A Response Propensity Based Evaluation of the Treatment of Unit Nonresponse for Selected Business Surveys

Katherine Jenny Thompson and Katrina T. Washington[1]
U.S. Census Bureau
4600 Silver Hill Road, Washington DC 20233
Katherine.J.Thompson@census.gov / Katrina.T.Washington@census.gov

Introduction

In the world of survey design, methodologists select a "representative" sample that achieves targeted reliability from a complete frame. In the world of survey collection, not all sample units respond (complete non-response), and those that do will not always provide data on every questioned characteristic (partial non-response). Business surveys publish totals. Consequently, complete-case analyses – analyses based solely on unadjusted respondent data – are biased, so the survey practitioner assumes a model for the nonresponse and develops adjustments to their survey estimates accordingly. Under an ignorable response mechanism, adjustment can be accomplished via weight adjustment or imputation. The choice of adjustment method and model(s) should be determined by the survey publication objectives and supported by analysis of the survey response set.

In this paper, we examine the **unit** nonresponse adjustment procedure used in the Services Annual Survey (SAS). Our analysis is considerably simplified from the production procedures outlined in Nelson (2011). These ongoing programs impute complete records for unit nonresponse using a hierarchy that imputes items in a pre-specified sequence determined by the expected reliability of available imputation models. Each item has its own imputation model hierarchy that maximizes the use of logical edits and direct substitution before attempting model imputation. Our analysis focuses entirely on the ratio model imputation used by these programs. The ratio imputation model utilizes a prediction model between the missing (dependent) variable and the auxiliary (independent) variable. Auxiliary variables differ by item. This approach allows for maximum flexibility in modeling and preserves the expected cell totals, but does not preserve multivariate relationships between items and creates variance estimation challenges. Generally, there is an abundance of reliable auxiliary data that are used to formulate the imputation models for the key totals, such as revenue and expenses. However, these programs also request detailed "breakdowns" of each total, and there is often little auxiliary information on the detail items.

An alternative approach is to develop a single unit nonresponse weight adjustment procedure. We consider two adjustment-to-sample weighting approaches, both described in Kalton and Flores-Cervantes (2003): one that adjusts the respondent units' final weights by the weighted inverse response rate, i.e. the so-called "quasi-randomization estimator" or "count estimator"(Oh and Scheuren, 1983); and an alternative that adjusts the respondent units' final weights by an unweighted inverse response rate, i.e. the InfoP estimator described in Särndal and Lundström (2005, Chapter 7.3) and endorsed by Little and Vartivarian (2005). We also examine the simplified version of the currently-used ratio imputation procedure, which can be rewritten as the InfoS ratio estimator described in Särndal and Lundström (2005, Chapter 7.5).

We use six years of historical data to empirically examine the ignorable response assumptions and the prediction models for each considered adjustment method. First, we describe our analysis approach. We follow with background information on the studied programs, including information on survey design, content, and data limitations. After this, we present and discuss our results, concluding with broad observations.

Developing Adjustment Procedures for Unit Nonresponse

Several factors must be considered before developing a treatment for unit nonresponse. First is the availability of "useful" auxiliary data for each sampled or population element, in the sense that the auxiliary data is related to nonresponse or is predictive of survey outcome. As mentioned above, survey totals that use only unadjusted

[1] This paper is released to inform interested parties of research and to encourage discussion. The views expressed are those of the authors and not necessarily those of the U.S. Census Bureau.

respondent data are biased. When implementing an adjustment procedure to reduce the magnitude of this bias, care must be taken to avoid large increases in variance as a consequence. Little and Vartivarian (2005) summarize the effects of adjustments on bias and variance of a mean or total by strength of association of the adjustment cell variables with nonresponse and outcome by Figure 1, with the optimal adjustment procedure being indicated by the (shaded) Cell 4.

		Association with Outcome (Prediction)	
		Low	High
Association with Response (Propensity)	Low	Cell One Bias: -- Variance: --	Cell Two Bias: -- Variance: ↓
	High	Cell Three Bias: -- Variance: ↑	Cell Four Bias: ↓ Variance: ↓

Figure 1: Effects of Adjustments on Bias and Variance With Respect to Auxiliary Variables
Little and Vartivarian (2005)

How is this balance achieved in practice? First, the population must be partitioned into P disjoint subpopulations (nonresponse adjustment cells) using $\underline{\boldsymbol{x}_p}$, a vector of auxiliary categorical variables available for all units in the population. This partitioning process assumes that $\underline{\boldsymbol{x}_p}$ is related to response propensity, and that the respondents comprise a random subsample within the nonresponse adjustment cells. After establishing the nonresponse adjustment cells, the unit nonresponse adjustment procedures can be evaluated.

Throughout the remainder of this document, we use the following notation. Each sample S has been partitioned into P disjoint subpopulations (indexed by $p = 1, 2,\ldots, P$), each containing n_p units. Within each subpopulation p, $s_{p,r}$ = the respondent portion of the sample containing r_p units and $s_{p,nr}$ = the nonrespondent portion of the sample containing m_p units. Each unit j within subpopulation p has a design weight w_j. Auxiliary data $\underline{\boldsymbol{x}}$ are available for all sampled units, and elements in $\underline{\boldsymbol{x}}$ may be either continuous (indicated by x_{pj}) or categorical (indicated by x_p). Characteristic data $\underline{y}$ are collected in the survey, and values in $\underline{\mathbf{y}}$ may or may not be reported. Each unit j has a response indicator I_{pj} whose value is assigned after data collection; each I_{pj} has a probability of response φ_{pj}.

Evaluation Procedures
Response Propensity Analysis

Under the "quasi-randomization model" (Oh and Scheuren, 1983; Särndal and Lundström , 2005, Chapter 6.1), unit nonresponse is treated as the second phase of a two phase sample, where each sampled unit j has an unknown probability of responding(φ_{pj}). Thus, r_p is a random variable with expected value $n_p\varphi_j$.

We consider two different ignorable response mechanisms: covariate-dependent and missing-at-random (MAR). The covariate-dependent response mechanism assumes that in each subpopulation p, $P(I_{pj}=1 \mid y_{pj}, x_{pj}) = P(I_{pj}=1 \mid x_{pj}) > 0$, $E(y_{pj} \mid x_{pj}) = \beta x_{pj}$, and $V(y_{pj} \mid x_{pj}) = \sigma_p^2 x_{pj}$. Thus, the probability of unit j in subpopulation cell p is related to the observed auxiliary value, not the characteristic of interest, and the expectation and variance of the characteristic of interest are likewise statistically related to the auxiliary variable. For example, x_{pj} could be the frame measure of size (MOS) for unit j. Because business populations tend to be positively highly skewed, the MOS is often highly correlated with probability of response i.e., the larger units are more likely to provide response data. Moreover, survey designers take great care to develop MOS variables that are highly correlated with characteristics of interest. In the applications presented below, the MOS is generally highly correlated with the collected totals items, less so with the detail items. Under this response mechanism, it can be shown that the best linear unbiased estimator (B.L.U.E.) for each y is the ratio estimator presented later in equation (4).

The MAR response mechanism assumes that in each subpopulation p, $P(I_{pj}=1 \mid y_{pj}, x_{pj}) = P(I_{pj}=1 \mid x_{pj}) = \varphi_p > 0$, i.e., φ_p is constant within sub-population p. Thus, the MAR unit nonresponse adjustment uses an "inverse response rate" adjustment to the design weights to produce an "unbiased" total estimate. Under this model, the literature (Kalton and Flores-Cervantes, 2003) has repeatedly shown that the minimum requirements for developing weighting cells for unit nonresponse adjustment are that:

i. The probability of response (φ_{pj}) for unit *j* in cell *p* is the same for all sampled units *j* (j = 1, 2, …, n_p) in cell *p* ($\varphi_{pj} = \varphi_p$) or $E(y_p) = \mu_p$ for all units *j* in cell *p*; and
ii. The probability of response in cell *p* differs from that of cell p' for all $p \neq p'$ or the cell mean for variable *y* differs by weighting cell.

We examine whether a given set of adjustment cells satisfies these two conditions via logistic regression analysis and via comparisons of unit response rates (URR). Response propensity modeling (Little and Rubin, 2002) uses logistic regression analysis to determine sets of explanatory covariates that result in adjustment cells with differing response propensities. Our response propensity model estimates $\hat{\phi}_p$ with the model $\phi_p = f_p(\beta)$, where $v_p = \log[f_p(\beta)/\{1-f_p(\beta)\}] = x_p'(\beta)$, where $\underline{x_p}$ is the vector of categorical values defined earlier, and β is a vector of unknown parameters estimated via maximum likelihood estimation with the appropriate "sandwich" estimator variance estimates. We fitted the logistic regression models with the SAS SURVEYLOGISTIC procedure, which incorporates the complex survey adaptations outlined in Roberts, Rao, and Kumar (1987). Certainty units are excluded from these analyses to facilitate appropriate complex survey adjustments to the test statistics (the SAS procedure excludes certainty units from the sandwich variance estimation computations).

For each fitted logistic regression model, we test the goodness-of-fit hypothesis (H_0: β=0). If we fail to reject the null hypothesis, we examine the marginal test results for each individual cell. With these analyses, rejecting the null hypothesis provides evidence that the variables used to construct adjustment cells are related to response propensity. We can add a continuous covariate to x_p such as the MOS to assess whether the size of the unit is a predictor of unit nonresponse.

The logistic regression analysis can indicate that the covariates used for the adjustment cells are predictors of response propensity. Alone, however, the model analysis does not provide empirical evidence of condition (ii), i.e. different response propensities by adjustment cell. For this, we compute unit response rates (URRs) within adjustment cells (see Särndal and Lundström , 2005, Chapter 3.3).

Unit Nonresponse Adjustment Procedures

The imputed estimator of the population total for characteristic *y* is given by

$$\hat{Y}_I = \sum_p \left[\sum_{j \in s_{p,r}} w_{pj} y_{pj} + \sum_{j \in s_{p,nr}} w_{pj} y_{pj}^* \right] = \hat{Y}_R + \hat{Y}_M \qquad (1)$$

where y_{pj}^* is the imputed value obtained for nonrespondent unit *j* in adjustment cell *p*. We present three imputation models, both of which are re-expressed as adjustment weighting models. We use the variance estimation procedure from Shao and Thompson (2009) to obtain reweighted estimate variances.

The first two models assume a MAR response mechanism, with the true unknown response probabilities estimated by the weighted and unweighted unit response rates, respectively.

Weighted Mean (Count): Imputed value $y_{pj}^{\bullet} = \dfrac{\sum_{j \in s} w_{pj} y_{pj} I_{pj}}{\sum_{j \in s} w_{pj} I_{pj}} = \hat{Y}_R^p \big/ \hat{N}_R^p$

This imputed estimate can be re-expressed as the reweighted **count** estimate

$$\hat{Y}^C = \sum_{j \in s_{pr}} \left\{ \frac{\hat{N}^p}{\hat{N}_R^P} \right\} w_j y_{pj} = \sum_{j \in s_{pr}} w_{pj}^c y_{pj} \qquad (2)$$

Unweighted Mean (Count_u): Imputed value $y_{pj}^{*} = \dfrac{\sum_{j \in s} y_{pj} I_{pj}}{r}$

This imputed estimate can be re-expressed as the reweighted **count _u** estimate

$$\hat{Y}^{C\_U} = \sum_{j \in s_{pr}} \left\{ \frac{n_p}{r_p} \right\} w_j y_{pj} = \sum_{j \in s_{pr}} w_{pj}^{c\_u} y_{pj} \qquad (3)$$

Both adjustments are designed to reduce or eliminate the bias caused by unit nonresponse and require that the adjustment cells satisfy conditions (i) and (ii). Besides the bias reduction properties, these adjustment methods have the advantage of computational simplicity. However, the additional stage of weighting will cause the variance to increase, especially when the adjustment factors are quite variable (Kish, 1992; Kalton and Flores-Cervantez, 2003; Little and Vartivarian, 2005). Indeed, when the categorical variables used to define the weighting adjustment cells are associated only with response propensity and are not predictive of characteristic outcome, then the bias amelioration of the reweighted estimates is minimal but carries a cost of increased variance.

