INFORMATION COLLECTION
SUPPORTING JUSTIFICATION
Survey of Northeast Regional and Intercity Household Travel Attitudes and Behavior

INTRODUCTION

This is to request the Office of Management and Budget's (OMB) one year approved clearance for the information collection entitled, "Survey of Northeast Regional and Intercity Household Travel Attitudes and Behavior"

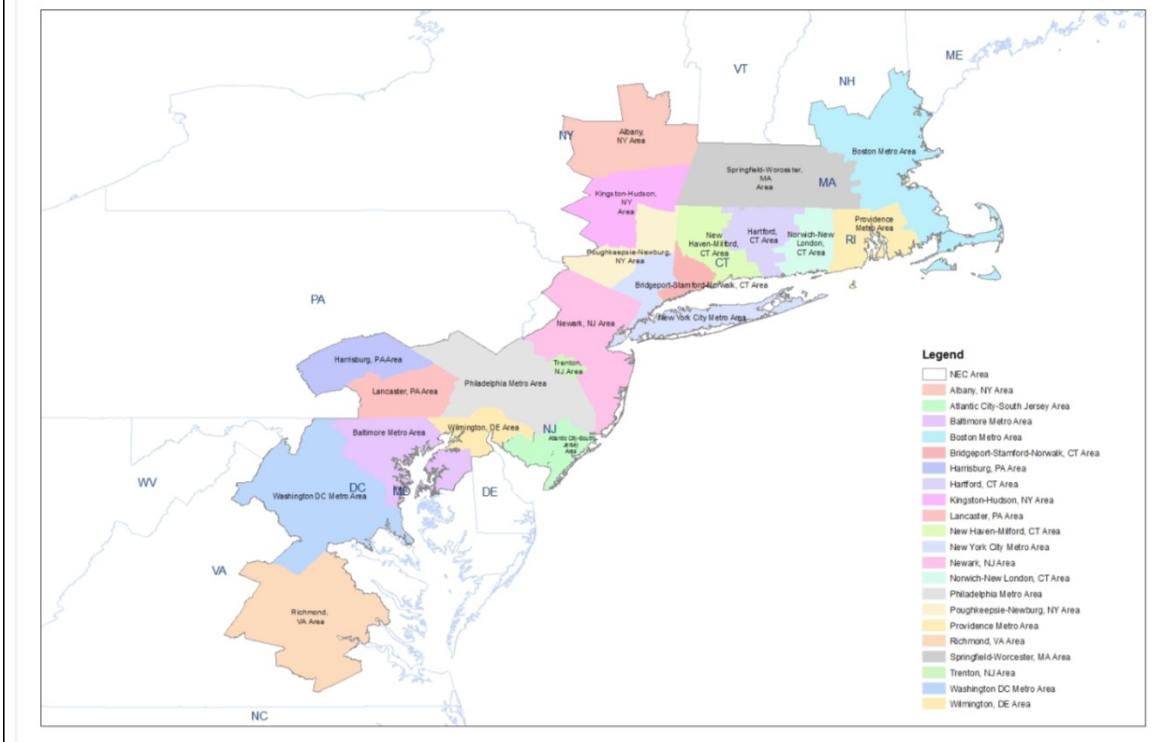## Part B. Collections of Information Employing Statistical Methods

The proposed study will employ statistical methods to analyze the information collected from respondents and draw inferences from the sample to the target population. It will use an overlapping dual frame design with respondents interviewed on landline phones and on cell phones. The following sections describe the procedures for respondent sampling and data tabulation.

1. **Describe the potential respondent universe and any sampling or other respondent selection method to be used.**

a. Respondent universe

The respondent universe consists of all persons aged 18 to 74 residing in households with a working telephone located within the Northeast Corridor of the United States. The Northeast Corridor is defined graphically in Figure B.1. Also, since interviews will be conducted in only English, any person who does not speak English will be excluded from this study. Based on the prevalence of the use of English in the United States and especially among target respondents, who are long distance travelers, excluding respondents who speak a language other than English will have minimal impact on the study.

FIGURE B.1 NORTHEAST CORRIDOR GEOGRAPHY

The current population of each of the areas shown in Figure B.1, based on 2010 US Census, is summarized Table B.1.

TABLE B.1 NORTHEAST CORRIDOR POPULATION BY AREA

NEC Survey Regions	2010 Total Population
New York City Metro Area	12,368,525
Newark, NJ Area	6,471,215
Philadelphia Metro Area	5,671,115
Washington DC Metro Area	5,582,170
Boston Metro Area	5,195,718
Baltimore Metro Area	2,730,686
Springfield-Worcester, MA Area	1,622,713
Providence Metro Area	1,600,852
Richmond, VA Area	1,258,251
Hartford, CT Area	1,212,381
New Haven-Milford, CT Area	1,052,404
Lancaster, PA Area	954,417
Albany, NY Area	920,935
Bridgeport-Stamford-Norwalk, CT Area	916,829
Wilmington, DE Area	705,670
Harrisburg, PA Area	683,043
Poughkeepsie-Newburg, NY Area	670,301
Atlantic City-South Jersey Area	528,712
Norwich-New London, CT Area	392,483
Trenton, NJ Area	366,513
Kingston-Hudson, NY Area	294,810
Total	**51,199,743**

It is not anticipated that the use of a telephone household sample will cause coverage bias. Research has shown that telephone service for many households is intermittent, and that gains or losses in service are associated with changes in their financial situation, (Keeter, 1995[1]). By implication, telephone coverage is not a static phenomenon, but rather, the telephone population at any given point includes households that were recently part of the non-telephone population. Households with intermittent telephone service are similar on a variety of important demographic characteristics to households who usually lack telephones (Keeter, 1995; Frankel et al, 2003[2]). Thus, weighting using information on interruptions in telephone service can ensure the correct coverage in the final sample of persons including those in non-telephone households.

_____

[1] Keeter, S. *Estimating non-coverage bias from a phone survey*. Public Opinion Quarterly, 1995; 59: 196-217.

