Final Codebook (Public Use) and Technical Documentation

Contract Number: DTTS59-99-D-00489
Deliverable Number: 0002AM

for

Survey Administration for the
Summer 2000 Bureau of Transportation Statistics
Omnibus Survey

to the

Bureau of Transportation Statistics

August 10, 2000

Battelle
. . . Putting Technology To Work

TABLE OF CONTENTS

LIST OF TABLES

1. INTRODUCTION AND BACKGROUND

The Bureau of Transportation Statistics (BTS) has a requirement to conduct a national survey about satisfaction with transportation across all transportation modes. The information derived from this survey will be used as a primary source of data on satisfaction with travel and transportation, with a particular emphasis on highway-related travel. This survey also will serve as an information source for the modal administrators, both to support congressional requests, and to provide performance indicators for internal use by the U.S. Department of Transportation (DOT).

This study collected data for the Bureau of Transportation Statistics (BTS) Omnibus Survey during the summer of 2000. Data were collected from households in the U.S. using a random-digit-dialed telephone survey. The final completed sample size is 2,030 cases, and the total number of variables in the data set is 207. The data were collected by Battelle with assistance from Mathematica Policy Research, Inc. (MPR), under contract with the BTS.

This codebook provides technical documentation for this BTS Omnibus survey. Its primary goal is to document background information, sampling procedures, data collection, data elements and survey variables, response rates, and final weights. It also provides guidance on the selection of Form A and Form B versions of the questionnaire and the appropriate use of weights (also referred to as Survey A and Survey B).

This codebook contains the following information:

- Background on the survey initiative;
- Overview of how sample members were selected for the survey;
- Information regarding the data collection period and the number of completed interviews;
- Information on the number of cases in the file and guidance on the use of weights for analyses;
- Information about the number of surveys received, and the decision rules used to remove records from the survey file;
- An annotated questionnaire that provides the names of survey variables, their respective values, and their codes;
- Index of data elements by position in the data set and alphabetically;
- A list of variables in the data file generated from the SAS data set;
- The sampling plan used for the survey effort;
- Interviewer training materials.

The codebook includes eight appendices, as follows:

1. **Appendix A: Annotated Questionnaire**. This includes detailed information on all screening questions, the questions for which data have been collected from respondents using Form A and Form B, demographic questions asked of all respondents, and questions asked of respondents who received answering machine messages from the interviewer.

2. **Appendix B: Index of Data Elements by Position Order.** This is output from the SAS Contents Procedure (SAS® Proc Contents) that shows 207 variables ordered by position.

3. **Appendix C: Index of Data Elements by Alphabetic Order.** This is output from the SAS Contents Procedure (SAS® Proc Contents) that shows 207 variables ordered alphabetically.

4. **Appendix D: Documentation of Data Elements.** This is output from the SAS Frequencies Procedure (SAS® Proc Freq) that shows the marginal frequency distribution (counts and percentages for categorical variables) and from the SAS Univariate Procedure (SAS® Proc Univariate) that shows selected descriptive statistics (ranges or measures of central tendency and variability and quartiles for continuous variables) for all 207 variables in the survey data file, including weighting variables, variables that reflect characteristics of the respondents' telephone exchange area (as derived from the U.S. Census by GENESYS), additional SAS file variables, and other survey control variables, such as a flag for Form A and Form B (SUR_FORM).

5. **Appendix E: SAS Format Library Program for Survey Data.** This appendix provides values for each of the questions in the codebook, along with the appropriate labels for the response categories.

6. **Appendix F: Final Sampling Plan.** This plan discusses procedures for selecting the sample and creating the sample weights and adjustments for non-response and undercoverage. This plan was prepared prior to data collection. The final sampling and weighting procedures are fully documented in this codebook.

7. **Appendix G: Telephone Interviewer Training Manual.** This is a copy of the final manual that was used by MPR to train their CATI interviewers for this survey. The manual covers everything from the purpose of the survey to how to conduct the interview and deal with difficult interview situations. Answers to commonly asked questions and objections are included.

8. **Appendix H: Final Survey Questionnaire.** This is a hard copy of the final survey questionnaire that was used to collect the data by CATI telephone interviews. Note that question D9 (Zip Code) is not reflected in the data set in conformance with Privacy Act provisions.

A few brief guidelines to users of this codebook follow:

