Results of Calibration Research for the 2015 American Housing Survey

Brian Shaffer, Stephen Ash, Ernest Lawley
U.S. Census Bureau, 4600 Silver Hill Road, Washington, DC 20233

Abstract
In 2015, the American Housing Survey is selecting a new sample cohort of housing units. The prior cohort was selected in 1985 and was interviewed every other year until 2013. With this new sample comes the opportunity to reexamine and revise the weighting methodology. The 2013 methodology included two sets of ratio adjustments that were combined in a raking procedure: one for known totals of housing units and another for population distributions. The adjustment for population distributions used the concept of the principal person to define the distributions. In our paper, we discuss the results of our research into two main questions. First, can we improve the principal-person methodology of the current ratio adjustments and replace it with a calibration weighting adjustment to population totals? Second, can we combine population and housing-unit ratio adjustments into one calibration adjustment? Here we examine whether one calibration adjustment can be employed to adjust for housing unit characteristics that include both housing unit and population characteristics.

Key words: Calibration, Raking, Weighting

1. Introduction

The American Housing Survey (AHS) is a national survey of the U.S. non-institutional housing stock that produces estimates for occupied and vacant units, as well as estimates of characteristics of occupied units. The AHS uses a multistage probability sample where the primary sampling unit (PSU) is a county or group of counties and the ultimate sampling unit is the housing unit. With this sampling design, the Horvitz-Thompson estimator produces an unbiased estimate of the population total of some characteristic of y_i. The estimator is given as

$$\hat{T} = \sum_{i=1}^{n} d_i y_i,$$

where d_i denotes the i^{th} respondent's design weight, the inverse of its probability of selection, adjusted for nonresponse, and n is the distinct set of respondents in the domain of interest.

Although the Horvitz-Thompson estimator is unbiased, one can improve estimates by adjusting the design weights so that the weighted sums equal some externally-provided set of control totals. These adjustments serve several purposes. First, coverage gaps in the sampling frame can introduce bias in the estimates. Second, adjusting sample estimates to control totals produces more powerful statistical tests through a reduction in variance estimates. Third, aligning estimates with those from a trusted source ensure consistency between the survey in question and the source of the totals, which can be a different survey that measures the same characteristic.

This paper will first provide a brief description of the AHS weighting methodology currently in use. Second, it discusses the Generalized Least Squares calibration weight. Third, it reviews existing research on the unification of person and housing unit weights. Fourth, the existing methods are applied to the 2013 AHS-National sample, and the models' impacts on the point and variance estimates are summarized. Lastly, concluding remarks are made.

2. Current methods

The AHS weighting methodology used in the 1980 design uses two separate weighting adjustments that incorporate external information; an adjustment that aligns the sample to externally-provided housing unit totals and an adjustment that aligns units to externally-provided demographic distributions.

The first ratio adjustment is a cell-based method that aligns the sample estimate to independently estimated housing unit totals provided by the Census Bureau's Population Division. Additional independent estimates for post-1980 new construction are provided by the Manufacturing and Construction Division. Estimates of conventional housing completions are calculated with the Survey of Construction, while mobile home totals are calculated with the Manufactured Homes Survey. These new construction totals are provided at the census region level for each year, with cells representing five-year built date increments. While the survey is longitudinal, the weighting methodology produces cross-sectional estimates. Therefore, the estimates of completions and placements must account for units built in those timeframes that were later lost from the housing stock. The demolished housing units are included in the factor calculation for new construction, but excluded from the final totals. The adjusted new construction estimates are subtracted from the total housing unit independent estimates to create fixed old-construction totals.

The second ratio adjustment is a cell-based method that ensures that certain head-of-household ethnicity, racial, marital status, tenure, and age distributions agree with distributions provided by the Current Population Survey (CPS). Additionally, the adjustment ensures the estimated unit vacancy types (for rent, for sale, seasonal, etc.) are distributed consistently with the Housing Vacancy Survey (HVS). Unlike other calibration totals based on actual survey measurements and estimates, these are 'synthetic' in the sense that control totals are obtained by distributing the AHS sample estimate of overall occupied units proportionate to the CPS, and AHS sample estimates of overall vacant units proportionate to HVS vacancy types. These adjustments were made at the census division level in 2013.

These two ratio adjustments are iterated several times, to a predefined tolerance level, to bring the sample estimates into closer agreement with the independent estimates provided in both stages. This methodology is also known as raking; Battaglia *et al.* (2004) provides a good overview of raking. Essentially, when population totals are known marginally for certain domains but not for the cross-classified domains, separate ratio adjustments can be made iteratively to ensure the aggregated sample weights agree to all marginal totals, within a given tolerance. Using the first marginal total(s), the analyst ratio adjusts the sample estimate corresponding with the marginal total(s). The weights calculated with the first ratio adjustment are used to calculate the sample estimate of the next marginal total(s). During adjustments, the marginal estimates from the prior rakes likely fall out of agreement with the control totals, so a second iteration is used. The final raked weight of the first iteration is used in the sample estimate for the next iteration's ratio adjustment to the first control total(s). This is repeated until all estimates agree with the control totals within a tolerance level. The final weight can be seen as the product of the design weight (adjusted for nonresponse) multiplied by all raking ratio adjustment factors.

