Task N

Weighting Report

Lou Rizzo
Susan Swain
Mark Freedman
Katie Hubbell
Jing Kang
David Morganstein
Joel Wakesberg

November 17, 2010

Table of Contents

Appendices

Tables

This weighting report documents the revised weights that were delivered to the Federal Highway Administration in August and September 2010. This report includes an overview of aspects of the weights which were done in 2009 and were not changed in the reweighting work of 2010. It also includes detailed documentation of the reweighting steps.

The reweighting was designed to address several issues found with the 2009 weights. The major issues were as follows:

- The distribution of weights was found to be too large (the numbers of weights with extreme values);

- There were not enough geographic variables in the control totals for many of the major domains (e.g., Florida);

- There were inconsistencies in demographic variables used as controls across the major domains.

The overall steps in the weighting process were as follows:

- Construction of base weights—the base weights are the reciprocals of the telephone frame sampling rates within each 'sample group' (the sample group was a particular telephone number sample for a particular study[1] at a particular point in time, with five sample groups being taken over the course of the year for most of the study areas).

- Construction of jackknife replicate base weights—the replicate weights are designed to allow the user to easily produce valid jackknife variance estimators based on the sample design[2];

- Household-level nonresponse adjustments (done within each sample group and study area separately);

- Composition of the household nonresponse-adjusted base weights by 'domain' (the domains are defined by the quarter in which the travel date occurred and by final geography as determined by geocoding each household's address);

---

[1] One of the 19 add-on studies, or the national sample.

[2] Each step below also processes the jackknife replicate weights

- **Household-level raking and trimming (using the composite weights);**

- **Person-level nonresponse adjustments (for nonresponding persons within useable households);**

- **Person-level weights (with person-level raking and trimming).**

The three bolded items are the components of the weighting process that were redone in 2010. This report discusses these parts of the weighting process in more depth and detail than the earlier portions of the weighting process not redone in 2010, though all aspects of the weighting process are discussed comprehensively in this report.

The listing of chapters roughly follows the order of the weighting process. Chapter 2 describes the process for computing household-level base weights within each stratum. Chapter 3 describes adjustments for nonresponse at the household level, done within the separate studies and sample groups. Chapter 4 discusses compositing the nonresponse-adjusted base weights from the sampling strata. Chapter 5 describes the raking procedure at the household level. Chapter 6 describes person-level nonresponse adjustments preceding person-level raking. Chapter 7 describes person-level raking adjustments. Chapter 8 describes special weights for children and vehicles.

Household Level Weights—Base Weights 2

The primary component of the base weight is the inverse of the probability of selection of the telephone number in the given sample group frame.

The sample design for each stratum is an equal probability sample of telephone numbers from the working banks designated for the stratum in that quarter (the strata for each study are defined in the May 20, 2008 sample documentation). The sampling rate for each stratum is simply the sample size divided by the total number of telephone numbers in the frame[3].

2.1   Base Weights at the Frame Level

The base weights are the inverse of the probability of selection for each telephone number from the sample group frame.

2.2   Replicate Base Weights at the Household Level

Replicate base weights were also computed for each frame and lowest level design stratum separately. The replicate variance strata are nested within the studies and sampling strata, which are based on the following hierarchy:

- Study areas (e.g., New York add-on study);

- Sample Groups (e.g., Sample Group 1 corresponding to a particular frame);

- Geographic strata within studies (e.g., Bronx NY within the New York add-on study, but not within the national study).

The variance strata were generated based on the original frame order of the systematic sample, and were generated as 'blocks' of sampled telephone numbers based on this original frame order. The

---

[3] Each sample is deduplicated against all of the previous samples which have been drawn in the same working banks. No telephone number can therefore be sampled twice. The frame count then excludes these previously sampled telephone numbers.

blocks had varying sizes across the sampling strata, based on the expected number of completed interviews per block. Each block was made large enough so that there would be at least one completed interview within each block.

Variance units were defined which split each variance stratum in half. This split into variance units was done randomly while keeping the selection order. A replicate was created for the variance stratum by deleting one variance unit randomly and retaining the other and doubling the weights for the retained one. For that variance stratum, all other replicate weights are equal to the base weight (only the designated replicate is 'perturbed' by setting half of the sample units' weights to zero (for that replicate) and doubling the other half of the sample units' weights for that replicate.

The last step was to combine the variance strata. The total number of replicate weights was 100. The variance strata were combined so that there were exactly 100 combined strata. In their original ordering within and across the sampling strata, variance strata 1 through 100 for example would be mapped to combined strata 1 through 100; variance strata 101 through 200 would be mapped to combined strata 1 through 100, etc. Replicate weight 1 would then correspond to combined stratum 1, corresponding to variance stratum 1, 101, 201, 301, etc. This guarantees that each replicate weight corresponds to 1/100 of the variance strata and 1/100 of the sample units, with that 1/100 spread as evenly as possible over the entire sample (all study groups and all sampling strata).

Household Level Nonresponse Adjustments $\boxed{3}$

Nonresponse unfortunately is a major and continuously growing problem with every RDD survey. We did extensive work analyzing nonresponse and the potential for bias, and built in adjustments based on this analysis. This analysis was commenced before the end of the study using three quarters of data. This allowed us to have analysis in time to inform our nonresponse adjustment process. The ¾ nonresponse analysis report is available as a separate document.

The nonresponse adjustments are based on a paradigm generally used in survey research (see for example Oh and Scheuren 1983). Under this paradigm, nonresponse is treated as a subsampling process within carefully selected nonresponse-adjustment cells. The nonresponse-adjustment cells are selected to be heterogeneous in response propensity (the probability of responding) across cells, and homogeneous in response propensity within cells. The nonresponse bias analysis informed this cell selection process by finding characteristics that were related to response propensity (propensity to be successfully contacted, propensity to cooperate at the screener level, propensity to cooperate at the interview level). The final nonresponse adjustments were equal to the inverse of the weighted response rates within the selected nonresponse adjustment cells. These nonresponse adjustment cells nested within the strata utilized in sample selection. The cells were not smaller than 15 sample units, as cells with limited numbers of sample units generate unreliable (highly variable) nonresponse adjustments. In addition, cells with very low weighted response rates were collapsed with other cells to avoid extreme weighting adjustments which can increase variability. The cutoff was 1/3: cells with weighted response rates lower than one-third of the overall response rate for the study area, stratum and sample group were collapsed with other neighboring cells[4].

Nonresponse adjustment was done separately at the screener level (screener nonresponse and noncontact) and the travel data retrieval (extended) interview level (interview nonresponse within completed screened households). The cell structure differed, as response patterns are different for the different stages of data collection. In particular, the final nonresponse adjustments for screener nonresponse includes an adjustment reflecting estimated percentages of residential noncontacts[5].

---

[4] Cells are defined as 'neighboring' based on their location in a tree generated by a binary search algorithm: see the discussion below.

[5] This needs to be estimated as we do not know whether individual ring no answers and answering machine only outcomes correspond to residential households or not, and can only estimate a percentage of residential households for these outcomes within each nonresponse adjustment cell.

There is a rich set of data available at the exchange level which was used for the formation of nonresponse adjustment cells. The nonresponse bias report shows that screener responses rates do in fact differ across different exchange groups. Based on these analyses, we have defined characteristics to use to define cells (those characteristics that are related to response propensity for a particular study).

For the interview nonresponse-adjustment cells, we can also include information from the screener questionnaire, such as household size, race/ethnicity of reference person, home ownership, location, home type, and number of vehicles in household. The nonresponse bias report shows that many screener questionnaire outcomes do in fact predict well propensity interview response, and therefore were used in nonresponse cell generation.

All of the nonresponse adjustments are fully replicated: the replicate nonresponse adjustment for a particular cell is the sum of the particular replicate weights for the numerator set divided by the sum of the replicate weights for the denominator set. The nonresponse-adjusted replicate weight is equal to the product of the replicate weight preceding the adjustment multiplied by the replicate nonresponse adjustment for the particular cell containing the sample unit.

3.1   Screener Nonresponse Adjustments within Sampling Strata

The screener nonresponse adjustment cells are intended to be homogeneous in screener response and contact propensity within the cells, and heterogeneous in screener response and contact propensity across cells. The cells were generated after a selection of cells using WESSEARCH (a binary search algorithm software routine), as discussed in Section 3.3 below. The nonresponse adjustments within each of these cells was equal to the summation of base weights over all 'residential' numbers in the cell, divided by the summation of base weights for all screener respondent households in the cell. The numerator included all sample units which were definitely identified as being residences (respondent or not), and excluded all sample units which were definitely identified as not being residences. For the set of numbers for which residency status was undefined, an estimated portion of residences was computed, and added to the numerator.

The two categories of unknown residency status numbers are the NMs (phone numbers where during the screening process an answering machine was reached at least once but residential status was still uncertain) and NAs (phone numbers where during the screening process no human contact

was made or answering machines were reached). For the NMs we compute an overall weighted eligibility rate among working numbers with known eligibility status within each stratum. We use working numbers in this benchmark as we know that NMs are working numbers. This rate is applied to the NMs, giving an estimated total of residential numbers among this set which is added then to the numerator of the nonresponse adjustment.

For the NAs we compute an overall weighted eligibility rate among all numbers with known eligibility status within each stratum. We use all numbers in this benchmark, as NAs may be residential or nonresidential numbers. This rate is applied to the NAs, giving an estimated total of residential numbers among this set which is added then to the numerator of the nonresponse adjustment.

The final screener nonresponse adjusted weight for each screener respondent is equal to the base weight multiplied by the screener nonresponse adjustment for the nonresponse cell containing the sample unit, multiplied by a further factor for households with multiple landline telephone numbers. Any households with multiple landline telephone numbers have multiple chances of being selected. To avoid bias (these households being overrepresented in the estimators), a factor of ½ is applied to the weights of these households indicating at least two household landline telephone numbers[6].

3.2　Extended Interview Completion Nonresponse Adjustments within Sampling Strata

A household was viewed as *useable* in terms of its completed retrieval interviews *if at least 50% of the adults enumerated within the household completed the extended (travel) interview*. Otherwise, the household was viewed as a nonrespondent (not useable) at this level.

To adjust for these excluded nonuseable households, each useable household received an adjustment for nonuseable households equal to the reciprocal of the weighted "usability rate" in its adjustment cell. These adjustment cells were selected to be as heterogeneous in usability rates across cells as possible, and as homogeneous in usability rates within cells as possible. The cells were selected following an analysis of household characteristics found to be correlated with usability rates (see Section 3.3 below).

---

[6] The best factor would be the reciprocal of the number of telephone numbers in the household, whether it is 2, 3, 4, etc. But more than 2 telephone numbers is rare, and it requires a battery of screener questions to fully discern true residential numbers from numbers used for business, modems, etc. Thus we used the simple ½ factor.

As compared to screener nonrespondents, there was more information available about households that do not have enough completed travel interviews. This extra information comes from the completed screener.

The nonusability adjustments for each cell were the weighted sum of screener respondents within the cell divided by the weighted sum of useable households within the cell. The weights were the screener nonresponse adjusted weights described in the previous section. The final weight for each useable household was the screener-nonresponse-adjusted weight multiplied by the usability adjustment for the usability cell containing the household.

3.3   Specification of Nonresponse Adjustment Cells

Westat's software routine WESSEARCH was utilized to define nonresponse cells within each sampling stratum for both screener nonresponse and household usability separately. In NHTS 2001 we used SPSS CHAID for the same purpose. WESSEARCH is operationally much simpler than the CHAID macro and can be run in batch mode (allowing for automated calls in the context of larger programs), which facilitates this process considerably. WESSEARCH is based on a search algorithm produced by and used with the permission of the University of Michigan (http://www.isr.umich.edu/src/smp/search/).

For screener nonresponse adjustment cells, the WESSEARCH algorithm will search within each study, sample group, and stratum separately. In some cases, the final cell will be the study—sample group—stratum combination alone (no breakdown of this basic sampling cell). The algorithm will avoid cells with a sample size smaller than 15 or a screener respondent sample size smaller than 3, and did not allow for adjustments larger than three times the mean adjustment for the sample group—stratum combination. In case of violation of these norms, the cells were collapsed, by collapsing 'up the tree'. Every terminal cell found to be deficient was linked to a parent cell. All child cells of this parent cell were collapsed making the parent cell the new terminal cell.

Potential cells were generated based on Genesys exchange level information. The nonresponse cells were dichotomous cells (above-median and below-median) using weighted medians of Genesys exchange characteristics within the study, sample group, and stratum. For example, one set of cells was above-median percentage of Hispanics and below-median percentage of Hispanics for a

particular stratum and sample group. The above-median percentage cell includes telephone numbers in exchanges that are above the cutoff (the weighted median) in percentage Hispanics, as estimated by Genesys. Not every set of cells was chosen: only those that registered as significantly correlated to response propensity within the stratum and sample group were chosen. And, not every Genesys characteristic was tested: only those that registered as important at the study level in the ¾ nonresponse bias analysis were tested. The listing of which Genesys characteristics passed the test during the ¾ nonresponse bias analysis and were tested for cell formation by WESSEARCH is given in Appendix A.

For useability, WESSEARCH searched using the characteristics given in the tables in Appendix B. Most of these are screener questionnaire responses, but some are Genesys exchange characteristics. The analysis leading to these tables was described previously in the ¾ nonresponse bias report.