- Due to the large number of questions that were included in the survey questionnaire and the desire to keep the interview time below 20 minutes for the respondents, the respondents were randomly assigned either Form A or Form B. All Form A questions are designated with the letter “A” in the variable name; likewise, all Form B questions are designated with the letter “B” in the variable name. While all screening questions are designated with the letter "S" in the variable name, all core demographic questions are designated with the letter “D”. They are included in both Form A and Form B questionnaires.
- This public use data set contains 207 variables and 2,030 observations. There are 1,015 Form A observations and 1,015 Form B observations. The user will note that question D9 (Appendix H) asked for Zip Code. This variable had been removed from the final public use data set in conformance with the Privacy Act provisions.
- In Appendix D, Documentation of Data Elements, the specification of Missing Values reflects the aggregate effect of both the sample split across Form A and Form B, and the effect of the skip patterns. For example, variable A1 shows 1,067 missing values. This is composed of 1,015 respondents who responded to Form B, 50 who skipped past A1 because they answered “NO” in S7 to “a”, “b”, and “c”, and 2 respondents who answered "DON'T KNOW" to S7. Likewise, in question A2 there are 1,136 missing values. Again, as for all Form A questions, 1,015 of these missing are Form B respondents, 52 skipped from S7, and the remaining 69 answered other than “YES” to A1 and therefore skipped to A17. There are no other forms of missing values.
- Appendix F presents the final sampling plan, prior to any data collection. The actual procedures that were followed are fully documented in the technical documentation in this codebook. These procedures include the development of the sampling and analysis weights, and their derivation is described in detail. The first set of weights (defined on the last page of Appendix D and discussed in the codebook) are the sample weights, which adjust for the probability of selection of the phone number of the respondent. The second set of weights adjust for non-response and do not include post-stratification adjustments. The third set of weights incorporates the above adjustments, as well as the post-stratification adjustments. This is the recommended set of weights to use for most analyses with WEIGHT_A for Form A, WEIGHT_B for Form B, and WEIGHT for the questions administered to all respondents. Note that the sample weights and the non-response weights are provided for users who would like to generate their own post-stratification procedures and weights.
- Appendix D shows the marginal frequency distribution on responses to the questionnaire, coupled with variables that have been created in SAS and variables from the GENESYS file. These latter variables are contextual variables that describe characteristics of the telephone exchange areas in which a respondent is located. The sole purpose of presenting the data in this appendix in this way is to provide the analyst with a verification of all the data in the SAS data set.

2. SURVEY METHODOLOGY

This section presents the sampling design and overall methodology for the survey. The sampling design is probability based so that study results can be used to make inferences about adults in the U.S. household population. Steps involved in sample design and implementation include: (1) definition of the target population, (2) construction of the sampling frame, (3) specification of sample selection procedures, (4) evaluation of the precision of estimates, and (5) creation of sampling weights and adjustment for nonresponse and undercoverage.

2.1 THE TARGET POPULATION

The survey's *target population* is the entire set of population units about which the survey data are to be used to make inferences (Cox & Cohen, 1985)[1]. For this survey, the target population was all adults eighteen years of age or older in the fifty U.S. states and the District of Columbia. Further, the target population was constrained to adults in the civilian noninstitutionalized population.

2.2 THE SAMPLING FRAME

A survey's *sampling frame* is the list or mechanism used to enumerate these population units for sample selection purposes. The survey's sampling frame was derived from a list-assisted, random-digit-dialed (RDD) telephone sample approach. Of course, telephone frames exclude those households without telephones, but this source of undercoverage has been steadily declining over time. In 1963, only eighty percent of American households had telephones; by 1988 about ninety-three percent of all households had telephone service (Thornberry & Massey, 1988)[2]. The 1998 Current Population Survey, March Supplement, measured household telephone coverage at ninety-four percent.

This list-assisted RDD sampling frame provides an innovative solution to the operational problems commonly encountered in the more traditional Mitofsky-Waksberg telephone sampling approach (Waksberg, 1978)[3]. Commercial vendors construct these list-based RDD sampling frames by first obtaining a list of all working area code/exchange combinations allocated for residential service (Kulp, 1994)[4]. Adding all combinations of digits from 00 to 99 to these six-digit area code/exchange combinations creates all residential-service hundred-number banks. (These banks are called *hundred-number banks* because they represent the first eight digits of the ten-digit phone number and hence can be linked to one hundred unique potential phone numbers.) In the "list-assisted" step of frame building, all possible hundred-number banks are

---

[1] Cox, Brenda G., and Steven B. Cohen (1985). Methodological Issues for Health Care Surveys, New York: Marcel Dekker Inc.
[2] Thornberry, Owen T., Jr., and Massey, James T. (1988). "Trends in United States Telephone Coverage Across Time and Subgroups," in R. M. Groves, P.. P. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nicholls, and J. Waksberg (eds.), Telephone Survey Methodology, New York: John Wiley & Sons, pp. 25-50.
[3] Waksberg, J. (1978). "Sampling Methods for Random Digit Dialing," *Journal of the American Statistical Association*, 73, 40-46.
[4] Kulp, Dale W. (1994). "Dynamics of 'List Assisted' Random Digit Dialing (RDD) Frame Coverage," *Proceedings of the American Statistical Association, Survey Research Methods Section*.

compared to a frame of listed telephone numbers, and the number of residential telephone listings associated with each hundred-number bank is recorded. Finally, geographic coordinates are used to associate location (such as county) and demographic characteristics (such as percent minority) to each hundred-number bank.

We included in the sampling frame all hundred-number banks that contained at least one listed residential telephone number. Hundred-number banks that had zero residential listings were excluded. This exclusion substantially reduced the incidence of nonworking numbers in the sampling frame, thereby increasing the efficiency of the RDD sampling process. Although some residential telephone numbers (for example, a few unlisted residential numbers) may be deleted by excluding hundred-number banks with no residential listings, studies have shown that excluding the zero listed hundred-number banks results in minimal undercoverage bias because few unlisted residential numbers are likely to arise in such banks (Brick, et al., 1995)[5].