3. Generalized Least Squares

One focus of this research is to apply Generalized Least Squares (GLS) weighting adjustments to the AHS. GLS weighting is the initial distance function generalized in Deville and Särndal (1992). They cite Zieschang (1986, 1990), who in turn cites Leury (1980) and Roman (1982). GLS weights are created by minimizing the distance function between the design weight d_k and the final weight w_k, $\sum_{k\in s}(w_k - d_k)^2/d_k q_k$; where q_k is an unrelated weight used to generalize the distance function, usually 1; subject to the constraint that the sum of final weights across a column vector of responses equals a row vector of externally-provided totals; $\sum_{k\in s} w_k \mathbf{x_k} = \mathbf{t}_{\boldsymbol{x}}$. Using Lagrange multipliers, this optimization finds one weight for each sample unit that satisfies the entire set of external totals.

$$w_k = d_k\left(1 + \frac{q_k \mathbf{x}_k'(\boldsymbol{t_x} - \boldsymbol{\hat{t}_{x\pi}})}{\sum_{k\in S} d_k q_k \mathbf{x}_k \mathbf{x}_k'}\right)$$

Through substitution, Deville and Särndal demonstrated that the generalized regression estimator of the total t_y is obtained with the given calibration weights.
The estimator of t_y is $\hat{t}_{yreg} = \sum_{k\in S} w_k y_k = \hat{t}_{y\pi} + (\boldsymbol{t_x} - \boldsymbol{\hat{t}_{x\pi}})'\widehat{\boldsymbol{\beta}}$, where

$$\widehat{\boldsymbol{\beta}} = \left(\sum_{k\in S} d_k q_k \boldsymbol{x}_k \boldsymbol{x}_k'\right)^{-1} \sum_{k\in S} d_k q_k \boldsymbol{x}_k y_k$$

Using a matrix algebra software application, the vector of final weights, $\boldsymbol{w}$, can be calculated as

$$\boldsymbol{w} = \boldsymbol{d}(\mathbf{1} + \left((\mathbf{t}_{\boldsymbol{x}} - \hat{\mathbf{t}}_{\boldsymbol{x\pi}})(\mathbf{X}'\mathbf{D}\mathbf{X})^{-1}\mathbf{X}'\right)')$$

where $\mathbf{t}_{\boldsymbol{x}} - \hat{\mathbf{t}}_{\boldsymbol{x\pi}}$ is a 1 x p row vector of differences between each control total and its Horvitz-Thompson estimate calculated with design weights, $\boldsymbol{d}$ is an n x 1 column of design weights for respondents, adjusted for nonresponse, $\mathbf{X}$ is an n x p matrix of survey responses associated with the p control totals, and $\mathbf{D}$ is an n x n diagonal matrix of the design weights.

With all constraints being met exactly and simultaneously, GLS can provide a more efficient alternative to weighting adjustments than raking, which (a) can be resource-intensive due to its iterative nature and (b) may still produce small discrepancies betweent the estimate and the known total (Bankier, 1990) after the convergence acceptance criteria have been met and raking has completed.

Zieschang (1990) commented that nothing in GLS prevents weights from being negative. This can be viewed as counterintuitive to the researcher, who would have to accept a unit representing less than zero units. Bankier (1990) commented that if either the sample is under-represented for a control total or too many constraints are in the matrix, the optimal solution could contain negative weights. Several options are available to remedy this phenomenon. The first option would be ratio estimation with raking. This option would never produce a negative factor. A second option would be through a change to the objective function. Deville and Särndal (1992) and Slud *et al.* (2013) produced penalty functions that, when added to the objective function, penalize extreme weights and ensure a positive weight. This becomes a numerical analysis problem, which can be plagued with nonconvergence. The third option, discussed by Bankier (1990), involves dropping, or collapsing, marginal totals until weights are positive.

A singular matrix is noninvertible and therefore does not have a unique solution. This poses a critical moment in the process of calibration, which is to select a set of external totals that do not produce linear combinations of other variables. If this perfect multicollinearity exists in the $\mathbf{X}$ matrix, $\mathbf{X}'\mathbf{D}\mathbf{X}$ has a determinant of 0 and therefore cannot be inverted. For instance, combining Hispanic persons and non-Hispanic persons, while combining elderly persons and non-elderly persons, produces two identical columns of persons. Therefore, removing one of these four sets of totals can produce a nonsingular matrix and the resulting calibration will produce a set of weights that still allows the analyst to implicitly produce an estimate of the removed group.

4. One weight for persons and housing units

Another focus of this research involves the use of person-level data in addition to housing unit-level data. In a housing unit context, the columns of $\mathbf{X}$ are indicator variables that group sample units so that they can be related to the associated control total and n equals the number of housing unit responses. However, Zieschang (1990) pointed out that $\mathbf{X}$ can consist of sample characteristics whose aggregate values are known with certainty. This shows that GLS weighting can produce calibrated person-level estimates of the number of individuals in the household population with a given characteristic, if all persons in the housing unit are represented.

The concept of producing person-level estimates from household/housing unit data is not novel. The methods discussed here adopt the terminology used in Zieschang (1990) for consistency. Common methodologies include principal-person ratio estimation (Alexander, 1987 – as cited in Jayasuriya and Valliant, 1995), (Zieschang, 1990 cites Hanson (1978, ch. V) and Alexander (1986)), and is used in the American Community Survey (ACS weighting methodology, January 2014); Generalized Least Squares (GLS) (Deville and Särndal (1992), Zieschang (1990) cites Leury (1980) and Roman (1982)), and GLS-Person weighting (Lemaitre and Dufour, 1987 and Zieschang, 1990, who cites Leury (1980, 1986) and Alexander (1987)). All methods seek an integrated weight that can be used for making housing unit and person-level estimates.

