Single-Stage Generalized Raking Application in the American Housing Survey

Brian Shaffer, Yang Cheng, and Eric Slud
U.S. Census Bureau[1], Washington, D.C. 20233-0001

Abstract

The American Housing Survey (AHS) produces many estimates regarding the U.S. housing market. In this paper, we examine the current AHS weighting procedures and apply a newly developed alternative single stage generalized raking method, which adjusts nonresponse, coverage, and calibration simultaneously. We use housing unit totals and head-of-household distributions as calibration constraints in a minimization of the objective function. After adjusting our weights, we develop replicate weights and calculate variance estimates to examine the discrepancies between the current and alternative weight adjustment methods.

Key Words: American Housing Survey, Calibration, Single-Stage Generalized Raking, Soft Constraints, Weighting

1. Introduction

Survey weights are an important element in the estimation of population characteristics. Researchers require each sample unit to provide an approximately unbiased estimate of the segment of the population from which it was drawn. With an inadequate file of survey weights, population estimates (especially for characteristics of smaller domains) are biased and/or variances are inflated to the point where one cannot make any inferences with the sample estimate.

The weighting process begins at sample selection. Each unit is selected from the population with some probability. The inverse of this probability is known as the design weight. If all units in the sample have an equal probability of selection, all units will have the same design weight. Sometimes, the survey sponsor can request an oversample of a specific segment of the population, and the design weight is reduced to account for the fact that the sample unit is representing fewer units in the population. Conversely, subsampling might occur when the interviewer determines more units exist than listed on the sampling frame. In this case, the design weight is inflated to account for the additional units the sampled unit is representing.

During the interview process, some of the eligible sample units (i.e., those in the universe of interest) do not respond, and some units are found by the interviewer to be ineligible. Common practice for ineligible units is to assume the population units represented by those sample units are also ineligible, and the overall population estimate is reduced by the weighted sum of the ineligible units. If an eligible unit does not respond, the units represented by the non-respondent should be distributed to the responses that share common traits with the non-respondent. A good distribution of non-response weight can

---

[1] This report is released to inform interested parties of research and to encourage discussion of work in progress. Any views expressed on statistical, methodological, or operational issues are those of the authors and not necessarily those of the U.S. Census Bureau.

produce smaller variance and mean-squared error than a haphazard distribution of that weight. After non-response adjustments are made, the weighted sum of response indicators equals the sum of design weights for all eligible units.

The non-response-adjusted weights of the respondents can further be calibrated so that the weighted sum of selected response indicators over demographic domains are equal to the corresponding known population totals.

The American Housing Survey (AHS) is a survey of the U.S. housing stock conducted by the U.S. Census Bureau for the Department of Housing and Urban Development (HUD). The AHS is conducted on a sample of housing units, occupied and vacant, every two years, following the same housing units (accounting for growth and demolition) since 1985. The AHS has one sample designed to represent the national housing stock, as well as approximately sixty metropolitan samples selected to represent the housing stocks of those individual areas. The survey is currently undergoing a redesign, and the Census Bureau is selecting an entirely new sample for 2015 interviewing. The redesign also includes a new list of metropolitan areas. This research focuses on the national sample.

2. AHS sample design

The 1985 sample design consists of two stages. The first stage is a selection of Primary Sampling Units (PSUs), each of which is a county or group of counties. Before selection, all PSUs were identified as self-representing (SR) or non-self-representing (NSR). A self-representing PSU is defined by its housing unit count; all SR PSUs had at least 100,000 housing units at the time of selection. These SR PSUs are only representing themselves at the national level. The non-self-representing PSUs in the country were grouped into strata developed to ensure all PSUs in each stratum shared certain predetermined characteristics. From each stratum, one NSR PSU was selected at random to represent the entire stratum. Probabilities of selection were calculated based on the measure of size in each PSU as a proportion of the measure of size of the entire stratum. The 1985 design consists of 170 SR PSUs and 224 NSR PSUs.