[2] Frankel M, Srinath KP, Hoaglin DC, Battaglia MP, Smith PJ, Wright RA, Khare M. *Adjustments for non-telephone bias in random-digit-dialing surveys*.- Statistics in Medicine, 2003: 22; 1611-1626

b. Statistical sampling methods

A sample of 12,500 completed surveys will be obtained. The sample design will employ a partially overlapping dual frame design with probability samples drawn from independent sampling frames for landline phones and for cell phones.

The landline sample will be drawn from telephone banks randomly selected from an enumeration of the Working Residential Hundred Blocks within the active telephone exchanges. The Working Hundreds Blocks are defined as each block of 100 potential telephone numbers within an exchange that includes one or more residential listings (i.e., this will be a list-assisted sample). A two-digit number will then be randomly generated for each selected Working Residential Hundred Block to complete the phone number to be called. By randomly generating these numbers, a process known as random digit dialing (RDD), every number in the sampling frame of Hundreds Blocks has an equal probability of selection regardless of whether it is listed or unlisted. The RDD sample of telephone numbers is dialed to determine which are currently working residential household telephone numbers. The systematic dialing of those numbers to obtain a residential contact should yield a probability sample of landline telephone numbers.

The next step is selecting the household member to interview. Once a household is contacted, the interviewer will introduce him/herself and the survey, and ask how many persons residing in the household are 18 to 74 years old. If there is only one person in that age range residing in the household, then the interviewer will seek to interview that person. If there are multiple people in that age range residing in the household, the interviewer will randomly select one of those household members by asking to speak with the age-eligible household member who had the most recent birthday. Only that randomly selected person will be eligible to participate in the survey; no substitution by other household members will occur. It was decided to obtain data from only one person in each household though the intracluster (household) correlation is likely to be low. Collecting data for more than one person imposes an increased response burden. Also, the average number of persons over 18 years in a household was less than 2.59 in 2010 according the U.S. Bureau of Census. Therefore, it is unlikely that the variance of the estimates would increase substantially because of sampling one person within a household. If the selected respondent is not available, then the interviewer will arrange a callback. In cases where no one residing within the household is in the eligible age range, the interviewer will thank the individual who answered the phone, terminate the call, and the number will be removed from the sample.

Although list-assisted landline RDD sampling provides only a small coverage error for landline telephone households within landline banks, the restriction of the sampling frame to only landline banks would introduce a much more serious coverage error in general population surveys. The increasing percentage of households that have abandoned their landline telephones for cell phones has significantly eroded the population coverage provided by landline-based surveys. The key group that is missing from landline RDD samples is the cell phone-only group. But there is also potential bias in landline samples from under-coverage of young people who tend to rely on their cell phones more than their landline phones. The cell phone sample for this survey will

therefore be composed of both cell phone only and cell phone users with a landline in their household.

Due to the higher cost of cell phone interviews compared to landline interviews, dual frame surveys are not usually designed with proportional allocation of sample between the landline and cell phone strata. The most recent data published from the National Health Interview Survey[3] shows 27.1% of adults residing in cell phone only households in the Northeast region of the United States during the first half of 2013. The sample allocation will consist of a two-stratum design (landline and cell phone) with 22% of the total sample (or 2,750 interviews) being obtained from the cell phone frame.

The cell phone sample will be randomly selected from 1,000 banks used exclusively for cell phones, using RDD. We are not stratifying by area code since area code is not a great predictor of geography for cell phone sample. Procedures for sample selection will be similar to those used in selecting the landline sample, except that the sample will not be list-assisted as it is in the landline sample. In addition, the cell phone will be treated as a single user device. This means the cell phone sample will not require the procedures used with the landline sample to select from multiple eligible household members. Upon contact and introduction, the interviewer will immediately ask questions to determine that the person on the phone is not in a situation that could pose a safety risk to that individual (e.g., driving at the time of the call). If the contacted individual is found to be in a situation that could pose a risk, the interviewer will terminate the call and call back another time. If it is safe for the contacted individual to proceed with the call, then the interviewer will move on to asking screening questions regarding age eligibility.

c. Precision of sample estimates

The objective of the sampling procedures described above is to produce a random sample of the target population. This means that with a randomly drawn sample, one can make inferences about population characteristics within certain specified limits of certainty and sampling variability.

The margin of error, d, of the sample estimate of a population proportion, P, equals:

$$d = (t_\alpha) se(P)$$

Where t_α equals 1.96 for *1-α* = 0.95, and the standard error of P equals:

$$se(P) = \sqrt{deff} \sqrt{\frac{P(1-P)}{n}}$$

Where:

[3] Blumberg, Stephen J. and Julian V. Luke. *Wireless Substitution: Early Release of Estimates From the National Health Interview Survey, January-June 2013.* U.S. Department Of Health And Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics. Released 12/2013.

deff = design effect arising from the combined impact of the random selection of one eligible individual from a sample household, and unequal weights from other aspects of the sample design and weighting methodology, and

n = the size of the sample (i.e., number of interviews)

Using these formulas, the margin of error at 95% confidence level for a sample size of 12,500 interviews is d =0.01073 to 0.01239 using an average deff of 1.5 to 2.0 and setting P equal to 0.50. The need for this sample size is driven by plans to examine key geographic and trip purpose subgroups, which is discussed later in this document. The Pew Research Center, the Kaiser Family Foundation and AP/GfK report average design effects for their overlapping dual frame sample designs of 1.39, 1.33 and 1.68, respectively (Lambert et al. 2010[4]), making it reasonable to assume that the average design effect will be around 1.5 to 2.0.

d. Sample weighting – Landline sample

Starting with the calculation of the weights for the landline RDD sample, the base sampling weight will be equal to the reciprocal of the probability of selecting the household's telephone number in its geographic sampling stratum. Thus, the base sampling weight will be equal to the ratio of the number of telephone numbers in the 1+ working banks for the number of telephone numbers drawn from those banks and actually released for use. Next, the reciprocal of the number of voice-use landline telephone numbers in the household (up to a maximum of three such numbers) will be calculated. It will compensate for a household's higher probability of selection when it has multiple voice-use residential telephone lines. As noted earlier we will compensate for the exclusion of non-telephone households by using an interruption in landline telephone service adjustment. A within-household weighting factor will then be created to account for the random selection of one respondent from the household. The design weight then equals the product of each of the weight components described above.