2.3 SAMPLE SELECTION

For this survey, sample selection procedures were developed and used in association with the truncated, list-based frame maintained by Genesys Sampling Systems[6]. The sample was selected systematically after sorting the frame by the nine Census divisions (New England, Middle Atlantic, East North Central, West North Central, South Atlantic, East South Central, West South Central, Mountain, and Pacific, which is divided between Alaska and Hawaii and all others) and by urban versus rural counties. Because the counties were divided by division and metropolitan status and carefully ordered, the systematic selection resulted in a sample that has the equivalent of 20 implicit strata. The underlying sampling frame structure can be conceptualized as a hierarchy. Within each Census division, urban counties were ordered from largest to smallest metropolitan area. Within each metropolitan area, exchanges were ordered by those serving the county containing the central city, followed by those serving the remaining non-central city counties. Within each division, rural counties were geographically ordered in serpentine fashion from north to south and from east to west. This implicit stratification imposed geographic representation and reduced the expected sampling variation for survey characteristics correlated with geography. The sample was created in-house on May 15, 2000.

A number of assumptions were made to determine the initial sample size. However, some of the assumptions may not be entirely accurate given the limited time frame for this project – a four-week data collection period. We present below the working residential hit rates and cooperation rates we encountered in this RDD telephone survey. Another unknown factor was the effect of screening half the sample using ID Plus, which could lead to a higher percent of working residential numbers. Therefore, we included a process that allowed us to test these assumptions and adjust the total sample size accordingly. First, an initial sample was selected based upon optimistic assumptions about response and eligibility rates. After Genesys prescreening excluded nonresidential, nonworking numbers, 9,089 potentially residential numbers remained. For methodological purposes, we divided 9,089 numbers into three approximately equal sized waves: Wave 1 contained 3,053 numbers, Wave 2 contained 2,984 numbers, and Wave 3 contained 3,052 numbers. Data collected from the first wave was used to refine our estimate for the

[5] Brick, J. Michael, Joseph Waksberg, Dale Kulp, and Amy Starer (1995). "Bias in List-Assisted Telephone Samples". *Public Opinion Quarterly*. Vol 59: 218-235.
[6] Zero-listed banks are not included in the truncated frame.

response rate and other required assumptions. We then determined how much additional sample was necessary to reach the desired number of completed interviews. We had anticipated the possibility of adding a fourth wave to make up for any short fall associated with less than optimistic response or eligibility rates, but found a fourth wave unnecessary.

This survey required 2,000 completed interviews: 1,000 interviews with Survey A and 1,000 interviews with Survey B. To achieve this result, we screened 9,089 telephone numbers to determine if the number was a working residential number. We were able to determine the residential status for approximately 69 percent of these numbers or 6,303 numbers. Roughly 61 percent of these 6,303 numbers were identified as residential numbers for a total of 3,814 identified residential numbers. Having identified a number as residential, we then "rostered" adult household members. Approximately 64 percent of the identified residential numbers (2,436 households) provided the roster information. Of the 2,436 households completing the roster more than 99 percent had an eligible adult. Having completed the roster, we then randomly selected an adult from the list and randomly assigned them to Survey A or Survey B. From the 2,429 eligible sampled adults, 84 percent cooperated with the interview to yield 2,030 completed interviews, that is, 1,015 completed interviews for Survey A and 1,015 completed interviews for Survey B.

2.4 SURVEY WEIGHTS

The sampling frame for this survey was derived from Genesys' list-assisted, random-digit-dialed (RDD) telephone sample approach. The sample was systematically selected after sorting the frame by the nine Census divisions (plus Alaska and Hawaii) and by urban versus rural counties. The sample of 12,008 selected numbers was selected in four equal-size replicates of 3,002 telephone numbers. Each replicate received a unique combination of questionnaire form and pre-screening procedure.

Two questionnaire forms (Survey A and Survey B) were used in data collection, with each form containing a common core of questions as well as questions unique to that form. They may be distinguished using the variable SUR_FORM. Two replicates were designated for administration of Survey A and two for Survey B. In addition to two questionnaire forms, two different Genesys pre-screening methods were used to identify business and nonworking numbers: ID and IDplus. One replicate from each survey's pair of replicates was randomly assigned to receive each pre-screening method. The assignment was as follows:

Replicate 1:	Survey A and ID pre-screening
Replicate 2:	Survey A and IDplus pre-screening
Replicate 3:	Survey B and ID pre-screening
Replicate 4:	Survey B and IDplus pre-screening

Sampled telephone numbers were ineligible if they were not associated with a household or if the household contained no adults age eighteen years or older. Households were asked to provide a roster of adult members, which was used to select one adult for interview.

Two types of weights are discussed in the next sections: analysis weights and sampling weights. The analysis weight reflects all nonresponse, post-stratification, and within household selection adjustments that have been made, and is the weight that should be used for the analysis of the data. The sampling weight reflects only the probability of selection; it is the inverse of the probability of selection.