The second stage is a selection of housing units from these 394 PSUs. Housing units were selected from the 1980 census so that an overall probability of selection (accounting for the first stage PSU probability of selection) was approximately 1 in 2,148. If units were inside permit issuing areas and had complete addresses, they were sorted and a sample of housing units which received long-form questionnaires in the census was selected from a list of all housing units in the sample PSUs that received long-form questionnaires. For areas that were outside permit issuing areas or had a high percentage of incomplete addresses, a sample of areas was selected. Those areas were divided into smaller areas of land called segments. A sample of segments was then selected, and finally a sample of long-form-receiving housing units in those segments was chosen. A special study was also conducted to identify areas of census undercoverage and select units from those areas. Post-1980 census housing unit growth is captured by selecting a sample from permit offices across the country. After the 2000 census, additional samples of mobile homes and assisted living facilities were selected to account for undercoverage in these two segments of the population.

3. Current AHS Weighting Procedures

The current AHS weighting procedures occur in several stages. First, a cell-based non-response adjustment is made to ensure the weighted sum of responses equals the weighted sum of all eligible units. Then, three sets of ratio adjustments are applied. The first stage adjusts the NSR PSUs to align with census totals in all NSR PSUs. The second stage adjusts the sample to externally-provided housing unit totals. The third stage adjusts the sample to externally-provided head-of-household demographic proportions, as well as to geographic distributions of vacant units.

The non-response adjustment distributes the weight of the eligible non-respondents to the units that did respond. Since the AHS is longitudinal, information from the prior interview is used to determine which interviews share the most in common with the non-interviews. The adjustment aggregates interviews and non-interviews based on the prior year tenure (owner vs. renter); number of units in structure; number of rooms; urban/rural; inside/outside central city; and whether the unit is a mobile home, assisted living facility, or a conventional housing unit. If the unit was previously vacant or if no prior year data are available (in the case of incoming units), less information is available. For units selected from 2000 census long form cases, sufficient information can be obtained from that source. If the sample size within a group is too small, or if the factor is too large, groups are combined using a predetermined set of rules.

The first ratio adjustment is a set of fixed factors that are applied in each enumeration's weighting procedures. After the census, the Census Bureau counted the number of housing units in all NSR PSUs within each census region. Then, the Bureau calculated regional estimates of the NSR housing unit totals by multiplying the census count within the sample PSUs by the NSR PSU weight (the inverse of the PSU's probability of selection). Within each region, the counts were further partitioned by tenure, vacancy status, race/ethnicity, inside/outside central city, and urban/rural groups (cells). Within each cell, the ratio of census totals to estimated totals was calculated (American Housing Survey, 1985).

The second 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. Totals for non-mobile homes are calculated with the Survey of Construction, while mobile home totals are calculated with the Survey of Mobile Home Placements. These new construction totals are provided at the census region level for each year. The current weighting procedures group these totals into 5-year increments ranging from 1980 to the current year of enumeration. The sample estimates corresponding with these cells are first adjusted to include demolitions, and the adjusted new construction estimates are subtracted from the total housing unit independent estimates to create fixed old-construction totals.

The third ratio adjustment ensures that certain head-of-household demographic distributions agree with distributions provided by the Current Population Survey (CPS). The demographic characteristics were believed to be correlated with other characteristics of interest (American Housing Survey, 1985). 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 domain totals are obtained by multiplying AHS sample estimates of occupied units by distributions obtained with the CPS. CPS-based numerators were calculated by multiplying the sample estimate of all occupied units containing the given demographic characteristic by the CPS percentage of households containing that characteristic. Denominators were calculated by obtaining the sample estimate of households containing the characteristic. Likewise, HVS-based numerators were calculated by multiplying the sample estimate of all vacant units by the HVS percentage of vacant units with a given vacancy status, separated by inside/outside central city or outside metropolitan statistical area. In 2009, these adjustments were made at the census region level.