Composite Nonresponse Adjusted Base Weights 4

The nonresponse adjustments were all done within the original sample groups and studies. This reflects the philosophy that nonresponse adjustment is treated as if it is another level of sampling following the original sampling process (the 'pseudorandomization paradigm for nonresponse adjustment', one can cite for example Oh and Scheuren 1983). At the completion of the nonresponse adjustments we put the various samples corresponding to each domain and each time period together, to result in a single set of weights which allows one to use all of the sample units from all of the sample groups and studies without the need for explicit compositing (of separate estimators) on the parts of users.

The purpose of this adjustment is to rescale the weights to reflect the fact that each realized sample could have come from multiple frames. We first defined an exhaustive set of mutually exclusive geographically contiguous domains, which are based on the sampling strata[7]. For states that do not overlap within any of the add-on areas, the states are the composition domains (e.g., ME01 for Maine, MA01 for Massachusetts, NH01 for New Hampshire, etc.). For states that do overlap add-on areas, the domains are roughly equivalent to strata within the add-on areas. Because of overlap between add-on areas, the mapping is not one-to-one. Domains for states that were add-on study states, or that overlapped add-on study areas, were as follows[8]:

- Alabama: AL01—Lee county, AL02—Russell county, AL03—remainder of state[9];

- Arizona: AZ01—Phoenix add-on area, AZ02—Tucson add-on area, AZ03—remainder of state;

- California: CA01—San Diego stratum, CA02—remainder of state;

- Florida: FL01 through FL32, the 32 strata of the Florida add-on study;

- Georgia: GA03 through GA09, the seven within-state strata of the Georgia add-on (the first and second strata of the Georgia add-on study are counties in South Carolina and Alabama).

---

[7] The sampling strata are also based on contiguous geographical areas, but are finally defined in terms of sets of telephone exchanges, which are not perfectly mapped to contiguous geographical areas.

[8] This is given on the delivery files as the field DOMAIN.

[9] Lee and Russell counties in Alabama were part of the Georgia add-on study.

- Indiana: IN01 through IN07, the seven strata of the Indiana add-on study;

- Iowa: IA01—Pottawatomie County[10], IA02—the Cedar Rapids add-on study area, IA03—the remainder of the state;

- Nebraska: NE01—Douglas and Sarpy counties (within the Omaha add-on study area), NE02—remainder of the state;

- New York: NY01 through NY25, the 25 strata of the New York add-on study;

- North Carolina: NC01—Alamance County, NC02—Forsyth county, NC03—Guilford County, NC04—portion of Davidson county in Piedmont add-on study, NC05—portion of Davie county in Piedmont add-on study (Bermuda Run), NC06—Portion of Orange county in Piedmont add-on study (Mebane), NC07—Portion of Randolph county in Piedmont add-on study (Trinity and Archdale), NC08—Portion of King county in Piedmont add-on study (Stokes), NC09—Triangle stratum area of North Carolina Add-on Study, NC10—Metrolina stratum area of North Carolina add-on study, NC11—Urban counties with HH survey stratum area of North Carolina add-on study, NC12—Other North Carolina stratum area of North Carolina add-on study[11];

- South Carolina: SC01—Greenville county, SC02—Spartansburg county, SC03—Beaufort county, SC04—Edgefield county, SC05—Aiken county, SC06—Jasper county, SC07—Remainder of state.[12]

- South Dakota: SD01—Rapid City add-on area, SD02—Sioux Falls add-on area, SD03—remainder of state;

- Tennessee: TN01 (one domain);

- Texas: TX01 (one domain);

- Vermont: VT01-VT05: five Vermont add-on study strata;

- Virginia: VA01 through VA13: thirteen Virginia add-on study strata, VA14—the remainder of the state;

- Wisconsin: WI01 (one domain).

The time period ranges from April 2008[13] through April 2009 (a total of thirteen months). The quarters $q=1,…,4$ consisted of the following[14]:

---

[10] Pottawatomie County was part of the Omaha, Nebraska add-on study.

[11] Domains NC09 through NC12 exclude any areas which were defined as being part of the North Carolina Piedmont study area.

[12] Greenville and Spartansburg Counties were strata for the South Carolina add-on study. Beaufort County was a strata for the South Carolina add-on study, and also a part of the Georgia add-on study. Edgefield, Jaspar, and Aiken Counties were part of the Georgia add-on study. Aiken county only was explicitly excluded from the South Carolina add-on study: the remaining two counties were included in the SC add-on study.

[13] There are also four days at the end of March 2008 which are included in the month April 2008.

- *q*=1: all travel dates up to June 30, 2008;

- *q*=2: all travel dates from July 1, 2008 through September 30, 2008;

- *q*=3: all travel dates from October 1, 2008 through December 31, 2008;

- *q*=4: all travel dates in 2009.

The geographic areas are strata from the add-on studies, or in the case of the national study in states without add-on studies, the states. For example, Maryland is a domain geographic area as Maryland is not an add-on study. Guilford County in North Carolina is also a domain geographic area as it is a stratum in the North Carolina Triad Add-on Study.

Each domain/quarter pair had a set of sampling strata that contributed to it. For example, the domain/quarter pair Guilford County NC/Quarter 3 will have Sample Groups 1, 2, and 3 of the North Carolina Add-on Study, North Carolina Triad Add-on Study, and the US national samples all contributing it. Each of these sampling strata has a set of sampling units with nonresponse-adjusted weights that fall into the domain. **Each of these sampling strata sets with their associated nonresponse-adjusted weights is assumed to produce an unbiased estimator of the domain mean value.** Composition essentially assigns a composite factor to each sampling strata, with the composite factors adding to 1. Whatever the composite factors, the overall estimator will be unbiased if the bolded assumption is valid. We select composite factors based on the relative variability of each sampling stratum set. This variability in turn is measured by the effective sample size (the sample size adjusted for the design effect from weighting). For example, suppose there are two sampling stratum sets for a domain with sample sizes 10 and 50. Suppose the design effects[15] from weight differences are 1.1 and 1.3 respectively, using the '1+CV-squared' measure of design effects from weights (see Kish 1992)[16]. Then the effective sample sizes would be 9 and 38 respectively. The first sampling stratum would receive a compositing factor of 9/47 and the second sampling stratum would receive a compositing factor of 38/47. The sampling stratum are given a share of the final weight set proportional to this effective sample size. The weights are designed then

---

[14] This is given on the delivery files as the field DOMQTR. For the Phoenix and Tucson AZ add-on studies there is a special definition for domain quarter: 1—to June 30, 2008, 2—July 1 to August 31, 2008, 3—September 1 to October 15, 2008, 4—October 16 to November 30, 2008, 5—December 1, 2008 through January 31, 2009, 6—February 1, 2009 to the end of the study. This was due to the specialized time intervals in which these studies took place.

[15] The design effect is defined to be the ratio of the variance of the weighted estimate to the corresponding variance of an unweighted estimate.

[16] The criterion measures the design effect as the coefficient of variation of the weights squared, plus 1. This would be the variance inflation from using differential weights using a model in which the observations are all independent and identically distributed.

to add to the sum of weights from the 'best' sampling stratum (the sampling stratum with the highest effective sample size).

The final composite weight is equal to the product of the nonresponse-adjusted household weight and the computed composite factor. The final composite replicate weight for each replicate is equal to the product of the nonresponse-adjusted household replicate weight and the computed composite factor (the composition process is not replicated).

The basic methodology for computing the compositing factor was done for example in the 1993 National Adult Literacy Survey (Burke et al. 1994). Most approaches to computing these compositing factors take account of the variance of each of the stratum estimator contributions to the composite domain estimator (see for example Lohr and Rao (2007)). Using the effective sample size is a proxy for these variances assuming implicitly that the population variances are in fact equal in the various sampling strata (and avoids differing weights for differing characteristics: this approach will generate one and only one set of weights and replicate weights).

It is well-known that RDD surveys do not cover the full population of households of interest, as not every household has a telephone. An estimated 2-4 percent of households in the United States do not have a telephone. In addition to nontelephone households, households with telephone numbers in zero banks (sets of 100 telephone numbers with the same prefix that have no listed residential numbers: these are not included on the sampling frame for efficiency reasons) and cellphone-only households are not covered in the main sample. There may be significant numbers of households in zero banks, and we know that there is a growing number of cellphone-only households (possibly as much as 20% now). These various sources of undercoverage can result in people of a certain demographic characteristic being underrepresented in the final sample (e.g., Hispanics or males). It is possible to adjust for undercoverage through a poststratification weighting process called "raking," where the weights are iteratively adjusted to independent controls totals for various demographic categories. The process has the effect of differentially adjusting the weights of the sampled households within groups of demographically similar households, so that the total sum of weights for the sampled households equals the corresponding independent control totals for all households (including those not covered by the RDD sample)[17].

Raking and trimming steps were performed iteratively at the household level. The starting point was the composite household weights described in Section 4 of this report. The trimming steps included a 'pre-trim' step preceding the first household raking step, and a 'post-trim' step following each household raking step. Each trimming step was applied within the composite domains (i.e., the domains used to construct the composite weights in Step 4), which after the compositing step are essentially equivalent to the lowest-level sampling strata. For example, New York add-on study stratum 19 corresponds to New York City. The composite domain consists of all sample units from New York City notwithstanding whether they were originally sampled for the national study or the New York add-on study. Thus, the sample units from the national sample were combined with the sample units from the add-on study sample within the composite domain in this trimming step. This is possible as the compositing brings the national sample unit weights and the composite sample unit

---

[17] The 'best' household summations for VA14 were 231,964 for Quarter 1, 6,826 for Quarter 2, 201,350 for Quarter 3, and 2,609 for Quarter 4. These summations are in effect estimators of the number of households for VA14. For Quarter 2 and 4 for DOMAIN=VA14, these contributors were from the Virginia Add-On Sample, and thereby represent 'spillover': households on the border. For Quarter 1 and 3, the summations come from the national sample, and are unbiased. We corrected this by replacing the Quarter 2 and 4 summations with the mean value of 231,964 and 201,350, in both cases.

weights into the same relative order of magnitude (whereas before compositing, the two sets of weights may have very different median weights). Compositing in effect eliminates or reduces the variation of the national study weights for the combined samples within each add-on area composite domain.

The pre-trim step for each major domain consisted of checking for weights that were more than 3.0 times the median weight for the particular composite domain. If less than 1% of the weights fell into this category, then all such weights were trimmed back to equal the cutoff (3.0 times the median weight for the composite domain). If more than 1% of the weights fell into this category[18], then the largest 1% set of the weights were trimmed back to equal the 99th percentile of the weights.

The trimming steps following the raking steps (the 'post-trim steps') were also done within each composite domain, and targeted for trimming any weights that were 4.5 times smaller or 4.5 times larger than the median weight for the composite domain. A maximum of 2.5% of the weights could be trimmed on the high side and a maximum of 2.5% of the weights could be trimmed on the low side for each post-trim step[19]. If more than 2.5% of the weights was greater than 4.5 times the median weight (less than 4.5 times the median weight), then the largest (smallest) 2.5% of the weights was trimmed back to the 97.5[th] percentile (the 2.5[th] percentile).

The iteration of raking and trimming steps was complete when all of the trimming factors for that final putative trimming step were between 0.99 and 1.01[20].

A flag indicating that a weight had been pre-trimmed was provided on the delivery files, and a flag indicating that a weight had been post-trimmed at least once has been provided on the delivery files. These flags will not count households as being trimmed at least once if their trimming factors are always within the trimming tolerance for the major domain run (either 0.99 through 1.01 or 0.999 through 1.001: see previous paragraph and footnote). For example, if the trimming tolerance is 0.01 and a household's trimming factors are always between 0.99 and 1.01, then the trimming flag indicates no trimming.

---

[18] The number of weights affected by this rule was 1% of the number of sample units, rounded up to the smallest larger integer. For example, if the sample size was 120, then the number of trimmed weights was 2 (1.2 rounded up). In particular, there was always at least one weight trimmed if any weights exceeded 3.0 times the median weight. The actual percentage of trimmed weights then could be slightly larger than 1%.

[19] This count of trimmed weights was 2.5% of the number of sample units, rounded up to the smallest larger integer. For example, if the sample size was 100, then the maximum high side or low side number of trimmed weights was 3 (2.5 rounded up). The actual percentage of trimmed weights on the high and low side could be slightly larger than 2.5%.

[20] 0.999 and 1.001 for Major Domains FL, AZ, CA, D2, D3, D5, D7, TN, WI. These were all done earlier in the process before we converted to the 0.99 to 1.01 tolerance range.

Each household raking step in the cycle was done to a tolerance of ±5,000 (i.e., the weighted household totals will be raked until they are within 5,000 of the household control totals).

Section 5.1 below describes the initial raking dimensions for household raking in the supplemental weighting plan. We attempted to achieve convergence using these dimensions first. If convergence failed, our first remedial step was to attempt to locate problematic cells and collapse categories of the marginal control totals selectively, a carefully conducted manual process. If convergence was still not reached we in some cases collapsed full dimensions. Most of these cases involved Black and Hispanic cells in which the sample sizes or populations sizes or both were very small. All of our decisions for household raking are documented in Appendix F of this weighting report.

5.1 Raking Dimensions for Households

We used the 2008 American Community Study (ACS) data to develop the control totals. The control totals for all dimensions were derived from one-year 2008 ACS estimates where possible, using 2006-08 three year estimates where necessary to impute distributions for areas for which 2008 one-year ACS estimates were not available. Appendix D describes the process of computing control totals.

5.1.1 Add-On States (Add On Area Encompasses The Entire State)

The dimensions for 'add-on states[21]' (CA, FL, GA[22], IA, IN, NC, NY, SC[23], TN, TX, VT, WI) were as follows[24].