Adjustment for unit non-response will be based on at least two cells. The non-response adjustment cells will be defined by whether the respondent (or person who was selected for an interview but did not complete it) resides in a directory-listed number household. Other telephone exchange variables may be considered in the development of the unit non-response adjustments. The non-response adjusted weight equals the design weight times the non-response adjustment factor.

e. Sample weighting – Cell phone sample

For the cell phone sample the first step is to assign a base sampling weight that equals the reciprocal of the probability of selection of the cellular telephone number. The cellular

---

[4] Lambert D, Langer G, McMenemy M. 2010. *"CellPhone Sampling: An Alternative Approach"* Paper presented at 2010 AAPOR Conference, Chicago IL.

phone will be treated as a personal device so there will be no random selection of a respondent. The design weights will then be adjusted for unit non-response separately for cell phone-only respondents and for cell mostly respondents. Once this is completed, we will have non-response adjusted weights for the cell phone sample. Note that the control totals for the post-stratification adjustment will come from the NHIS.

f.   Sample weighting – Combined sample

The first step involves compositing the dual user respondents from the landline sample with the dual user respondents from the cell phone sample. The compositing factor, λ, will be calculated to minimize the variance of the estimate of the population total. The details are shown under the section on variance estimation.

In the final file used for weighting, all contacted households from the landline cross-section sample and the cell phone sample will be included. It is important to include all contacted households, rather than only completed interviews since the contacted households provide useful information for weighting purposes.

The population control totals for the raking of the combined sample will be based primarily on the latest Census Bureau population estimates (age by gender by race/ethnicity) and the American Community Survey (ACS) to develop some of the population control totals for socio-demographic variables such as education. Battaglia et al. (2008)[5] have shown the potential for strong bias reduction in RDD samples from the inclusion of socio-demographic variables in the raking. We will also use the results of the nonresponse bias analysis to inform which variables should be included in the raking.

One might rake the combined sample to the socio-demographic control totals discussed above, but it is important to also have raking control totals related to telephone usage: number of persons age 18 to 74 years old that only have landline telephone service, number of persons age 18 to 74 years old that only have cellular telephone service, and number of persons age 18 to 74 years old that have dual (landline and cellular) telephone service. Such estimates can be obtained for the U.S. from the National Health Interview Survey (NHIS).

To develop the final weights, the non-response adjusted weights will be raked using gender by age, race/ethnicity, education, and telephone usage group as margins... The latest version of the Izrael, Battaglia, Frankel (IBF) SAS raking macro (Izrael et al. 2004)[6] will be used to implement the raking. The raking macro will allow both raking and trimming of large weights. Median weight plus six times the interquartile range will be used as the limit for identifying large weights. The possibility of using additional margins will be examined. These weights will be used when estimating the mode share model from the survey data.

---

[5] Battaglia, M.P., Frankel, M.R. and Link, M.W. 2008**.** Improving Standard Poststratification Techniques for Random-Digit-Dialing Telephone Surveys**.** *Survey Research Methods*. Vol 2, No 1.

[6] Izrael, D., Hoaglin, D.C., and Battaglia, M.P.. 2004. "To Rake or Not To Rake Is Not the Question Anymore with the Enhanced Raking Macro " *Proceedings of SUGI 2*9 2004, Montreal, Canada

g. Variance estimation

Variance estimation will need to rely on statistical software such as SUDAAN, however, existing software will not directly calculate standard errors for overlapping dual frame designs. Rather, it will be necessary to calculate the variance for each of the two samples and use the results to calculate the variance for the combined sample.

The dual frame estimator commonly used for overlapping dual frame telephone samples is:

$$\hat{y} = \hat{y}_{LL-O} + \hat{y}_{CELL-O} + \lambda \hat{y}_{DUAL}^{LL} + (1 - \lambda)\hat{y}_{DUAL}^{CELL}$$

LL designates the landline sample and CELL designates the cellular sample. LL-O, CELL-O, and DUAL identify the landline-only, cell-only, and dual service domains, respectively. λ is the compositing factor for the overlap component.

The composting factor λ that minimizes the variance of $\hat{y}$ is given by

$$\lambda = \frac{v\left(\hat{y}_{Dual}^{Cell}\right)}{v\left(\hat{y}_{Dual}^{LL}\right) + v\left(\hat{y}_{Dual}^{Cell}\right)}.$$

The estimates are based on a random sample of households in the landline frame and persons in the cellular fame. Therefore, the variances are functions of effective sample sizes. The effective sample sizes are computed by taking the sample sizes (respondents) and dividing by the design effects of the estimates. The design effect is based on the coefficient of variation of the sampling weights in each sample. Let n be the effective sample size of cell phone users from the landline fame and let $n^{\dot{\iota}}$ be the sample size of persons having landline from the cell phone frame. Then, the value of λ is approximated by

$$\lambda = \frac{n}{n + n^{\dot{\iota}}}.$$

If we treat λ as fixed, then the variance of $\hat{y}$ can be written as:

$$\mathrm{var}(\hat{y}) = \mathrm{var}(\hat{y}_{LL-O} + \lambda \hat{y}_{DUAL}^{LL}) + \mathrm{var}(\hat{y}_{CELL-O} + (1 - \lambda)\hat{y}_{DUAL}^{CELL})$$

8

The first component will be estimated for the landline sample using SUDAAN by taking the complex nature of the sample design into account. The second component will be estimated for the cellular sample also using SUDAAN by taking the complex nature of the sample design into account. SUDAAN will take into account both the sampling weights and the design used for the selection of the sample while computing variance estimates.

2. Describe the procedures for the collection of information.

a) Call/Contact strategy

The NEC Future survey will be administered via CATI to a randomly selected sample of 12,500 respondents age 18 to 74 years old residing in the Northeast Corridor who have made longer distance trips in the Northeast Corridor within the past 12 months. Approximately 9,750 interviews will be conducted with a landline sample, and 2,750 interviews conducted with a cell phone sample.