2.4.1 Analysis Weights

Analysis weights were developed for each replicate. These replicate analysis weights were combined to create:

- analysis weights for the combined sample (by dividing the replicate weights by 4), and

- analysis weights for separate analysis of the A Survey and the B Survey samples (by dividing the appropriate replicate weights by two).

Creation of these replicate analysis weights involved the following steps:

- calculation of the sampling weight for the telephone number,
- recognition of the results of pre-screening,
- adjustment to account for loss of information on residential status,
- adjustment to account for inability to collect roster information,
- calculation of a sampling weight for subsampled adults,
- adjustment for nonresponding adults, and
- standardization to the civilian, noninstitutionalized population.

The remainder of this section discusses these weighting steps.

2.4.2 Calculation of the Sampling Weight

The first step in weighting the sample was to calculate the sampling weight for each sampled telephone number in each replicate. The sampling weight is the inverse of the telephone number's probability of selection for the replicate. The sample design used for the BTS Omnibus Survey was a replicated systematic sample. The four replicates were randomly assigned to the two questionnaires and to the two pre-screening procedures. The sampling weight $W_s\,(ri)$ for telephone number i in replicate r was calculated as the inverse of its probability of selection for replicate r or:

$$W_s(ri) = \frac{N}{n_r}$$

where N is the total number of telephone numbers and n_r is the total number of sampled telephone numbers for replicate r. For this survey, the total number of telephone numbers in the sampling frame was 2,412,401,000.[7] The number of sampled telephone numbers was 12,008, divided into four sample replicate of 3,002 numbers each.

2.4.3 Nonresponse Adjustments

The next step was to adjust for the various levels of nonresponse. Nonresponse leads to differing amounts of data loss. Complete response for a sampled telephone number implies that we collected the following data:

- Residential Status: data that determined whether the telephone number was associated with a residence;

- Household Eligibility: for residential numbers, data that determined whether one or more adults were members of the household;

- Household Roster: for those households containing one or more adults, data listing adult members for use in sampling one adult member for interview; and

- Questionnaire: from sampled adults from each household who completed a roster, interview data.

Nonresponse adjustments were made to account for nonresponse at each step. These adjustments were made within weighting classes formed by the cross-classification of the replicate r with a classing variable c based upon metropolitan status (rural versus urban) and geographic area (Census division for urban areas and Census region for rural areas).

2.4.4 Genesys Pre-screening

By definition, no nonresponse occurred at this stage because pre-screening was completed for all numbers. Numbers identified in pre-screening as nonworking or nonresidential were ineligible for further data collection. The remaining sampled numbers were included in the next data collection step. Of the 12,008 sampled cases, Genesys prescreening removed 2,919 nonresidential, nonworking telephone numbers, leaving 9,089 numbers to be further screened and interviewed in CATI.

2.4.5 Residential Status Determination

The first step in data collection was to identify the status of the telephone numbers remaining after Genesys pre-screening operations were complete. For this adjustment, response was considered to have been obtained for the ith number from the rth replicate [$\delta_{res}(ri)=1$] when we determined whether the number was either residential, nonresidential, or nonworking.

---

[7] The total number of telephone numbers stated here is 10 times larger than the actual number of telephone numbers. This factor of 10 has no impact on the calculation of the final weight.

Nonresponse at this stage implied that we could not determine whether the number was working or residential [$\delta_{res}(ri)$=0].

The residential-status, nonresponse adjustment adjusted the sampling weights to account for those sampled cases for which residential status could not be determined. The adjustment was done within classes formed by the cross of replicate r by a classing variable c_{res}. By definition, nonresponse could not occur for sampled numbers identified as ineligible during Genesys pre-screening. Such cases were segregated from the remaining cases in forming weighting classes for this adjustment (c_{res} = 100 – eliminated in Genesys pre-screening). The remaining telephone numbers in each replicate r were assigned to weighting classes defined by metropolitan status and geographic area.

For weighting class rc_{res}, the residential-status nonresponse adjustment factor $ADJ_{res}(rc_{res})$ was defined as:

$$ADJ_{res}(rc_{res}) = \frac{\sum_{i=1}^{n(rc_{res})} W_s(ri)}{\sum_{i=1}^{n(rc_{res})} d_{res}(ri)\ W_s(ri)}$$

where $\delta_{res}(ri)$ is equal to 1 for those cases where status was asked and determined, and 0 otherwise. Because response status is automatically known for all sampled numbers eliminated in Genesys pre-screening, the residential-status adjustment factor for these weighting classes (c_{res} = 100) is 1.

Next, the sampling weight of the ith telephone number from the rth replicate and the c_{res}-th weighting class was multiplied by the residential-status response indicator and the adjustment factor to derive the residential-status, nonresponse-adjusted weight $W_{res}(ri)$ or

$$W_{res}(ri) = d_{res}(ri)\ \mathrm{ADJ}_{res}(rc_{res})\ W_s(ri)\ .$$

Note that this approach resulted in adjusted weights of zero for all telephone numbers where residential status was unknown [$\delta_{res}(ri)$=0]. These telephone numbers were excluded from subsequent adjustments and instead had adjustment factors of 0 and weights of 0 assigned.