- (1) CBSACAT * Subregion (as defined below for each State)

- (2) Subregion * Black

- (3) Subregion * Hispanic

---

[21] States for which the entire state was an add-on study area.

[22] GA augmented with 1 South Carolina county (Aiken) and 2 Alabama counties

[23] Aiken County was excluded from the South Carolina sample

[24] These dimensions are ordered by importance. A higher number in the ordering indicates the dimension would be collapsed sooner if there were convergence problems.

- ■ (4) Subregion * Owner/Renter

- ■ (5) Subregion * Number of Vehicles

- ■ (6) Month

- ■ (7) Day of Week

- ■ (8) HH size * worker

In the event that convergence failed, the collapsing of cells was conducted in a stepwise, automated manner using the rules provided by FHWA. Collapsing begins in the highest numbered domain containing problematic cells. For example, if convergence failed, the program identified the cells that were problematic – let's say (4) Subregion * Owner/Renter and (2) Subregion * Black. Because Subregion * Owner/Renter is dimension 4 and Subregion * Black is dimension 2 we followed the FHWA rules for collapsing dimension 4 to determine if convergence could be achieved. If not, we checked to see if another dimension has become problematic, if not we collapsed dimension 2.

These rakings were done separately for each add-on state.

For California the Subregions defined for the first raking dimension were San Diego County vs remainder of state. For Florida we used the seven transportation regions[25]. For Georgia and Indiana the Subregions were the sampling strata (seven for Georgia which includes part of South Carolina and Alabama, seven for Indiana). For South Carolina the Subregions were the sampling strata. For Iowa the Subregions were Pottawattamie county (from the Omaha add-on), Linn County (Cedar Rapids add-on), and the Remainder of the State. For New York the Subregions were 24 of the total 25 sampling strata. Saratoga County was divided into two strata for sampling purposes, but Saratoga County and Saratoga Springs were considered one Subregion in the weighting steps. For North Carolina the Add-on Area was Piedmont (three counties Alamance, Forsyth, Guilford). The Subregions were the four strata for the North Carolina Study. For Vermont the Subregions were Chittendon County and all other counties. For Tennessee, Texas, and Wisconsin there were no Subregions.

5.1.2 Add-On Areas (Add On Area Encompasses Less Than Full State)

Dimensions for raking the 'add-on area' states (AZ, NE, SD, and VA) were the following.

---

[25] The original weighting attempted to use the individual stratum and when convergence failed we collapsed across state.

- (1) CBSACAT * add-on-area*Subregion

- (2) Add-on area*Black

- (3) Add-on area *Hispanic

- (4) Add-on area *Owner/Renter

- (5) Add-on area *Number of Vehicles

- (6) Month

- (7) Day of Week

- (8) HH size * worker

These rakings were done separately for each state.

For Arizona the Add-on Areas were Phoenix, Tucson and rest of state. For Nebraska the Add-on Areas were Omaha area (two counties: Douglas and Sarpy) and the Rest of State (no subregions). For South Dakota the Add-on Areas were Rapid City (two counties: Meade and Pennington), Sioux Falls (two counties: Minnehaha and Lincoln), and the Rest of State (no subregions). For Virginia there was one Add-On Area vs. Rest of State. The Subregions were the 13 strata for the Virginia Add-on Study.

5.1.3   Remainder of U.S.

For the Remainder US domain (all other states not specified above), the recommended dimensions were as follows:

- (1) CBSACAT * State

- (2) State*Black

- (3) State *Hispanic

- (4) State *Owner/Renter

- (5) State *Number of Vehicles

- (6) Month

- (7) Day of Week

- (8) HH size * worker

Because the sample sizes for many of the states in the US remainder are small (e.g., 250 HHs), the US remainder domain was divided into and raked to the following nine Census Divisions as listed below:

- ■ Division 1: New England

- ■ Division 2: Mid-Atlantic

- ■ Division 3: East North Central

- ■ Division 4: West North Central

- ■ Division 5: South Atlantic

- ■ Division 6: East South Central

- ■ Division 7: West South Central

- ■ Division 8: Mountain

- ■ Division 9: Pacific.

Using Census Division rather than Region produced smaller cell sizes and although more collapsing may have been required, it retained more granularity. When convergence failed, the design called for the collapsing of cells (e.g., day to weekday/weekend), then dimensions, and finally the elimination of a dimension.

Note that this means that Census Division was implicitly a crossing variable for every dimension.

5.2   Construction of Control Totals for Time Dimensions

The time dimension control totals all add to the domain household total for each domain. Each month was allocated 1/12 of the household total, so that each month had an equal weight in the final estimator (the sum of household weights with travel dates corresponding to one month were made equal to 1/12 of the household total for each major domain). Each weekday was allocated 1/7 of the household total.

Person-Level Nonresponse Adjustments within Useable Households 6

The aim of the survey was that retrieval interviews were completed for each eligible person within each of the screened households. In terms of sampling, this means that every person in the household had a probability of selection equal to that of the household. In principle, then, each person's base weight was equal to the household base weight. It is appropriate that the base person-level weights include the household nonresponse adjustments.

A household was defined as useable if at least 50% of the eligible adults followed through in completing the extended interview. Within the subset of useable households, we therefore had a considerable pool of nonresponding adults. The interviews of the responding adults needed to be adjusted to account for these nonresponding adults. One approach might be to do this adjustment within each useable household. For example, if there are two adults and one of the two responds, then the weight of the responding adult is doubled to account for the nonresponding adult. A second approach is to generate cells across useable households, so that a particular nonresponding adult is adjusted for in the weights of a number of responding adults across a number of useable households.

The second approach was used, as the first approach leads to more variable adjustments, which in turn leads to higher variances. The nonresponding adult can be adjusted for with a pool of responding adults who are not in the same household but have many of the same characteristics as the nonresponding adult (sex, age, driver status, etc.). The cells were selected based on response propensity (as is generally done under the quasirandomization approach).

The adjustment cells were generated using WESSEARCH in a similar way as has been described for the screener nonresponse and useability adjustment cells. The characteristics used in the WESSEARCH runs are given in Appendix C[26]. These characteristics were chosen following a nonresponse bias analysis of conditional person-level nonresponse. Details of this analysis are given in the nonresponse bias analysis report. Cells were generated separately for adults and for children, as the nonresponse processes for children were different from those of adults (in particular, for

---

[26] The WESSEARCH runs were done using the 2009 household weights. If we did these runs anew using the 2010 household weights, a slightly different set of cells would have resulted.

younger children an adult in the household generally completed the interview for the child, thus nonresponse for children was related to both child and adult characteristics).

The final weighted adjustment factors were computed for each nonresponding person in each adjustment cell as the ratio of the total summation of person-level base weights for all persons responding or nonresponding in the cell to the summation of person-level base weights for responding persons only (those completing travel interviews). Replicated adjustment factors were computed in a similar way using the replicate person-level base weights.

It should be noted that for adult interview adjustments, households with only one adult were excluded, as conditional adult interview response for these households is 100% (if the household is useable, then by definition the one adult has in fact completed a travel interview). Households with only one adult were not excluded from nonresponse adjustment calculations for children, as the response of the one adult does not necessarily determine response for any children in the household.

Person-Level Raking Adjustments 7

A person-level raking adjustment was conducted using the final questionnaire items from completed extended interviews. In the person-level raking adjustments, we utilized the Census county-level population estimates for total persons and by sex, age, and race/ethnicity. We used Census estimates because Census estimates are available for every county, so that there would be no need to distinguish between counties with estimates and counties without, as would be the case with the ACS.

Personal level trimming and raking followed an iterative process. The starting point was the nonresponse adjusted person-level weights coming from the person-level nonresponse adjustments described in Section 5. An initial trimming was done, followed by an initial raking, followed by cycles of trimming and raking and trimming to convergence.

The pre-trim step for each major domain consisted of checking for weights that were more than 3.0 times the median weight for the particular composite domain. If less than 1% of the weights fell into this category, then all such weights were trimmed back to equal the cutoff (3.0 times the median weight for the composite domain). If more than 1% of the weights fell into this category[27], then the largest 1% set of the weights were trimmed back to equal the 99th percentile of the weights.

The trimming steps following the raking steps (the 'post-trim steps') were also done within each composite domain, and targeted for trimming any weights that were 4.5 times smaller or 4.5 times larger than the median weight for the composite domain. A maximum of 2.5% of the weights could be trimmed on the high side and a maximum of 2.5% of the weights could be trimmed on the low side for each post-trim step[28]. If more than 2.5% of the weights was greater than 4.5 times the median weight (less than 4.5 times the median weight), then the largest (smallest) 2.5% of the weights was trimmed back to the 97.5th percentile (the 2.5th percentile).

---

[27] The number of weights affected by this rule was 1% of the number of sample units, rounded up to the smallest larger integer. For example, if the sample size was 120, then the number of trimmed weights was 2 (1.2 rounded up). In particular, there was always at least one weight trimmed if any weights exceeded 3.0 times the median weight. The actual percentage of trimmed weights then could be slightly larger than 1%.

[28] This count of trimmed weights was 2.5% of the number of sample units, rounded up to the smallest larger integer. For example, if the sample size was 100, then the number of trimmed weights was 3 (2.5 rounded up). The actual percentage of trimmed weights on the high and low side could be slightly larger than 2.5%.

The cycle of raking and trimming steps was complete when all of the trimming factors (the adjustments to the weight during the trimming step) for that potentially final trimming step were between 0.99 and 1.01[29].

Two flags indicating trimming status were provided on the delivery files. One flag indicated whether or not a weight was pre-trimmed. Another flag indicated whether or not a weight was post-trimmed at least once. These flags identify a person-level record as being trimmed at least once if their trimming factors were always within the trimming tolerance for the major domain run (either 0.99 through 1.01 or 0.999 through 1.001: see previous paragraph and footnote).

The person-level raking process was done to a tolerance of $\pm 10,000$ (i.e., the weighted person totals were raked until the sum of weight was within 10,000 of the person control totals).

Section 7.1 below describes the initial raking dimensions for person-level raking. These dimensions are what we tried first. In the event that convergence failed, the collapsing of cells was conducted in a stepwise, automated manner using the rules jointly developed by Westat and FHWA. All of our decisions for collapsing cells are documented in Appendix G of this weighting report.

7.1. Person-Level Raking Dimensions

The control totals for all dimensions used in the person-level raking came from the Census population estimates for July 2008. These are Census model based estimates coming from Census 2000.

7.1.1 Add-On States (Add On Area Encompasses The Entire State)

The dimensions for person-level raking adjustments for the 'add-on states' (CA, FL, GA[30], IA, IN, NC, NY, SC[31], TN, TX, VT, and WI) were the following.

- ■ (1) CBSACAT crossed with Subregion (as defined in the paragraph below)
- ■ (2) Subregion*Race

---

[29] 0.999 and 1.001 for Major Domain FL. This was done earlier in the process before we converted to the 0.99 to 1.01 tolerance range.

[30] GA augmented with 1 South Carolina county (Aiken) and 2 Alabama counties

[31] Aiken County was excluded from the South Carolina sample

- ■ (3) Subregion * Ethnicity

- ■ (4) Subregion * Sex * Age group

- ■ (5) Month Pair

- ■ (6) Day of week

These rakings were done separately for each state.

For California the Subregions defined for the first raking dimension were San Diego County vs remainder of state. For Florida we used the seven transportation regions[32]. For Georgia and Indiana the Subregions were the sampling strata (seven for Georgia which includes part of South Carolina and Alabama, seven for Indiana). For South Carolina the Subregions were the sampling strata. For Iowa the Subregions were Pottawattamie county (from the Omaha add-on), Linn County (Cedar Rapids add-on), and the Remainder of the State. For New York the Subregions were 24 of the total 25 sampling strata. Saratoga County was divided into two strata for sampling purposes, but Saratoga County and Saratoga Springs are considered one Subregion in the weighting steps. For North Carolina the Add-on Area was Piedmont (three counties Alamance, Forsyth, Guilford). The Subregions were the four strata for the North Carolina Study. For Vermont the Subregions were Chittendon County and all other counties. For Tennessee, Texas, and Wisconsin there were no Subregions.

Race and ethnicity were collected as household level variables in the NHTS. Thus it was necessary to assign the race/ethnicity of the screener (recruitment) respondent to each person in the household in order to rake on race and ethnicity at the person-level. For example, a respondent was defined as Black if and only if the screener respondent indicated his/her race as Black (a multiracial designation which includes Black was viewed as nonBlack). The control totals for this dimension also included persons who designated a single race with that single race being Black. Similarly, a respondent was defined as Hispanic if and only if the screener respondent indicates themselves as Hispanic in the ethnicity question or in the follow-up question about race. The control totals also only included persons who designate themselves as Hispanic (regardless of race).

---

[32] The original weighting attempted to use the individual stratum, and when convergence failed, we collapsed across state. We expected that some precision would be gained by using the seven transportation regions compared to collapsing the whole state into one geographic stratum.

7.1.2 Add-On Areas (Add On Area Encompasses Less Than Full State)

The dimensions for person-level raking adjustments for the 'add-on area states' (AZ, NE, SD, and VA) were the following.

- ■ (1) CBSACAT crossed with Add-On Area crossed with Subregion
- ■ (2) Add-On Area crossed with race
- ■ (3) Add-on area crossed with ethnicity
- ■ (4) Add-On Area crossed with sex and age group
- ■ (5) Month Pair
- ■ (6) Day of week

These rakings adjustments were done separately for each state.