Ratio imputation models are designed to reduce the nonresponse bias and increase the estimate precision of a given characteristic y. These models require a strictly positive auxiliary variable x.

Ratio Model (Ratio): Imputed value $y_{pj}^{*} = \left[\dfrac{\sum_{j=1}^{n_p} w_j y_{pj} I_{pj}}{\sum_{j=1}^{n_p} w_j x_{pj} I_{pj}} \right] x_{pj}$

This imputed estimate can be re-expressed as the reweighted **ratio** estimate

$$\hat{Y}^{R} = \sum_{j \in s_{pr}} \left\{ \frac{\hat{X}^{p}}{\hat{X}_{R}^{P}} \right\} w_j y_{pj} = \sum_{j \in s_{pr}} w_{pj}^{R} y_{pj} \qquad (4)$$

Under the covariate-dependent response mechanism, this ratio estimator is the B.LU.E of β, and the resultant ratio estimate will have increased precision over the **count** or **count_u** adjusted estimate. If the prediction model is not valid or if the strength of association between x and y is weak, then the bias induced by the ratio estimator increases the MSE over the corresponding **count** or **count_u** estimates. This situation can occur with the ancillary survey values -- such as detail items -- that are not primary characteristics of interest.

Our studied business programs utilize two ratio imputation models:

Trend $\quad y_{tj} = \beta y_{t-1,j} + \varepsilon_{tj},\ \varepsilon_{tj} \sim (0, y_{t-1,j}\sigma^2)$

Auxiliary $\quad y_{tj} = \beta x_{tj} + \varepsilon_{tj},\ \varepsilon_{tj} \sim (0, x_{tj}\sigma^2)$

where t indexes the statistical period. Within imputation cell, the imputation base used to estimate β is restricted to complete **respondent** data[2] after outlier detection and treatment. This imputation procedure is referred to in-house as the "ratio-of-identicals" procedure. We assess the goodness-of-fit of these models to our survey data using the SAS SURVEYREG procedure within the six-digit industry cells, again excluding certainty cases.

[2]The **trend** model parameters use reported or imputed data from respondents.

Random Subsample Analysis

For each considered set of adjustment cells, we used a contingency table analysis to determine whether respondents comprise a random subsample within adjustment cell p. For nonresponse adjustment cells that are not equivalent to sampling strata (e.g., industry cells), we constructed the 3 x 2 contingency table shown in Figure 2 within each imputation cell. For these analyses, the design weight (w_j) is the analysis variable, and P_{33} and P_{66} represent the 33rd and 66th percentiles, respectively, of the imputation cell's distribution of w_j. We chose these percentile values to ensure that each cell in the contingency table contains at least five noncertainty units; we did not perform any type of "optimality" analysis.

The null hypothesis of interest is that response status is independent of unit size. Failing to reject the null hypothesis provides evidence of a random subsample within imputation cell. To perform the chi-squared tests for independence with the Rao-Scott adjustment procedure to account for complex survey sampling (Rao and Scott, 1992), we used the SAS SURVEYFREQ procedure. As with the logistic regression analyses presented above, certainty units are excluded from this analysis.

	Respondent	Nonrespondent	
$0 <= w_j < P_{33}$	n_{11}	n_{21}	$n_{1\bullet}$
$P_{33} <= w_j < P_{66}$	n_{12}	n_{22}	$n_{2\bullet}$
$P_{66} <= w_j$	n_{13}	n_{23}	$n_{3\bullet}$
	$n_{\bullet 1}$	$n_{\bullet 2}$	$n_{\bullet\bullet}$

Figure 2: Contingency Table for Tests of Independence

When using sampling strata as adjustment cell, we performed a similar analysis using a 1 x 2 contingency table because the noncertainty units in the same stratum generally have the same design weight.

Survey Background

The Service Annual Survey (SAS) is a mandatory scientifically designed sample survey of approximately 75,000 employer businesses having one or more locations that provide services to individuals, businesses, and governments. The SAS surveys companies in NAICS sectors 22, 48-49, 51, 52, 53, 54, 56, 61, 62, 71, and 81. The SAS collects aggregate and detailed revenues and expenses, e-commerce, exports and inventories data from a sample of business firms with paid employees. Additionally, North American Product Classification System (NAPCS) data are collected in selected industries. SAS uses a stratified random sample design with one certainty stratum and multiple non-certainty strata assigned for each industry. In the initial sampling, companies are stratified by their major kind of business (determined by the industry containing the largest portion of total receipts for the company), and then sub-stratified by estimated annual receipts or revenue. All companies with total receipts above applicable size cutoffs for each kind of business are included in the survey as part of the certainty stratum and are asked to report for all their service industry locations. For companies with receipts below the applicable size cutoff, the Employer Identification Numbers (EINs) of these companies are then stratified by major kind of business and sub-stratified by total annual receipts or revenue. Within each noncertainty size stratum, a simple random sample without replacement of (EINs) is selected. Thus, the sampling units are either companies or EINs. Each sampling unit represents one or more establishments/locations owned or controlled by the same firm. The initial sample is updated quarterly to reflect births and deaths, adding new employer businesses identified in the Business and Professional Classification Survey and dropping firms and EINs that are no longer active.

The SAS collects total revenue, total and detailed expenses, and e-commerce for all industries, both taxable and tax-exempt; sources of revenue and expenses by type, as well as export and inventory data for selected industries; operating expenses for tax-exempt firms; and other selected industry-specific items. The key items collected by SAS are total revenue and total expenses. For both revenue and expenses, there are many detail revenue and expense items collected that sum up to their respective totals. Collected detail revenue items vary across industries within NAICS sectors. Expense detail items are primarily the same for all sectors, with an occasional additional expense detail or two collected for select industries. Appendix One presents item relationships in SAS.

Data collection and nonresponse adjustment for total revenue and total expenses are much less problematic than for the detail items. Companies are usually able to proportion out their "bottom line" revenue or expenses in a number of ways. However, methods of bookkeeping, limited staffing, company structure, and smaller company size may make it difficult for some respondents to calculate or even estimate values for a number of requested detailed revenue or expense items. For example, companies may do all their accounting by region, as opposed to by types of industries in which they do business. Similarly, a company's line item in its bookkeeping may hold their expenses for **all** computer needs -- both hardware and software -- together. However, the SAS collects hardware and software expenses separately. For nonrespondents, missing values are replaced by predicted values obtained using appropriate imputation models for nonresponse. Imputation methodology is used to account for both unit and item nonresponse in the SAS. The imputation models use auxiliary survey and administrative records data as input. Survey and administrative records data for the detailed receipts and expenses are often very sparse or nonexistent. Thus, the imputation models for these variables are not as reliable nor predictive as desired. The research outlined in this paper was conducted to determine if weight adjustment could be used as an alternative nonresponse adjustment methodology for SAS.

For processing purposes, the SAS is divided into five subsurveys, each covering one or more NAICS service sectors. This research is focused on the SAS subsurveys covering the transportation and health industries (SAS-T and SAS-H, respectively). These subsurveys cover NAICS sectors 48-49 and 62.

Results

All of the presented analyses use six years of historical data. In the two studied data sets, the six-digit industry code (NAICS), tax-exempt status (all units in SAS-T are in taxable industries), certainty status, weights, and sampling stratum are available for all sampled units, and frame measure of size (MOS) is available for most units.

Response Propensity Analysis (Logistic Regression Analysis)
The first step of our evaluation was to determine candidate sets of auxiliary variables that could be used to form adjustment cells. Currently, SAS-T uses six-digit industry to define adjustment cells; SAS-H uses six-digit industry and tax-status (taxable or tax-exempt) to define adjustment cells. Certainty and noncertainty units are not considered separately and are jointly used to develop ratio imputation parameters. For these analyses, we fit four separate logistic regression models using noncertainty cases:

Model 1: x_p = Current procedure adjustment cells
Model 2: x_p = Current adjustment cells and frame unit MOS (a continuous variable)
Model 3: x_p = frame unit MOS nested within current adjustment cells
Model 4: x_p = sampling strata (noncertainty)

Model 1 evaluates the current adjustment cells. Model 2 addresses whether the existing adjustment cells need to be further refined to incorporate unit size by testing whether MOS contributes to response propensity, given the existing covariates. In all cases, the MOS was significant, and consequently, Model 3 addresses whether unit size **within** adjustment cell yields an improved propensity model. Lastly, Model 4 uses sampling strata, which incorporates both primary industry and unit size in the cell definitions [Note: sample sizes within a stratum can be quiet small due to fine stratification and limited numbers of "large" units in the population.]

Appendix Two provides counts of significant logistic regression goodness-of–fit tests and individual parameter tests for the six years of historical data for SAS-H and SAS-T for Models 1, 3, and 4; the MOS covariate in Model 2 was always significant (6 of 6 years), and the other parameter tests are – as expected -- identical. Each tallied count represents the results from one statistical period, with a maximum of six possible tests. Recall that significant tests results provide evidence that $\beta \neq 0$, i.e. that the auxiliary vector is a predictor of response propensity.

For both SAS-H and SAS-T, the logistic regression analysis provides evidence that industry classification is not strongly related to response propensity (Model 1 results), but the unit size (MOS) and MOS nested within industry are (Models 2 and 3). It also provides evidence that stratum classification is related to response propensity. Given these results, we decided to evaluate two alternative sets of adjustment cells in addition to the currently-used six digit industry cells:

1. Six-digit industry with four separate unit size class adjustment cells defined within industry (one certainty industry and the other three defined using the design weight classification parameters defined for the contingency table analyses presented in 4.1.2)
2. Sampling stratum with two separate adjustment cells defined by certainty and noncertainty status whenever the within-cell sample size was five or greater; some form of collapsed stratum meeting the sample size and common sector criteria otherwise. This collapsing was necessary because several strata comprised two or fewer certainty units.

Appendix Three presents average unit response rates (URR) for each considered set of adjustment cells for SAS-H and SAS-T. With each set of adjustment cells, unit response rates appear to be quite different, providing some indication that response propensities do differ by imputation cell. However, no statistical tests were performed to validate this observation.

Prediction Model Analysis

Ratio estimation improves estimate precision if and only if the auxiliary (independent) variable is highly positively correlated with the (imputed) dependent characteristic. Appendix Four summarizes the regression analysis results for SAS-H and SAS-T. We present two separate diagnostics: the results of the goodness-of-fit test (H_0: β=0) and average R^2. Note that some results are missing. For both programs, trend model evaluations cannot be performed with the 2005 data sets. With SAS-T, there was insufficient data from noncertainty units to fit some detail regression models in the 2006 data sets. With SAS-H, the complete set of collected detail items differs sometimes by statistical period.

Figure 2 summarizes the predictive model analysis by plotting average model-R^2 values obtained from fitting the weighted list squares auxiliary ratio-of-imputation models for SAS-H; trend models are not considered. Blue diamonds indicate where null hypothesis (β=0) was rejected in 80-percent or more of the tests. The blue vertical asymptote separates the total items (182100/Total Payroll and 190000/Total Expenses) from the details. The grey horizontal asymptote is located at R^2 = 0.75. We define any observation that falls above this horizontal asymptote as a “highly predictive” model. Notice that the regression models used for the total items are significant and highly predictive, whereas the majority of the detail item models are not significant. The two apparently predictive “not-significant” models for detail items 4061A00 and 189900 are attributed to random sampling error.