The second and third stage ratio adjustments are repeated several times alternately to bring the sample estimates into closer agreement with the independent estimates provided in both stages.

4. Proposed Single Stage Weight Adjustment Method

Deville and Särndal (1992) and Deville, Särndal, and Sautory (1993) use the term "generalized raking" to refer to calibration methods subject to linear calibration constraints which optimize one of a class of loss functions for the distance between design weights (w_i^o) and adjusted weights (w_i). Their allowed loss functions sum up terms across sampled and responding units which (as functions of w_i, locally near w_i^o) look like $\frac{(w_i-w_i^o)^2}{2w_i^o}$. These papers remark that conventional raking, in which iterative proportional fitting allows two or more sets of marginal totals to be restricted without constraining the cross-classified totals, can be viewed as a special case of generalized raking.

We propose to apply a single-stage generalized raking method, as described in Slud, Grieves, and Rottach (2013), to the AHS. The single stage weight adjustment techniques were developed by Slud and Thibaudeau (2010). Those techniques involved weight optimization with respect to a loss function in the spirit of Deville and Särndal (1992), subject to population-control constraints, with additive penalty terms for discrepancies between weight-adjusted survey totals and corresponding known or base-weighted estimated totals for certain survey attributes, and with an additional nonlinear penalty term designed to force weights not to be too different from the design weights. The novel elements of the research include: defining several appropriate additive quadratic penalty terms corresponding to the current multistage AHS nonresponse adjustment; developing a methodology to define penalty multipliers by tracking properties of current AHS weights across weighting stages; enforcing weight compression by a penalty term in place of the current AHS approach based on cell collapsing; and implementing the method on AHS data for detailed comparison with the weights as currently adjusted in AHS.

Finding agreement in all constraints can be difficult with large sets of constraints. In a large government survey such as the AHS, the sample size is of the order 70,000 while the numbers of controls can be of the order 100–200 and the intermediate constraints can reach the 1000s. Moreover, the survey design is put in place with the idea that final weights should be maintained as similar as possible to the design weights. Starting from the seminal paper of Deville and Särndal (1992), ratio adjustment, raking and linear calibration have been viewed as weight adjustment methods in which (hard) constraints are met while the final weights are determined as close as possible to the initial weights with respect to a loss function.

As described in Kott (2006), Slud and Thibaudeau (2010), and Slud, Grieves, and Rottach (2013), there is a stream of papers from 1992 to the present in which final-stage survey weights are determined by optimizing a loss function with a set of constraints, possibly also including a penalty term enforcing that the ratios of final weight to initial weight never or rarely depart from a bounded interval (*a*, *b*) containing 1. Many of these papers, from Deville and Särndal (1992) up through Slud and Thibaudeau (2010), show that under some superpopulation regularity conditions guaranteeing that the response propensities can be consistently estimated and that the great majority of changes are quite small, the survey estimators are design-consistent. Many of these same papers establish asymptotic normality and provide asymptotic variance formulas based on joint inclusion probabilities based on Taylor linearization.

A key feature of the stream of survey methodology papers cited by Slud and Thibaudeau (2010) and Kott (2006) is that any two of the three goals of nonresponse adjustment, population controls, and weight compression had already been seen to be achievable simultaneously in a single optimization step, and Slud and Thibaudeau showed that one can actually accomplish all three. The method of Slud and Thibaueau allowed the choice of a tuning parameter to control how closely the balance relation used in nonresponse adjustment would be satisfied by the final adjusted weights. Slud, Grieves, and Rottach (2013) expanded on that research to allow for the possibility tested here that there might be a series of k approximate equalities to be satisfied simultaneously - to differing degrees, which might be chosen by the survey analyst or survey client.