Interviewing will be conducted according to a schedule designed to facilitate successful contact with sampled households and complete interviews with the designated respondent within those households. Initial telephone contact will be attempted during the hours of the day and days of the week that have the greatest probability of respondent contact based on the call history of previous surveys conducted at Abt SRBI. This means that much of the interviewing will be conducted between 5:30 p.m. and 9:30 p.m. on weekdays; between 9:00 a.m. and 9:30 p.m. on Saturdays; and between 12:00 noon and 9:30 p.m. on Sundays.

The NEC Future survey will include daytime calling within its calling algorithm. Research has demonstrated that a variation in calling times can reach subsamples disproportionately. For the general population, evenings have the highest contact rate as opposed to weekdays and weekends. But for low income groups and minorities, daytime calls increase the probability of contacting the respondent. Lower income respondents are more likely to have irregular work hours (e.g. shift work, etc.) and sometimes have lower labor market participation, which places them at home during the day.

The NEC Future survey will employ a 10 call strategy for landline and cell phone numbers where up to 10 call attempts will be made to ringing but unanswered numbers before the number is classified as a permanent no answer. This change from the pilot protocol of 5 call attempts is being enacted to help improve the survey response rate. Callbacks to unanswered numbers will be made on different days over a number of weeks according to a standard callback strategy. If contact is made but the interview cannot be conducted at that time, the interviewer will reschedule the interview at a time convenient to the respondent. If someone picks up the phone, but terminates the call before in-house

selection of a designated respondent can be made, then the Contractor will apply a protocol to maximize participation while minimizing respondent annoyance.

b) *Protocol for answering machines and voice mail*

When contact is made with an answering machine or voice mail, a message will be left according to a set protocol. For landline numbers, a message will be left on the 3rd attempt to contact a household member. The message will explain that the household had been selected as part of a national USDOT study, ask that the Contractor's toll-free number be called to schedule an interview, and include reference to the FRA web site which will include information about the survey so that prospective respondents can verify the survey's legitimacy. For the follow-up survey phone contacts that reach an answering machine or voice mail, a message will be left that we tried to reach them and that will offer the option of having them call our toll-free number to schedule an appointment.

c) *Initiating the interview*

When a household is reached in the landline sample on the recruit survey, the interviewer will screen for age eligibility. If only one household member is age eligible, then the interviewer will seek to interview that individual. If there is more than one eligible household member, then the interviewer will randomly select one respondent from among them using the last birthday method and seek to interview that person. Appointments will be set up with respondents if it is inconvenient for them to be interviewed at the time of contact. If the randomly selected respondent is not available at the time of contact, then the interviewer will ask what would be a good time to call back to reach that person.

For the cell phone sample on the recruit survey, respondents will first be asked if they are in a situation where it would be unsafe to speak with the interviewer. If the respondent says "Yes," then the interviewer will say that he or she will call back at another time, and immediately terminate the call. Once a cell phone user is reached at a safe time, the interviewer will first screen for age eligibility. If the cell phone user is eligible to participate, then the interviewer will seek to proceed with the interviewer. If it is an inconvenient time for the respondent, then the interviewer will try to set up an appointment.

d) *Interviewer monitoring*

Each interviewer will be monitored throughout the course of the project. The Monitor will evaluate the interviewer on his or her performance and discuss any problems that an interviewer is having with the Shift Supervisor. Before the end of the interview shift, the Monitor and/or Shift Supervisor will discuss the evaluation with the interviewer. If the

interviewer cannot meet the Contractor's standards, he or she will be removed from the project.

All interviewers on the project will undergo two types of monitoring. The Study Monitor will sit at a computer where s/he can see what the interviewer has recorded, while also audio-monitoring the interview. The audio-monitoring allows the Supervisor to determine the quality of the interviewer's performance in terms of:

1) Initial contact and recruitment procedures;

2) Reading the questions, fully and completely, as written;

3) Reading response categories, fully and completely, (or not reading them) according to study specifications;

4) Whether or not open-ended questions are properly probed;

5) Whether or not ambiguous or confused responses are clarified;

6) How well questions from the respondent are handled without alienating the respondent;

7) Avoiding bias by either comments or vocal inflection;

8) Ability to persuade wavering, disinterested or hostile respondents to continue the interview; and,

9) General professional conduct throughout the interview.

The Supervisor will also monitor the interviewer's recording of survey responses as the Supervisor's screen emulates the interviewer's screen. Consequently, the Supervisor can see whether the interviewer enters the correct code, number or verbatim response.

e) Survey content and structure

As described in detail in Supporting Statement A, Question #2, the survey includes questions about current longer distance trips and Stated Preference (SP) questions about alternative choices in response to different mode availability and modal service characteristics. Respondents should be able to recall the details of these less frequent and likely more memorable long distance trips.

The specific SP trade-off questions reflect an experimental design that addresses a cross section of all of the potential mode availability and service characteristic combinations so that each respondent is not asked to address too complex a choice task or is unnecessarily burdened by a longer interview. Specifically, each respondent is presented with choice questions addressing three of the following six (6) modes of travel within the NEC:

- High Speed Train

- Regional Train
- Commuter Train
- Metropolitan Train (a new service type)
- Passenger Car/Truck/Van
- Plane
- Bus

For a specific respondent, the selected modes used in the stated preference questions will include the respondent's first choice mode, (i.e., the mode they currently use for travel in the NEC) and two randomly selected modes. Thus, each respondent will be asked to choose from among three modes of travel. Based on the 2006 Amtrak NEC Traveler Surveys and the ACS CTPP Journey to Work (JTW) flow data, which provide a basis for estimating the incidence of longer commuter trips, and on an assessment of where new services could be introduced as part of the NEC FUTURE project, we have estimated the expected number of respondents that will be exposed to each of the NEC modes is as follows:
- High Speed Train: 5,000 respondents
- Regional Train: 7,200 respondents
- Commuter Train: 2,700 respondents
- Metropolitan Train: 4,700 respondents
- Passenger Car/Truck/Van: 10,300 respondents
- Plane: 3,900 respondents
- Bus: 3,700 respondents

Note that the above total to 37,500, or three times the sample size, because each respondent will be exposed to three modes.