Principal-person methodology involves assigning the housing unit weight to each person within the household and adjusting weights to externally-known person totals. Jayasuriya and Valiant (1995) and Zieschang (1990) used the design weight, adjusted for nonresponse. The ACS weighting methodology (2014) includes raking to known housing unit totals, assigning that housing unit weight to each person in the housing unit, then raking to known person totals. After the weighting adjustments are made, each person could have a different weight if the household composition is diverse in such a way that individuals are grouped to produce the sample estimates used in the ratio estimation. To ensure one weight works for housing unit and person estimates, the weight associated with the principal person is assigned as the housing unit weight. Zieschang (1990) assigned the weight associated with the person type with the best coverage, while Jayasuriya and Valiant (1995) remarked that defining a principal person has an element of arbitrariness when deciding which person within the unit is the principal person.

The GLS-Person weight consists of defining a person-level matrix **Z** of size N x p, where N equals the number of persons in the housing units, assuming all persons within the housing units are listed. Within each housing unit, the elements of **Z** are calculated as an average of all persons in the housing unit,

$$Z_{ij} = \frac{U_{hj}}{n_h},$$

where U_{hj} is the total response of characteristic j for household h and n_h is the number of persons in housing unit h. Since each member in the housing unit has the same row value in **Z**, all members will have the same calibrated weight, thereby harmonizing person and housing unit weights.

5. Results

Table 1 shows the independent estimates used as calibration constraints. Housing-unit and household-population totals were produced at the census-division level, while new construction and HUD totals were produced at the census-region level. Coverage ratios were calculated as the ratio of the estimate to the calibration constraint. Non-Black and non-Hispanic household population totals, as denoted with an asterisk, were excluded from GLS methodology to produce a nonsingular matrix. Overall, the sample's national coverage was 95% for housing units and 89% for household population.

Table 1: Calibration constraints by geographic level and summaries of coverage ratios.

Totals, Division-Level	Min Coverage	Max Coverage	Totals, Region-Level	Min Coverage	Max Coverage
HUs Metropolitan	0.88	1.06	Public Housing	0.62	1.36
HUs Micropolitan	0.62	1.48	Vouchers	1.00	1.19
HUs Outside	0.51	1.14	Multi-unit	0.87	0.96
Age 65+	0.87	1.02	Built 05-06	0.85	1.02
LT Age 65	0.83	0.96	Built 07-08	0.97	1.12
Black	0.76	0.93	Built 09-10	0.85	0.93
Non-Black*	0.86	0.99	Built 11-12	0.67	1.05
Hispanic	0.82	1.07			
Non-Hispanic*	0.84	0.96			

*Excluded from GLS

With these constraints, we calculated weights using the GLS and raking methods. Forty-five iterations of raking were needed for factors to converge. Table 2 provides estimates of the marginal totals calculated with GLS and raked weights, as well as estimates calculated with weights adjusted only for non-interviews. Since no external totals were used for vacant units, conformity to the vacancy rate calculated from the sample was deemed an attractive attribute of a candidate method; neither method preserved the vacancy rate calculated from the sample. Using both methods, weight associated with occupied units increased while weight associated with vacant units decreased; this caused a decrease in the estimated vacancy rate. A contributing factor to this reduction in weight associated with vacant units is the difference in coverage between housing units and household population. Since weights of occupied units had to be increased more to meet person-level constraints than housing unit-level constraints, and there was not a constraint imposed specifically on vacant units, weight associated with vacant units was allocated to ensure both housing unit and person constraints were met.

Table 2: Marginal estimates calculated using initial GLS and raked weights.

	Known Totals	Non-Interview Adjusted	GLS	Raking, i=45
Total HUs	132,802,859	125,552,921	132,802,859	132,802,859
Total Occupied HUs		109,144,192	117,835,764	121,913,795
Total Vacant HUs		16,408,728	14,967,095	10,889,064
Total Pop	308,099,169	273,734,627	308,099,169	307,944,082
Persons per Occupied HU		2.51	2.61	2.53
Vacancy Rate		13.1%	11.3%	8.2%

To preserve the estimated vacancy rate, 13.1%, three options were explored using GLS: (1) calibrate only to housing unit totals, (2) calibrate to housing unit totals and household population totals in two stages, and (3) produce ‘synthetic’ occupancy status-based constraints by apportioning the housing unit calibration constraints based on percentages calculated from the non-adjusted sample estimate. We also applied raking methodology using the synthetic constraints. The first two options are related. In the first option, the sample estimates are forced to adhere only to housing unit and new construction calibration constraints. The second option uses the housing unit-adjusted weight as the initial weight for person-level weight adjustments. Both GLS principal-person and GLS-Person adjustments were made. Marginal estimates are given in table 3.

Table 3. Marginal estimates calculated using alternative methods.