For Arizona the Add-on Areas were Phoenix, Tucson and the rest of state. For Nebraska the Add-on Areas were Omaha area (two counties Douglas and Sarpy) and the Rest of State (no subregions). For South Dakota the Add-on Areas were Rapid City (two counties Meade and Pennington), Sioux Falls (two counties Minnehaha and Lincoln), and the Rest of State (no subregions). For Virginia there was one Add-On Area vs. Rest of State. The Subregions were the 13 strata for the Virginia Add-on Study.

7.1.3 Remainder of the US

For the Remainder US domain (all other states not specified above, plus the Remainder of Arizona), the dimensions that we used for person-level taking were as follows.

- ■ (1) CBSACAT crossed with State
- ■ (2) State crossed with race
- ■ (3) State crossed with ethnicity
- ■ (4) State crossed with sex and age group
- ■ (5) Month pair
- ■ (6) Day of week

These rakings were done separately for each Census Division. Note that this means that Census Division is implicitly crossed with each dimension.

The control totals for all dimensions came from the Census population estimates for July 2008. Note that race/ethnicity of each person in the household was assumed equal to the race/ethnicity of the screener reference person. Someone was defined as Black if and only if the screener respondent indicated themselves as single-race Black (a multiracial designation was viewed as nonBlack). The control totals also only included persons who designated a single race with that single race being Black (whether or not they also are indicated as being Hispanic). Someone was defined as Hispanic if and only if the screener respondent indicated themselves as Hispanic on the Hispanic question. The control totals also only included persons who designated themselves as Hispanic (regardless of race).

Other Weights 8

8.1. Child Weights

The final child weight for most characteristics is equal to the final household weight (as every child is included if the household is in the sample: there is no subsampling factor for most characteristics). There was a subsample of children age 5 to 15 within the household selected for the purpose of administering a specific section of the interview. A special child-level weighting factor was generated for analysis of this section of the interview. This factor was equal to the number of eligible children in the household. For items in this section of the interview, the weight was this final child weight times the special factor.

If weights are needed for children age 0 to 4, the household weight can be used for this purpose.

8.2 Vehicle Weights

The raked household weight can be used to analyze characteristics of vehicles reported by households.

8.3. Day Weights

The day weights are equal to 365 times the person-level weights. These are the appropriate weights for counting trips for the year (e.g., for annual travel estimates).

References

Burke, J., Mohadjer, L., Green, J., and Waksberg, J. (1994). Composite estimation in national and state surveys, *Proceedings of the Section on Survey Research Methods, American Statistical Association,* 873-878.

Kish, L. (1992). Weighting for unequal P_i. *Journal of Official Statistics 8*, 183-200.

Lohr, S. L., and Rao, J. N. K. (2007). Estimation in multiple-frame surveys, *Journal of the American Statistical Association*, 101, 1019-1030.

Oh, H. L. and Scheuren, F. J. (1983). Weighting adjustment for unit nonresponse, in W.G. Madow, I. Olkin, D. B. Rubin (eds.), *Incomplete Data in Sample Surveys, Vol. 2*, New York: Academic Press, pp. 143-184.

Appendix A.   Genesys Characteristics to be Used to Form Screener Nonresponse Adjustment Cells

Table A-1 below presents for each study the Genesys exchange characteristics that were used to define possible screener nonresponse cells. For example, for the national study, an 'X' with 'Median Rent' indicates that we generated above-median and below-median cells for each stratum in the national study based on the weighted median for median rent within the stratum. WESSEARCH was used to test these cells, and they become part of the final set of nonresponse cells if the above-median and below-median cells for the particular Genesys exchange characteristic registered as significant within the stratum.

Table A-1.   Dichotomous Cells for Potential Nonresponse Adjustment Cells By Study (Part 1 of 2).

	00-National Study	01-PH	02-TU	03-CA	04-FL	05-GA	06-Cedar Rapids, IA	07-IA	08-IN	09-Omaha, NE
Central City with Heavy Rail	X			X	X	X			X	
Central City without heavy rail		X						X		
nonMetro	X			X	X			X		X
Median income			X				X			
Median rent	X				X	X		X		
Median home value			X							
Median years education			X	X				X		X
Percent Homeowners	X	X	X	X	X	X	X	X	X	X
Percent 0 to 44		X	X							
Percent 18 to 24								X		
Percent 25 to 44	X			X		X				
Percent 35 to 54					X					
Percent 35 to 64							X			
Percent 45+	X		X	X						X
Percent 45 to 54								X	X	
Percent 55+		X			X	X				
Percent Asian	X			X						
Percent Black	X	X		X						
Percent Hispanic	X	X		X	X					
Percent White	X	X	X	X	X	X	X	X	X	X
Percent College Grads			X		X					
Percent Income $0 to $25K			X				X			
Percent Income $0 to $35K										X
Percent Income $0 to $50K								X		
Percent Income $25K to $75K	X									
Percent Income $35K to $75K		X								
Percent Income $15K to $75K					X					
Percent Income $50K and up			X							
Percent Income $75K to $100K							X			
Percent Income $100K up					X					

Table A-1.   Dichotomous Cells for Potential Nonresponse Adjustment Cells By Study (Part 2 of 2).

	10-NY	11-NC	12-NC Pied	13-SC	14—Rapid City, SD	15-Sioux Falls, SD	16-TN	17-VA	18-VT	19-WI	20-TX
Central City with Heavy Rail											
Central City without heavy rail							X	X	X		X
nonMetro				X				X	X		X
Median income						X				X	
Median rent	X	X				X	X		X		X
Median years education									X		
Percent Homeowners	X		X			X	X	X		X	X
Percent 18 to 34										X	
Percent 45+	X		X					X			X
Percent 55 to 64							X				
Percent 55+										X	
Percent Asian	X										X
Percent Black	X	X	X								X
Percent Hispanic	X	X	X								X
Percent White	X		X		X	X	X	X		X	X
Percent Income $0 to $25K			X	X							
Percent Income $0 to $35K									X	X	
Percent Income $0 to $50K					X	X					
Pct Income $25K to $100K	X										
Percent Income $50K up			X				X			X	
Percent Income $75K up									X		

Appendix B. Genesys and Screener Questionnaire Characteristics Used to Form Household Useability-Nonuseability Adjustment Cells

Table B-1 below presents for each study the Genesys exchange and screener questionnaire characteristics that were used to define possible usability-nonusability cells, similar to those of Appendix A. WESSEARCH was used to test these cells within each study-sample group-stratum combination, and those that best distinguished usability rates were retained for cell construction for that study-sample group-stratum triplet.

Table B-1. Dichotomous Cells for Potential Usability Nonresponse Adjustment Cells By Study (Part 1 of 2).

	00-National Study	01-PH	02-TU	03-CA	04-FL	05-GA	06-Cedar Rapids, IA	07-IA	08-IN	09-Omaha, NE
Screener respondent age 18 to 24	X	X	X	X	X	X	X	X	X	X
Screener respondent age 25 to 34	X	X	X	X	X	X	X	X	X	X
Screener respondent age 35 to 44		X	X	X	X	X	X	X	X	
Screener respondent age 35 to 54	X									X
Screener respondent age 45 to 54		X	X	X	X	X	X	X	X	
Screener respondent age 55 to 64		X	X	X		X			X	
Screener respondent age 55 and	X				X		X	X		X
Screener respondent age 65 and		X	X	X		X			X	
Scr resp less than high schl diploma	X	X	X	X	X		X	X	X	X
Scr resp high school diploma only		X	X	X	X		X			
Screener respndnt Bachelor's Degree	X		X		X					
Screener resp BS or graduate degree		X				X			X	
Screener respondent graduate	X		X	X	X			X		
Household has children	X	X	X	X	X	X	X	X	X	X
Household has trucks		X	X							
Household has workers	X	X	X	X	X	X	X	X	X	X
Screener respondent Hispanic	X	X	X	X	X	X			X	X
Scr respondent nonHispanic White	X	X	X	X	X	X	X	X	X	X
Scr respondent nonHispanic Black	X	X	X	X	X	X			X	X
Household in single family house	X			X					X	
Household with one adult	X				X	X	X	X	X	X
Household with more than 2 adults	X	X	X	X	X	X	X	X	X	X
HH has fewer drivers than adults	X	X	X	X	X	X	X	X	X	
HH has fewer vehicles than drivers	X	X	X	X	X	X			X	
HH has more vehicles than drivers			X		X					
Household has two workers		X	X		X	X	X	X		
Household has three or more	X	X	X	X	X	X	X	X	X	X
Homeowner vs renter	X	X	X	X	X	X	X	X	X	X
Median home value (exchange level)			X	X			X			X
Exchange percent age 0 to 17		X	X						X	
Exchange percent age 45+			X							X
Exchange percent age 45-64							X		X	
Exchange percent age 55+					X					
Exchange percent age 65+							X			
Exchange percent homeowners				X					X	X
Exchange percent nonHisp Whites	X	X			X	X	X	X	X	X
Exch pct annl income less than $50K			X	X			X		X	X

Table B-1. Dichotomous Cells for Potential Useability Adjustment Cells By Study (Part 2 of 2).

	10-NY	11-NC	12-NC Pied	13-SC	14—Rapid City, SD	15-Sioux Falls, SD	16-TN	17-VA	18-VT	19-WI	20-TX
Screener respondent age 18	X	X	X	X	X	X	X	X	X	X	X
Screener respondent age 25	X	X	X			X		X	X		X
Screener respondent age 25				X	X		X			X	
Screener respondent age 35	X	X	X			X		X	X		X
Screener respondent age 35											
Screener respondent age 45	X	X	X	X	X	X	X	X	X	X	X
Screener respondent age 55	X			X			X	X	X		X
Screener respondent age 55		X	X		X	X				X	
Screener respondent age 65	X			X			X	X	X		X
Scr resp less than high schl	X	X	X	X		X	X	X		X	X
Scr resp high school diploma		X									
Screener respndnt Bachelor's							X	X		X	X
Screener resp BS or graduate			X								
Screener respondent	X	X		X			X	X		X	X
Household has children	X	X	X	X	X	X	X	X	X	X	X
Household has trucks		X		X							
Household has workers	X	X	X	X	X	X	X	X	X	X	X
Screener respondent Hispanic	X	X	X	X		X	X	X	X	X	X
Scr respondent nonHispanic	X	X	X	X		X	X	X	X	X	X
Scr respondent nonHispanic	X	X		X						X	X
Household in single family	X	X	X	X			X	X	X		X
Household has one adult		X	X	X	X	X	X	X			X
Household with more than 2	X	X	X	X	X		X	X	X	X	X
HH has less drivers than age	X	X	X	X	X	X	X	X	X	X	X
HH has less vehicles than	X	X	X	X			X		X	X	
HH has more vehicles than	X	X							X		
Household has two workers	X		X	X			X	X	X		X
HH has three or more workers	X		X	X	X	X	X	X	X	X	X
Homeowner vs renter	X	X	X	X	X		X	X	X	X	X
Median home value	X	X									
Exchange percent age 0 to 17											
Exchange percent age 45+	X	X	X				X	X		X	X
Exchange percent age 45-64											
Exchange percent age 55+											
Exchange percent age 65+											
Exchange percent	X	X	X				X	X		X	X
Exchange percent nonHisp	X	X	X	X		X	X	X			X
Exch pct annl inc less than	X										
Exch pct annl inc less than			X						X		
Exch pct annl inc less than		X						X		X	X
Exch pct annl inc $50K–	X										
Exch pct annl inc $75K and			X					X			

Appendix C.   Genesys and Screener Questionnaire Characteristics Used to Form Person-level Interview Nonresponse Adjustment Cells

Table C-1 presents for each study the Genesys exchange and screener questionnaire characteristics that were used to define possible adult interview nonresponse adjustment cells, similar to those of Appendix B. WESSEARCH was used to test these cells within each study-sample group-stratum combination, and those that best distinguished interview rates were retained for cell construction for that study-sample group-stratum triplet. These potential cells were only for multi-adult households (households with two or more adults). Households with single adults did not receive this adjustment (as response is 100% at this level for useable households).

Table C-1.   Dichotomous Cells for Potential Adult Interview Nonresponse Adjustment Cells By Study (within multi-adult households) (Part 1 of 2).

	00-National Study	01-PH	02-TU	03-CA	04-FL	05-GA	06-Cedar Rapids IA	07-IA	08-IN	09-Omaha NE
Person is screener respondent	X	X	X	X	X	X	X	X	X	X
Person is scr rspndt's spouse,	X	X	X	X	X	X	X	X	X	X
Person is scr rspndt's relative	X	X	X	X	X	X	X	X	X	X
Person is not related to scr rspndt	X	X	X	X	X	X	X	X	X	X
Person age 18 to 24	X	X			X	X	X	X	X	X
Person age 18 to 34			X	X						
Person age 25 to 34	X	X			X	X				
Person age 25 to 44							X	X	X	X
Person age 35 to 44			X	X						
Person age 35 to 54	X	X			X	X				
Person age 45 to 54				X						X
Person age 45 and over			X				X	X	X	
Person age 55 to 64										
Person age 55 and over	X	X			X	X	X			X
Person age 65 and over										
1-Person has less than high schl	X	X	X	X	X	X	X	X	X	X
2-Person has high school diploma		X	X	X	X		X	X	X	X
3-Person has some postsecondary		X	X	X						
4-Person has BS with no graduate				X						
5-Person has graduate degree	X			X		X		X	X	X
Person is worker	X	X	X	X	X	X	X		X	X
Person is male or female	X	X			X	X	X	X	X	X
Family annual income less than	X		X	X	X	X				X
Family annual income less than		X								
Family annual income $10K to	X		X	X	X	X				X
Household has children	X	X	X	X	X	X	X	X	X	X
Screener interview in Spanish	X	X	X	X	X	X	X	X	X	X
Screener rspndt Hispanic, Black,	X	X	X	X	X	X	X			
Scr respondent White, Asian	X	X	X	X	X	X	X			
Household in single family house	X	X	X	X		X				
Homeowner vs renter	X	X	X	X	X	X	X	X	X	X
In Heavy Rail Central City area						X				
In nonHeavy Rail Central City area										
Above-median home-ownership		X	X							X
Above-median white percentage	X	X	X	X	X					

Table C-1.   Dichotomous Cells for Potential Adult Interview Nonresponse Adjustment Cells By Study (within multi-adult households) (Part 2 of 2).