2.4.6 Roster Completion

Next, adjustments were made to account for nonresponse to the roster of household members from telephone numbers identified as residential. Rosters were considered to have been completed [$\delta_{ros}(ri)$=1] when we obtained a listing of adult household members or determined that the household contained no adult members. Nonresponse at this stage [$\delta_{ros}(ri)$=0] implied that a roster of adult members was not obtained from a telephone number identified as residential and that we did not know whether the household contained one or more adult members.

The roster-completion, nonresponse adjustment adjusted the residential-status, nonresponse-adjusted weights to account for data loss from identified residences. By definition, nonresponse could not occur for sampled numbers identified as ineligible during Genesys pre-screening and for sampled numbers determined to be nonresidential during interviewing. Such cases were segregated from the remaining cases in forming weighting classes for this adjustment. Telephone numbers determined to be ineligible during Genesys pre-screening (c_{ros}=100) and numbers identified as nonresidential during interviewing (c_{ros}=101) had roster-completion indicators [$\delta_{\mathrm{ros}}(\mathrm{ri})$] of 1. Those remaining telephone numbers identified as residential were assigned to weighting classes defined by metropolitan status and geographic area.

For sampled telephone numbers within a given weighting class c_{ros}, the roster-completion adjustment factor $ADJ_{ros}(rc_{ros})$ for weighting class rc_{ros} was defined as:

$$ADJ_{\mathrm{ros}}(\mathrm{rc}_{\mathrm{ros}}) = \frac{\sum_{i=1}^{n(\mathrm{rc}_{\mathrm{ros}})} \mathrm{W}_{\mathrm{res}}(\mathrm{ri})}{\sum_{i=1}^{n(\mathrm{rc}_{\mathrm{ros}})} d_{\mathrm{ros}}(\mathrm{ri})\, \mathrm{W}_{\mathrm{res}}(\mathrm{ri})}$$

where $\delta_{\mathrm{ros}}(ri)$ is equal to 1 for those cases where the roster was completed and 0 otherwise. Because complete roster information was obtained by definition from nonworking and nonresidential numbers, the roster-completion adjustment factor for these weighting classes (c_{ros} = 100, 101) is 1. Next, the residence-status, nonresponse-adjusted weight of the *i*th telephone number from the *r*th replicate and the c_{ros}-th weighting class was multiplied by the roster-completion indicator and this adjustment factor to derive the roster-completion, nonresponse-adjusted weight $W_{ros}(ri)$ or

$$W_{\mathrm{ros}}(\mathrm{ri}) = d_{\mathrm{ros}}(\mathrm{ri})\ \mathrm{ADJ}_{\mathrm{ros}}(\mathrm{rc}_{\mathrm{ros}})\ \mathrm{W}_{\mathrm{res}}(\mathrm{ri})\ .$$

Note that this approach resulted in adjusted weights of zero for all residential telephone numbers not completing rosters [$\delta_{\mathrm{ros}}(ri)=0$]. These telephone numbers were excluded from subsequent adjustments and instead had adjustment factors of 0 and weights of 0 assigned.

2.4.7 Selection of an Adult for Interview

Only adults eighteen years or older were selected for interview from each household completing the roster of adult members. Thus, the conditional probability of selection of the *j*th adult member of responding household *ri* is $1/N(ri)$, where $N(ri)$ is the number of adults rostered for household *ri*. The sampling weight for sampled adults[8] was calculated as the product of the roster-completion, nonresponse-adjusted weight times the conditional weight $CW_{ind}(rij)$ associated with within-household selection from household *ri*, or

[8] This sampling weight also includes adjustments to account for loss of roster data for nonresponding households, but does not include multiplicity adjustments needed for households with multiple phone lines.

$$W_{ind}(rij) = W_{ros}(ri)CW_{ind}(rij)$$

where

$$CW_{ind}(rij) = N(ri)$$

2.4.8 Unit NonResponse

The next step in creating analysis weights for sampled adults was to adjust for nonresponse to the questionnaire. This nonresponse adjustment was made using the same weighting classes as for the roster-completion adjustment, except that households with no adult members (c_{ind}=102) were not included in the adjustment. For weighting class c_{ind}, the questionnaire-completion adjustment factor $ADJ_{\text{QC}}(rc_{\text{ind}})$ was defined as:

$$ADJ_{\text{QC}}(\text{rc}_{\text{ind}}) = \frac{\sum_{i=1}^{n(\text{rc}_{\text{ind}})} \text{W}_{\text{ind}}(\text{rij})}{\sum_{i=1}^{n(\text{rc}_{\text{ind}})} d_{\text{QC}}(\text{rij})\ \text{W}_{\text{ind}}(\text{rij})}$$

where $d_{QC}(rij)$ is equal to 1 for those cases where the questionnaire was completed and 0 otherwise.