The goal of this single-stage weight adjustment method is to show how a linear-calibration loss function can be combined with quadratic penalty terms quantifying inequality and possibly also a penalty term. Objective functions for weight adjustment with penalty terms like the soft constraint terms with coefficients have previously been considered in a simplified survey calibration setting by Fuller (2009, p. 164) and by Datta et al. (2011) in a technique known as “Bayesian benchmarking”. For simplicity of notation, from now on we adopt the convention that adjusted weights are nonzero only for responding sampled units.

The objective function given in Reid, Grieves, and Rottach (2013),

$$J(\underline{\mathbf{w}}) \equiv \sum_{i\in S} r_i \frac{\left(w_i - w_i^o\right)^2}{2w_i^o} + \sum_{k=1}^{K} \frac{\alpha_k}{2} \left\| \sum_{i\in S} r_i w_i \mathrm{x}_i^{(k)} - t_k^* \right\|^2 + \sum_{i\in S} r_i w_i^o Q\left(\frac{w_i}{w_i^o}\right)$$

is minimized such that $\sum_{i\in S} r_i w_i \mathrm{z}_i = t_{\mathrm{z}}^*$. The first quantity is a distance function measuring the difference between the final weight (w_i) and the initial weight (w_i^o). Given the constraint, this is a calibration method; since we are summing over only the responses ($r_i = 1$), this also incorporates nonresponse adjustment by ensuring the weighted sum of z_i values among responders adds to externally-provided totals, t_{z}^* (these totals are referred to as ‘hard totals.’). While the initial weight is traditionally thought of as the pure design weight, the analyst can also apply various adjustments to the weights before using them in the objective function. The second quantity in the objective function calculates the difference between the weighted sum of responses $\mathrm{x}_i^{(k)}$ and a corresponding set of externally-provided totals, t_k^* (these totals are referred to as ‘soft totals.'). The researcher can specify k independent sets, or batches, of soft controls. The α_k parameter is used to control the order of magnitude this part of the objective function contributes to $J(\underline{\mathbf{w}})$. The final summation in the objective function penalizes for extreme ratios of final weight to initial weight. The function $Q(x)$ is defined as

$$Q(x) = A_1 I_{[x \le c_1]} \frac{(c_1 - x)^2}{x - L} + A_2 I_{[x \ge c_2]} \frac{(x - c_2)^2}{U - x},$$

where A_1 and A_2 control the order of magnitude this part of the function contributes to $J(\underline{\mathbf{w}})$, c_1 and c_2 define where the ratio is extreme and the function is called, and U and L define where the function approaches infinity. The objective function is minimized with the resubstitution algorithm proposed in Slud, Grieves, and Rottach (2013). The form of $Q(x)$ used in this paper rapidly increases to infinity near U and L. Similar functions have been used by Deville and Särndal (1992) in their "Case 6" loss function, and also in penalty functions by Singh and Mohl (1996).

The proposed methodology provides several advantages. First, it allows the researcher to impose strict calibration constraints for some attributes while allowing controlled tolerance for discrepancies for other attributes. Next, it allows for clarity of documentation, so that users can see which constraints are 'hard' and what the level of tolerance in the soft constraints is, as provided by the final adjusted weights. Finally, the proposed methodology allows for the possibility of an alternate linearized variance formula, as a cross-check to the balanced replication variance estimation method. As with the current method, estimates calculated from weights using the proposed method should be approximately design-unbiased, with similar variances if penalty-term coefficients are appropriately chosen.

Fuller (2002, 2009 p. 164) remarked that weight compression to satisfy linear inequalities can be accomplished by quadratic programming with linear equality and inequality constraints, under objective functions that are quadratic forms in the weights. A great deal is known about the solution of quadratic programming problems with linear equality and inequality constraints. The topic of numerical optimization via quadratic programming is admirably treated in the book of Nocedal and Wright (1999, Ch. 16). However, not all methods are suitable for extremely large problems, in which the best methods are determined by the special structure of the problem.