	10-NY	11-NC	12-NCpd	13-SC	14-RapCty	15-SxFlls	16-TN	17-VA	18-VT	19-WI	20-TX
Person is screener respondent	X	X	X	X	X	X	X	X	X	X	X
Person is scr rspndt's spouse,	X	X	X	X	X	X	X	X	X	X	X
Person is scr rspdnt's relative	X	X	X	X	X	X	X	X	X	X	X
Person is not related to scr rspndt	X	X	X	X	X	X	X	X	X	X	X
Person age 18 to 24	X	X	X			X	X	X	X	X	X
Person age 18 to 34				X							
Person age 25 to 34		X	X			X	X	X	X	X	
Person age 25 to 44	X										X
Person age 35 to 44											
Person age 35 to 54		X	X	X		X	X	X	X	X	
Person age 45 to 54											X
Person age 45 and over	X										
Person age 55 to 64			X			X	X		X		X
Person age 55 and over		X		X				X		X	
Person age 65 and over			X			X	X		X		X
1-Person has less than high schl	X	X	X	X				X	X		X
2-Person has high school diploma	X	X	X						X		X
3-Person has some postsecondary		X	X								
4-Person has BS with no graduate											
5-Person has graduate degree				X				X	X		
Person is worker	X	X	X	X	X	X	X	X	X	X	X
Person is male or female	X	X	X	X				X	X	X	X
Family annual income less than	X	X	X	X		X	X	X	X	X	X
Family annual income less than											
Family annual income $10K to		X		X			X			X	X
Household has children	X	X	X	X		X		X		X	X
Screener interview in Spanish	X	X	X	X			X	X		X	X
Screener rspdnt Hispanic, Black,	X	X	X	X			X	X		X	X
Scr respondent White, Asian	X	X	X	X			X	X		X	X
Household in single family house	X	X	X	X		X	X	X	X		X
Homeowner vs renter	X	X	X	X	X	X	X	X	X	X	X
In Heavy Rail Central City area	X										
In nonHeavy Rail Central City area		X		X						X	X
Above-median home-ownership	X		X								
Above-median white percentage	X		X				X				X

Table C-2 presents for each study the Genesys exchange and screener questionnaire characteristics that were used to define possible child interview nonresponse adjustment cells. WESSEARCH was used to test these cells within each study-sample group-stratum combination, and those that best distinguished interview rates will be retained for cell construction for that study-sample group-stratum triplet.

Table C-2.   Dichotomous Cells for Potential Child Interview Nonresponse Adjustment Cells By Study (Part 1 of 2).

	00-National Study	01-PH	02-TU	03-CA	04-FL	05-GA	06-Cedar Rapids, IA	07-IA	08-IN	09-Omaha, NE
Child 14 to 17	X	X	X		X	X				
Reference adult age 18 to 24			X							X
Reference adult age 25 to 34			X							
Reference adult age 65 and										X
Ref adlt has less than HS					X					X
Reference adult has HS										
Ref adlt has some post HS				X					X	
Reference adult has BS				X					X	
Reference adult has graduate										
Child related to scrnr	X		X	X	X		X		X	X
Family annual income less	X		X			X				
Family annual income less		X			X					
Family annual income less				X						
Family annual income $10K to						X				
Family annual income $30K to				X						
Family annual income $50K to				X						
Screener interview in Spanish	X	X		X	X	X			X	X
Scrnr rspdnt Black, Hispanic,	X	X		X	X	X			X	
Scrnr rspdnt White, Asian	X	X		X	X	X			X	
Household in single family	X	X	X	X	X	X	X		X	
Homeowner vs renter	X			X		X	X			
Three or more children in				X						
In Heavy Rail Central City area		X						X		
In nonHeavy Rail Central City										
Above-median home-	X	X	X	X	X					
Above-median white	X	X	X	X	X				X	

	10-NY	11-NC	12-NC pd	13-SC	14-Rap Cty	15-Sx Flls	16-TN	17-VA	18-VT	19-WI	20-TX
Child age 14 to 17			X	X							X
Reference adult age 18 to 24	X							X			
Reference adult age 25 to 34											
Reference adult age 65 and over	X										
Ref adlt has less than high schl diploma		X	X				X	X		X	X
Reference adult has HS diploma only			X								
Reference adult has some postsecondary											
Reference adult has Bachelor's Degree											
Reference adult has graduate degree		X									X
Child related to scrnr respondent	X	X	X		X		X	X	X	X	X
Family annual income less than $10,000			X	X							
Family annual income less than $20,000							X				X
Family annual income less than $30,000		X									
Family annual income $10K to $20K											
Family annual income $30K to $50K											
Family annual income $50K to $70K											
Screener interview in Spanish	X	X	X	X				X		X	X
Screener rspdnt Black, Hispanic, Other								X		X	X
Scrnr rspdnt White, Asian								X		X	X
Household in single family house			X				X				
Homeowner vs renter				X	X	X				X	X
Three or more children in household											
In Heavy Rail Central City area	X										
In nonHeavy Rail Central City area											
Above-median home-ownership area							X	X			
Above-median white percentage area	X	X	X		X	X	X	X	X	X	X

Appendix D.   Construction of Control Totals from the American Community Survey

Geographic raking cells were always defined at the county level (either single counties or sets of counties).

For the larger counties, the ACS had estimates of control totals that are based on the one year of data alone (the most recent year). This was the ideal situation. For many other counties, the ACS had an estimate, but this estimate is a moving average based on the three most recent years. ACS moving averages were used only in such a way that they did not contradict the one-year estimates for any given area. The three-year estimates were only used to 'fill in' whatever was missing from the one-year estimates (only percentages were utilized). Finally, there were counties for which there was no estimate.

For a geographic raking cell to have a one-year ACS estimate, it was necessary that all counties within the raking cell had one-year ACS estimates. The raking cell estimate was the summation of the one-year ACS county estimates. If all counties within a raking cell had three-year ACS estimates, but only some had one-year ACS estimates, then a three-year estimate was computed for these domains, by taking the summation of the county-level three-year estimates. There was no 'mixing' of one-year and three-year estimates in any summations. If a particular raking cell had at least one county with no ACS estimate at all (one- or three-year), then no estimate was generated for the raking cell.

As an example, we explicate control totals for Indiana (we choose Indiana as it has a relatively simple strata structure, but not a trivial one). Indiana has three strata (01-Northwest Indiana, 02-Marion County, and 05-Small Metropolitan) for which one-year ACS estimates are available for 2007 for every county within the stratum, and three strata (03-Suburban counties, 04-Medium Metropolitan, 06-Small Cities) for which three-year estimates are available for 2007 for every county within the stratum. One stratum (07-Rural) does not have any ACS estimates for at least some of its counties, and is a complement stratum (viewed as having no ACS estimate).

Tables D-1 and D-2 present what the compiled 2007 ACS data look like for counts of households by home ownership tenure (Dimension 2).

Table D-1. 2007 one-year ACS estimates, Indiana strata

	Number of households	Number of households who own their home	Number of households who rent (or other)
Indiana	2,462,278	1,763,734	698,544
		71.63%	28.37%
Stratum 1 Northwest IN	289,415	208,210 71.94%	81,205 28.06%
Stratum 2 Marion County	356,238	219,970 61.75%	136,268 38.25%
Stratum 5 Small Metro IN	313,778	203,013 64.70%	110,765 35.30%
Stratum 3 Suburban counties	n/a	n/a	n/a
Stratum 4 Medium Metro, IN	n/a	n/a	n/a
Stratum 6 Small cities	n/a	n/a	n/a
Comp Stratum Rural counties	n/a	n/a	n/a

Table D-2. 2007 three-year ACS estimates, Indiana strata

	Number of households	Number of households who own their home	Number of households who rent (or other)
Indiana	2,447,887	1,764,680	683,207
		72.09%	27.91%
Stratum 1 Northwest IN	286,972	208,714 72.73%	78,258 27.27%
Stratum 2 Marion County	357,445	218,903 61.24%	138,542 38.76%
Stratum 5 Small Metro IN	309,006	201,654 65.26%	107,352 34.74%
Stratum 3 Suburban counties	365,297	285,626 78.19%	79,671 21.81%
Stratum 4 Medium Metro, IN	401,220	285,405 71.13%	115,815 28.87%
Stratum 6 Small cities	342,702	259,187 75.63%	83,515 24.37%
Comp Stratum Rural counties	n/a	n/a	n/a

It should be noted that the one-year and three-year estimates do not agree with each other as to overall counts or percentages, as we had to expect in general.

Our first step was to generate total-household control totals for each stratum, taking the one-year estimates for the study area and the three one-year strata, and imputing household cell totals for

Stratum 3, Stratum 4, Stratum 6, and the Complement Stratum. The initial estimate for the complement stratum was the three-year ACS total minus the 3-year ACS totals for the six strata with ACS totals available (a value of 385,245, in italics). The final adjusted ACS household totals utilized the one-year estimates for the study area and the three strata that have one-year household total estimates, and then 'allocated' the balance of households for Indiana (state total minus the three strata) to the three other strata using the percentages from the three-year estimates. The three-year estimates provide input into the system only by providing percentages, which were used to allocate the balance (2,462,278 minus 289,415+356,238+313,778).

Table D-3.   Imputations for 2007 Strata Household Totals in Indiana

	Number of households 3-year ACS	Stratum percent of study area	Adjusted ACS household totals	Stratum percent of study area
Indiana State Level	2,447,887	100.00	2,462,278	100.00
Stratum 1-Northwest IN	286,972	11.72	289,415	11.75
Stratum 2-Marion County	357,445	14.60	356,238	14.47
Stratum 5-Small Metro	309,006	12.62	313,778	12.74
Stratum 3-Suburban counties	365,297	14.92	367,346	14.92
Stratum 4-Medium Metro	401,220	16.39	403,471	16.39
Stratum 6-Small Cities	342,702	14.00	344,624	14.00
Complement—Rural counties	385,245	15.74	387,406	15.73

The 367,346 for Stratum 3 for example is computed as 365,297 divided by (365,297+401,220+342,702+385,245) multiplied by {2,462,278-(289,415+356,238+313,778)}.

The next step was to allocate across the household tenure cells (homeowners and renters). For the full state and strata 1, 2, and 5, the household tenure cell totals were equal to the one-year estimates. For Strata 3, 4, and 6, the percentages from the three-year estimates were used along with the just-imputed stratum-level household totals to get new imputations for home owning households and renter households. These imputations were given in Table D-4.

Table D-4.   Imputations for Strata 3, 4, and 6 household composition cells.

	Number of households	Number of households who own their home	Number of households who rent (or other)
Indiana	2,462,278	1,763,734	698,544
		71.63%	28.37%
Stratum 1 Northwest IN	289,415	208,210	81,205
		71.94%	28.06%
Stratum 2 Marion County	356,238	219,970	136,268
		61.75%	38.25%
Stratum 5 Small Metro, IN	313,778	203,013	110,765
		64.70%	35.30%
Stratum 3 Suburban counties	367,346	287,228	80,118
		78.19%	21.81%
Stratum 4 Medium Metro, IN	403,471	287,006	116,465
		71.13%	28.87%
Stratum 6 Small cities	344,624	260,641	83,983
		75.63%	24.37%

Note that the three-year estimates were set aside for Strata 3, 4, and 6, and were replaced with imputed counts. However, the three-year ACS percentages of homeowners and renters were utilized fully for these three strata (this is the best information about this that is available, and it doesn't contradict any of the information derived from the one-year estimates). The final step was to generate imputed counts for the complement stratum by subtracting the four strata totals from the study area total, for both homeowners and renters separately. This is given in Table D-5.

Table D-5.   Imputations for Complement Stratum household composition cells.

	Number of households	Number of households who own their home	Number of households who rent (or other)
Indiana	2,462,278	1,763,734	698,544
		71.63%	28.37%
Stratum 1 Northwest IN	289,415	208,210	81,205
		71.94%	28.06%
Stratum 2 Marion County	356,238	219,970	136,268
		61.75%	38.25%
Stratum 5 Small Metro, IN	313,778	203,013	110,765
		64.70%	35.30%
Stratum 3 Suburban counties	367,346	287,228	80,118
		78.19%	21.81%
Stratum 4 Medium Metro, IN	403,471	287,006	116,465
		71.13%	28.87%
Stratum 6 Small cities	344,624	260,641	83,983
		75.63%	24.37%
Comp Stratum Rural counties	387,406	297,666	89,740
		76.84%	23.16%

Appendix E. Imputation Report

There were missing values for almost all of the screener and extended interview variables. Although the proportion missing of items was very low, even such a small number of imputations was still needed for poststratification, as the weighted totals have to represent the full population to the last percentage point to match the Census and ACS control totals. Only variables that were used in poststratification were imputed.