We then adjust the individual sampling weight by this factor to derive the nonresponse-adjusted person weight by multiplying the individual sampling weight by the questionnaire-completion indicator and the questionnaire-completion adjustment factor to derive the questionnaire-completion, nonresponse-adjusted weight $W_{\text{QC}}(rij)$ or

$$W_{\text{QC}}(\text{rij}) = d_{\text{QC}}(\text{rij})\ \text{ADJ}_{\text{QC}}(\text{rc}_{\text{ind}})\,\text{W}_{\text{ind}}(\text{rij})\ .$$

Note that this approach resulted in adjusted weights of zero for all sampled adults who did not complete the questionnaire [$\delta_{\text{QC}}(rij)$=0]. These telephone numbers were excluded from subsequent adjustments and instead had adjustment factors of 0 and weights of 0 assigned hereafter.

2.4.9 Adjustment for Multiple Phone Numbers

The next adjustment factor considered was for multiple selection probabilities that occurred for households with more than one residential telephone number. The information indicating whether the household had multiple telephone lines was located in questions D9a and D9b, which were inadvertently excluded from the original questionnaire. These questions were added to the instrument midway through data collection. However, we were able to gather this information for only one-third of the respondents. A test was conducted for the significance of the effect with those households for which this information was obtained. First, we examined the distribution of the average number of residential telephone lines of urban households within Census divisions and rural households within Census regions. We found that rural households tended to have fewer residential phone lines than urban areas. We also found that rural households in the Midwest and West had slightly more residential phone lines than did rural households in the Northeast and South. For urban households, no clear geographic pattern was found. The more important question was whether we would observe significant differences in selected key variables that could be attributed to the effect of having multiple phone lines in the household. For all households who received this question, we constructed weights with the phone line effect included and without that effect, and observed no significant differences in responses on such variables as overall satisfaction with highways, number of miles traveled, or respondent education. As a result, the presence of multiple phone lines was not considered further in the development or adjustment of final weights.

2.4.10 Standardization Adjustment

The last step was to standardize the weights for Survey A and Survey B so that they summed to national projections of the civilian, noninstitutionalized population within poststrata *p* formed the cross-classification of age group by race by sex. We also accounted for nonresponse to the classification variables. We randomly imputed for each nonrespondent proportional to the respondents' distribution. For example, if 75 percent of respondents are white, then 75 percent of nonrespondents will be assigned as white. Moreover, respondents could select more than one race, but current population totals do not account for multiple race individuals. Therefore, if the respondent selected "black" and some other race, then we regarded this respondent as black. If the respondent selected any other combination of races, then we regarded the respondent as "Other race." This adjustment was independently made for each replicate *r*.

For responding persons from poststrata *p* and replicate r, the standardization adjustment factor $ADJ_{st}(\mathrm{r}p)$ was defined as:

$$ADJ_{st}(\mathrm{rp}) = \frac{\mathrm{N(p)}}{\sum_{\mathrm{i=1}}^{\mathrm{n(rp)}} \mathrm{W}_{QC}(\mathrm{rij})}$$

where $N(p)$ is the population count for poststratum p.

We then multiplied the questionnaire-completion, nonresponse-adjusted person weight by this factor to derive the standardized person weight $W_{\mathrm{st}}(\mathrm{rij})$ or

$$W_{\mathrm{st}}(\mathrm{rij}) = \mathrm{ADJ}_{\mathrm{st}}(\mathrm{rp})\, \mathrm{W}_{QC}(\mathrm{rij})\,.$$

The standardized person weight was the final analysis weight for replicate *r*. These standardized replicate weights were used to derive the analysis weights for Survey A and Survey B and for the combined sample.

2.5 RESPONSE RATES

The procedure for response rate calculation are based on the guidelines established by the Council of American Survey Research Organizations (CASRO 1982)[9] in defining a response rate. The final response rate for the survey was obtained as the product of the residential determination completion rate, the roster completion rate, and the interview completion rate, or:

$$RR = CR_{residence} \times CR_{roster} \times CR_{\text{int}erview}$$

We calculated the *residential determination completion rate* $CR_{residence}$ as:

$$CR_{residence} = \frac{\text{Total Residentia l Status Determined}}{\text{Total Number Dialed}} = \frac{6{,}303}{9{,}089} = 0.69$$

that is, we completed the residential determination process for 69 percent of the numbers dialed.

We calculated the *roster completion rate* CR_{roster} as:

$$CR_{roster} = \frac{\text{Total Roster Completed}}{\text{Total Number Residentia l}} = \frac{2{,}436}{3{,}814} = 0.64$$

that is, we completed the roster with 64 percent of residences. We calculated the *interview completion rate* $CR_{interview}$ as:

$$CR_{\text{int}erview} = \frac{\text{Completed Intereview s}}{\text{Total Roster Completed and Adult Present}} = \frac{2{,}030}{2{,}429} = 0.84$$

that is, we completed the interview with someone at 84 percent of rostered households. The count of completed interviews includes those households with no adults, which are not eligible for the survey. Therefore, the overall response rate for the survey was about 37 percent. Table 1 presents a summary of response rates by various subgroups.

[9] CASRO (Council of American Survey Research Organizations), Report of the CASRO Completion Rates Task Force, New York, Audits and Surveys, Inc., unpublished report, 1982.