5. Applying the Proposed Methodology

We tested the proposed methodology with the 2009 AHS. In this year of enumeration, there were 62,135 sample units interviewed. Of these, there were 53,448 responses, 6,249 eligible nonresponses, and 2,438 ineligible units (e.g., demolitions or incomplete construction).

We created a file of initial weights for the 53,448 responses. Initial weights were calculated by multiplying the base weight (inverse of the probability of selection) by the first stage NSR PSU adjustment factor. We scaled up our initial weights with regional housing unit totals so that the weighted sum of responses equaled our housing unit control totals (23,316,060, 29,403,380, 49,371,526, and 28,020,641 for the Northeast, Midwest, South, and West regions, respectively).

We minimized the objective function given in Slud, Grieves, and Rottach (2013) such that the final weights summed over 24 externally-provided (hard) totals. Within each census region, we developed six controls; two for conventional and mobile post-1980 new construction, and one for old construction. Conventional and mobile post-1980 construction were provided externally, as well as overall housing units. Old construction controls were calculated by subtracting our combined new construction totals from the overall housing unit totals. Our new construction totals are annual estimates, and we therefore had to account for demolitions in the post-1980 growth before subtracting our new construction totals from the overall. Within these three groups, we created synthetic occupied and vacant controls by calculating the proportion occupied from our sample data and applying those percentages to each total. We assumed eligible non-respondents were occupied units when calculating proportion occupied, which is reasonable since most eligible non-responses come from housing units where the occupant refuses to participate in the interview.

Within the objective function, we provided two batches of (soft) totals, which we used in our distance minimization from our final weighted totals. Each batch contained 84 separate synthetic totals, 21 in each census region. Our first batch contained, within each region, eighteen totals corresponding with occupied Hispanic/non-Hispanic heads-of-household based on tenure (owner/renter) and age. We calculated these totals by multiplying the CPS-provided head-of-household proportions by the occupied totals from the objective function's constraints. The final three soft totals within each region were also synthetic, as we multiplied the HVS-provided vacancy proportions based on metropolitan status (inside/outside central city, outside MSA).

Table 1 provides a comparison between the AHS controls and the controls developed for this research. Where synthetic controls were developed, estimates differ from the AHS. Since these totals do not have a true external source, the analyst must still develop them if the proposed weighting procedure is used.

Table 1. Comparison of control totals.

Control	2009 AHS	Hard (t_z^*)	Soft (t_k^*)
Total	130,111,607	130,111,607	
Occupied	111,805,795	111,837,303	
Vacant	18,305,812	18,274,304	
New Conventional	42,806,081	42,806,081	
New Mobile	5,839,049	5,839,049	
Old Construction	81,466,477	81,466,477	
Owner	76,427,983		75,856,942
Renter	35,377,813		35,980,361
Black* HOH	14,470,957		14,786,992
Hispanic HOH	12,728,894		12,739,527

* Black combined with other races

6. Results

6.1 Optimization

We implemented the optimization algorithm in several stages. In our first stage, we set α_k, A_1, and A_2 to equal 0, thereby only including the first quantity in the objective function. This allowed us to see the order of magnitude of the objective function attributed to the first quantity alone and determine acceptable values for α_k, A_1, and A_2 in subsequent stages. $J(\underline{\mathbf{w}})$ equaled 305,612 when minimized. Next, we provided positive values for α_k to allow differences between our weighted totals and soft controls to penalize our objective function. Influence of the soft totals corresponding to Hispanic/non-Hispanic head-of-household are controlled with the α_1 parameter, while those corresponding with Black head-of-household are controlled with the α_2 parameter. We specified α_1=α_2 to provide equal contribution from our Hispanic/non-Hispanic head-of-household totals and Black/non-Black head-of-household totals. Starting with $\alpha_k = 1$, we increased the parameter until $\alpha_k = 7$. Beyond this value, the numerical algorithm would not converge. For α_1=$\alpha_2 = 7$, $J(\underline{\mathbf{w}})$ equaled 366,532. Table 2 shows the ratios of optimized weights, as calculated in this stage, to the initial weights. The optimization never calculated a weight that was extremely different from the initial weight, and therefore $Q(x)$ was never needed.