The imputation methodology was hot deck imputation. Under this approach, each record that requires an imputation for a particular item is called a 'beggar'. Each beggar was randomly assigned a 'donor': a record that did have that particular item nonmissing. This random selection guarantees that the imputed records had roughly the same distribution as the donor records, preventing for example a set of imputed values that had less variability because they are generated from means. The donors were selected within an imputation cell. The imputation cells were defined by levels of characteristics that were correlated to the item being imputed. This helped reduce the variability of the imputation, and guaranteed that the imputations successfully passed edit checks (e.g., no seven-year old drivers).

Some characteristics had no missing values. These include geographic variables, Genesys exchange characteristics, and a few variables that were collected for all screener respondents at the screener level.

The following variables were imputed in the following order (order was important in that later steps can use the imputations from the earlier steps):

- Age of person;

- Homeowner vs. renter or other at the household level;

- Worker status for all adults (full-time/part-time worker vs. not worker);

- Sex of person;

- Race/ethnicity of household reference person.

The following characteristics were used for formation of imputation cells. Missingness is one category. For the imputation of age of person the following cell generators were used:

- Age of person, categorized into 0-17 and 18+;

- Driver status;

- Worker status;

- Recoded primary activity last week (working age, going to school, retired, and other);

- Educational status of person;

- Respondent relationship to household respondent;

- Recoded race/ethnicity of person (Hispanic, nonhispanic White, nonhispanic Black, nonhispanic others);

For the imputation of homeowner vs. renter or other the following cell generators were used:

- Poststratification domain;

- Domain quarter;

- Genesys exchange percentage of home owners in exchange (categorized into five categories by quintiles within domain and domain quarter);

- Family income, categorized into less than $10,000 annual income, $10K-$20K annual income, $20K-$40K annual income, $40K-$60K annual income, $60K-$80K annual income, $80K-$100K annual income, missing family income.

- Number of adults in household;

- Presence of children in household;

- Age of reference person, categorized into 18-24, 25-34, 35-44, 45-54, 55-64, 65+;

- Home type (single family home, etc.);

For imputation of working status of adults[33], the following cell generators were used:

- Poststratification domain;

- Domain quarter;

- Age of person, categorized into 0-15, 16-64, 65+;

- Educational status of adult;

- Race/ethnicity of reference person in household;

---

[33] When missing, we impute working status for children as not working.

- Family income (same categories as in the previous list);

- Home-owner vs. renter status;

- Presence of children in household;

- Number of adults in household, and working status of other adults (no other adults, one other adult who is working, one other adult who is not working, one other adult with unknown working status, more than one other adult);

For imputation of sex, we used the following list:

- Poststratification domain;

- Domain quarter;

- Number of adults in household, presence of children, and sex of other adults;

- Age of person, categorized into 0-17, 18-64, 65+;

For imputation of race/ethnicity, we imputed each person as Black nonHispanic, Hispanic, Black Hispanic, or nonHispanic other race (Whites, Asians, American Indians, Pacific Islanders). All persons in the household 'inherited' the race/ethnicity of the screener reference person. Hotdeck imputation was only needed then to fill in race/ethnicity for the screener reference person where it was missing. Multirace persons were viewed as nonBlack. The following characteristics were used (at the household level) to set up boundaries for imputation of race/ethnicity:

- Poststratification domain;

- Domain quarter;

- Genesys exchange percentage of Whites in exchange (categorized into six categories by percentiles[34] within domain);

- Genesys exchange percentage of Blacks in exchange (categorized into six categories by percentiles[35] within domain);

- Genesys exchange percentage of Hispanics in exchange (categorized into six categories by percentiles[36] within domain);

- Screener interview completed in Spanish;

---

[34] 10th percentile, 25th percentile, median, 75th percentile, 90th percentile.

[35] 10th percentile, 25th percentile, median, 75th percentile, 90th percentile.

[36] 10th percentile, 25th percentile, median, 75th percentile, 90th percentile.

■   Education status of screener reference person;

■   Presence of children in household;

■   Family income, categorized into less than $10,000 annual income, $10K-$20K annual income, $20K-$40K annual income, $40K-$60K annual income, $60K-$80K annual income, $80K-$100K annual income, missing family income.

Appendix F. Summarization of Household Raking Steps

Arizona

- Obtained full convergence to the control totals (with 60 being the largest observed difference between weight sum and control total)

- All dimensions retained

- Collapsing of cells was required.

  In order to retain Blacks versus nonBlacks, Hispanics versus nonHispanics, homeowners versus renters, and all four vehicle categories we had to sacrifice geography (collapsing the Arizona Add-on and Arizona Remainder). There was just too small a sample size in the Arizona Remainder cells (89 households all together) to allow for retention of this cell except in Dimension 1 itself.

  Also Month 10 and Month 11 (January and February) had to be collapsed, again due to low sample sizes (these months are only contributed to from the national sample: AZ add-on is in September through November).

- Process required 16 rake-trim cycles

- A total of 3.8% of the records were trimmed at least once on the high side, and 2.5% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

California

- Obtained full convergence to the control totals (with 3004 being the largest observed difference between weight sum and control total)

- All dimensions retained

- Collapsing of cells was not required

- The process required 8 rake-trim cycles

- A total of 3.2% of the records were trimmed at least once on the high side, and 0.41% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

Florida

- Obtained full convergence to the control totals (with 409 being the largest observed difference between weight sum and control total)

- All dimensions retained

- Collapsing was unnecessary

- The process required 11 rake-trim cycles

- A total of 3.6% of the records were trimmed at least once on the high side, and 0.9% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

Georgia

- Obtained full convergence to the control totals (with 1881 being the largest observed difference between weight sum and control total)

- All dimensions retained

- Limited collapsing was necessary in dimensions 2 through 5:

  - **Dimension 2**: (stratum by Black/NonBlack) collapsed Aiken County, SC with Lee & Russell Counties, AL within Black/NonBlack.

  - **Dimension 3**: (stratum by Hispanic/NonHispanic) collapsed Aiken County, SC with Lee & Russell Counties, AL and Strata 03, 04 within Hispanic/NonHispanic.

  - **Dimension 4**: (stratum by Owner/Renter) collapsed Aiken County, SC with Lee & Russell Counties, AL within Owner/Renter.

  - **Dimension 5**: (stratum by #vehicle categories) collapsed Aiken County, SC with Lee & Russell Counties, AL and Strata 03 within 0 Vehicles.

- The process required 7 rake-trim cycles

- A total of 3.4% of the records were trimmed at least once on the high side, and 1.9% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

Iowa

- Obtained full convergence to the control totals (with 447 being the largest observed difference between weight sum and control total)

- All dimensions retained

- Limited collapsing was necessary in dimensions 2 through 5:

  - **Dimension 2**: (stratum by Black/NonBlack) collapsed Linn County, Pottawattamie County, and remainder of Iowa and Black/NonBlack

  - **Dimension 3**: (stratum by Hispanic/NonHispanic) collapsed Linn County, Pottawattamie County, and remainder of Iowa

  - **Dimension 4**: (stratum by owner/renter) collapsed Linn County, Pottawattamie County, and remainder of Iowa

  - **Dimension 5**: (stratum by vehicle categories) collapsed Linn County, Pottawattamie County, and remainder of Iowa

  - **Dimension 5**: (month) collapsed May and June.

- The process required 8 rake-trim cycles

- A total of 3.14% of the records were trimmed at least once on the high side, and 1.71% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

Indiana

- Obtained full convergence to the control totals (with 4809 being the largest observed difference between weight sum and control total)

- All dimensions retained

- Limited collapsing was necessary

  - **Dimension 1**: (CBSA status by stratum) collapsed CBSA 2 and 4 in Small Cities

  - **Dimension 2**: (stratum by Black/NonBlack) collapsed Northwest Indiana, Marion County and Suburban Counties, and collapsed Medium Metropolitan, Small Metropolitan, Small Cities, and Rural

- **Dimension 3**: (stratum by Hispanic/NonHispanic) collapsed Northwest Indiana, Marion County and Suburban Counties, collapsed Medium Metropolitan, Small Metropolitan, Small Cities, and Rural

- **Dimension 5**: (stratum by vehicle categories) collapsed Northwest Indiana, Marion County and Suburban Counties for the 0 vehicle category

■ The process required 28 rake-trim cycles

■ A total of 3.01% of the records were trimmed at least once on the high side, and 0.52% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

North Carolina

■ Obtained full convergence to the control totals (with 539 being the largest observed difference between weight sum and control total)

■ All dimensions retained

■ Collapsing was necessary:

- **Dimension 2**: (CBSA categories by stratum) collapsed North Carolina Metrolina stratum and North Carolina Urban with household survey stratum

- **Dimension 3**: (stratum by Black/NonBlack) collapsed North Carolina Metrolina stratum and North Carolina Urban with household survey stratum

- **Dimension 4**: (stratum by Hispanic/NonHispanic) collapsed North Carolina Metrolina stratum and North Carolina Urban with household survey stratum

- **Dimension 5**: (stratum by owner/renter) collapsed North Carolina Metrolina stratum and North Carolina Urban with household survey stratum

■ The process required 6 rake-trim cycles

■ A total of 3.17% of the records were trimmed at least once on the high side, and 0.84% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

Nebraska

- Obtained full convergence to the control totals (with 1264 being the largest observed difference between weight sum and control total)

- All dimensions retained

- Limited collapsing was necessary in dimensions 2 through 5:

    - **Dimension 2**: (stratum by Black/NonBlack) collapsed Omaha with Remainder of NE within Black/NonBlack.

    - **Dimension 3**: (stratum by Hispanic/NonHispanic) collapsed Omaha with Remainder of NE within Hispanic/NonHispanic.

    - **Dimension 4**: (stratum by Owner/Renter) collapsed Omaha with Remainder of NE within Owner/Renter.

    - **Dimension 5**: (stratum by #vehicle categories) collapsed Omaha with Remainder of NE within 0 Vehicles.

- The process required 11 rake-trim cycles

- A total of 3.3% of the records were trimmed at least once on the high side, and 0.62% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

New York

- Obtained full convergence to the control totals (with 6 being the largest observed difference between weight sum and control total)

- All dimensions retained

- Collapsing was necessary:

    - **Dimension 2**: (stratum by Black/NonBlack) collapsed the Capital District with Saratoga County and Glens Falls, Utica-Rome with Syracuse, Ithaca with Elmira with Binghamton, Kingston with Newburgh-Middletown, Bronx with Manhattan, Brooklyn with Queens, Nassau with Suffolk, and Putnam with Rockland

    - **Dimension 3**: (stratum by Hispanic/NonHispanic) collapsed the Capital District with Saratoga County and Glens Falls, Utica-Rome with Syracuse, Ithaca with Elmira with Binghamton, Rochester with Buffalo, Bronx with Manhattan, Brooklyn with Queens, and Nassau with Suffolk

-     – **Dimension 4**: (stratum by owner/renter) collapsed Bronx with Manhattan, Brooklyn with Queens, and Nassau with Suffolk

-     – **Dimension 5**: (stratum by vehicle categories) collapsed Bronx with Manhattan, Brooklyn with Queens, and Nassau with Suffolk

- The process required 15 rake-trim cycles

- A total of 3.1% of the records were trimmed at least once on the high side, and 0.6% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

South Caronia

- Obtained full convergence to the control totals (with 3930 being the largest observed difference between weight sum and control total)

- All dimensions retained

- Limited collapsing was necessary:

  -     – **Dimension 3**: (state by Hispanic/NonHispanic) collapsed Beaufort and Spartanburg Counties.

  -     – **Dimension 5**: (state by vehicle categories) collapsed Beaufort and Spartanburg Counties for the 0 vehicle category

- The process required 16 rake-trim cycles

- A total of 3.6% of the records were trimmed at least once on the high side, and 0.72% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

South Dakota

- Obtained full convergence to the control totals (with 400 being the largest observed difference between weight sum and control total)

- Unable to retain Blacks versus nonBlacks

- Collapsing was necessary

- **Dimension 3**: (stratum by Hispanic/NonHispanic) we had to do full collapsing (collapsed Rapid City with Sioux Falls with the Remainder of SD).

- **Dimension 5**: (stratum by vehicle categories) collapsed Rapid City with the Remainder of SD for the 0 vehicle category

■ The process required 18 rake-trim cycles

■ A total of 3.7% of the records were trimmed at least once on the high side, and 2.4% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

Tennessee

■ Obtained full convergence to the control totals (with 48 being the largest observed difference between weight sum and control total)

■ All dimensions retained

■ No collapsing of cells was required

■ The process required 11 rake-trim cycles

■ A total of 3.33% of the records were trimmed at least once on the high side, and 0.55% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

Texas

■ Obtained full convergence to the control totals (with 545 being the largest observed difference between weight sum and control total)

■ All dimensions retained

■ No collapsing of cells was required

■ The process required 5 rake-trim cycles

■ A total of 3.3% of the records were trimmed at least once on the high side, and .63% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

Virginia

■ Obtained full convergence to the control totals (with 2413 being the largest observed difference between weight sum and control total)

■ All dimensions retained

■ Limited collapsing was necessary in dimensions 2 through 5:

– **Dimension 2**: (stratum by Black/NonBlack) collapsed across all strata within Black/NonBlack.

– **Dimension 3**: (stratum by Hispanic/NonHispanic) collapsed across all strata within Hispanic/NonHispanic.

– **Dimension 4**: (stratum by Owner/Renter) collapsed across all strata within Owner/Renter.