	GLS HUs only	GLS Principle-Person	GLS-Person	GLS Synthetic	Raking Synthetic, i=80
Total HUs	132,802,859	139,475,031	139,687,323	132,802,859	132,802,859
Total Occupied HUs	115,425,777	122,097,949	122,310,241	115,426,475	115,426,475
Total Vacant HUs	17,377,082	17,377,082	17,377,082	17,376,384	17,376,384
Total Pop	289,169,272	306,788,849	308,099,169	308,099,169	299,428,237
Persons per Occupied HU	2.51	2.51	2.52	2.67	2.59
Vacancy Rate	13.1%	12.5%	12.4%	13.1%	13.1%

Applying GLS only to housing unit totals preserved the estimated vacancy rate and produced housing unit estimates that matched the housing unit calibration constraints. However, with no constraint on household population, the undercoverage associated with persons in the housing units continued to cause the sample to underestimate the household population. Using GLS principal-person methodology on the occupied units after calibrating only to housing units; when person totals are calibrated separately, the housing unit constraints are not enforced in the second stage of this method. Therefore, weights of occupied units increased to meet person constraints, with only the person constraints as a ceiling. However, the person totals do not match the controls because the head of household weight was used to represent all individuals in the housing unit, and diversity within the housing unit produced diverse within-housing unit weights. The GLS-Person method produced a common person-level weight within each housing unit, and therefore household population estimates matched the calibration constraints. Since this method also did not enforce housing unit constraints in the second stage, the estimate for total housing units was larger than the housing unit constraints. Using GLS on synthetic occupancy-status constraints produced weights that matched housing unit and household population constraints, while maintaining the vacancy rate estimated from the sample. Raking with the synthetic occupancy-status constraints required 80 iterations until factors showed little change across iterations. While household population estimates did not exactly match the constraints, this method preserved the estimated vacancy rate.

GLS-Person and GLS principal-person are similar methods. Since the GLS-Person method produces a common weight for each person in a unit, person-level estimates can be made that match the calibration constraints. Therefore, GLS principal-person calibration was excluded from additional consideration among candidate methods. All methods produced similar, sometimes equivalent weights for vacant units. However, occupied units' weights differed across methods. Figure 1 plots the weights calculated for the four candidate methods were plotted against the initial weight for occupied units.

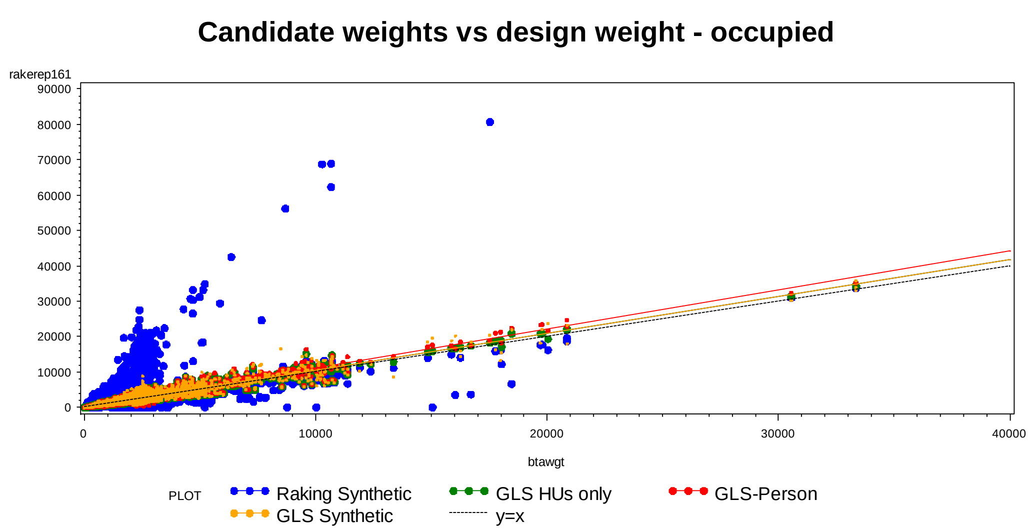

Figure 1: Changes in occupied units' weights after applying adjustments.

Since the GLS methods minimize the difference between the initial and final weights, these methods produced similar weights. GLS-Person weights were slightly larger because the housing unit constraints were not enforced during person-level calibration. Raking, though, produced extreme weights for many occupied units.

To find an explanation on how, although similar, the synthetic method produces person-level estimates that match the constraints while the HU-only method does not, estimates of characteristics related to the household population were calculated with the different sets of weights. Table 4 provides estimates using the candidate weighting methods.

Table 4: Publication estimates (in 1,000s) using candidate weights.

Characteristic	AHS Publication	GLS HUs only	GLS-Person	GLS Synthetic	Raking Synthetic, i=80
1 person in unit	32,268	31,840	33,272	27,803	27,635
2 persons in unit	38,677	38,836	40,985	37,706	39,951
3 persons in unit	18,134	17,948	19,299	18,678	19,264
4 persons in unit	15,288	15,288	16,414	17,038	16,395
5 persons in unit	7,182	7,198	7,713	8,535	7,621
6 persons in unit	2,703	2,702	2,900	3,435	2,855
7+ persons in unit	1,601	1,613	1,727	2,232	1,706
Hispanic HOH	14,675	14,843	15,551	14,716	12,892
Black HOH	15,015	13,938	16,289	14,733	15,169
Elderly HOH	26,784	28,173	28,487	27,735	27,111

Estimates were calculated using the four candidate methods. The AHS publication estimates are also presented. While the GLS HUs only method produces estimates similar to the publication, the GLS Synthetic method produces smaller estimates for housing units containing 1 or 2 persons, and larger estimates for housing units with 3 or more persons. This suggests that the synthetic method is adding more weight to housing units with more persons to meet the household population calibration constraints. However, head-of-household (HOH) estimates calculated with the GLS Synthetic method produce similar estimates to the publication. The Raking Synthetic method behaved similarly for these characteristics, although the lower Hispanic head-of-household estimate provides a good starting point in understanding the extreme weights calculated with this methodology.