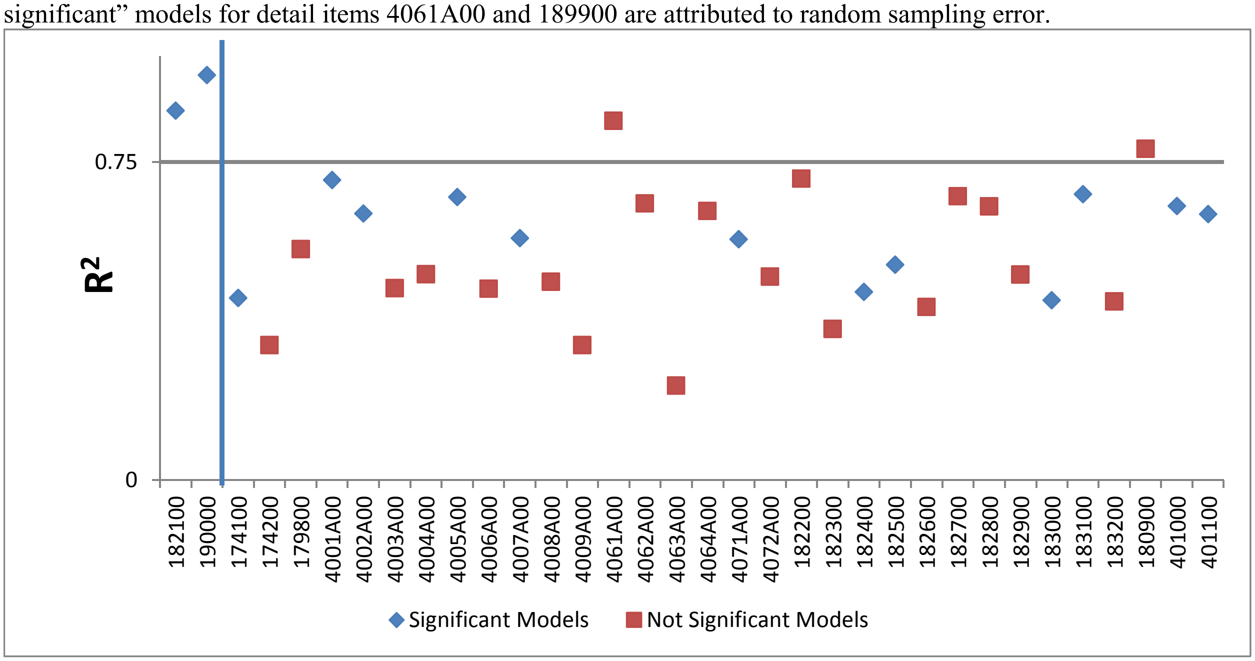

Figure 2: Prediction Model Analysis for SAS-H

Figure 3 summarizes the predictive model analysis by plotting average model-R^2 values obtained from fitting the weighted list squares auxiliary ratio-of-imputation models for SAS-T.

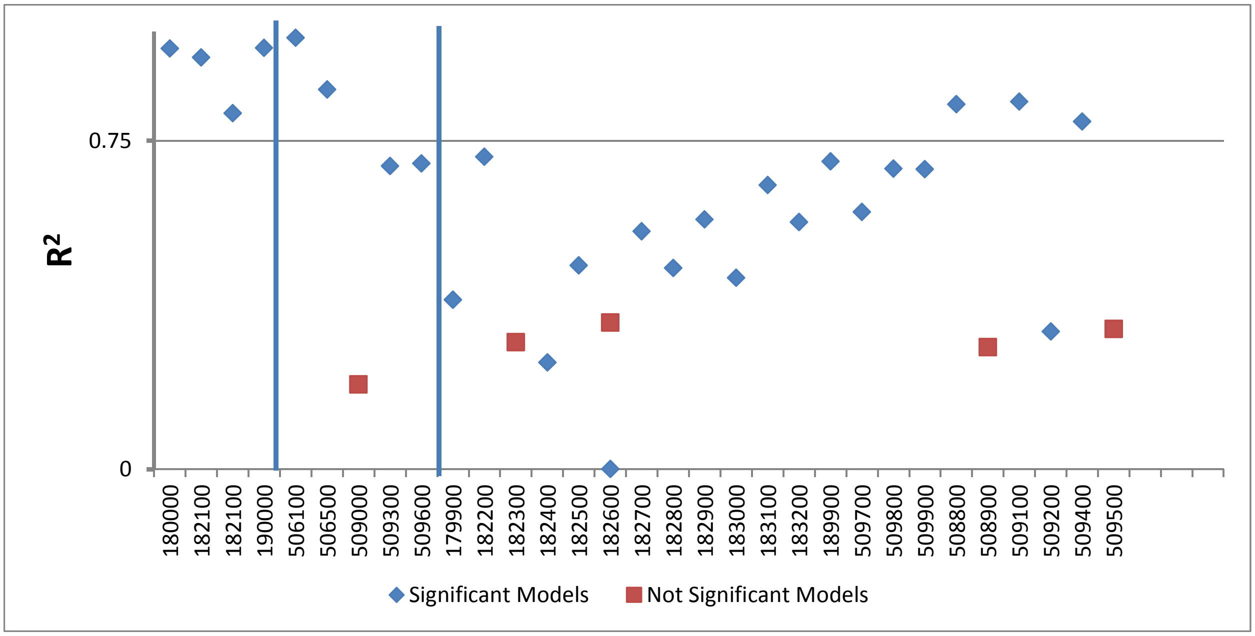

Figure 3: Prediction Model Analysis for SAS-T

As with the SAS-H models, the prediction model results are quite strong for the total items (180000/Total Payroll, 182100 and 190000), providing some validation of the ratio adjustment procedure. For the subtotals, (506100, 506500, 509000, 509300, and 509600), this strong relationship between independent variable and prediction variable holds for two of the five items. For the detail items, however, the majority of the models items do very little to explain the overall variance in the dependent variable, even though the regression parameters are significant.

Random Subsample Analysis

Appendix Five summarizes the results of the chi-squared tests for independence. These tests evaluate whether response is independent of size for the industry cells (Current Adjustment Cells columns) and whether the response is a random subsample within stratum (Strata Adjustment Cells columns). Consequently, rejecting the null hypothesis provides evidence against a random subsample.

Recall that the Pearson chi-squared test for independence requires a minimum of five units per cell. The Rao-Scott modification for sample survey data requires that each cell contain an effective sample size[3] of five and generally sets a minimum value of five sampled units per cell in addition. These sample size constraints limited our analysis.

For SAS-H, 11 of the 74 industry by tax status imputation cells were excluded from the 3*x*2 contingency cell analysis; 44 of the 63 remaining cells had "good fits," providing evidence that the currently used adjustment cells could be "improved" by adding within-cell size categories. In SAS-H, each primary stratum is further classified by size category. Sixteen of the 79 primary strata were excluded from the analysis due to sample size limitations; another 20 were excluded due to small within-size-class counts. Thirty-five of the remaining 43 strata had good fits, however.

For SAS-T, only four of the 12 industry imputation cells had good fits, providing evidence that subdividing industry by size does not result in more representative samples This result is not unexpected, as the sampling weights for non-certainty SAS-T units are not very variable, especially when compared to SAS-H. Ten of the 22 SAS-T strata were excluded from the analysis due small sample size, but the remaining 12 strata had good fits. Thus, for SAS-T, the test results provide evidence against a random subsample of units within industry cells, but provide evidence of a random subsample within noncertainty strata.

[3] Actual sample size divided by the design effect

For both SAS-H and SAS-T, a visual inspection of the within-industry or within-industry/tax status contingency tables provided further indication that larger units (i.e., units with smaller design weights) are more likely to respond.

Overall, the response propensity and prediction model results reinforce the concerns about the currently-used ratio imputation procedures for unit nonresponse. The logistic regression analysis provides little evidence for nonresponse bias reduction, given the current set of industry adjustment cells. The prediction models appear to be valid, but the SAS-H and SAS-T respondent samples consist of the larger sampled units and are not representative subsamples. Thus, the ratio imputation procedure tends to create units that resemble the larger sampled units.

Estimate and Variance Estimate Effects

In this section, we examine the effects of the alternative adjustment methods on estimates and variance estimates. For this, we computed three sets of estimates per set of adjustment cells and item: a **ratio** estimate that uses the currently used ratio-of-identicals model for each applicable item (equation 5); the **count** estimate that uses the weighted inverse response rate within adjustment cell as adjustment factor (equation 3); and the **count_u** estimate that uses the unweighted inverse response rate within adjustment cell as adjustment factor (equation 4). The auxiliary variables for ratio estimation differ by item; the same inverse response rate adjustment is applied to each unit within adjustment cell for the count and count_u procedures. The Taylor linearization estimates of sampling variance were obtained using the SAS SURVEYMEANS procedure.

All of our comparisons are presented with respect to the ratio estimates. These comparisons give indications of estimation and variance estimation effects caused entirely by changes in adjustment cell and/or adjustment methodology. They **do not** provide any indication of the effects of these changes on the production totals, which use ratio imputation for unit nonresponse as a last resort in all cases. Moreover, our variance estimation results cannot be extrapolated to the production estimates. SAS uses the method of random groups for variance estimation; the variance estimator does not account for unit nonresponse or imputation effects on variances. Shao and Steel (1999) derives the appropriate variance estimator for a survey that uses composite imputation such as SAS, but is very difficult and expensive to implement; the Shao and Thompson method (2009) is easy to implement for either version of the count estimator, but is only accurate for reweighted estimators or imputed estimators that use one mean or ratio imputation model and requires that the auxiliary variable used in the ratio estimator be available for all sampled units.

Appendix Six presents average ratios of estimates and sampling variances computed with each adjustment procedure (count, count_u, and ratio) within the same set of adjustment cells for SAS-H and SAS-T. These statistics examine the effects on estimation and variance estimation due to a change in adjustment procedure. From these results, we see that changing the adjustment model from a ratio model to an inverse response rate (count or count_u) tends to increase the estimates and the associated sampling variances. The estimation results are completely in-line with the earlier analyses, since the majority of the responding units are large and using the ratio-of-identicals models force the imputed values to look like "large values" whose ratios tend to be more stable and are often smaller than those obtained from smaller units. With the sampling variance calculations, using the inverse response rates increase the variance by a multiplicative factor $\left(1/\phi_p^2\right)$. The sampling variances are, however, underestimates.

To gauge whether these large increases in variance are “reasonable,” we computed the variance inflation factor (VIF) proposed in Kish (1992, equation 4.2), which provides a rough estimate of the increased variance in the sample mean obtained using the **count** or **count_u** reweighting procedure. The VIF’s obtained for the q = c (**count)** and $q=c\_u$ (**count_u)** adjusted estimates for a given set of adjustment cells are given as

$$VIF^{q} \approx 1 + cv^{2}(w_{pj}^{q}) = n\sum_{j}\left(w_{pj}^{q}\right)^{2} / \left(\sum_{j} w_{pj}^{q}\right)^{2} \qquad (5)$$

Appendix Seven presents the average, minimum, and maximum VIFs for these two procedures by adjustment cell for SAS-H and SAS-T. These factors tend to be quite large, with approximately a 13-percent increase for SAS-H and a 27-percent increase for SAS-T in expected variance regardless of adjustment cells. This provides supporting evidence of the large variance increases with the count or count_u procedures over the ratio models that have strong predictive power (e.g., totals, subtotals).

Lastly, Appendix Eight presents average ratios of estimates and sampling variances by adjustment procedure (count, count_u, and ratio) between sets of adjustment cells. These tables examine the effects on estimation and variance estimation due to a change in adjustment cell, given the same adjustment procedure. Surprisingly, the change in adjustment cell doesn't appear to have much of an effect with the ratio estimators, but does come into play with the two inverse response rate adjustments (count and count_u). Adding a unit size classification within industry greatly increases the count and count_u total estimates generally, while reducing the sampling variance by quite a bit.

Discussion

This research was motivated by the need to validate the currently-used procedures implemented by several ongoing business programs at the U.S. Census Bureau. Survey sampling literature discusses methods of weight adjustment to account for unit nonresponse; the studied programs impute a complete record instead. If properly implemented, weight adjustment can reduce nonresponse bias and well as estimate variance while preserving multivariate relationships.

So, why impute? Ultimately, the implemented procedures are designed to fulfill survey objectives to publish timely and accurate estimates of key survey data items. With the SAS data, it is often possible to build excellent unit level imputation models for the **totals** items, by logically deriving a replacing value from other reported items, by using historic data from the same unit, or by using administrative or even census data. For these programs, ratio imputation is a last resort. Unfortunately, the same wealth of auxiliary data is not available for many detail items, and the individual item response rates tend to be fairly low. Especially for these items, there is little evidence that the respondent values are "representative," and it is difficult to develop strong predictive models.