The detailed trip information obtained before the stated preference trade-off questions provide the context for the respondent's travel choices and a basis for defining trip-relevant service characteristics in the trade-off questions. The stated preference questions vary the values of a randomly selected subset of service characteristics for each mode using an experimental design that minimizes the correlation among independent variables. Respondents will be randomly assigned to one of three (3) subgroups who will see changes in these variables:

- Group 1: travel time and cost
- Group 2: travel time and schedule
- Group 3: cost and schedule

This survey design limits the number of changing variables that any one respondent will need to react to and thus makes the task more manageable. Across the entire sample, however, all of the above combinations will be addressed. Taken in combination with the expected respondent samples by mode, the smallest cells would contain 900 respondents in each of the above groups exposed to the Commuter Train mode choice.

Respondents will be presented with a total of six (6) SP questions that addresses three choices of mode of travel with varying characteristics within one of the three pairs of variables listed above. These six (6) questions will be selected for a given respondent from a total of eight (8) possible, using an experimental design that includes five variables at two levels each. A sixth variable, which will be present unless one of the modes is auto and one of the variables is schedule (which is not defined for auto), will not vary. Having a sixth variable vary would have made the task too difficult for respondents. Thus, there will be up to a total of 6 variables, 2 for each of the three modes, and the experimental design will change 5 of the 6 variables as follows across the questions:

SP	Mode 1		Mode 2		Mode 3	
Question	Variable 1	Variable 2	Variable 3	Variable 4	Variable 5	Variable 6 (if needed)
1	[V1-Base]	[V2-Base]	[V3-Base]	[V4-Base]	[V5-Base]	[V6-Base]
2 or 3	[V1-High]	[V2-High]	[V3-Low]	[V4-Base]	[V5-Base]	[V6-Base]
2 or 3	[V1-Base]	[V2-High]	[V3-Low]	[V4-Low]	[V5-Base]	[V6-Base]
2 or 3	[V1-High]	[V2-Base]	[V3-Base]	[V4-Low]	[V5-Base]	[V6-Base]
4, 5 or 6	[V1-Base]	[V2-High]	[V3-Base]	[V4-Low]	[V5-Low]	[V6-Base]
4, 5 or 6	[V1-High]	[V2-Base]	[V3-Low]	[V4-Low]	[V5-Low]	[V6-Base]
4, 5 or 6	[V1-Base]	[V2-Base]	[V3-Low]	[V4-Base]	[V5-Low]	[V6-Base]
4, 5 or 6	[V1-High]	[V2-High]	[V3-Base]	[V4-Base]	[V5-Low]	[V6-Base]

The above experimental design offers two levels of five variables (a 2x2x2x2x2 design). It represents a compromise between the desire for a lot of variation and the need to keep to a reasonable number of questions and minimize the correlation among the variables.

Base values of all variables are pre-determined for each mode for each possible origin and destination in the market area. The values for a particular respondent are determined by the respondent's self-reported origin and destination from the randomly selected "reference trip" (see "Origin_City" and "Destination_City" derived from screening questions in Survey) as well as respondent-provided fare information (see Questions 12A-C & 13A-B in the survey. If the respondent does not remember the fare paid, default values will be used based on published fares for travel by the chosen mode between the origin and destination). In the pilot survey phase, most self-reported rail, air and bus fares by respondents were reasonable when compared to published fares. The project team has developed transportation network models of all modes which produce the relevant base values for travel times and costs, which scale with trip length and geography. High and low values will be computed as follows:

- Total Travel Time High: randomize among +15%, +30% over base values
- Total Travel Time Low: randomize among -15%, -30% under base values
- Total Cost High: randomize among +15%, +30% over base values
- Total Cost Low: randomize among -15%, -30% under base values
- Schedule High: randomize among next two higher amounts over base (e.g., "Every Two Hours", "Every Three Hours")

- Schedule Low: randomize among next three lower amounts under base (e.g., "Every 30 minutes", "Every 20 minutes", "Every 15 minutes")

The responses to these above SP survey questions ultimately provide the basis for estimating key sensitivities to changes in the service characteristics in the new model. This is described in further detail Supporting Statement A.

Considerable care has been taken the design of the NEC FUTURE survey to avoid presenting or showing any bias for/against any specific mode to respondents. First and foremost, the survey is household-based and mode neutral in recruiting respondents and identifying specific trips they have taken. Key elements of this include:

- Identification of the survey as USDOT-sponsored, not FRA, when speaking with respondents
- Randomized order of presenting candidate markets to select a specific market where the respondent made a trip
- Random selection of specific mode and trip purpose taken in this market
- Use of this randomly selected mode as one of the available modes shown in the stated preference questions
- Using respondent's reported alternative mode (second choice) as one of the available modes shown in the stated preference questions
- Randomly selecting from remaining available modes to complete the list of modes shown to each respondent in the stated preference questions
- Randomizing the order that these modes are shown to respondents
- Using similar language throughout the survey where mode-specific information is presented to respondents

In addition, the FRA has provided briefings on the survey approach during expert panel meetings of an OST-sponsored, FHWA-managed long distance travel demand modeling project. FHWA, FAA, FTA and BTS are all participants in the project, which FRA has been involved in since 2011. FHWA has specifically requested and reviewed the draft survey instrument, and has pursued active discussions with FRA on how to gain the most DOT-wide value from FRA's current efforts. We fully anticipate sharing data and results from this travel survey with other modes, both for Northeast Corridor specific understanding, and also to inform national travel demand models

NEC FUTURE study includes additional outreach to constituencies representing both rail and non-rail modes of travel within the NEC. Non-rail agencies on the stakeholder correspondence list include:

- Federal Highway Administration (headquarters and regions within the NEC)
- Federal Aviation Administration (headquarters and offices within the NEC)
- Federal Transit Administration (headquarters and region offices within the NEC)
- Bradley International Airport
- Islip MacArthur Airport
- Newcastle County Airport
- Rhode Island Airport Corporation

- T F Green Airport
- Tweed New Haven Regional Airport Maritime Administration
- Delaware River & Bay Authority
- Delaware River Joint Toll Bridge Commission
- Maryland Transportation Authority
- Massachusetts Department of Transportation
- Metropolitan Transportation Authority - Bridges and Tunnels
- New Jersey Turnpike Authority
- New York State Bridge Authority
- New York State Thruway Authority
- Pennsylvania Turnpike Commission
- Rhode Island Turnpike and Bridge Authority
- Port Authority of New York & New Jersey
- New Haven Parking Authority
- Delaware River Port Authority
- Diamond State Port Corporation (Port of Wilmington)
- MassPort (Port of Boston)
- New Haven Port Authority
- Philadelphia Regional Port Authority

f) Survey sample requirements

The key dimension driving the survey sample size requirements is trip purpose. Prior survey research and model estimation analysis, including the previously referenced Amtrak survey and model as well as most other intercity and regional surveys/models nationwide, have consistently shown trip purpose to be a significant determinant of travel behavior with respect to key sensitivities to service characteristics. For example, travelers on business trips typically show a higher value of time than travelers on non-business trips. As such, the new model will be stratified into the following three trip purposes:

- Commute Trips, including only the daily commute to or from the usual place of work
- Business Trips, which include all non-commute trips associated with a business purpose such as company meetings, sales trips, etc.
- Non-Business Trips, which includes all other non-commute and non-business trips

Another important dimension is geography. The survey sample needs to address a cross-section of trips in different markets that is representative of the NEC. In the NEC Future study, initial review and analysis of available market data and conceptual future rail alternatives has identified and confirmed the importance of the following key geographic stratification of the NEC:

- Travel North of New York

- Travel South of New York
- Travel through the New York Area (between points north of and south of New York)

The above stratification is particularly important with respect to business and non-business "intercity" trips, where there are important differences in the characteristics and availability of different modes of travel. The specific type of longer commute trip between regions addressed by the new model is in itself a unique market. However, it is a much smaller market that does not lend itself to similar geographic stratification.

Since the new model will be stratified into business, non-business and commute trips, it is important to ensure the overall survey sample provides sufficient numbers of completed surveys within each of these segments. It is a given that larger samples will provide more precision, but there is a diminishing return on this relationship and requirements must be properly balanced with the need to efficiently and effectively use available resources.

As with any sampling plan, there is always some uncertainty as to whether the data will actually reflect the universe of travelers until the survey itself is completed. However, there are available sources of data to estimate what could be reasonably expected from an NEC sample. Two key sources were examined to assess the expected distribution of the NEC sample – (1) the 2006 Amtrak NEC Traveler Surveys, which used a similar survey approach, and (2) the 2006 to 2008 3-year ACS CTPP Journey to Work (JTW) flow data, which provide a basis for estimating the incidence of longer commuter trips between NEC regions. Table B.2 below provides the estimated size of the key trip purpose and geography subsamples for within a total sample of 12,500 completed surveys.

TABLE B.2 SURVEY SAMPLE BY TRIP PURPOSE & GEOGRAPHY

	Business	Non-Business	Commute	Total
North of NY	517	2,462	*	
South of NY	1,242	5,816	*	
Through NY	371	1,367	*	
Total	**2,130**	**9,645**	**725**	**12,500**

* stratification of commuter travel by these markets will not be examined independently

The margin of error at 95% confidence level for an estimated population percentage of 50% based on the sample sizes given under Business in Table B.2 will be between plus or minus 2.6 percentage points and plus or minus 6.2 percentage points under the assumption that the design effect is 1.5. The margin of error at 95% confidence level for non-business will be between plus or minus 1.2 and plus or minus 3.2 percentage points.

Another important consideration is the intersection of the survey experimental design with trip purpose. Table B.3 below shows the expected number of respondents that will be exposed to each of the NEC modes by trip purpose, using the same Amtrak survey

data sources. Again, note that Table B.3 totals to 37,500, or three times the sample size, because each respondent will be exposed to three modes.

TABLE B.3 SURVEY RESPONDENT EXPOSURE TO MODES BY TRIP PURPOSE

	Business	Non-Business	Commute	Total
High Speed Train	990	3,860	150*	5,000
Regional Train	1,300	5,615	285	7,200
Commuter Train	400	1,765	535	2,700
Metropolitan Train	790	3,595	315	4,700
Passenger Car/Truck/Van	1,470	8,190	640	10,300
Plane	845	3,055	**	3,900
Bus	595	2,855	250	3,700
Total	**6,390**	**28,935**	**2,175**	**37,500**

* Because of its limited-stop operation, high-speed train service (like Amtrak's Acela) is not available in many commuter markets
** Plane service does not exist in most commuter markets; it is not a relevant mode of travel for commuters because of price.

Based on the available data, the above subsamples can be achieved from a random sample of the NEC population without any oversampling. Both business and non-business trips can be easily segmented geographically, and have separate targets based on the NEC geography. With one exception, these subsamples would also be achieved with random selection of a specific respondent trip among those reported. The exception is for the commute purpose, where the very low incidence of this trip requires selecting that particular trip from among all respondents who report it.

g) Quality Assurance and Quality Control of Data

The JV (Joint Venture consisting of Parsons Brinckerhoff and AECOM) and Abt SRBI, the JV's market research contractor, will implement a comprehensive quality assurance and quality control system to ensure the delivery of a clean database with a maximum degree of consistency, completeness and accuracy.

The process begins with the software Abt SRBI uses for telephone surveys. Abt SRBI's CATI system is able to program loops, rotations, randomization and extremely complex skip patterns. It also includes automatic range checks for data entry. It can be programmed to conduct complicated calculations. And it can also carry forward earlier responses, which can be integrated into later questions. All these functions ensure the quality of the data collected.