Aside from consistency in estimates, another topic of interest in this research is how the different methods impact the precision of the estimates. The AHS uses Successive Differences Replication and Balanced Repeated Replication (Fay & Train, 1995) to produce variance estimates. The methodology simulates repeated sampling from the population and produces a file of replicate weights that, in turn, is used to produce standard error estimates. Wolter (1985) commented on an observed relationship between a survey estimate and its sampling variance estimate. Within a domain, many characteristics exist. Point estimates and variance estimates can be calculated for these characteristics. With an underlying relationship, a least-squares regression curve can be calculated that allows the analyst to obtain a variance estimate for a given population estimate, $\hat{X}$. This curve is known as a generalized variance function (GVF). While the preferred method of calculating a variance estimate is with a direct method, i.e. replication, using the GVF produces a close estimate when computing resources are limited. Many models can be fit to produce a GVF. The model y=a+b/X is used because it produces a good fit to the direct variance estimates and with this model, var(p)=var(1-p). The GVF is helpful in this research, as it allows one to simultaneously evaluate how a given calibration scenario impacts the domain point estimates and variance estimates.

Figure 2 shows that the use of housing unit constraints reduces variance estimates related to total housing units. Most candidate methods produced similar GVFs. However, GLS-Person calibration produced higher variances for larger characteristics and an overall higher estimate of total housing units. Additionally, the sum of principal persons' weights exceeded the externally-provided housing unit totals. This is not surprising, as the assignment of each principal person's weight as the housing unit weight eliminates the property that each replicate-weighted sum of eligible housing units matched the total housing unit control. All methods reduced variances from just using the weights adjusted only for non-interviews.

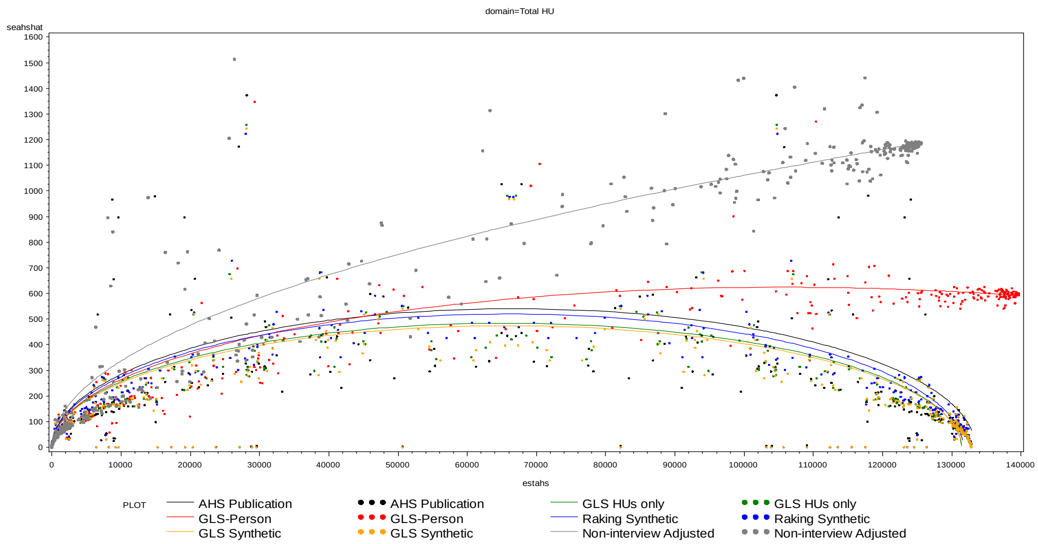

Figure 2: GVFs – Total Housing Units

Figure 3 shows that all candidate methods produced similar variance estimates for the 'total occupied units' domain, and those variance estimates were smaller on average than those calculated with unadjusted weights. Since housing unit totals were unconstrained in this method, the GLS-Person method produced slightly larger estimates of occupied units than the other methods.

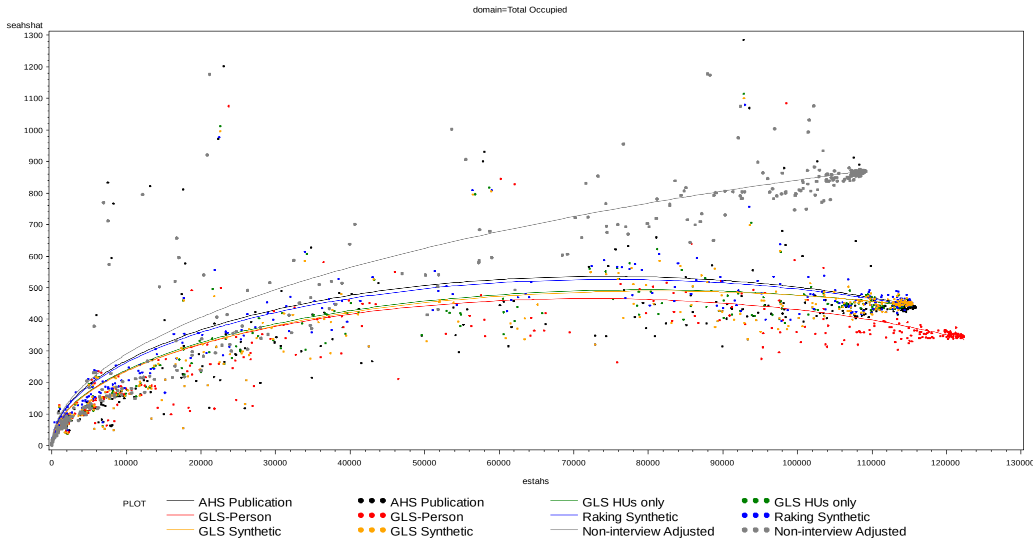

Figure 3: GVFs – Total Occupied Units

Figure 4 shows that all candidate methods produced similar variance estimates for the 'total vacant units' domain, and those variance estimates were smaller on average than the weights adjusted only for non-interviews.