Our analysis of the SAS-H and SAS-T ratio models provides some validation of the ratio imputation procedure for totals, but does draw some "red flags." The adjustment cells do not appear to be predictive of unit nonresponse, so there are bias concerns. The respondent sample is comprised mostly of larger sample units. Certainty and noncertainty units are used jointly to develop imputation parameters, even though their response propensities appear to differ substantially within industry imputation cell. It appears that nonresponse bias effects could be ameliorated using the same ratio imputation models, but modifying the adjustment cells, perhaps subdividing industry cells by unit size categories or by using some combination of sampling strata as adjustment cell.

In contrast, we have presented evidence of estimation benefits for detail items by using reweighting instead of imputation. It should be noted, however, that the item-level response rates for many of the detail items are quite low and tend to be reported by the larger units. Consequently, the reweighted detail items will still suffer from some model bias, and the extent of this bias cannot be determined without additional data collection/non-response follow-up. However, the reweighting does preserve multivariate relationships between items. An alternative approach – not discussed in the paper – would be to publish unadjusted detail items along with a "not specified by kind" remainder.[4]

In general, this analysis highlights several of the major challenges that business surveys encounter in addressing unit nonresponse. Respondents often do not comprise a random subsample, as larger units are more likely to provide data than smaller units. This phenomenon is an artifact of several factors, including the perceived benefits of the survey by the business community and the existing analyst nonresponse follow-up procedure, which focus on obtaining the most accurate estimated totals. Developing a more representative set of adjustment cells that contain sufficient respondents is equally challenging, as there are considerably fewer "large" units in the population than small units. Finally, there are data challenges, as several of the detail items that the survey would like to collect may not be available from the majority of the sampled units. Again, the respondent sample size issues for the detail items are compounded by collecting different sets of detail items by industry or sector.

Acknowledgements

[4] This procedure aggregates the weighted residual difference between the summed reported data values and the corresponding total.

The authors thank William Davie Jr., Xijian Liu, Philip Steel, and Julianne Zamora for their useful comments on earlier versions of this manuscript.

References

Kalton, G. and Flores-Cervantes, I. (2003). Weighting Methods. *JOS*, **19**, pp. 81-97.

Kish, L. (1992). Weighting for unequal P_i .*JOS*, **8**, 183-200.

Little, R.J.A. and Rubin, D.R. (2002). Statistical Analysis with Missing Data. New York: John Wiley and Sons.

Little., R.J.A. and Vartivarian, S. (2005). Does Weighting for Nonresponse Increase the Variance of Survey Means? *Survey Methodology*, **31**, pp. 161-168.

Nelson, M. (2011). StEPS Imputation Specifications for the Transportation Portion of the 2010 Service Annual Survey (SAS-T). U.S. Census Bureau internal memorandum, EDMS #38537, available upon request.

Oh, H.L. and Scheuren, F.J. (1983). Weighting Adjustment of Unit Nonresponse. Incomplete Data in Sample Surveys (Vol.20). New York: Academic Press, 143-184.

Rao, J.N.K. and Scott, A.J. (1992). A Simple Method of the Analysis of Clustered Binary Data. *Biometrika*, **74**, pp. 577-585.

Roberts, G., Rao, J.N.K., and Kumar, S. (1987). Logistic Regression Analysis of Sample Survey Data. *Biometrika*, **74**, pp. 1 - 12.

Särndal, C.E. and Lundström, S. (2005). Estimation in Surveys with Nonresponse. New York: John Wiley & Sons, Inc.

Shao, J. and Steel, P. (1999). Variance Estimation For Survey Data With Composite Imputation and Non-negligible Sampling Fractions. *Journal of the American Statistical Association*, **93**, pp. 254-265.

Shao, J. and Thompson, K.J. (2009). Variance Estimation In The Presence Of Nonrespondents and Certainty Strata. *Survey Methodology*, **35**, pp. 215-225.

Definition of Item Relationships in SAS-H and SAS-T

Total	Details:	For Sector/subsectors:
190000 (Total Expenses)	182100, 182200, 182300, 182400, 182500, 182600, 182700, 182800, 182900, 183000, 183100, 183200, 189900	49, 62 (2008 onward)
	182100, 182200, 182300, 182400, 182500, 182600, 182700, 182800, 509700, 182900, 509800, 183000, 509900, 183100, 183200, 189900	48
1800000 (Total Revenue)	174100, 174200, 179800	624
	4001A00, 4002A00, 4003A00, 4004A00, 4005A00, 4006A00, 4007A00, 4009A00, 4061A00, 4062A00, 4063A00, 4064A00	6215
	4001A00, 4002A00, 4003A00, 4004A00, 4005A00, 4006A00, 4008A00, 4071A00, 4072A00, 174100, 174200, 180900,	Tax-exempt 623
	4001A00, 4002A00, 4003A00, 4004A00, 4005A00, 4006A00, 4008A00, 4071A00, 4072A00, 180900,	Taxable 623
	4001A00, 4002A00, 4003A00, 4004A00, 4005A00, 4006A00, 4007A00, 4008A00, 4009A00	621, excluding six-digit industries that begin with '6214,' '6215,' '6216,' and '6219'
	4001A00, 4002A00, 4003A00, 4004A00, 4005A00, 4006A00, 4007A00, 4008A00, 174100 (selected industries), 174200 (selected industries), 180900	Tax-exempt 622 and six-digit industries that begin with '6214,' '6216,' and '6219'
	506100, 179900	484
509000 (Truck Inventory)	508800, 508900	484
509300 (Truck-tractor Inventory)	509100, 509200	484
509600 (Trailer Inventory)	509400, 509500	484

Counts of Significant Propensity Analyses (Full Model and Individual Parameters) -- χ^2 Test

SAS-H

	Model 1				Model 3				Model 4			
	Industry by Tax Status Adjustment Cells								**Sampling Strata Adjustment Cells**			
Full Sample	6				6				6			
Marginal Models 180900χ^2 only	Cell	Count	Cell	Count	Cell	Count	Cell	Count	Cell	Count	Cell	Count
	621111T0	1	622218E0	4	621111T0	6	622218E0	5	H11A	6	H23P	4
	621112T0	0	622219E0	1	621112T0	0	622219E0	3	H12A	5	H31A	6
	621210T0	6	622219T0	2	621210T0	1	622219T0	1	H13A	6	H31Z	3
	621310T0	2	622318E0	2	621310T0	2	622318E0	4	H13B	5	H32A	4
	621320T0	4	622319E0	0	621320T0	0	622319E0	0	H13C	6	H32B	5
	621330T0	5	622319T0	6	621330T0	1	622319T0	2	H13D	6	H32Y	4
	621340T0	1	623110E0	6	621340T0	0	623110E0	6	H13E	6	H32Z	1
	621391T0	3	623110T0	1	621391T0	4	623110T0	3	H13F	6	H33A	5
	621399T0	6	623210E0	6	621399T0	5	623210E0	5	H14A	5	H33B	6
	621410E0	1	623210T0	0	621410E0	5	623210T0	0	H14B	6	H33Y	5
	621410T0	0	623220E0	0	621410T0	0	623220E0	0	H14C	6	H33Z	3
	621420E0	1	623220T0	5	621420E0	5	623220T0	0	H14D	6	H39A	6
	621420T0	0	623311E0	6	621420T0	1	623311E0	6	H14E	6	H39Z	5
	621491E0	1	623311T0	0	621491E0	1	623311T0	0	H14F	6	H41A	6
	621491T0	3	623312E0	3	621491T0	2	623312E0	0	H14U	4	H41B	6
	621492E0	0	623312T0	3	621492E0	1	623312T0	1	H14V	4	H41C	6
	621492T0	1	623990E0	1	621492T0	0	623990E0	0	H14W	3	H41X	3
	621493E0	1	623990T0	6	621493E0	0	623990T0	4	H14X	3	H41Y	0
	621493T0	1	624110E0	4	621493T0	2	624110E0	0	H14Y	5	H41Z	1
	621498E0	1	624110T0	3	621498E0	4	624110T0	0	H14Z	1	H42A	5
	621498T0	4	624120E0	6	621498T0	2	624120E0	6	H15A	6	H42B	6
	621511T0	5	624120T0	6	621511T0	0	624120T0	1	H15B	6	H42C	6
	621512T0	5	624190E0	3	621512T0	1	624190E0	0	H16A	6	H42X	2
	621610E0	2	624190T0	0	621610E0	6	624190T0	1	H16Z	3	H42Y	2
	621610T0	2	624210E0	4	621610T0	2	624210E0	6	H19A	6	H42Z	2
	621910E0	3	624210T0	4	621910E0	3	624210T0	0	H19B	6	H43A	6
	621910T0	5	624221E0	6	621910T0	0	624221E0	2	H19Y	3	H43Z	1
	621991E0	6	624221T0	3	621991E0	2	624221T0	6	H19Z	3	H44A	6
	621991T0	1	624229E0	6	621991T0	2	624229E0	1	H21A	6	H44Z	5
	621999E0	1	624229T0	6	621999E0	0	624229T0	1	H21G	5	O33Z	2
	621999T0	4	624230E0	2	621999T0	1	624230E0	3	H21P	0	Y150	6
	622118E0	3	624230T0	6	622118E0	6	624230T0	0	H22A	6		
	622119E0	6	624310E0	5	622119E0	6	624310E0	0	H22G	6		
	622119T0	0	624310T0	1	622119T0	5	624310T0	0	H22P	5		
			624410E0	0			624410E0	5	H23A	6		

SAS-T

		Model 1	Model 3	Model 4			
	Industry Adjustment Cells			Sampling Strata Adjustment Cells			
Full Sample		3	6	6			
Marginal Models χ^2 only	Cell	Count	Count	Cell	Count	Cell	Count
	484110	4	6	E14C	6	T42B	6
	484121	1	6	L11B	6	T42C	5
	484122	0	5	L31A	6	T85A	6
	484210	1	5	P16A	6	U12A	6
	484220	1	6	T11A	6	W21A	5
	484230	2	4	T11B	5	W22A	6
	492110	0	3	T12A	5	W31A	5
	492210	0	0	T411	5	W31B	4
	493110	0	0	T41A	6	W31C	4
	493120	1	2	T41B	6	W31D	6
	493130	1	6	T41C	3	W31C	4
	493190	3	1	T42A	5	W31D	6

Average Unit Response Rates (URR) Using Current Adjustment Cells Subdivided by Certainty Status for SAS-H

Imputation Cell/ Certainty Status		URR	Imputation Cell/ Certainty Status		URR	Imputation Cell/ Certainty Status		URR	Imputation Cell/ Certainty Status		URR
621111T0	C	60.50	621493E0	C	57.95	622218E0	C	74.51	624110E0	C	46.60
621111T0	N	76.62	621493E0	N	52.94	622218E0	N	67.66	624110E0	N	81.62
621112T0	C	38.59	621493T0	C	69.01	622219E0	C	60.92	624110T0	C	59.14
621112T0	N	72.82	621493T0	N	74.57	622219E0	N	67.74	624110T0	N	62.93
621210T0	C	79.44	621498E0	C	35.50	622219T0	C	59.05	624120E0	C	45.69
621210T0	N	80.52	621498E0	N	86.08	622219T0	N	61.30	624120E0	N	87.41
621310T0	C	64.68	621498T0	C	58.18	622319E0	C	53.29	624120T0	C	64.16
621310T0	N	67.04	621498T0	N	61.03	622319E0	N	77.84	624120T0	N	64.03
621320T0	C	67.31	621511T0	C	49.59	622319T0	C	58.66	624190E0	C	41.04
621320T0	N	73.69	621511T0	N	66.82	622319T0	N	55.57	624190E0	N	82.72
621330T0	C	62.35	621512T0	C	48.57	623110E0	C	62.60	624190T0	C	65.83
621330T0	N	73.16	621512T0	N	66.18	623110E0	N	85.41	624190T0	N	66.35
621340T0	C	49.83	621610E0	C	61.45	623110T0	C	62.63	624210E0	C	64.51
621340T0	N	70.48	621610E0	N	85.24	623110T0	N	74.45	624210E0	N	82.25
621391T0	C	79.87	621610T0	C	60.56	623210E0	C	59.53	624210T0	C	63.48
621391T0	N	73.12	621610T0	N	67.42	623210E0	N	88.33	624210T0	N	48.56
621399T0	C	58.46	621910E0	C	79.75	623210T0	C	48.72	624221E0	C	42.86
621399T0	N	56.31	621910E0	N	80.48	623210T0	N	72.89	624221E0	N	87.00
621410E0	C	61.79	621910T0	C	62.96	623220E0	C	52.63	624221T0	C	48.27
621410E0	N	81.80	621910T0	N	64.11	623220E0	N	77.92	624221T0	N	55.86
621410T0	C	70.68	621991E0	C	88.51	623220T0	C	52.80	624229E0	C	46.07
621410T0	N	70.92	621991E0	N	87.30	623220T0	N	62.11	624229E0	N	81.62
621420E0	C	52.71	621991T0	C	54.31	623311E0	C	53.21	624229T0	C	52.51
621420E0	N	84.20	621991T0	N	53.04	623311E0	N	94.11	624229T0	N	41.24
621420T0	C	58.38	621999E0	C	35.54	623311T0	C	39.25	624230E0	C	67.49
621420T0	N	68.91	621999E0	N	76.77	623311T0	N	69.80	624230E0	N	81.14
621491E0	C	43.06	621999T0	C	36.01	623312E0	C	52.23	624230T0	C	66.45
621491E0	N	77.78	621999T0	N	53.75	623312E0	N	68.57	624230T0	N	39.51
621491T0	C	45.99	622118E0	C	90.94	623312T0	C	47.69	624310E0	C	52.83
621491T0	N	43.13	622118E0	N	74.83	623312T0	N	66.15	624310E0	N	86.13
621492E0	C	29.41	622119E0	C	77.78	623990E0	C	46.15	624310T0	C	54.43
621492E0	N	90.24	622119E0	N	87.27	623990E0	N	76.64	624310T0	N	62.82
621492T0	C	35.31	622119T0	C	73.77	623990T0	C	57.90	624410E0	C	44.35
621492T0	N	65.70	622119T0	N	70.50	623990T0	N	47.50	624410E0	N	79.18
									624410T0	C	55.20
									624410T0	N	61.10

