistics Meeting on "ACS Group Quarters".

Contact Information

Michael.A.Beaghen@census.gov

For more information on the ACS, visit

<http://www.census.gov/acs/www>