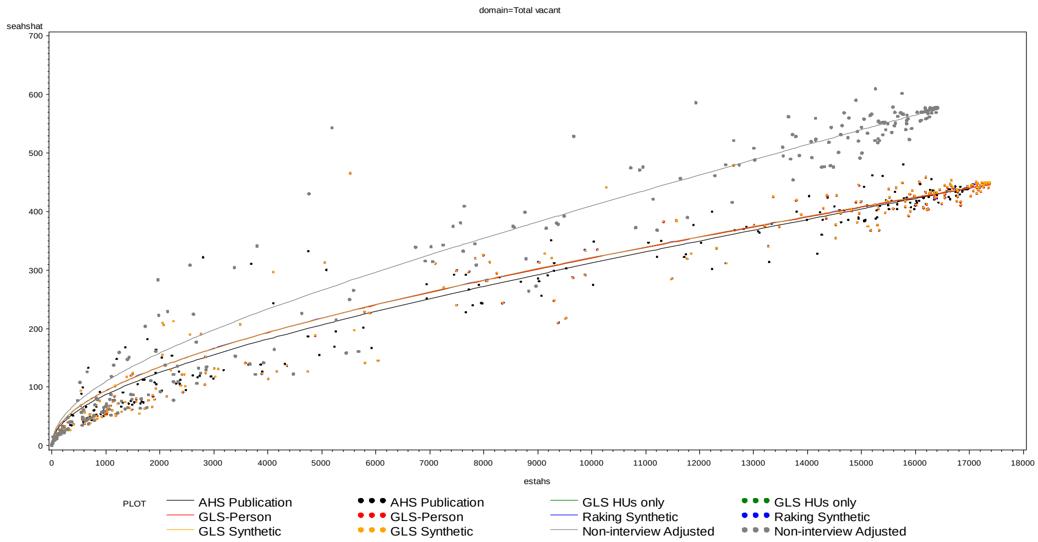

Figure 4: GVFs – Total Vacant Units

Figure 5 shows variances slightly increased from the weights adjusted only for non-interviews when only housing-unit totals were used. For methods incorporating person-level constraints, variance estimates were reduced.

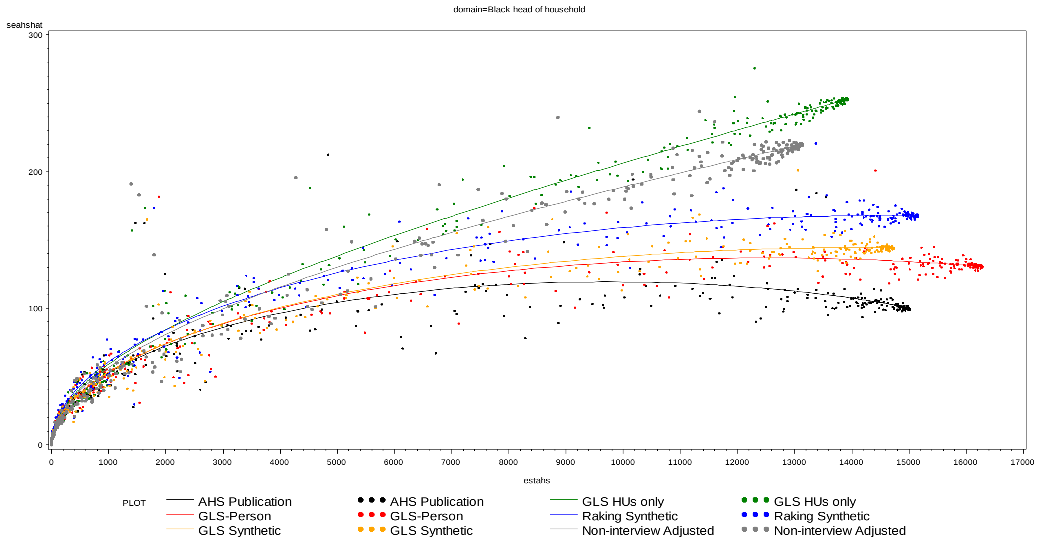

Figure 5: GVFs – Units with Black Head of Household

Figure 6 shows a similar increase in the variance from the weights adjusted only for non-interviews for the Hispanic head-of-household domain when only housing unit constraints are used. However, the raking method produced inflated variances over all methods in this domain.