Average Unit Response Rates (URR) Using Current Adjustment Cells Subdivided by Size Category for SAS-H

Imputation Cell	Size Class	URR	Imputation Cell	Size Class	URR	Imputation Cell	Size Class	URR	Imputation Cell	Size Class	URR
621111T0	C	60.50	621410E0	C	61.79	621498T0	C	58.18	621999T0	C	36.01
621111T0	N1	82.62	621410E0	N1	90.15	621498T0	N1	64.40	621999T0	N1	57.32
621111T0	N2	80.17	621410E0	N2	75.14	621498T0	N2	60.70	621999T0	N2	55.29
621111T0	N3	68.13	621410E0	N3	75.62	621498T0	N3	58.53	621999T0	N3	49.09
621112T0	C	38.59	621410T0	C	70.68	621511T0	C	49.59	622118E0	C	90.94
621112T0	N1	78.83	621410T0	N1	81.48	621511T0	N1	72.40	622118E0	N1	82.10
621112T0	N2	69.14	621410T0	N2	67.15	621511T0	N2	73.65	622118E0	N2	69.08
621112T0	N3	70.12	621410T0	N3	62.86	621511T0	N3	57.27	622118E0	N3	71.97
621210T0	C	79.44	621420E0	C	52.71	621512T0	C	48.57	622119E0	C	77.78
621210T0	N1	82.05	621420E0	N1	88.45	621512T0	N1	72.50	622119E0	N1	87.43
621210T0	N2	82.79	621420E0	N2	84.27	621512T0	N2	73.05	622119E0	N2	87.58
621210T0	N3	77.82	621420E0	N3	80.62	621512T0	N3	56.22	622119E0	N3	86.93
621310T0	C	64.68	621420T0	C	58.38	621610E0	C	61.45	622119T0	C	73.77
621310T0	N1	65.96	621420T0	N1	69.29	621610E0	N1	90.67	622119T0	N1	72.43
621310T0	N2	68.64	621420T0	N2	64.61	621610E0	N2	84.32	622119T0	N3	67.71
621310T0	N3	67.19	621420T0	N3	71.58	621610E0	N3	81.02	622218E0	C	74.51
621320T0	C	67.31	621491E0	C	43.06	621610T0	C	60.56	622218E0	N1	79.70
621320T0	N1	70.52	621491E0	N3	77.78	621610T0	N1	70.73	622218E0	N2	75.00
621320T0	N2	77.05	621491T0	C	45.99	621610T0	N2	67.53	622218E0	N3	54.94
621320T0	N3	73.68	621491T0	N3	43.13	621610T0	N3	64.56	622219E0	C	60.92
621330T0	C	62.35	621492E0	C	29.41	621910E0	C	79.75	622219E0	N1	81.57
621330T0	N1	74.33	621492E0	N1	94.44	621910E0	N1	84.45	622219E0	N2	50.00
621330T0	N2	68.84	621492E0	N3	88.19	621910E0	N2	80.27	622219E0	N3	59.84
621330T0	N3	75.24	621492T0	C	35.31	621910E0	N3	76.52	622219T0	C	59.05
621340T0	C	49.83	621492T0	N1	66.95	621910T0	C	62.96	622219T0	N1	77.84
621340T0	N1	75.94	621492T0	N2	86.48	621910T0	N1	68.59	622219T0	N2	70.24
621340T0	N2	67.80	621492T0	N3	49.25	621910T0	N2	65.99	622219T0	N3	38.85
621340T0	N3	67.82	621493E0	C	57.95	621910T0	N3	58.90	622319E0	C	53.29
621391T0	C	79.87	621493E0	N1	62.50	621991E0	C	88.51	622319E0	N1	97.62
621391T0	N1	76.37	621493E0	N2	75.00	621991E0	N1	88.20	622319E0	N2	100.00
621391T0	N2	72.28	621493E0	N3	45.55	621991E0	N3	86.18	622319E0	N3	61.27
621391T0	N3	70.94	621493T0	C	69.01	621991T0	C	54.31	622319T0	C	58.66
621399T0	C	58.46	621493T0	N1	82.15	621991T0	N1	43.33	622319T0	N1	76.17
621399T0	N1	58.29	621493T0	N2	78.27	621991T0	N2	100.00	622319T0	N2	60.04
621399T0	N2	55.38	621493T0	N3	67.36	621991T0	N3	38.61	622319T0	N3	38.45
621399T0	N3	55.33	621498E0	C	35.50	621999E0	C	35.54	623110E0	C	62.60
			621498E0	N1	89.12	621999E0	N1	95.83	623110E0	N1	84.36
			621498E0	N2	85.88	621999E0	N3	57.22	623110E0	N2	85.84
			621498E0	N3	83.42				623110E0	N3	86.19

Average Unit Response Rates (URR) Using Current Adjustment Cells Subdivided by Size Category for SAS-H (Continued)

Imputation Cell	Size Class	URR	Imputation Cell	Size Class	URR	Imputation Cell	Size Class	URR	Imputation Cell	Size Class	URR
623110T0	C	62.63	623990E0	C	46.15	624210T0	C	63.48	624410E0	C	44.35
623110T0	N1	71.09	623990E0	N1	78.57	624210T0	N1	69.52	624410E0	N1	83.48
623110T0	N2	86.04	623990E0	N2	84.66	624210T0	N2	37.14	624410E0	N2	83.83
623110T0	N3	68.95	623990E0	N3	67.90	624210T0	N3	41.15	624410E0	N3	71.65
623210E0	C	59.53	623990T0	C	57.90	624221E0	C	42.86	624410T0	C	55.20
623210E0	N1	87.05	623990T0	N1	56.48	624221E0	N1	87.32	624410T0	N1	67.86
623210E0	N2	88.94	623990T0	N2	53.45	624221E0	N2	89.78	624410T0	N2	60.23
623210E0	N3	89.10	623990T0	N3	37.43	624221E0	N3	85.43	624410T0	N3	55.41
623210T0	C	48.72	624110E0	C	46.60	624221T0	C	48.27			
623210T0	N1	82.73	624110E0	N1	79.26	624221T0	N1	66.67			
623210T0	N2	75.10	624110E0	N2	83.79	624221T0	N3	51.82			
623210T0	N3	64.91	624110E0	N3	82.27	624229E0	C	46.07			
623220E0	C	52.63	624110T0	C	59.14	624229E0	N1	83.02			
623220E0	N1	77.66	624110T0	N1	77.97	624229E0	N2	80.75			
623220E0	N2	75.00	624110T0	N2	53.64	624229E0	N3	81.14			
623220E0	N3	79.22	624110T0	N3	57.53	624229T0	C	52.51			
623220T0	C	52.80	624120E0	C	45.69	624229T0	N1	55.09			
623220T0	N1	72.87	624120E0	N1	85.50	624229T0	N2	19.64			
623220T0	N2	75.78	624120E0	N2	89.45	624229T0	N3	35.83			
623220T0	N3	48.66	624120E0	N3	87.58	624230E0	C	67.49			
623311E0	C	53.21	624120T0	C	64.16	624230E0	N1	85.27			
623311E0	N1	96.20	624120T0	N1	76.41	624230E0	N2	77.69			
623311E0	N2	93.50	624120T0	N2	66.09	624230E0	N3	80.07			
623311E0	N3	92.52	624120T0	N3	51.31	624230T0	C	66.45			
623311T0	C	39.25	624190E0	C	41.04	624230T0	N1	34.08			
623311T0	N1	78.23	624190E0	N1	82.38	624230T0	N2	54.96			
623311T0	N2	71.28	624190E0	N2	88.79	624230T0	N3	32.49			
623311T0	N3	60.65	624190E0	N3	79.41	624310E0	C	52.83			
623312E0	C	52.23	624190T0	C	65.83	624310E0	N1	89.00			
623312E0	N1	82.33	624190T0	N1	69.22	624310E0	N2	87.71			
623312E0	N2	75.04	624190T0	N2	66.13	624310E0	N3	82.90			
623312E0	N3	50.10	624190T0	N3	64.23	624310T0	C	54.43			
623312T0	C	47.69	624210E0	C	64.51	624310T0	N1	61.02			
623312T0	N1	72.52	624210E0	N1	84.04	624310T0	N2	68.99			
623312T0	N2	65.92	624210E0	N2	78.73	624310T0	N3	60.23			
623312T0	N3	60.26	624210E0	N3	82.12						

Average Unit Response Rates (URR) Using Pseudo-Stratum for SAS-H

Pseudo-strata	URR	Pseudo-strata	URR	Pseudo-strata	URR	Pseudo-strata	URR
62	0.00	H14V N	55.33	H3	63.12	H42A N	49.45
A1	34.79	H14X C	16.67	H31A C	55.42	H42B C	16.67
A12A C	1.67	H14Y C	55.16	H31A N	75.69	H42B N	48.66
C39W C	19.87	H14Y N	85.64	H31Z C	66.67	H42C C	100.00
E16Z C	30.56	H14Z C	66.67	H31Z N	85.09	H42C N	40.64
F11A C	22.22	H14Z N	84.02	H32A C	46.30	H42X C	63.76
F41B C	50.13	H15A C	69.95	H32A N	71.75	H42X N	82.30
H1	67.15	H15A N	66.84	H32B C	50.00	H42Y C	55.38
H11A C	100.00	H15B C	83.33	H32B N	59.99	H42Y N	84.31
H11A N	76.51	H15B N	66.21	H32Y C	34.88	H42Z C	66.67
H12A C	83.33	H16A C	66.67	H32Y N	79.78	H42Z N	81.46
H12A N	80.62	H16A N	66.91	H32Z C	33.33	H43A C	67.82
H13A C	66.67	H16Z C	74.02	H32Z N	88.36	H43A N	64.34
H13A N	66.59	H16Z N	87.09	H33A C	50.21	H43Z C	33.33
H13B C	16.67	H19A C	16.67	H33A N	73.30	H43Z N	85.45
H13B N	73.69	H19A N	64.14	H33B C	51.12	H44A C	57.03
H13C C	83.33	H19B C	51.54	H33B N	67.28	H44A N	60.80
H13C N	73.17	H19B N	54.47	H33Y C	74.12	H44Z C	44.06
H13D C	50.00	H19Y C	78.33	H33Y N	71.10	H44Z N	79.52
H13D N	70.45	H19Y N	90.10	H33Z C	59.12	L11A C	52.69
H13E C	100.00	H19Z C	100.00	H33Z N	93.61	L22D C	48.57
H13E N	73.28	H19Z N	80.95	H39A C	0.00	O1	41.68
H13F C	100.00	H2	50.15	H39A N	48.54	O32Z C	31.46
H13F N	57.40	H21A C	50.15	H39Z C	83.33	O33Z C	25.45
H14A C	78.47	H21A N	71.57	H39Z N	77.41	O34Z C	39.29
H14A N	73.24	H21G C	83.68	H4	61.93	P17A C	30.95
H14B C	61.65	H21G N	75.61	H41A C	60.96	X160 C	91.11
H14B N	69.43	H21P C	0.00	H41A N	63.01	XP10 C	43.19
H14C C	83.33	H21P N	87.55	H41B C	83.33	XY	50.21
H14C N	53.31	H22A C	56.63	H41B N	63.96		
H14D C	100.00	H22A N	61.43	H41C C	83.33		
H14D N	63.29	H22G C	86.43	H41C N	66.24		
H14E C	83.33	H22G N	79.19	H41X C	0.00		
H14E N	74.73	H22P C	83.33	H41X N	83.51		
H14F C	0.00	H22P N	78.19	H41Y C	56.25		
H14F N	61.12	H23A C	55.58	H41Y N	88.16		
H14U C	57.83	H23A N	55.58	H41Z C	51.80		
H14U N	86.62	H23P C	41.94	H41Z N	81.97		
H14V C	66.02	H23P N	97.62	H42A C	100.00		