– **Dimension 5** (stratum by #vehicle categories) collapsed across all strata within 0 Vehicles.

■ The process required 19 rake-trim cycles

■ A total of 3.1% of the records were trimmed at least once on the high side, and 0.28% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

Vermont

■ Obtained full convergence to the control totals (with 50 being the largest observed difference between weight sum and control total)

■ Unable to retain Blacks versus nonBlacks and Hispanics versus nonHispanics dimensions

■ Collapsing was necessary:

– **Dimension 2**: (stratum by Black/NonBlack) we had to do full collapsing (collapsing Chittenden County and Remainder Vermont)

– **Dimension 3**: (stratum by Hispanic/NonHispanic) we had to do full collapsing (collapsing Chittenden County and Remainder Vermont)

- The process required 8 rake-trim cycles

- A total of 2.01% of the records were trimmed at least once on the high side, and 0.3% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

Wisconsin

- Obtained full convergence to the control totals (with 58 being the largest observed difference between weight sum and control total)

- All dimensions retained

- Collapsing of cells was not required

- The process required 16 rake-trim cycles

- A total of 3.57% of the records were trimmed at least once on the high side, and 0.06% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

New England Census Division Remainder

- Obtained full convergence to the control totals (with 2,137 being the largest observed difference between weight sum and control total)

- All dimensions retained

- Limited collapsing was necessary in the first five dimensions:

  - **Dimension 1**: (CBSA status by state) collapsed CBSA 3 and 4 in Massachusetts

  - **Dimension 2**: (state by Black/NonBlack) collapsed all 5 states (Massachusetts, New Hampshire, Rhode Island, Connecticut, Maine) to the census division level

  - **Dimension 3**: (state by Hispanic/NonHispanic) collapsed all 5 states to the census division level

  - **Dimension 4**: (state by owner/renter) collapsed all 5 states to the census division level

- **Dimension 5**: (state by vehicle categories) collapsed Massachusetts with New Hampshire and Connecticut with Maine for the 0 vehicle category

- The process required 15 rake-trim cycles

- A total of 3.8% of the records were trimmed at least once on the high side, and 1.9% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

Mid-Atlantic Census Division Remainder

- Obtained full convergence to the control totals (with 73 being the largest observed difference between weight sum and control total)

- All dimensions retained

- Limited collapsing was necessary in Dimension 3 only:

  - **Dimension 3**: (state by Hispanic/NonHispanic) collapsed New Jersey with Pennsylvania

- The process required 16 rake-trim cycles

- A total of 4.48% of the records were trimmed at least once on the high side, and 2.13% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

East North Central Census Division Remainder

- Obtained full convergence to the control totals (with 89 being the largest observed difference between weight sum and control total)

- All dimensions retained

- Limited collapsing was necessary in dimension 3 only:

  - **Dimension 3**: (state by Hispanic/NonHispanic) collapsed Michigan with Ohio

- The process required 10 rake-trim cycles

- A total of 3.86% of the records were trimmed at least once on the high side, and 0.83% were trimmed at least once on the low side [remember that the 2.5% of the records

trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

West North Central Census Division Remainder

- Obtained full convergence to the control totals (with 1,258 being the largest observed difference between weight sum and control total)

- All dimensions retained

- Limited collapsing was necessary in dimensions 2 through 5:

    - **Dimension 2**: (state by Black/NonBlack) collapsed Missouri, Minnesota, North Dakota, and Kansas

    - **Dimension 3**: (state by Hispanic/NonHispanic) collapsed Missouri, Minnesota, North Dakota, and Kansas

    - **Dimension 4**: (state by owner/renter) collapsed Missouri, Minnesota, North Dakota, and Kansas

    - **Dimension 5**: (state by vehicle categories) collapsed North Dakota with Kansas and Missouri with Minnesota

- The process required 8 rake-trim cycles

- A total of 3.69% of the records were trimmed at least once on the high side, and 1.92% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

South Atlantic Census Division Remainder

- Obtained full convergence to the control totals (with 210 being the largest observed difference between weight sum and control total)

- All dimensions retained

- Collapsing was necessary

    - **Dimension 1**: (CBSA status by state) collapsed CBSA 1 and 3 in West Virginia

    - **Dimension 2**: (state by Black/NonBlack) collapsed Delaware with West Virginia

- **Dimension 3**: (state by Hispanic/NonHispanic) we had to do full collapsing (collapsed Maryland with DC with West Virginia with Delaware).

- **Dimension 5**: (state by vehicle categories) collapsed DC with Maryland, and West Virginia with Delaware for the 0 vehicle category

■ The process required 32 rake-trim cycles

■ A total of 5.0% of the records were trimmed at least once on the high side, and 1.6% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

East South Central Census Division Remainder

■ Obtained full convergence to the control totals (with 803 being the largest observed difference between weight sum and control total)

■ Unable to retain Hispanics versus nonHispanics

■ Collapsing was necessary

- **Dimension 2**: (state by Black/NonBlack) we had to do full collapsing (collapsed Alabama with Kentucky with Mississippi)

- **Dimension 4**: (state by owner/renter) we had to do full collapsing (collapsed Alabama with Kentucky with Mississippi)

- **Dimension 5**: (state by vehicle categories) we had to do full collapsing of states (collapsed Alabama with Kentucky with Mississippi) and collapsed vehicle categories (1 vehicle with 2 vehicles with 3 or more vehicles)

■ The process required 7 rake-trim cycles

■ A total of 3.9% of the records were trimmed at least once on the high side, and 0.7% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

West South Central Census Division Remainder

■ Obtained full convergence to the control totals (with 49 being the largest observed difference between weight sum and control total)

- ■ All dimensions retained

- ■ Collapsing was necessary

  - – **Dimension 1**: (CBSA status by state) collapsed CBSA 2, 3, and 4 in Arkansas

  - – **Dimension 2**: (state by Black/NonBlack) collapsed Arkansas with Oklahoma

  - – **Dimension 3**: (state by Hispanic/NonHispanic) we had to do full collapsing (collapsed Arkansas with Louisiana with Oklahoma)

  - – **Dimension 4**: (state by owner/renter) collapsed Arkansas with Oklahoma

  - – **Dimension 5**: (state by vehicle categories) we had to do full collapsing of states (collapsed Arkansas with Louisiana with Oklahoma) and collapsed vehicle categories (1 vehicle with 2 vehicles with 3 or more vehicles)

- ■ The process required 12 rake-trim cycles

- ■ A total of 4.3% of the records were trimmed at least once on the high side, and 1.5% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

Mountain Census Division Remainder

- ■ Obtained full convergence to the control totals (with 1,283 being the largest observed difference between weight sum and control total)

- ■ All dimensions retained

- ■ Limited collapsing was necessary in dimensions 2 through 5:

  - – **Dimension 2**: (state by Black/NonBlack) collapsed Nevada, Idaho, New Mexico, Colorado, Utah, Wyoming and Montana

  - – **Dimension 3**: (state by Hispanic/NonHispanic) collapsed Nevada, Idaho, New Mexico, Colorado, Utah, Wyoming and Montana

  - – **Dimension 4**: (state by owner/renter) collapsed Nevada, Idaho, New Mexico, Colorado, Utah, Wyoming and Montana

  - – **Dimension 5**: (state by vehicle categories) collapsed Nevada, Idaho, New Mexico, Colorado, Utah, Wyoming and Montana

- ■ The process required 6 rake-trim cycles

- A total of 3.62% of the records were trimmed at least once on the high side, and 0.59% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

Pacific Census Division Remainder

- Obtained full convergence to the control totals (with 4,408 being the largest observed difference between weight sum and control total)

- All dimensions retained

- Collapsing was necessary:

  - **Dimension 2**: (state by Black/NonBlack) collapsed to a single cell

  - **Dimension 3**: (state by Hispanic/NonHispanic) had to collapse all 4 states to preserve Hispanic versus NonHispanic

  - **Dimension 5**: (state by vehicle categories) collapsed Alaska and Hawaii, Washington/Oregon for the 0 vehicle category. The 1, 2 and 3 plus vehicles categories for the four states were collapsed into a single cell.

- The process required 7 rake-trim cycles

- A total of 3.46% of the records were trimmed at least once on the high side, and 1.1% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

Appendix G.   Summarization of Person-Level Raking Steps

Arizona

- Obtained full convergence to the control totals (with 2,658 being the largest observed difference between weight sum and control total)

- All dimensions retained

- Collapsing was necessary

    - **Dimension 3:** (stratum by Black/NonBlack) we had to do full collapsing (collapsed Arizona Add-on areas Phoenix/Tucson with Arizona Remainder)

    - **Dimension 4:** (stratum by Sex and Age group) collapsed Arizona Add-on areas Phoenix/Tucson with Arizona Remainder for 5-17 year old Males, 5-17 year old Females, 18-24 Males, and 18-24 Females

    - **Dimension 5:** (Months) collapsed months into month pairs (April 2008-May 2008, June-July, August-September, October-November, December 2008-January 2009, February 2009-March 2009)

- The process required 9 rake-trim cycles

- A total of 3.3% of the records were trimmed at least once on the high side, and 2.6% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

California

- Obtained full convergence to the control totals (with 253 being the largest observed difference between weight sum and control total)

- All dimensions retained

- Collapsed according to the weighting plan

    - **Dimension 5**: (Months) collapsed months into month pairs (April 2008-May 2008, June-July, August-September, October-November, December 2008-January 2009, February 2009-March 2009)

- The process required 6 rake-trim cycles

- A total of 2.91% of the records were trimmed at least once on the high side, and 0.99% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

Florida

- Obtained full convergence to the control totals (with 681 being the largest observed difference between weight sum and control total)

- All dimensions retained

- Collapsing was unnecessary

- The process required 30 rake-trim cycles

- A total of 3.8% of the records were trimmed at least once on the high side, and 1.9% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

Georgia

- Obtained full convergence to the control totals (with 1215 being the largest observed difference between weight sum and control total)

- All dimensions retained

- Collapsing was necessary

  - **Dimension 3**: (stratum by Hispanic/NonNispanic) collapsed Aiken county, SC with Lee and Russell counties, AL and Strata 03, 04 within Hispanic and NonHispanic

  - **Dimension 4**: (stratum by Sex and Age group) collapsed Aiken county, SC with Lee and Russell counties, AL and Strata 03 within 18-24 Males and 18-24 Females

- The process required 20 rake-trim cycles

- A total of 3.3% of the records were trimmed at least once on the high side, and 2.6% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

Iowa

- Obtained full convergence to the control totals (with 95 being the largest observed difference between weight sum and control total)

- All dimensions retained

- Collapsing was necessary

  - **Dimension 2**: (stratum by Black/NonBlack) collapsed Linn County, Pottawattamie County and remainder of Iowa and Black/NonBlack

  - **Dimension 3**: (stratum by Hispanic/NonHispanic) collapsed Linn County, Pottawattamie County and Remainder of Iowa

  - **Dimension 4**: (stratum by Sex and Age group) collapsed Linn County, Pottawattamie County and Remainder of Iowa for 5-17 year old Males and 5-17 year old Females. Collapse Linn County, Pottawattamie County, remainder of Iowa, 18-24 years old, 25-44 years old, 44-64 years old for Males and Females

- Collapsed according to the weighting plan

  - **Dimension 5**: (Months) collapsed months into month pairs (April 2008-May 2008, June-July, August-September, October-November, December 2008-January 2009, February 2009-March 2009)

- The process required 7rake-trim cycles

- A total of 2.77% of the records were trimmed at least once on the high side, and 2.59% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

Indiana

- Obtained full convergence to the control totals (with 4,326 being the largest observed difference between weight sum and control total)

- All dimensions retained

- Collapsing was necessary

- **Dimension 2**: (stratum by Black/NonBlack) we had to collapse Marion County and Suburban Counties. Also collapse Medium Metropolitan, Small Metropolitan, Small Cities, and Rural

- **Dimension 3**: (stratum by Hispanic/NonHispanic) we had to collapse Marion County and Suburban Counties. Also collapse Medium Metropolitan, Small Metropolitan, Small Cities, and Rural

- **Dimension 4**: (stratum by Sex and Age group) collapsed Northwest Indiana, Marion County and Suburban Counties for 18-24 year old Males, and 18-24 year old Females. Also collapse Small Metropolitan, Small Cities, and Rural for 18-24 year old Males, and 18-24 year old Females.

- Collapsed according to the weighting plan

  - **Dimension 5**: (Months) collapsed months into month pairs (April 2008-May 2008, June-July, August-September, October-November, December 2008-January 2009, February 2009-March 2009)

- The process required 14 rake-trim cycles

- A total of 3.32% of the records were trimmed at least once on the high side, and 1.89% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

North Carolina

- Obtained full convergence to the control totals (with 7,531 being the largest observed difference between weight sum and control total)

- All dimensions retained

- Collapsing was necessary:

  - **Dimension 4**: (stratum by Sex and Age group) collapsed North Carolina Metrolina stratum and North Carolina Urban with household survey stratum for 18-24 Males, and 18-24 Females

- Collapsed according to the weighting plan

  - **Dimension 5**: (Months) collapsed months into month pairs (April 2008-May 2008, June-July, August-September, October-November, December 2008-January 2009, February 2009-March 2009)

- The process required 14 rake-trim cycles

- A total of 3.22% of the records were trimmed at least once on the high side, and 2.29% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

Nebraska

- Obtained full convergence to the control totals (with 3,867 being the largest observed difference between weight sum and control total)

- All dimensions retained

- Collapsing was necessary

  - **Dimension 2**: (stratum by Black/NonBlack) Omaha was collapsed with Nebraska Remainder within Black, NonBlack.