Table 2. Ratios of final weight to initial weight for each census region.

	$\boldsymbol{\alpha_1 = \alpha_2 = 0}$, $\boldsymbol{A_1 = A_1 = 0}$		$\boldsymbol{\alpha_1 = \alpha_2 = 7}$, $\boldsymbol{A_1 = A_1 = 0}$	
Region	**Ratio_min**	**Ratio_max**	**Ratio_min**	**Ratio_max**
1	0.7783	1.4149	0.7733	1.4217
2	0.8752	1.1600	0.8617	1.1747
3	0.8396	1.1975	0.8243	1.2178
4	0.8594	1.1527	0.8468	1.1717

After the soft controls were used in the optimization, the range of ratios of final weight to initial weight increased slightly, but not to the extent where $Q(x)$ would be needed.

The weights calculated from these two runs were added to create estimates of the control totals we developed earlier, and those estimates were compared to the controls (Table 3). Characteristics for status (occupied/vacant) and pre/post-1980 construction were specified as optimization constraints (hard totals) and therefore the estimates equaled the control totals.

Table 3. Estimates of control totals before and after use of soft totals in optimization.

Characteristic	**Control**	**Before ($\boldsymbol{\alpha_1}$=0, $\boldsymbol{\alpha_2}$=0)**	**After ($\boldsymbol{\alpha_1}$=7, $\boldsymbol{\alpha_2}$=7)**
Total	130,111,607	130,111,607	130,111,607
Occupied	111,837,303	111,837,303	111,837,303
Vacant	18,274,304	18,274,304	18,274,304
New Conventional	42,806,081	42,806,081	42,806,081
New Mobile	5,839,049	5,839,049	5,839,049
Old Construction	81,466,477	81,466,477	81,466,477
Owner	75,856,942	75,246,284	75,187,468
Renter	35,980,361	36,591,019	36,649,835
Black* HOH	14,786,992	13,496,135	13,571,899
Hispanic HOH	12,739,527	13,477,200	13,499,298

* Black combined with other races

During the optimization, weight shifted from owners to renters. The estimate of total Black head-of-household was less than the control total, while the estimate of total Hispanic head-of-household was greater than the control total. Additionally, implementation of the soft controls had mixed effects; the estimate for total Black head-of-household increased closer to the control, while the Hispanic head-of-household estimate continued to increase when it should have decreased.

When including only one batch of soft controls, the algorithm converged for larger values of α_k. When each batch was used independently, convergence occurred for α_k = 13, as opposed to α_k = 7 when both batches were included in the optimization. However, weight ratios did not become more extreme with the larger values of α_k (table 4), and therefore $Q(x)$ was not needed.

Table 4. Ratios of final to initial weights for each independently-used batch of soft totals.

	$\boldsymbol{\alpha_1 = 13, \alpha_2 = 0,}$ $\boldsymbol{A_1 = A_1 = 0}$		$\boldsymbol{\alpha_1 = 0, \alpha_2 = 13,}$ $\boldsymbol{A_1 = A_1 = 0}$	
Region	**Ratio_min**	**Ratio_max**	**Ratio_min**	**Ratio_max**
1	0.7684	1.4189	0.7717	1.4245
2	0.8726	1.165	0.8477	1.183
3	0.8284	1.205	0.8219	1.2257
4	0.8484	1.1568	0.8317	1.1825

Independently applied, the Hispanic head-of-household soft totals had little impact on the estimate of total Black head-of-household units; while the Black head-of-household soft totals had little impact on the estimate of Hispanic head-of-household units. The Black head-of-household soft totals improved the estimate of Black head-of-household units, while the Hispanic head-of-household soft totals hurt the estimate of Hispanic head-of-household units.