Average Unit Response Rates (URR) By Candidate Sets of Adjustment Cells for SAS-T

Industry by Certainty		Industry by Unit Size						Stratum or Collapsed Stratum			
Industry/ Certainty	URR	Industry	Size Category	URR	Industry	Size Category	URR	Strata/ Certainty	URR	Strata/ Certainty	URR
484110 C	65.59	484110	C	65.59	492110	C	13.81	484B	19.10	W31B C	78.92
484110 N	56.38		N1	67.43		N1	80.20	492B	12.31	W31B N	73.16
484121 C	49.05		N2	56.78		N2	67.88	493B	19.68	W31C C	69.96
484121 N	58.14		N3	45.22		N3	64.42	T31A C	2.78	W31C N	60.68
484122 C	40.72	484121	C	49.05	492210	C	67.51	T32A C	12.50	W31D C	70.83
484122 N	64.13		N1	71.22		N1	66.68	T41A C	68.09	W31D N	60.89
484210 C	72.72		N2	61.26		N2	56.36	T41A N	56.33		
484210 N	58.91		N3	44.77		N3	55.27	T41B C	57.80		
484220 C	67.17	484122	C	40.72	493110	C	34.88	T41B N	58.14		
484220 N	56.28		N1	74.31		N1	70.54	T41C C	68.82		
484230 C	55.13		N2	69.44		N2	78.23	T41C N	64.13		
484230 N	58.34		N3	50.75		N3	48.76	T42A C	64.80		
492110 C	13.81	484210	C	72.72	493120	C	53.13	T42A N	58.91		
492110 N	70.75		N1	70.39		N1	76.71	T42B C	67.97		
492210 C	67.51		N2	59.16		N2	82.14	T42B N	56.31		
492210 N	59.03		N3	47.91		N3	68.03	T42C C	58.65		
493110 C	34.88	484220	C	67.17	493130	C	75.61	T42C N	58.40		
493110 N	64.26		N1	65.53		N1	65.63	T85A C	37.14		
493120 C	53.13		N2	53.38		N2&3	56.51	W21A C	48.41		
493120 N	73.94		N3	48.97	493190	C	39.04	W21A N	71.47		
493130 C	75.61	484230	C	55.13		N1	68.65	W22A C	66.77		
493130 N	60.68		N1	60.97		N2	66.14	W22A N	59.03		
493190 C	39.04		N2	66.43		N3	51.00	W31A C	54.44		
493190 N	60.38		N3	50.04				W31A N	65.31		

Appendix Four

Regression Analysis Results for SAS-H with Industry by Tax Status Adjustment Cells

	Numerator	Denominator	Imputation Model	Number of Imputation Cells	Average Percentage of Significant Cells	Mean R^2					
						2005	2006	2007	2008	2009	2010
TOTAL	182100	182101	Trend	70	99.43		0.93	0.97	0.97	0.98	0.96
	180000	180001	Trend	70	99.71		0.94	0.96	0.96	0.98	0.97
	182100	190000	Auxiliary	70	89.52	0.83	0.88	0.88	0.88	0.88	0.88
	190000	180000	Auxiliary	70	90.95	0.95	0.96	0.95	0.96	0.96	0.96
SUBTOTAL	185600	190000	Auxiliary	32	81.25	0.44					
DETAIL	174100	180000	Auxiliary	32	82.29	0.43	0.36	0.45	0.46	0.46	0.43
	174200	180000	Auxiliary	33	77.78	0.35	0.34	0.34	0.32	0.28	0.29
	179800	180000	Auxiliary	32	36.98	0.81	0.59	0.48	0.48	0.46	0.45
	4001A00	180000	Auxiliary	52	86.22	0.69	0.69	0.71	0.73	0.72	0.71
	4002A00	180000	Auxiliary	52	84.29	0.65	0.63	0.62	0.64	0.62	0.62
	4003A00	180000	Auxiliary	52	79.17	0.48	0.45	0.46	0.43	0.46	0.44
	4004A00	180000	Auxiliary	52	62.82	0.70	0.42	0.44	0.48	0.46	0.42
	4005A00	180000	Auxiliary	52	83.65	0.58	0.70	0.68	0.68	0.69	0.67
	4006A00	180000	Auxiliary	49	55.44	0.64	0.39	0.45	0.42	0.39	0.41
	4007A00	180000	Auxiliary	41	83.33	0.58	0.56	0.60	0.56	0.56	0.56
	4008A00	180000	Auxiliary	50	78.67	0.55	0.49	0.46	0.43	0.44	0.44
	4009A00	180000	Auxiliary	54	32.72	0.47	0.41	0.38	0.30	0.19	0.16
	4061A00	180000	Auxiliary	26	7.69		0.87	0.83	0.83	0.87	0.84
	4062A00	180000	Auxiliary	26	7.69		0.66	0.66	0.66	0.65	0.64
	4063A00	180000	Auxiliary	26	6.92		0.11	0.16	0.19	0.33	0.32
	4064A00	180000	Auxiliary	26	7.69		0.52	0.63	0.64	0.68	0.71
	4071A00	180000	Auxiliary	12	88.89	0.47	0.54	0.60	0.61	0.60	0.60
	4072A00	180000	Auxiliary	12	77.78	0.41	0.53	0.49	0.51	0.44	0.49
	182200	190000	Auxiliary	70	60.00	0.81	0.83	0.37	0.84		
	182300	190000	Auxiliary	70	76.43	0.34	0.39	0.37	0.36	0.34	0.33
	182400	190000	Auxiliary	70	82.62	0.42	0.43	0.45	0.47	0.42	0.47
	182500	190000	Auxiliary	70	82.38	0.47	0.52	0.52	0.52	0.51	0.52
	182600	190000	Auxiliary	70	75.95	0.36	0.41	0.39	0.42	0.44	0.44
	182700	190000	Auxiliary	70	79.05	0.66	0.66		0.68		
	182800	190000	Auxiliary	70	57.86	0.52	0.59	0.85	0.62		
	182900	190000	Auxiliary	70	56.07	0.58	0.58	0.18	0.60		

	Numerator	Denominator	Imputation Model	Number of Imputation Cells	Average Percentage of Significant Cells	Mean R^2					
						2005	2006	2007	2008	2009	2010
	183000	190000	Auxiliary	70	80.48	0.36	0.42	0.44	0.45	0.45	0.43
	183100	190000	Auxiliary	70	87.62	0.68	0.67	0.65	0.69	0.68	0.67
	183200	190000	Auxiliary	70	79.76	0.34	0.45	0.44	0.43	0.43	0.44
	180900	190000	Auxiliary	70	61.79	0.77	0.76	0.82	0.77		
	401000	190000	Auxiliary	53	84.28	0.64	0.64	0.64	0.66	0.65	0.64
	401100	190000	Auxiliary	52	81.09	0.61	0.62	0.59	0.65	0.65	0.64

Appendix Four

Regression Analysis Results for SAS-T with Industry Adjustment Cells

Item Type	Dependent Variable	Independent Variable	Imputation Model	Number of Imputation Cells	Average Percentage of Significant Cells	Average R^2					
						2005	2006	2007	2008	2009	2010
TOTAL	180000	180001	Trend	12	100.00		0.95	0.96	0.98	0.97	0.95
	190000	180000	Auxiliary	12	100.00	0.96	0.98	0.98	0.98	0.98	0.89
	506500	506100	Auxiliary	6	100.00	0.87	0.88	0.88	0.86	0.86	0.84
SUBTOTAL	506100	180000	Auxiliary	6	100.00	0.99	0.99	0.98	0.99	0.98	0.99
	509000	506100	Auxiliary	6	63.89	0.27	0.21	0.17	0.17	0.20	0.14
	509300	506100	Auxiliary	6	83.33	0.71	0.73	0.70	0.71	0.72	0.57
	509600	506100	Auxiliary	6	100.00	0.71	0.71	0.71	0.67	0.72	0.68
DETAIL	179900	180000	Auxiliary	6	100.00	0.28	0.44	0.42	0.41	0.39	0.38
	182100	182101	Trend	12	100.00		0.89	0.97	0.98	0.98	0.89
	182100	190000	Auxiliary	12	98.61	0.81	0.83	0.84	0.82	0.81	0.78
	182200	190000	Auxiliary	12	98.33	0.68	0.74		0.75	0.72	0.69
	182300	190000	Auxiliary	12	75.00	0.31	0.26	0.26	0.27	0.31	0.34
	182400	190000	Auxiliary	12	84.72	0.28	0.29	0.25	0.25	0.21	0.19
	182500	190000	Auxiliary	12	87.50	0.50	0.44	0.46	0.49	0.47	0.43
	182600	190000	Auxiliary	12	75.00	0.28	0.26	0.45	0.34	0.37	0.30
	182700	190000	Auxiliary	12	91.67	0.54	0.56		0.52	0.48	0.61
	182800	190000	Auxiliary	12	90.00	0.43	0.45		0.45	0.46	0.51
	182900	190000	Auxiliary	12	91.67	0.50	0.62		0.60	0.56	0.57
	183000	190000	Auxiliary	12	93.06	0.41	0.51	0.51	0.41	0.38	0.40
	183100	190000	Auxiliary	12	97.22	0.64	0.64	0.64	0.63	0.66	0.68
	183200	190000	Auxiliary	12	88.89	0.58	0.57	0.63	0.58	0.55	0.47
	189900	190000	Auxiliary	12	100.00	0.66	0.70		0.67	0.71	0.77
	509700	190000	Auxiliary	6	88.89	0.56	0.58	0.50	0.63	0.64	0.62
	509800	190000	Auxiliary	6	91.67	0.71	0.65	0.68	0.78	0.70	0.60
	509900	190000	Auxiliary	6	94.44	0.71	0.71	0.71	0.67	0.70	0.62
	508800	509000	Auxiliary	6	100.00	0.87	0.82	0.82	0.87	0.83	0.79
	508900	509000	Auxiliary	6	58.33	0.25	0.40	0.26	0.27	0.22	0.27
	509100	509300	Auxiliary	6	97.22	0.86	0.86	0.83	0.84	0.81	0.83
	509200	509300	Auxiliary	6	83.33	0.30	0.31	0.28	0.33	0.37	0.30
	509400	509600	Auxiliary	6	100.00	0.82	0.81	0.78	0.79	0.77	0.80
	509500	509600	Auxiliary	6	75.00	0.26	0.31	0.32	0.33	0.38	0.32

Random Subsample Assessment for Considered SAS-H Adjustment Cells: Current Cells (3 *x* 2 Contingency Table)