In addition to programming, ad-hoc manual inspections of collected data will be conducted as further checks.   This primarily includes doing sense checks on the frequencies of questions, and examining the distribution of key survey variables (e.g., chosen mode).

The initial 307 responses were tested for general operational and content issues with the survey as well as respondent fatigue in the stated preference questions.   These initial 307 responses constituted the "pilot phase" of the survey. Pilot results and the quality checks conducted are discussed later in this document in Question #4b.

3. Describe methods to maximize response rates and to deal with issues of non-response.

a) Initial contact

The initial contact with the designated respondent is crucial to the success of the project. Most refusals take place before the interviewer has even completed the survey introduction (usually within the first 30 seconds of the call).   Numerous studies have shown that an interviewer's manner of approach at the time of the first contact is the single most important factor in convincing a respondent to participate in a survey. Many respondents react more to the interviewer and the rapport that is established between them than to the subject of the interview or the questions asked. This positive first impression of the interviewer is key to securing the interview.

While the brief introduction to the study concerning its sponsorship, purpose and conditions are sufficient for many respondents, others will have questions and concerns that the interviewer must address.   A respondent's questions should be answered clearly and simply. The interviewers for the NEC Future Survey will be trained on how to answer the most likely questions for this survey. A hard copy of Frequently Asked Questions will be provided to each interviewer at their station (see attached file NEC FUTURE FAQ's). The interviewers also will be trained to answer all questions in an open, positive and confident manner, so that respondents are convinced of the value and legitimacy of the study. If respondents appear reluctant or uncertain, the interviewer will provide them with a toll-free number to call to verify the authenticity of the survey. Above all, the interviewer will attempt to create a rapport with the study subject and anyone in the household contacted in the process of securing the interview.

b) Refusal documentation and tracking

Higher response rates can be achieved through procedures built on careful documentation of refusal cases. Whether the initial refusal is made by the selected respondent, a third party or an unidentified voice at the end of the phone, the CATI system used for this survey will be programmed to move the interviewer to a non-interview report section. The interviewer will ask the person refusing to tell him/her the main reason that he or she doesn't want to do the interview. The interviewer will enter the answer, if any, and any details related to the refusal into the program to assist in understanding the strength and the reason for the refusal. This information, which will be entered verbatim in a specific

field in the CATI program, will be part of the formal survey reporting record, and will be reviewed on a daily basis by the Field Manager and, on a weekly basis, by the Project Director.

The Survey Manager and the Project Director will review the information about refusals and terminations on the CATI system on an ongoing basis to identify any problems with the contact script, questionnaire or interviewing procedures that might contribute to non-participation. For example, they will scrutinize the distribution of refusals by time of contact, as well as any comments made by the respondent, in order to determine whether calling too early or too late in the day is contributing to non-participation. Also, the refusal rate by interviewer will be closely monitored.

In addition to relying on the CATI data records, the Project Director and Survey Manager will also consult with the interviewing Shift Supervisor, who has monitored the interviewing and debriefed the interviewers. The information from these multiple sources provides solid documentation of the nature and sources of non-response.

   c) *FRA web site information*

It is not uncommon for surveys such as these to generate incoming calls or emails to the U.S. Department of Transportation from people who have been contacted by the interviewer to conduct the survey. In a large percentage of the cases, it's an inquiry to check its legitimacy. People are wary about attempts to extract information from them, wondering if the caller is truthful in identifying the source of the survey and the use of their information.

To help allay these fears, FRA will place on its web site information that prospective respondents can access to verify the survey's legitimacy. The interviewers will provide concerned respondents with the web address for FRA's home page. There will be a link on the home page that will direct the respondents to information on the source of the survey and why it is important for them to participate. An 800 number to reach the survey Contractor will also be provided so they can schedule an interview.

   d) *Non-response bias analysis*

A non-response bias study will be conducted for the NEC Future Survey. Non-response is the failure to obtain survey measures on all sampled individuals. There are two main potential consequences from non-response: (1) non-response bias and (2) underestimation of the standard errors. The main challenge posed by non-response is that without a separate study (i.e., a non-response bias analysis), one can never be sure of whether the non-response has introduced bias or affected the estimation of the standard errors. The mere presence of non-response, even high non-response, does not necessarily imply bias or non-response error.

The demographic composition of the NEC Future Survey landline sample will be compared to the demographics of non-respondents as determined by auxiliary data providing demographic information for a proportion of sampled landline households.

The auxiliary data will be purchased from an external vendor like Marketing Systems Group (MSG) or Survey Sampling International (SSI) and will represent the latest and most recent information on demographic characteristics such as household income, presence of children in the household, number of people residing in the household, race of those living in the household, and age and gender of household members. The quality of this information varies since most of this is gathered though credit bureaus and applications for services. The younger and more mobile population is not as likely to be represented in the auxiliary data. Neither are minorities or lower income folks. This data does, however, give us a more robust analysis of who completed and who did not at the case level, rather than at the aggregate level, which is undertaken when comparing Census data to the sample.

The auxiliary data will be used in two different ways. First, the auxiliary data will be compared to the demographic data obtained from the surveys for those respondents who completed the interview. The analysis will assess whether the landline non-respondents are demographically different from those who completed the interview. If some of these characteristics are correlated with the propensity to respond, there is a chance that there might be some bias in those estimates. Secondly, the auxiliary data for survey respondents will also be compared to the auxiliary data for non-respondents.

Additionally, there will be a non-response follow-up (NRFU) survey with a sub-sample (n=500) of households that did not respond to the interview. The NRFU has two main objectives: 1) investigate the reasons for non-response, and 2) assess the risk of non-response error to model estimates. Because these are people who refused to participate, there are limitations on how many respondents one can expect to participate in the NRFU. However, the sample size of 500 households is large enough to estimate factors that contributed to non-response, assess the potential for non-response bias in key substantive variables, and determine key socio-demographic variables to include in the final post-stratification adjustments.