Table 5. Estimates of soft-total-related controls using weights calculated with each independently-used batch of soft totals.

Characteristic	**Control**	$\boldsymbol{\alpha_1 = 0, \alpha_2 = 0}$	$\boldsymbol{\alpha_1 = 13, \alpha_2 = 0}$	$\boldsymbol{\alpha_1 = 0, \alpha_2 = 13}$
Owner	75,856,942	75,246,284	75,153,594	75,227,204
Renter	35,980,361	36,591,019	36,683,710	36,610,099
Black* HOH	14,786,992	13,496,135	13,503,673	13,620,525
Hispanic HOH	12,739,527	13,477,200	13,519,486	13,476,294

* Black combined with other races

6.2 Variance Estimation

Files containing replicate weights have recently become an important deliverable for the AHS. Historically, analysts only had access to smoothed variance estimates from generalized variance functions. With the replicate weight file, the analyst can now produce direct variance estimates with ease. Therefore, this research must also produce a replicate weight file as part of its scope.

The Census Bureau calculated replicate factors for the AHS sample using methodology discussed in Wolter (1985) and Fay and Train (1995). In our research, we multiplied each replicate factor by the initial weight (i.e., base weight x NSR PSU adjustment factor) and scaled them up to regional housing unit totals. Due to the synthetic element in our control totals, we also applied replicate weights to our sample estimates when calculating proportion occupied for our hard and soft totals. We developed a file of replicate weights by applying the proposed algorithm to each set of initial weights and corresponding sets of control totals.

The variance estimates we produced with the proposed method were similar in magnitude to those produced with the current methodology for hard-total-related estimates (total occupied, total vacant). For items directly related to our soft totals (renter-occupied, race, and ethnicity), our variance estimates were over twice as large as the AHS variances. Our variance estimates for inside/outside Metropolitan Statistical Area were consistent with the AHS; since the replicate factors create more variation for non-self-representing PSUs than self-representing PSUs, this suggests that the proposed method is not shifting weight between the two PSU types (table 6).

Table 6. Coefficients of Variation (CVs) for select characteristics, 2009 AHS, optimization before using soft totals, and optimization after using soft totals.

Characteristic	AHS CV	$\alpha_1 = 0,\ \alpha_2 = 0$ CV	$\alpha_1 = 7,\ \alpha_2 = 7$ CV
Total	0.0%	0.0%	0.0%
Total Occupied	0.4%	0.4%	0.4%
Owner Occupied	0.4%	0.6%	0.6%
Renter Occupied	0.4%	1.0%	0.9%
Vacant	2.5%	2.5%	2.5%
Year-round Vacant	2.5%	1.9%	1.9%
Seasonal Vacant	2.9%	8.5%	8.3%
Manufactured/Mobile	1.9%	1.9%	1.9%
Inside MSA	1.3%	1.4%	1.4%
Outside MSA	4.9%	5.0%	5.0%
Black Alone	0.5%	1.9%	1.7%
Hispanic	0.4%	2.2%	2.2%

Stronger enforcement of the minimization of the distance between the weighted sums and the soft totals should create more stabilization across replicate-weighted sums. This is evidenced by the fact that the coefficient of variation, CV, decreased for Black-only head-of-household units as the soft totals were included in the optimization. For instance, the replicate-weighted sum of Hispanic head-of-household units had a range of 13,032,064 – 13,878,774. A stronger enforcement of the Hispanic head-of-household control totals would ensure all replicates were closer to the total 12,739,527.

7. Conclusions and Future Research

The single-stage weighting adjustment used in this research can produce reasonable estimates of items associated with optimization constraints (hard totals). Some totals used in the optimization function's minimization (i.e., Black head-of-household soft totals) proved somewhat effective in calculating weights that added back to those totals, but not all soft totals behave the same (i.e., Hispanic head-of-household soft totals). Future research should be conducted to determine why Hispanic head-of-household estimates diverged from the control totals as the influence of the soft totals was increased in the objective function.