Industry / Tax Status	Number of Statistical Periods	Number of Significant Tests	Percentage of Statistical Periods with Significant Tests	Industry/ Tax Status	Number of Statistical Periods	Number of Significant Tests	Percentage of Statistical Periods with Significant Tests
621111T0	6	6	100.00	622219T0	6	4	66.67
621112T0	6	0	0.00	622319T0	6	4	66.67
621210T0	6	6	100.00	623110E0	6	0	0.00
621310T0	6	2	33.33	623110T0	6	6	100.00
621320T0	6	0	0.00	623210E0	6	0	0.00
621330T0	6	1	16.67	623210T0	6	0	0.00
621340T0	6	2	33.33	623220E0	6	0	0.00
621391T0	6	0	0.00	623220T0	6	4	66.67
621399T0	6	3	50.00	623311E0	5	2	40.00
621410E0	5	0	0.00	623311T0	6	1	16.67
621410T0	6	1	16.67	623312E0	6	5	83.33
621420E0	6	4	66.67	623312T0	6	2	33.33
621420T0	6	0	0.00	623990E0	6	5	83.33
621492T0	6	5	83.33	623990T0	6	5	83.33
621493E0	4	0	0.00	624110E0	6	0	0.00
621493T0	6	5	83.33	624110T0	6	0	0.00
621498E0	6	2	33.33	624120E0	6	1	16.67
621498T0	6	1	16.67	624120T0	6	5	83.33
621511T0	6	6	100.00	624190E0	6	3	50.00
621512T0	6	6	100.00	624190T0	6	0	0.00
621610E0	6	3	50.00	624210E0	6	1	16.67
621610T0	6	3	50.00	624210T0	6	4	66.67
621910E0	6	0	0.00	624221E0	6	0	0.00
621910T0	6	4	66.67	624229E0	6	0	0.00
621991E0	2	0	0.00	624229T0	6	2	33.33
621991T0	2	0	0.00	624230E0	6	0	0.00
621999T0	6	2	33.33	624230T0	6	1	16.67
622118E0	6	0	0.00	624310E0	6	3	50.00
622119E0	6	0	0.00	624310T0	6	1	16.67
622119T0	6	0	0.00	624410E0	6	4	66.67
622218E0	6	6	100.00	624410T0	6	3	50.00
622219E0	5	2	40.00				

Random Subsample Assessment for Considered SAS-H Adjustment Cells: Current Cells (2 *x*1 Contingency Table)

Industry	Number of Statistical Periods	Number of Significant Tests	Percentage of Statistical Periods with Significant Tests	Industry	Number of Statistical Periods	Number of Significant Tests	Percentage of Statistical Periods with Significant Tests
621410E0	1	1	100.00	621999E0	5	0	0.00
621491T0	5	1	20.00	622219E0	1	0	0.00
621492E0	4	3	75.00	622319E0	6	3	50.00
621493E0	2	0	0.00	623311E0	1	1	100.00
621991E0	4	4	100.00	624221T0	5	4	80.00
621991T0	4	0	0.00				

Random Subsample Assessment for Considered SAS-H Adjustment: Sampling Stratum (3 *x* 2 Contingency Table)

Strata	Number of Statistical Periods	Number of Significant Tests	Percentage of Statistical Periods with Significant Tests	Strata	Number of Statistical Periods	Number of Significant Tests	Percentage of Statistical Periods with Significant Tests
H12A	6	6	100.00	H14U	1	0	0.00
H13A	6	6	100.00	H14V	6	1	16.67
H13B	3	1	33.33	H15A	6	3	50.00
H13C	5	3	60.00	H16A	2	1	50.00
H13D	3	2	66.67	H19A	6	2	33.33
H13E	3	0	0.00	H19B	3	0	0.00
H13F	6	1	16.67	H19Y	2	0	0.00
H14A	5	2	40.00	H19Z	2	0	0.00
H14B	6	4	66.67	H21A	1	0	0.00
H14E	1	0	0.00	H21G	1	0	0.00
H14F	3	1	33.33	H22A	4	1	25.00
H14U	1	0	0.00	H22P	1	0	0.00
H14V	6	1	16.67	H23A	6	2	33.33
H15A	6	3	50.00	H31A	4	3	75.00
H16A	2	1	50.00	H32A	2	1	50.00
H19A	6	2	33.33	H33A	2	0	0.00
H19B	3	0	0.00	H33B	6	4	66.67
H19Y	2	0	0.00	H33Y	1	1	100.00
H19Z	2	0	0.00	H39A	1	0	0.00
H21A	1	0	0.00	H41A	5	0	0.00
H21G	1	0	0.00	H41B	1	0	0.00
H12A	6	6	100.00	H41C	3	1	33.33
H13A	6	6	100.00	H41Y	2	1	50.00
H13B	3	1	33.33	H41Z	3	0	0.00
H13C	5	3	60.00	H42A	2	1	50.00
H13D	3	2	66.67	H42B	4	0	0.00
H13E	3	0	0.00	H42C	4	1	25.00
H13F	6	1	16.67	H42Y	5	0	0.00
H14A	5	2	40.00	H43A	4	0	0.00
H14B	6	4	66.67	H43Z	1	0	0.00
H14E	1	0	0.00	H44A	6	1	16.67
H14F	3	1	33.33	H44Z	6	3	50.00

Random Subsample Assessment for Considered SAS-T Adjustment Cells

Current Adjustment Cells				Strata Adjustment Cells			
Industry	Number of Statistical Periods	Number of Significant Tests	Percentage of Statistical Periods with Significant Tests	Strata**	Number of Statistical Periods	Number of Significant Tests	Percentage of Statistical Periods with Significant Tests
484110	6	6	100.00	T41A	6	3	50.00
484121	6	5	83.33	T41B	6	1	16.67
484122	6	1	16.67	T41C	6	0	0.00
484210	6	5	83.33	T42A	6	0	0.00
484220	6	3	50.00	T42B	6	1	16.67
484230	6	6	100.00	T42C	6	1	16.67
492110	6	1	16.67	W21A	6	1	16.67
492210	6	1	16.67	W22A	6	3	50.00
493110	6	5	83.33	W31A	6	0	0.00
493120	6	1	16.67	W31B	6	1	16.67
493130	6	3	50.00	W31C	6	1	16.67
493190*	6	3	60.00	W31D	6	0	0.00

* There was insufficient sample to use the 3 *x* 2 contingency table test in this adjustment cell with the 2006 data. The p-value for the1*x* 2 contingency table analysis was 0.1026.

** There were several strata that were excluded from this analysis due to prohibitively small samples of nonrespondents within adjustment cell.

Effect of Unit Nonresponse Adjustment by Type of Adjustment cell (Average Ratios): SAS-H

Item Type	Item	Industry				Industry by Size				Pseudo-Strata			
		Estimates		Sampling Variances		Estimates		Sampling Variances		Estimates		Sampling Variances	
		C/R	C_U/R	C/R	C_U/R	C/R	C_U/R	C/R	C_U/R	C/R	C_U/R	C/R	C_U/R
TOTAL	182100	1.00	0.96	0.90	0.99	0.99	1.12	0.89	0.93	1.00	1.10	0.94	0.95
	190000	1.00	0.97	0.93	1.02	1.00	1.13	0.91	0.92	1.00	1.12	0.98	0.98
DETAIL	174100	0.99	0.98	0.98	0.95	1.08	1.10	0.91	0.88	0.96	0.97	0.80	0.76
	174200	1.00	1.00	1.04	0.99	1.04	1.23	0.96	0.88	1.00	1.16	0.93	0.86
	179800	1.00	1.00	0.91	0.87	1.10	1.13	0.87	0.86	0.96	0.98	0.66	0.60
	4001A00	1.00	0.98	0.86	0.94	1.00	1.17	0.83	0.95	1.02	1.17	0.82	0.89
	4002A00	1.00	0.98	0.98	0.98	1.00	1.12	0.92	0.89	1.00	1.10	0.92	0.87
	4003A00	1.02	0.97	1.11	1.12	0.96	0.99	1.03	0.90	0.99	1.02	1.21	1.09
	4004A00	0.99	0.94	0.80	0.80	0.96	1.10	0.82	0.96	0.99	1.08	0.83	0.84
	4005A00	1.00	0.98	0.87	0.93	0.99	1.18	0.83	0.96	1.01	1.17	0.82	0.87
	4006A00	1.00	0.94	0.95	0.89	0.98	1.09	0.93	0.95	1.00	1.07	0.96	0.88
	4007A00	0.98	0.95	0.95	0.98	0.98	1.08	0.93	0.96	1.00	1.07	1.01	0.97
	4008A00	1.01	0.95	0.94	0.95	1.01	1.10	0.89	0.95	1.02	1.10	0.91	0.90
	4009A00	1.01	0.93	0.88	0.90	0.98	1.08	0.80	0.96	1.01	1.06	0.90	0.92
	4061A00	1.00	0.90	0.90	0.79	1.05	0.89	0.67	0.78	1.14	0.91	0.78	0.71
	4062A00	1.00	0.93	0.93	0.72	1.06	0.91	0.71	0.70	1.15	0.95	0.81	0.63
	4063A00	1.02	0.97	1.05	0.79	0.93	0.98	0.83	0.77	0.99	0.94	0.95	0.70
	4064A00	1.03	0.95	1.16	0.50	1.03	0.92	0.83	0.47	1.11	0.91	0.94	0.42
	4071A00	1.00	0.97	0.78	0.78	1.03	1.04	0.78	0.88	1.03	1.00	0.71	0.71
	4072A00	1.00	0.98	0.95	0.74	1.00	1.05	0.92	0.76	0.98	1.00	0.86	0.67
	182300	1.01	0.96	0.90	0.88	1.00	1.11	0.90	0.95	1.01	1.09	0.91	0.87
	182400	0.98	0.95	0.92	0.89	0.99	1.07	0.94	0.96	1.00	1.05	0.98	0.94
	182500	0.99	0.96	0.90	0.90	1.03	1.11	0.91	0.94	1.02	1.08	0.88	0.86
	182600	1.00	0.97	0.98	0.99	0.99	1.11	0.94	0.96	1.00	1.09	0.98	0.94
	183000	1.00	0.94	0.85	0.87	0.98	1.07	0.80	0.96	0.99	1.05	0.85	0.87
	183100	1.00	0.98	0.85	0.97	1.01	1.20	0.83	0.93	1.01	1.18	0.83	0.93
	183200	0.99	0.94	0.89	0.88	0.99	1.08	0.91	0.97	1.01	1.07	0.91	0.88
	401000	1.00	0.97	0.96	0.99	0.99	1.13	0.91	0.97	1.01	1.12	0.99	0.96
	401100	1.00	0.99	0.87	0.99	1.00	1.20	0.83	0.95	1.01	1.19	0.87	0.95

Effect of Unit Nonresponse Adjustment by Type of Adjustment cell (Average Ratios): SAS-T