The NRFU features an abbreviated survey instrument in which the household is asked to provide a few basic socio-economic and demographic variables, reasons for non-participation, and responses to a few basic survey items.   NRFU households will each be offered $20 for their participation, as this amount has been effective in past non-response follow-up studies conducted by the Contractor. The results of the non-response follow-up survey will provide valuable information for assessing the extent to which non-response bias affects the overall survey respondent sample and for developing weighting adjustments that minimize the impact of non-response bias on sample statistics and model parameters.

The analysis will include comparison of key variables from the NEC Future Survey respondents with the same variables for the NRFU respondents. This will provide insights about the direction and magnitude of possible non-response bias. The project team will investigate whether any differences remain after controlling for major weighting cells (e.g., within race and education groupings). If weighting variables eliminate any differences, this suggests that weighting adjustments will reduce non-

response bias in the final survey estimates. If, however, the differences persist after controlling for weighting variables, then this would be evidence that the weighting may be less effective in reducing non-response bias.

Based on the pilot phase results, the modified one-phase phone survey is expected to have a 12-15% response rate as described in Supporting Statement A, Question #2 (Data Collection for Model). In terms of coverage, the pilot results showed that the dual-frame telephone sample can closely match the Census geographic distribution across the study regions.

Table B.4 below summarizes the expected response rates for different sections of the survey based on results from the pilot. We will count a survey as a "complete" if the respondent provided a response to at least one stated preference question.

TABLE B.4 SURVEY QUESTION ITEM NON-RESPONSE ESTIMATES

Question Topic	Question Number(s)	Expected Range of Non-Response (DK/Ref)	How Non-Response Will Be Handled in Data Analysis
Travel Mode and Frequency	S2D to S2E; S3E to S3F	<1%	Acceptable, no special protocols needed.
Specific One-Way Trip & Mode	1 to 24B	<5% for most questions: 5-20% for trip fare or cost	Acceptable, no special protocols needed for most questions. For an out-of-range fare or cost, a pre-determined typical fare or cost for travel in the city pair will be used.
Main Mode Choice Trade-Off Questions	25 (A-D) to 30 (A-D)	<1%	Acceptable, no special protocols needed.
Demographics	D-1 to D-12	<1% for all demographic questions except income question (<5%)	Acceptable, no special protocols needed.

4. Describe any tests of procedures or methods to be undertaken.

a) Testing the programming

The project team will test the CATI program thoroughly in test mode – running the interviewing program through multiple loops. The analytical staff will attempt to test all possible response categories for each question in order to identify embedded logic errors, as well as obvious skip problems. Several analysts may test the program simultaneously to identify problems quickly, and to double check the comprehensiveness of the testing.

After initial testing and corrections, the questionnaire program is also run through our autopilot program. This program tests the interview program by initiating the CATI interview and then by generating a dummy database of random responses as the questions

appear. This database permits us to track the response pattern compared to the hard copy questionnaire, in order to further identify skip or other programming errors.

b) *Pilot test*

As mentioned before, a pilot test of 307 respondents was conducted from 8/28/13 through 11/4/13. The purpose of the pilot was to consider the effectiveness of the stated preference exercises, test the strength of the model and assess response rates and burden estimates. In summary, while the pilot effort were generally successful in obtaining the necessary information for modeling, the cumulative 2-phase response rate garnered was low.

The specific findings of the pilot were:
- The recruitment survey was 7 minutes and the follow-up survey was 15 minutes for the phone portion of the mail/phone version and 18 minutes for the Internet version.
- The cumulative response rate was 4%. It was 9% in the recruitment and 49% in the follow-up.
- The geographic areas captured by the survey generally reflected the relative populations of the areas within the Northeast Corridor. Trip purpose (mostly leisure) and chosen travel mode (mostly car) results were also as expected given the qualifications of the survey (long distance travel).
- The demographics of pilot respondents skewed less Hispanic, older and higher income. Note, however, that it is possible that long distance travelers who qualify for this study may also skew in the same direction on these demographic dimensions.
- Test model estimations showed that coefficients behaved as expected.
- Stated preference results did not show signs of respondent fatigue as they progressed through the choice exercises.
- 48% of respondents did not indicate they would switch travel modes in the stated preference section. Most of these were Auto travelers.
- Illogical switching was tested for responses in the pivot questions section. Transfer and access modes had reasonable results but station availability showed a greater number of illogical responses.

In response to the pilot findings, the following changes are proposed to increase the response rate, and obtain additional information for the non-response bias analysis, while preserving the necessary elements for stated preference analysis.
- Revise the two-phase data collection approach to be a one-phase telephone survey.
- Increase the number of call attempts per sampled case from 5 to 10.
- Increase the incentive from $5 to $10 for those who complete the survey.
- Restructure the questionnaire to reduce overall survey length and improve flow of questions, by doing the following:
  - o Reduce the number of stated preference choice exercises from 12 to 6.
  - o Remove sections on alternatives and pivot questions.

- o Remove questions on origin and destination location type.
- o Streamline screener to ask about states traveled to (instead of areas) and to ask all commute and non-commute questions in their respective sections
- o Obtain demographics from all respondents who live within the area, but do not make long distance trips. This will be used for the non-response bias analysis and allow us to compare the sample to Census estimates for the Northeast region.

5. Provide the name and telephone number of individuals consulted on statistical aspects of the design.

The following individuals have reviewed technical and statistical aspects of procedures that will be used to conduct the NEC Future Survey:

Bruce Williams
Senior Consulting Manager
AECOM Transportation
2101 Wilson Boulevard, 8th Floor
Arlington, VA 22201
(703) 340-3078

Mindy Rhindress, Ph.D.
Senior Vice President
Abt SRBI, Inc.
275 Seventh Ave, Ste 2700
New York, NY 10001
(212) 779-7700

Paul Schroeder
Vice President
Abt SRBI, Inc.
8405 Colesville Road, Suite 300
Silver Spring, MD 20910
(301) 608-3883

K.P. Srinath, Ph.D.
Senior Sampling Statistician
Abt SRBI, Inc.
8405 Colesville Road, Suite 300
Silver Spring, MD 20910
(301) 628-5527