  - **Dimension 3**: (stratum by Hispanic/NonHispanic) Omaha was collapsed with Nebraska Remainder within Hispanic/NonHispanic.

  - **Dimension 4**: (stratum by Sex and Age group) Omaha was collapsed with Nebraska Remainder for 18-24 Males and 18-24 Females.

- The process required 24 rake-trim cycles

- A total of 3.7% of the records were trimmed at least once on the high side, and 1.3% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

New York

- Obtained full convergence to the control totals (with 1,271 being the largest observed difference between weight sum and control total)

- All dimensions retained

- Collapsing was necessary:

  - **Dimension 2**: (stratum by Black/NonBlack) collapsed the Capital District with Saratoga County with Glens Falls with Utica-Rome with Syracuse, Ithaca with Elmira with Binghamton, and Putnam with Rockland

- **Dimension 3**: (stratum by Hispanic/NonHispanic) collapsed the Capital District with Saratoga County with Glens Falls, Rochester with Buffalo, and Ithaca with Elmira with Binghamton

- **Dimension 4**: (stratum by Sex and Age group) collapsed Ithaca with Elmira with Binghamton, and Bronx with Manhattan for 18-24 Males, and 18-24 Females

- The process required 38 rake-trim cycles

- A total of 3.57% of the records were trimmed at least once on the high side, and 1.60% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

South Carolina

- Obtained full convergence to the control totals (with 924 being the largest observed difference between weight sum and control total)

- All dimensions retained

- Collapsing was necessary

  - **Dimension 3**: (stratum by Hispanic/NonHispanic) collapsed Beaufort and Spartanburg Counties.

  - **Dimension 4**: (stratum by Sex and Age group) collapsed Beaufort and Spartanburg Counties for 18-24 years old Males, and 18-24 years old Females

- Collapsed according to the weighting plan

  - **Dimension 5**: (Months) collapsed months into month pairs (April 2008-May 2008, June-July, August-September, October-November, December 2008-January 2009, February 2009-March 2009)

- The process required 19 rake-trim cycles

- A total of 3.43% of the records were trimmed at least once on the high side, and 2.69% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

South Dakota

- Obtained full convergence to the control totals (with 896 being the largest observed difference between weight sum and control total)

- All dimensions retained

- Collapsing was necessary:

  - **Dimension 2**: (stratum by Black/NonBlack) we had to do full collapsing (collapsed Rapid City with Sioux Falls with the Remainder of SD)

  - **Dimension 3** (stratum by Hispanic/NonHispanic) collapsed Rapid City with the Remainder of SD

  - **Dimension 4**: (stratum by Sex and Age group) collapsed Rapid City with the Remainder of SD for 18-24 Males, and 18-24 Females

- Collapsed according to the weighting plan

  - **Dimension 5**: (Months) collapsed months into month pairs (April 2008-May 2008, June-July, August-September, October-November, December 2008-January 2009, February 2009-March 2009)

- The process required 30 rake-trim cycles

- A total of 4.07% of the records were trimmed at least once on the high side, and 2.39% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

Tennessee

- Obtained full convergence to the control totals (with 377 being the largest observed difference between weight sum and control total)

- All dimensions retained

- No collapsing was necessary, but Dimension 5 (Months) was collapsed to month pairs as according to the weighting plan (April 2008-May 2008, June-July, August-September, October-November, December 2008-January 2009, February 2009-March 2009)

- The process required 7 rake-trim cycles

- A total of 3.20% of the records were trimmed at least once on the high side, and 2.33% were trimmed at least once on the low side [remember that the 2.5% of the records

trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

Texas

- Obtained full convergence to the control totals (with 1,389 being the largest observed difference between weight sum and control total)

- All dimensions retained

- No collapsing was necessary

- The process required 6 rake-trim cycles

- A total of 2.99% of the records were trimmed at least once on the high side, and 2.50% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

Virginia

- Obtained full convergence to the control totals (with 4,730 being the largest observed difference between weight sum and control total)

- All dimensions retained

- Collapsing was necessary

  - **Dimension 4**: (stratum by Sex and Age group) collapsed Virginia Add-on area with Virginia Remainder for 18-24 Males, and 18-24 Females

- The process required 9 rake-trim cycles

- A total of 3.2% of the records were trimmed at least once on the high side, and 0.81% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

Vermont

- Obtained full convergence to the control totals (with 1,690 being the largest observed difference between weight sum and control total)

- Unable to retain Blacks versus nonBlacks dimension

- Collapsing was necessary:

  - **Dimension 2**: (stratum by Black/NonBlack) we had to do full collapsing (collapsing Chittenden County and Remainder Vermont)

  - **Dimension 3**: (stratum by Hispanic/NonHispanic) collapsed Chittenden County, Hispanic with Remainder Vermont , Hispanic

- Collapsed according to the weighting plan

  - **Dimension 5**: (Months) collapsed months into month pairs (April 2008-May 2008, June-July, August-September, October-November, December 2008-January 2009, February 2009-March 2009)

- The process required 6 rake-trim cycles

- A total of 2.51% of the records were trimmed at least once on the high side, and 0.79% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

Wisconsin

- Obtained full convergence to the control totals (with 1,520 being the largest observed difference between weight sum and control total)

- All dimensions retained

- Collapsing was unnecessary

- The process required 13 rake-trim cycles

- A total of 3.37% of the records were trimmed at least once on the high side, and 0.68% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

New England Census Division Remainder

- Obtained full convergence to the control totals (with 4,953 being the largest observed difference between weight sum and control total)

- All dimensions retained

- Collapsing was necessary

  – **Dimension 1**: (CBSA status by state) collapsed CBSA 3 and 4 in Massachusetts

  – **Dimension 2**: (state by Black/NonBlack) collapsed Massachusetts, New Hampshire, with Rhode Island and collapsed Connecticut with Maine

  – **Dimension 3**: (state by Hispanic/NonHispanic) collapsed Massachusetts, New Hampshire, with Rhode Island and collapsed Connecticut with Maine

  – **Dimension 4**: (state by Sex and Age group) collapsed Massachusetts, New Hampshire, with Rhode Island and collapsed Connecticut with Maine for 18-24 Males, and 18-24 Females

- Collapsed according to the weighting plan

  – **Dimension 5**: (Months) collapsed months into month pairs (April 2008-May 2008, June-July, August-September, October-November, December 2008-January 2009, February 2009-March 2009)

- The process required 19 rake-trim cycles

- A total of 3.67% of the records were trimmed at least once on the high side, and 2.43% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

Mid-Atlantic Census Division Remainder

- Obtained full convergence to the control totals (with 2,528 being the largest observed difference between weight sum and control total)

- All dimensions retained

- Collapsed according to the weighting plan

  – **Dimension 5**: (Months) collapsed months into month pairs (April 2008-May 2008, June-July, August-September, October-November, December 2008-January 2009, February 2009-March 2009)

- The process required 6 rake-trim cycles

- A total of 3.31% of the records were trimmed at least once on the high side, and 2.38% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

East North Central Census Division Remainder

- Obtained full convergence to the control totals (with 1,148 being the largest observed difference between weight sum and control total)

- All dimensions retained

- Collapsed according to the weighting plan

  - **Dimension 5**: (Months) collapsed months into month pairs (April 2008-May 2008, June-July, August-September, October-November, December 2008-January 2009, February 2009-March 2009)

- The process required 8 rake-trim cycles

- A total of 3.35% of the records were trimmed at least once on the high side, and 2.32% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

West North Central Census Division Remainder

- Obtained full convergence to the control totals (with 9,736 being the largest observed difference between weight sum and control total)

- All dimensions retained

- Collapsing was necessary

  - **Dimension 2**: (state by Black/NonBlack) collapsed Missouri with Minnesota, collapsed North Dakota with Kansas

  - **Dimension 3**: (state by Hispanic/NonHispanic) collapsed Missouri with Minnesota, collapsed North Dakota with Kansas

  - **Dimension 4**: (state by Sex and Age group) collapsed Missouri, Minnesota, North Dakota, and Kansas for 18-24 Males, and 18-24 Females

- Collapsed according to the weighting plan

  - **Dimension 5**: (Months) collapsed months into month pairs (April 2008-May 2008, June-July, August-September, October-November, December 2008-January 2009, February 2009-March 2009)

- The process required 11 rake-trim cycles

- A total of 3.89% of the records were trimmed at least once on the high side, and 2.59% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

South Atlantic Census Division Remainder

- Obtained full convergence to the control totals (with 3,939 being the largest observed difference between weight sum and control total)

- All dimensions retained

- Collapsing was necessary

  - **Dimension 2**: (state by Black/NonBlack) collapsed Delaware with West Virginia

  - **Dimension 3**: (state by Hispanic/NonHispanic) collapsed DC with Delaware with West Virginia

  - **Dimension 4**: (state by Sex and Age group) collapsed DC with Maryland and collapsed Delaware with West Virginia for 18-24 Males, and 18-24 Females

- We collapsed according to the weighting plan

  - **Dimension 5**: (Months) collapsed months into month pairs (April 2008-May 2008, June-July, August-September, October-November, December 2008-January 2009, February 2009-March 2009)

- The process required 15 rake-trim cycles

- A total of 4.09% of the records were trimmed at least once on the high side, and 2.30% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

East South Central Census Division Remainder

- Obtained full convergence to the control totals (with 8,322 being the largest observed difference between weight sum and control total)

- All dimensions retained

- Collapsing was necessary

- **Dimension 2**: (state by Black/NonBlack) we had to do full collapsing (Alabama with Kentucky with Mississippi)

- **Dimension 3**: (state by Hispanic/NonHispanic) we had to do full collapsing (Alabama with Kentucky with Mississippi)

- **Dimension 4**: (state by Sex and Age group) collapsed Alabama with Mississippi for 18-24 Males, and collapsed Alabama with Kentucky with Mississippi for 18-24 Females

■ Collapsed according to the weighting plan

- **Dimension 5**: (Months) collapsed months into month pairs (April 2008-May 2008, June-July, August-September, October-November, December 2008-January 2009, February 2009-March 2009)

■ The process required 7 rake-trim cycles

■ A total of 3.80% of the records were trimmed at least once on the high side, and 2.65% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

West South Central Census Division Remainder

■ Obtained full convergence to the control totals (with 8,871 being the largest observed difference between weight sum and control total)

■ All dimensions retained

■ Collapsing was necessary

- **Dimension 1**: (CBSA status by state) collapsed CBSA 2 and 3 in Arkansas

- **Dimension 2**: (state by Black/NonBlack) collapsed Arkansas with Oklahoma

- **Dimension 3**: (state by Hispanic/NonHispanic) we had to do full collapsing (Arkansas with Louisiana with Oklahoma)

- **Dimension 4**: (state by Sex and Age group) collapsed Arkansas with Louisiana with Oklahoma for 18-24 Males, and 18-24 Females

■ Collapsed according to the weighting plan

- **Dimension 5**: (Months) collapsed months into month pairs (April 2008-May 2008, June-July, August-September, October-November, December 2008-January 2009, February 2009-March 2009)

- The process required 9 rake-trim cycles

- A total of 4.07% of the records were trimmed at least once on the high side, and 3.08% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

Mountain Census Division Remainder

- Obtained full convergence to the control totals (with 1,319 being the largest observed difference between weight sum and control total)

- All dimensions retained

- Collapsing was necessary

    - **Dimension 2**: (state by Black/NonBlack) collapsed Nevada, Idaho, New Mexico, Colorado, Utah, Wyoming and Montana

    - **Dimension 3**: (state by Hispanic/NonHispanic) collapsed Nevada, Idaho, New Mexico, Colorado, Utah, Wyoming and Montana

    - **Dimension 4**: (state by Sex and Age group) collapsed Nevada, Idaho, New Mexico, Colorado, Utah, Wyoming and Montana for 5-17 Males, 5-17 Females, 18-24 Males, 18-24 Females , 25-44 Males, 25-44 Females, 45-64 Males, 45-64 Females , 65 years old or older Males, and 65 years old or older Females

- Collapsed according to the weighting plan

    - **Dimension 5**: (Months) collapsed months into month pairs (April 2008-May 2008, June-July, August-September, October-November, December 2008-January 2009, February 2009-March 2009)

- The process required 7 rake-trim cycles

- A total of 3.69% of the records were trimmed at least once on the high side, and 2.06% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]

Pacific Census Division Remainder

■   Obtained full convergence to the control totals (with 6,278 being the largest observed difference between weight sum and control total)

■   All dimensions Retain

■   Collapsing was necessary:

– **Dimension 2**: (state by Black/NonBlack) had to collapse all 4 states to preserve Black versus NonBlack

– **Dimension 3**: (state by Hispanic/NonHispanic) had to collapse all 4 states to preserve Hispanic versus NonHispanic

– **Dimension 4**: (state by Sex and Age group) collapse all 4 states for Age 18-24 Male and Age 18-24 Female

■   Collapsed according to the weighting plan

– **Dimension 5**: (Months) collapsed months into month pairs (April 2008-May 2008, June-July, August-September, October-November, December 2008-January 2009, February 2009-March 2009)

■   The process required 7 rake-trim cycles

■   A total of 3.97% of the records were trimmed at least once on the high side, and 2.19% were trimmed at least once on the low side [remember that the 2.5% of the records trimmed is per rake-trim cycle not in total and each new cycle may trim a few previously untrimmed weights, which is ok]