The inability of the optimization function to converge for larger values of α_k also seems to hinder the progress in this research. It is possible that Hispanic head-of-household estimates will revert back to control totals for larger values of α_1, as the parameter will add more influence from the penalty function to the overall objective function. Additionally, larger values of α_k may produce final weights that will necessitate the use of $Q(x)$. Therefore, the choice of input parameters for $Q(x)$ would be an interesting topic for future research. However, there is a known approach called gradient projection, different from the resubstitution algorithm used in Grieves, Rottach, and Slud (2013) and in this research, to obtain convergent solutions in the quadratic-formulation of the single-stage optimization problem. This approach is currently under development and should be tested in the AHS application.

The key characteristics as determined by the Department of Housing and Urban Development are: tenure, vacancy status, units in structure, unit type, household population (Black, White, Hispanic, Elderly, and Married), number of rooms in unit, housing unit value, and rent burden (2011). After resolving differences between soft controls and weighted estimates of those controls, review of smaller characteristics should be conducted.

Variance estimation is a critical need for analysts. Should the proposed method's estimates of variance be larger than the current method's estimates even after closer agreement with the soft controls is found, this may provide a future research topic for improving upon the current method. Development of an alternative method of variance estimation based on Taylor linearization for this weight-adjustment method is also a topic for further research.

Since the AHS is longitudinal, the effect of this weighting procedure should be viewed over other years to ensure weights behave relatively consistently within housing units across years of enumeration.

Acknowledgements

The authors would like to thank Stephen Ash, James Farber, and Ruth Ann Killion of Census for their helpful comments on the research.

References

Datta, G., Ghosh, M., Steorts, R. and Maples, J. (2011) Bayesian benchmarking with applications to small area estimation, *Test* **20**, 574-588.

Deville, J.-C. and Särndal, C.-E. (1992), Calibration estimators in survey sampling, *Jour. Amer. Statist. Assoc.* **87**, 376-382.

Deville, J.-C., Särndal, C.-E., and Sautory, O. (1993), Generalized raking procedures in survey sampling, *Jour. Amer. Statist., Assoc.* **88**, 1013-1020.

Fay, Robert, and Train, George (1995), "Aspects of Survey and Model Based Postcensal Estimation of Income and Poverty Characteristics for States and Counties," Proceedings of the Section on Government Statistics, American Statistical Association, Alexandria, VA, pp. 154 159.

Fuller, W. (2002), Regression estimation for survey samples. *Survey Methodology* **28**, 5-23.

Fuller, W. (2009) *Sampling Statistics*. Hoboken: John Wiley.

Killion, R. (2011), Plan for the 2010 Sample Redesign of the American Housing Surveys, U.S. Census Bureau

Kott, P. (2006), Using calibration weighting to adjust for nonresponse and coverage errors. *Survey Methodology* **32**, 133-142.

Nocedal, J. and Wright, S. (1999) *Numerical Optimization*. Springer: New York.

Singh, A. and Mohl, C. (1996), Understanding calibration estimators in survey sampling. *Survey Methoology* **22**, 107-115.

Slud, E. and Thibaudeau, Y. (2010), Simultaneous Calibration and Nonresponse Adjustment, *Census Bureau CSRM Research Report* RRS2010/03.

Slud, E., Grieves, C., and Rottach, R. (2013), Single Stage Generalized Raking Weight Adjustment in the Current Population Survey, *2013 Joint Statistical Meetings paper*.

U.S. Departement of Housing and Urban Development and U.S. Census Bureau. *American Housing Survey for the United States: 1985*.

Wolter, Kirk (1985), Introduction to Variance Estimation, New York: Springer Verlag New York Inc.