Item Type	Item	Industry				Industry by Size				Pseudo-Strata			
		Estimates		Sampling Variances		Estimates		Sampling Variances		Estimates		Sampling Variances	
		C/R	C_U/R	C/R	C_U/R	C/R	C_U/R	C/R	C_U/R	C/R	C_U/R	C/R	C_U/R
TOTAL	190000	1.36	1.25	1.76	1.27	2.04	2.03	1.36	1.34	1.37	1.27	2.25	1.43
SUBTOTAL	506100	1.36	1.14	1.78	1.26	1.19	1.17	1.41	1.36	1.25	1.12	2.37	1.45
	506500	1.12	1.16	1.88	1.44	1.14	1.19	1.46	1.40	1.16	1.16	2.40	1.52
	509000	1.30	1.05	1.80	1.23	1.18	1.13	1.47	1.42	1.33	1.11	2.36	1.48
	509300	1.32	1.12	1.63	1.14	1.21	1.18	1.51	1.42	1.30	1.15	2.09	1.38
	509600	1.36	1.34	1.66	1.26	1.21	1.16	1.48	1.42	1.31	1.33	2.12	1.46
DETAIL	179900	1.39	1.11	1.92	1.38	1.10	1.08	1.32	1.28	1.07	0.96	2.25	1.40
	182100	1.41	1.32	1.93	1.55	2.20	2.19	1.58	1.51	1.39	1.30	2.52	1.60
	182200	1.39	1.38	1.88	1.51	2.44	2.43	1.49	1.46	1.48	1.41	2.50	1.60
	182300	1.35	1.27	1.65	1.53	1.91	1.89	1.77	1.61	1.39	1.27	2.35	1.46
	182400	1.40	1.31	1.77	1.27	2.40	2.38	1.58	1.54	1.51	1.35	2.41	1.53
	182500	1.39	1.42	1.71	1.28	3.40	3.38	1.55	1.49	1.58	1.49	2.23	1.48
	182600	1.39	1.37	1.80	1.32	2.52	2.50	1.65	1.45	1.54	1.44	2.11	1.61
	182700	1.39	1.37	1.77	1.31	1.92	1.91	1.62	1.45	1.47	1.37	2.41	1.97
	182800	1.39	1.31	1.87	1.43	1.81	1.80	1.34	1.37	1.34	1.20	2.48	1.65
	182900	1.34	1.22	1.90	1.35	2.01	1.99	1.59	1.53	1.51	1.40	2.55	1.57
	183000	1.43	1.37	1.91	1.29	2.74	2.72	1.54	1.40	1.51	1.41	2.40	1.72
	183100	1.40	1.28	1.86	1.29	2.06	2.04	1.47	1.41	1.49	1.39	2.56	1.61
	183200	1.41	1.27	1.89	1.46	1.68	1.66	1.49	1.44	1.48	1.36	2.47	1.64
	180900	1.48	1.47	2.03	1.67	2.59	2.56	1.46	1.40	1.62	1.45	2.26	1.60
	509700	1.41	1.17	1.93	1.33	1.15	1.14	1.49	1.44	1.20	1.10	2.55	1.59
	509800	1.39	1.16	1.87	1.31	1.22	1.20	1.56	1.53	1.36	1.19	2.64	1.61
	509900	1.39	1.16	1.86	1.33	1.22	1.20	1.59	1.54	1.37	1.19	2.47	1.59
	508800	1.27	1.02	1.66	1.19	1.14	1.09	2.12	1.97	1.31	1.09	2.32	1.60
	508900	1.28	1.02	1.58	1.11	1.02	0.99	2.05	1.85	0.85	0.73	2.19	1.49
	509100	1.44	1.22	1.70	1.23	1.29	1.27	1.52	1.47	1.41	1.24	2.38	1.51
	509200	1.29	1.06	1.60	1.10	1.08	1.06	1.67	1.57	0.94	0.83	2.31	1.47
	509400	1.33	1.14	1.85	1.41	1.19	1.18	1.23	1.19	1.26	1.12	2.47	1.53
	509500	1.29	1.06	1.64	1.11	1.08	1.06	1.59	1.50	0.93	0.83	2.18	1.41

C/R = Count adjusted estimate/Ratio estimate
C_U/R = Count_u adjusted estimate/Ratio estimate

Variance Inflation Factors for Reweighted Estimates Using Inverse Response Rates:

SAS-T

Adjustment Cell	Weighted Inverse Response Rate (count)			Unweighted Inverse Response Rate (count_u)		
	Average	Minimum	Maximum	Average	Minimum	Maximum
Industry/Tax Status	12.92	12.32	14.36	12.74	12.01	13.99
Industry/Tax Status by Unit Size	13.70	12.87	15.48	13.59	12.81	15.37
Pseudo-Strata	12.98	12.36	14.38	12.60	11.84	13.82

SAS-H

Adjustment Cell	Weighted Inverse Response Rate (count)			Unweighted Inverse Response Rate (count_u)		
	Average	Minimum	Maximum	Average	Minimum	Maximum
Industry	25.87	22.25	35.80	26.24	22.84	36.06
Industry by Unit Size	28.12	24.04	39.11	28.34	25.06	38.92
Pseudo-Strata	25.91	22.28	35.88	25.22	21.88	34.61

Effect of Adjustment Cells by Type of Estimator (Average Ratios): SAS-H

Item Type	Item	Estimates						Sampling Variances					
		Ratio		Count		Count_u		Ratio		Count		Count_u	
		IS/I	PS/I	IS/I	PS/I	IS/I	PS/I	IS/I	PS/I	IS/I	PS/I	IS/I	PS/I
TOTAL	182100	1.00	0.96	0.99	1.12	1.00	1.10	0.90	0.99	0.89	0.93	0.94	0.95
	190000	1.00	0.97	1.00	1.13	1.00	1.12	0.93	1.02	0.91	0.92	0.98	0.98
DETAIL	174100	0.99	0.98	1.08	1.10	0.96	0.97	0.98	0.95	0.91	0.88	0.80	0.76
	174200	1.00	1.00	1.04	1.23	1.00	1.16	1.04	0.99	0.96	0.88	0.93	0.86
	179800	1.00	1.00	1.10	1.13	0.96	0.98	0.91	0.87	0.87	0.86	0.66	0.60
	4001A00	1.00	0.98	1.00	1.17	1.02	1.17	0.86	0.94	0.83	0.95	0.82	0.89
	4002A00	1.00	0.98	1.00	1.12	1.00	1.10	0.98	0.98	0.92	0.89	0.92	0.87
	4003A00	1.02	0.97	0.96	0.99	0.99	1.02	1.11	1.12	1.03	0.90	1.21	1.09
	4004A00	0.99	0.94	0.96	1.10	0.99	1.08	0.80	0.80	0.82	0.96	0.83	0.84
	4005A00	1.00	0.98	0.99	1.18	1.01	1.17	0.87	0.93	0.83	0.96	0.82	0.87
	4006A00	1.00	0.94	0.98	1.09	1.00	1.07	0.95	0.89	0.93	0.95	0.96	0.88
	4007A00	0.98	0.95	0.98	1.08	1.00	1.07	0.95	0.98	0.93	0.96	1.01	0.97
	4008A00	1.01	0.95	1.01	1.10	1.02	1.10	0.94	0.95	0.89	0.95	0.91	0.90
	4009A00	1.01	0.93	0.98	1.08	1.01	1.06	0.88	0.90	0.80	0.96	0.90	0.92
	4061A00	1.00	0.90	1.05	0.89	1.14	0.91	0.90	0.79	0.67	0.78	0.78	0.71
	4062A00	1.00	0.93	1.06	0.91	1.15	0.95	0.93	0.72	0.71	0.70	0.81	0.63
	4063A00	1.02	0.97	0.93	0.98	0.99	0.94	1.05	0.79	0.83	0.77	0.95	0.70
	4064A00	1.03	0.95	1.03	0.92	1.11	0.91	1.16	0.50	0.83	0.47	0.94	0.42
	4071A00	1.00	0.97	1.03	1.04	1.03	1.00	0.78	0.78	0.78	0.88	0.71	0.71
	4072A00	1.00	0.98	1.00	1.05	0.98	1.00	0.95	0.74	0.92	0.76	0.86	0.67
	182300	1.01	0.96	1.00	1.11	1.01	1.09	0.90	0.88	0.90	0.95	0.91	0.87
	182400	0.98	0.95	0.99	1.07	1.00	1.05	0.92	0.89	0.94	0.96	0.98	0.94
	182500	0.99	0.96	1.03	1.11	1.02	1.08	0.90	0.90	0.91	0.94	0.88	0.86
	182600	1.00	0.97	0.99	1.11	1.00	1.09	0.98	0.99	0.94	0.96	0.98	0.94
	183000	1.00	0.94	0.98	1.07	0.99	1.05	0.85	0.87	0.80	0.96	0.85	0.87
	183100	1.00	0.98	1.01	1.20	1.01	1.18	0.85	0.97	0.83	0.93	0.83	0.93
	183200	0.99	0.94	0.99	1.08	1.01	1.07	0.89	0.88	0.91	0.97	0.91	0.88
	401000	1.00	0.97	0.99	1.13	1.01	1.12	0.96	0.99	0.91	0.97	0.99	0.96
	401100	1.00	0.99	1.00	1.20	1.01	1.19	0.87	0.99	0.83	0.95	0.87	0.95

IS/I = Unit Size within Industry/Tax-Status Adjustment Cells/Industry/Tax Status Adjustment Cells
PS/I = Pseudo-Strata Within Industry/Tax Status Adjustment Cells

Effect of Adjustment Cells by Type of Estimator (Average Ratios): SAS-T

Item Type	Item	Estimates						Sampling Variances					
		Ratio		Count		Count_u		Ratio		Count		Count_u	
		IS/I	PS/I	IS/I	PS/I	IS/I	PS/I	IS/I	PS/I	IS/I	PS/I	IS/I	PS/I
TOTAL	190000	1.00	1.00	1.50	1.01	1.62	1.02	0.71	0.75	0.66	1.01	0.93	0.91
SUBTOTAL	506100	1.00	0.99	0.87	0.91	1.03	0.96	0.70	0.76	0.57	1.00	0.76	0.89
	506500	0.99	0.92	0.85	0.91	1.01	0.96	0.72	0.74	0.61	1.00	0.82	0.89
	509000	0.98	0.95	0.88	0.97	1.06	1.00	0.90	0.79	0.82	1.00	1.12	0.96
	509300	0.98	0.95	0.89	0.93	1.03	0.97	0.79	0.80	0.72	1.00	0.95	0.93
	509600	1.00	0.93	0.89	0.89	1.03	0.94	0.72	0.78	0.57	1.00	0.71	0.83
DETAIL	179900	0.99	1.19	0.79	0.92	0.97	1.04	0.81	0.85	0.59	1.00	0.80	0.90
	182100	1.01	1.03	1.58	1.02	1.67	1.01	0.73	0.77	0.59	1.00	0.73	0.82
	182200	1.07	0.96	1.82	1.02	1.81	0.97	0.66	0.76	0.53	1.00	0.65	0.83
	182300	0.98	1.02	1.39	1.05	1.45	1.02	0.75	0.72	0.84	1.00	0.96	0.79
	182400	0.99	0.95	1.71	1.02	1.78	0.98	1.00	0.75	0.85	1.00	1.14	0.90
	182500	1.00	0.95	2.43	1.08	2.39	1.00	0.76	0.76	0.74	1.00	0.90	0.86
	182600	0.96	0.99	1.76	1.10	1.71	1.04	0.67	0.77	0.66	1.02	0.75	0.93
	182700	1.02	0.99	1.41	1.05	1.45	1.01	1.05	0.91	0.92	1.31	1.10	1.51
	182800	1.00	1.07	1.30	1.03	1.40	0.99	0.76	0.71	0.56	1.00	0.76	0.85
	182900	1.05	0.88	1.49	1.00	1.60	1.01	0.82	0.75	0.69	1.00	0.93	0.91
	183000	1.01	1.02	1.94	1.08	2.01	1.05	1.02	1.42	0.78	1.43	1.01	1.95
	183100	1.01	0.94	1.48	1.00	1.60	1.02	0.83	0.72	0.69	1.00	0.92	0.89
	183200	1.01	0.93	1.20	0.97	1.32	1.00	0.80	0.77	0.61	1.01	0.78	0.87
	180900	0.99	0.98	1.76	1.07	1.78	0.98	0.72	0.81	0.73	1.02	0.98	0.91
	509700	1.02	1.06	0.83	0.90	1.00	1.00	0.65	0.75	0.53	1.00	0.72	0.89
	509800	0.99	0.95	0.87	0.93	1.02	0.97	0.75	0.72	0.64	1.00	0.86	0.89
	509900	0.99	0.94	0.87	0.93	1.03	0.97	0.83	0.75	0.71	1.00	0.95	0.90
	508800	0.99	0.94	0.89	0.97	1.06	1.00	0.65	0.72	0.82	1.00	1.11	0.96
	508900	1.05	1.47	0.84	0.97	1.03	1.04	0.69	0.71	0.84	1.00	1.18	0.97
	509100	1.01	0.94	0.90	0.92	1.04	0.96	0.74	0.76	0.69	1.00	0.90	0.91
	509200	0.97	1.35	0.81	0.98	0.97	1.05	0.65	0.71	0.68	1.00	0.91	0.93
	509400	1.00	0.94	0.90	0.89	1.04	0.92	0.80	0.77	0.56	1.00	0.70	0.82
	509500	0.98	1.30	0.82	0.93	0.98	1.01	0.67	0.75	0.64	1.00	0.86	0.92

IS/I = Unit Size within Industry Adjustment Cells/Industry Adjustment Cells
PS/I = Pseudo-Strata Within Industry Adjustment Cells/Industry Adjustment Cells