Table 1. Summary of Response Rates by Population Subgroups

Subgroup	Completion Rates			Response Rate
	Residence	Roster	Interview	
Survey Group				
Survey A (1015 completes)	69%	63%	86%	37%
Survey B (1015 completes)	70%	65%	82%	37%
Census				
New England	67%	68%	82%	37%
Mid-Atlantic	67%	59%	83%	33%
East North Central	68%	65%	84%	38%
West North Central	73%	76%	85%	47%
South Atlantic	68%	62%	83%	35%
East South Central	71%	75%	89%	48%
West South Central	73%	66%	81%	39%
Mountain	75%	67%	84%	42%
Pacific	68%	56%	82%	31%
Metropolitan Status Code				
In center city of an MSA	69%	60%	84%	35%
Outside center city inside county with center city	68%	62%	83%	35%
Inside a suburban county of the MSA	67%	64%	83%	36%
In an MSA that has no center city	67%	53%	85%	30%
Not in an MSA	75%	75%	84%	47%

3. SUMMARY OF SURVEY PROCEDURES

3.1 DATA COLLECTION SCHEDULE

This survey required that 2,000 interviews be completed in a four-week timeframe: 1,000 interviews with one version of the survey (Survey A) and 1,000 interviews with a second version of the survey (Survey B). Data collection began on May 31, 2000 at 4:00 p.m. EST and was completed on June 25, 2000, with 2,030 completed interviews (1,015 from Survey A and 1,015 from Survey B). At the start of data collection, (May 30 - 31, 2000), a total of forty-one interviewers were trained. A week later (June 7, 2000) an additional fifteen interviewers were trained for the project. Thus, a total of fifty-six interviewers were trained for the study.

3.2 INTERVIEW PROCEDURES

3.2.1 Pretest

As part of the instrument design phase, seven interviewers were trained to conduct a pretest. The pretest allowed us to test the usability and timing of the two versions of the instrument and to test the training materials. We placed 189 random digit dial calls and completed thirty cases, fifteen from each questionnaire version. The pretest determined that Survey A took twenty-one minutes to administer while Survey B took seventeen minutes. The pretest report recommended shortening the questionnaires to fifteen minutes. The pretest report also recommended wording changes to questions that had not been fielded in previous BTS studies. No changes were recommended for questions that had been used on previous BTS surveys. Next, the pretest raised a sampling issue: should students living away from home be considered members of their parents' households or be eligible for sampling from their dormitory room telephone numbers?[10] The pretest also raised the issue of how best to obtain income information. An unfolding question sequence was decided upon to minimize the number of income questions a respondent was required to answer. Finally, the pretest raised the issue of the averting breakoffs during the roster, and we made a wording change that stressed the anonymity of the study.

3.2.2 Interviewer Training

Each of the fifty-six telephone interviewers received a minimum of four hours of study-specific training that included instructions for administering Survey A and B. Recruiting and training qualified, diversified interviewers was a critical component to ensuring overall data quality and success for this project. Interviewers played a major role in encouraging the respondents to cooperate. The purposes of the training activities were to:

- Standardize the quality of the data collection techniques and procedures from the outset
- Increase the accuracy, quality, and relevance of data collected

[10] Ultimately, it was determined that students residing away from home would not be eligible under their parents telephone number but would be eligible if they had their own phone.

- Provide explicit, nonjudgmental procedures for the data collection staff to follow

During and after training, we evaluated each interviewer's performance. Each interviewer had to display proficiency with the data collection instrument and procedures. Failure to exhibit the required skills resulted in additional training. We did not assign interviewers to this project until they demonstrated their ability to perform at an acceptable level. New-to-MPR interviewers attended MPR's standard twelve-hour general interviewer training program, conducted in three four-hour sessions. General training topics include gaining cooperation, understanding and avoiding bias, using appropriate probing methods, using the CATI software, and administrative issues. A variety of media and other methods were used in training, including videotape on the role of the interviewer. Also included were role playing and written exercises.

The BTS project-specific training agenda included:

- An introduction to the study, client, and sample

- Review of both survey instruments

- Review of special skills needed to conduct RDD surveys

- Review of introduction , screener, and refusal avoidance materials

- Hands-on practice and role-plays

As a final training and quality assurance step, MPR staff monitored each interviewer before he or she was permitted to begin live interviewing. Additionally, the first two interviews of each interviewer were unobtrusively monitored by project staff and telephone center staff. Monitoring involved listening to an interview while simultaneously observing on the monitoring screen the way the interviewer recorded a respondent's answers. This procedure permitted us to find and correct deficiencies quickly. MPR continued to monitor interviewers' work throughout the field period.

3.2.3 Scheduling Calls and Tracking Cases

All survey data were collected using the computer-assisted interviewing program, CASES. In addition, the tracking software was customized to track the sample, produce daily cost and production reports, and a specialized sample report. These reports allowed the survey team to monitor survey results on a daily basis. The CATI computer-assisted scheduler controlled telephone number release for interviewing according to industry-standard scheduling algorithms. The scheduling program randomly assigned sampled telephone numbers to interviewers. Calls were scheduled based on optimal calling patterns, dispersed over different times of the day and different days of the week. Firm appointments were scheduled within a twenty-minute window; other 'soft' appointments were scheduled within a sixty-minute time period, based on information provided by the interviewer. We limited follow-up efforts to nine calls to determine whether a telephone number was residential (that is, we had never had verbal or voice mail contact with a resident), and to sixteen calls to complete an interview (when we had ever had verbal or voice mail contact with a resident).