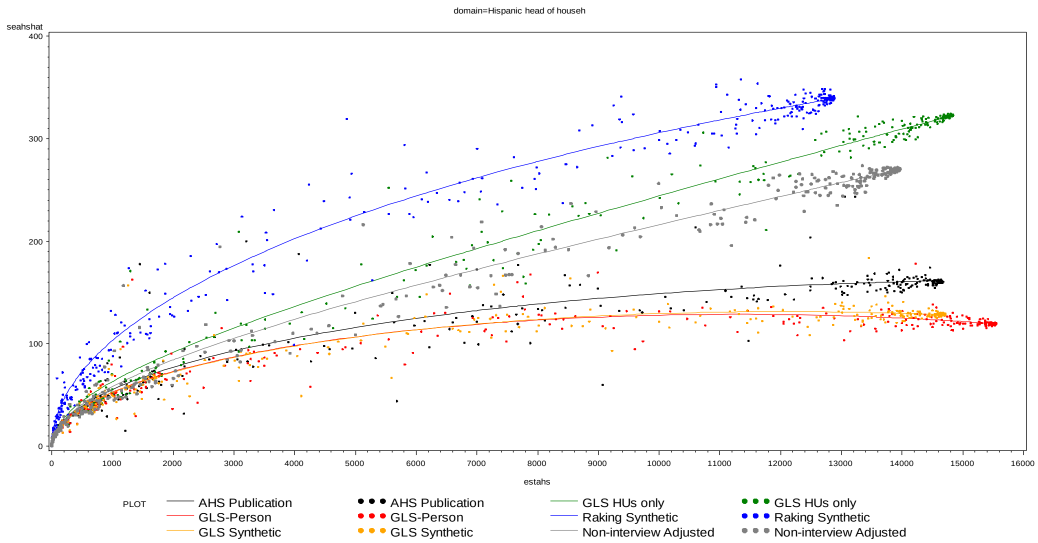

Figure 6: GVFs – Units with Hispanic Head of Household

Figure 7 shows all methods performing consistently for the elderly head-of-household domain. All methods caused variance reduction in this domain.

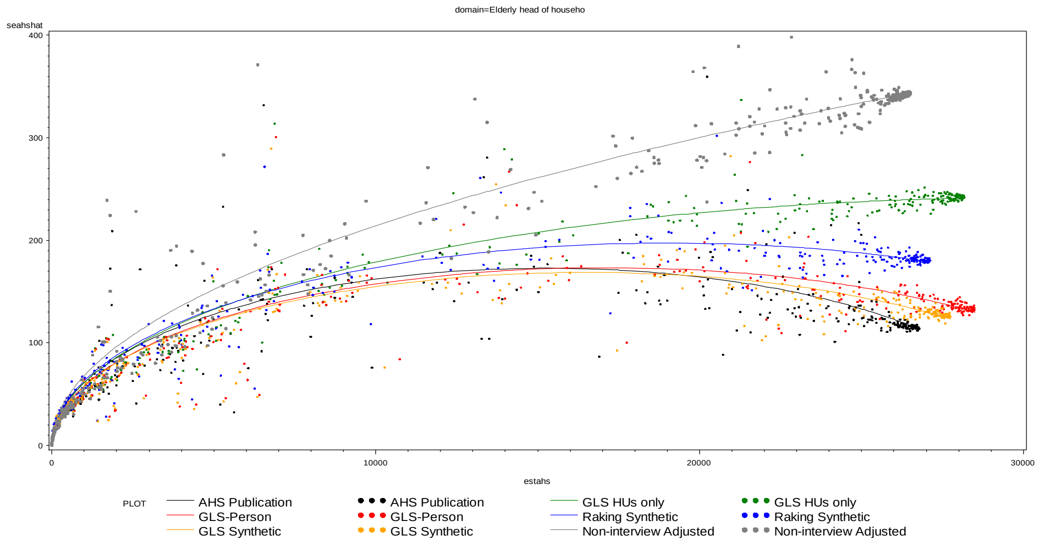

Figure 7: GVFs – Units with Elderly Head of Household

Figure 8 shows all methods performing consistently for the owner-occupied domain. All methods caused variance reduction in this domain.

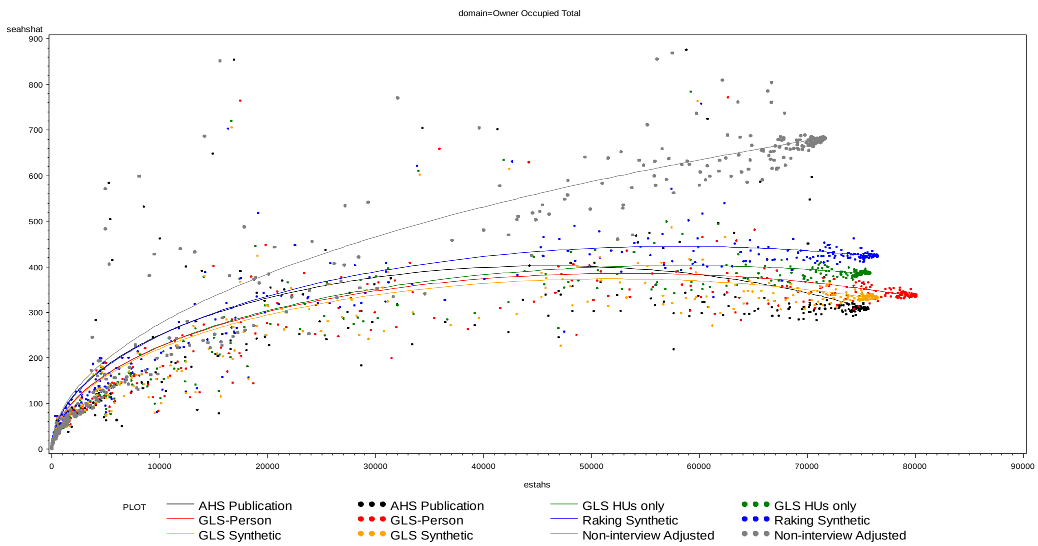

Figure 8: GVFs – Owner-occupied Units

Figure 9 shows all methods performing consistently with the weights adjusted only for non-interviews for the renter-occupied domain. All methods resulted in variance increase over the existing method, which incorporates synthetic tenure-based head-of-household constraints. The raking method produced variance estimates that were larger than those calculated with the weights adjusted only for non-interviews.