3.2.4 Household Screening and Rostering

Once contact was made with the individuals at a dialed telephone number, interviewers screened for eligibility by verifying that the number belonged to a residence (not a business or institution) and that the residence contained at least one individual eighteen years or older. Adults (eighteen or older) in multi-person households were "rostered" and a respondent randomly selected by computer program. In one-adult households, that adult was automatically selected. After respondent selection, the interviewer attempted to conduct the interview with the selected household respondent.

3.2.5 Interviewing

No incentives were offered to respondents for completing the interview. MPR conducted the survey only in English. The average length of the completed interview (based on 2,030 completes) was fourteen minutes. If the selected household member refused the interview, the interviewer recorded the reason for refusal. Often, this information helps a refusal converter to convert the case later on.

One strategy for decreasing nonresponse in RDD surveys is to leave a message on potential respondents' machine answering devices ("MAD"s) in hopes of enhancing the level of cooperation. Within this overall project, a small experiment was conducted to systematically test the results of this procedure. Households were randomly divided into a control group, not receiving answering machine messages and an experimental group receiving answering machine messages. The experimental group was left a MAD message on the first and fourth times an answering machine was encountered. The message for the experimental group was the following:

> *Hello, I'm [INSERT INTERVIEWER FULL NAME] calling on behalf of the Department of Transportation. We are calling to invite you to participate in a brief survey to determine your satisfaction with your local community's transportation system. Could you please call our toll-free number 888-633-8349 and ask to speak with the study supervisor, Barbara Taylor. We look forward to speaking with you. Again, that toll free number is 888-633-8349. Thank you, good-bye.*

3.2.6 Refusal Conversion

Refusal cases were assigned to a subgroup of particularly skilled interviewers known as "refusal converters." Refusal converters called refusing sample members as early as one week after the initial refusal. Refusal converters used information about the reason and intensity of the prior refusals (from the earlier interviewer's comments) in planning their calls. Due to the short field period, we retired a case after one refusal conversion attempt. Cases were coded as a final refusal if a second refusal was obtained when the refusal conversion attempt was made.

3.3 QUALITY CONTROL PROCEDURES AND REPORTING

Interviewer performance was evaluated on the basis of production reports and regular on-line monitoring. Interviewer conduct during interviews was evaluated primarily by supervisory monitoring of actual calls, supplemented by review of interviewer notes maintained in the CATI system (all calls and notes recorded about those calls are maintained by the CATI system).

MPR's silent monitoring system enables supervisors to listen to interviews without either the interviewers' or respondents' knowledge; it also allows supervisors to view interviewers' screens while the interview is in progress. Interviewers are informed they will be monitored but do not know when observations will take place. During monitoring, supervisors identify behavioral problems involving incorrect study presentation, errors in reading questions, biased probes, inappropriate use of feedback in answering respondent questions, and other unacceptable behavior, such as interrupting the respondent or offering a personal opinion about specific questions or about the survey. Supervisors review monitoring results with interviewers after the interviewer completes her or his shift.

Daily production reports provided information on several performance indicators, including completed interviews, calls made, refusals, refusal conversions, time per call, time per interview, and the ratio of completed interviews to time charged to interviewing. Progress reports were available daily to MPR staff to enable supervisors and project management staff to monitor production and performance continuously. Several reports were produced, including:

- ***Status disposition reports.*** These report daily and cumulative distributions of interim and final survey status codes (completions, various nonresponse and ineligibility dispositions, and current statuses for active cases). In these reports, the total sample is broken down by replicate wave releases. We also produced reports on the completion of Survey A and Survey B.

- ***Specialized weekly reports.*** These were used to monitor the results of the two experiments described above – 1) the mechanical answering devises (MAD) messages on response rates and 2) the Genesys ID or ID Plus sample on response rates and number of calls required before determining the number to be a residence.

- ***Daily interviewer performance reports*** to monitor last-day and cumulative performance statistics, including completions, refusals, number of calls per completed interview, number of calls per refusal, time per call, and time per completed interview.

3.4 SUMMARY OF DATA CLEANING

One of the most important advantages of computer-assisted surveys is that errors can be identified and averted during the interview by building logic, range, and consistency checks into the program. MPR's CATI system permits interviewers to back up and correct erroneous answers to previous questions without violating instrument logic.

Because of differences in design, separate instrument programs were written for the different survey versions A and B. Separate cleaning programs were written for each of the two survey instruments. The instrument cleaning programs enforced questionnaire logic strictly. An interview could not be certified as "clean" until all appropriate questions had either been answered or assigned an acceptable nonresponse value, and until the data record for each interview was consistent with the instrument program logic. Survey questions were all close-ended. Thus, no questions required manual coding.

A program was written to reformat the cleaned instrument responses. Analysis files were then prepared in SAS, and additional edits performed. The additional edits included checks on the number of missing values, assignment of additional nonresponse values, and some constructed variables. Weights were applied to the data files.