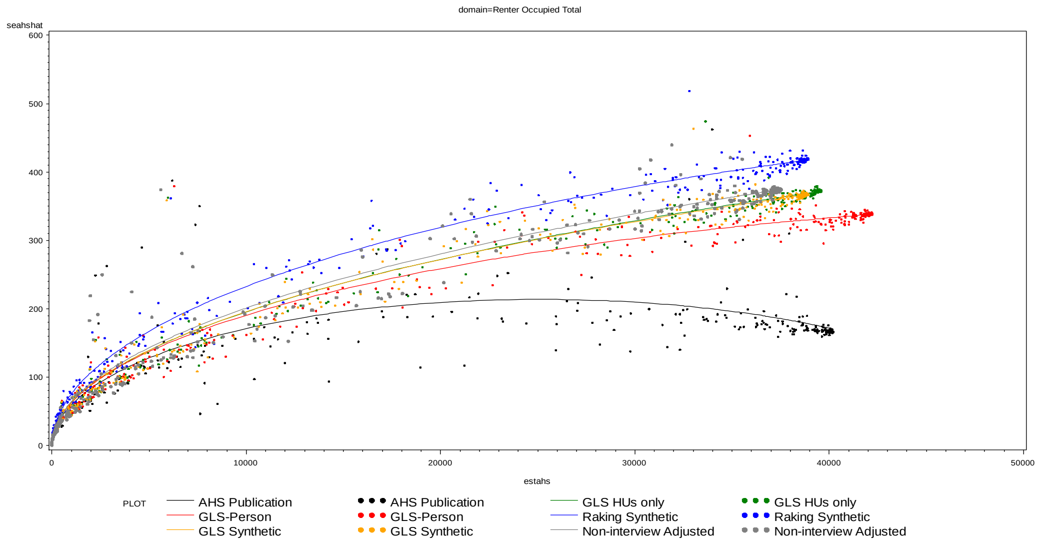

Figure 9: GVFs – Renter-occupied Units

6. Conclusions

The Generalized Least-Squares method inherently finds a vector of weighting adjustments that minimize the difference between the initial and final weights while satisfying many independent calibration constraints. This research combined housing unit-level calibration constraints with person-level constraints. Since the American Housing Survey is a housing-unit level survey, person-level information was summarized both at the housing unit level, and as multiple instances of the housing unit. With either summarization method, a single housing unit-level weight can be calculated that meets person-level constraints.

The AHS has varying levels of undercoverage between housing unit-level and person-level information. Without a method to constrain the number of vacant units, weight from those vacant units is allocated to occupied units to meet all the housing unit and person constraints. Several methods were presented to provide some constraint on the sample estimate of vacant units. When the number of housing units and the number of persons in the housing units are both constrained, these methods rely on housing units with larger numbers of persons to ensure all constraints are met.

The raking methodology is an alternative to GLS. However, its iterative nature makes it a less efficient alternative. Additionally, the use of person-level and housing unit-level constraints produces extreme weights and larger variances than its GLS counterpart. Additional research is needed to understand the exact cause of these extreme weights.

Views expressed in this paper are those of the authors and do not reflect the views or policies of the U.S. Census Bureau.

References

Alexander, C.H. (1987). A Model-Based Justification for Survey Weights, *Proceedings of the Section on Survey Research Methods, American Statistical Association,* pp. 183-188.

Bankier, M.D., (1990). Generalized Least Squares Estimation Under Poststratification, Statistics Canada Working Paper.

Battaglia, M., Izrael, D., Hoaglin, D., and Frankel, M. (2004). Tips and Tricks for Raking Survey Data (a.k.a. Sample Balancing), *59th Annual Conference of the American Association for Public Opinion Research*, pp. 4740-4745.

Deville, J.-C. and Särndal, C.-E. (1992). Calibration estimators in survey sampling, *Journal of the American Statistical Association*, 87, 376-382.

Fay, R., and Train, G. (1995). Aspects of Survey and Model-Based Postcensal Estimation of Income and Poverty Characteristics for States and Counties, *1995 Proceedings of the Section on Government Statistics, American Statistical Association,* pp. 154-159.

Jayasuriya, B. and Valliant, R. (1995). An Application Of Regression And Calibration Estimation To Post-Stratification In A Household Survey, *1995 Proceedings of the Section on Survey Research Methods, American Statistical Association*, pp. 902-907.

Lemâitre, G. and Dufour, J. (1987). An Integrated Method for Weighting Persons and Families, *Survey Methodology*, 13, 199-207.

Slud, E., Grieves, C., and Rottach, R. (2013). Single Stage Generalized Raking Weight Adjustment in the Current Population Survey, *2013 Proceedings of the Section on Survey Research Methods, American Statistical Association,* pp. 195-207.

U.S. Census Bureau (2014). *Design and Methodology,* American Community Survey.

Wolter, K.M. (1985). *Introduction to Variance Estimation*. New York: Springer-Verlag.

Zieschang, K. D. (1990). Sample Weighting Methods and Estimation of Totals in the Consumer Expenditure Survey, *Journal of the American Statistical Association*, 85, 986-1001.