Department of Transportation
Federal Aviation Administration Office of Environment and Energy

SUPPORTING STATEMENT
Supporting Statement for a Collection RE:
Neighborhood Environmental Study
OMB Control Number 2120-XXXX

INTRODUCTION

This information collection is submitted to the Office of Management and Budget (OMB) to request a three-year approval clearance for the information collection entitled Neighborhood Environmental Study, (OMB Control No. 2120-762)

Part B.

This request is being made as a renewal to the prior request. Over the last three years, the FAA has used the existing collection to gather information regarding the public's opinion on aircraft noise. FAA has analyzed this collected data and is currently reviewing the results and draft report. FAA believes that there may be a need to conduct additional analysis and/or collection as FAA continues to examine the data.

Table of Contents

B.1   Respondent Universe and Sampling Methods

Currently, the FAA defines significant noise as a Day-Night Average Sound Level (DNL) of 65 decibels (dB) or more. Over the last three years, the FAA has used the existing collection to gather information regarding the public's opinion on aircraft noise. FAA has analyzed this collected data and is currently reviewing the results and draft report. FAA believes that there may be a need to conduct additional analysis and/or collection as FAA continues to examine the data.

So far, FAA used the collected data to update the relationship between aircraft noise exposure and its effects on communities (annoyance) around U.S. airports. The survey collected data from a representative sample of airports and households surrounding each of the airports, and related the annoyance level to the noise exposure for each household address. The selected participants for the study represented a wide range of conditions with respect to number of operations, nighttime operations, temperature, population in proximity to the airport, and fleet mix, and the results from the study are generalized to the relevant population of U.S. airports. The same survey instruments and data collection procedures was used for all of the airports, and the survey was conducted during the same time period at all airports. Further application of these uniform procedures for additional collection, if necessary, will result in data that can be compared across airports and that can be used to construct a national dose-response curve relating annoyance levels to aircraft noise exposure.

The main purpose of the collection is to collect information to update previously estimated dose-response relationships in Schultz (1978) and FICON (1992) between aircraft noise exposure and the self-reported annoyance of residents for the nation as a whole. As FAA is still reviewing the data from the current analysis, FAA is seeking to renew the collection in case more information is needed. This additional information can either be surveying additional airports or surveying additional people at the same airports.

Noise exposure is measured using the day-night average sound level (DNL), measured in decibels (dB). FICON (1992) defines the DNL metric and gives

justifications for its use as a measure of noise exposure. DNL is an energy-averaged sound level that is integrated over a specified time period, with a 10 dB penalty for nighttime exposure to noise. "In 1981, FAA formally adopted DNL as the primary metric to evaluate cumulative noise effects on people due to aviation activities" (FAA, 2007, Chapter 17, page 1).

The primary interest is in developing a regression model relating the percentage "highly annoyed" (percent HA) to the noise exposure. Percent HA is used as the response in the logistic regression equation relating noise exposure to annoyance in FICON (1992): that equation has the form

$$Percent\ HA = \frac{100 \exp\left(\beta_0 + \beta_1 DNL\right)}{1 + \exp\left(\beta_0 + \beta_1 DNL\right)}. \tag{1}$$

For the FICON (1992) curve, the parameters were estimated as $\hat{\beta}_0 = -11.3$ and $\hat{\beta}_1 = 0.14$. The FICON (1992) relationship is used as the current basis for establishing significant impact (currently defined as exposure to DNL levels of 65 dB or higher). The primary research question for this study is whether there has been a shift over time in populations' annoyance to aviation noise at the same sound levels.

B1.1 Respondent Universe and Sample Frame

While FAA is unable to release any specific information regarding the 2015 collection, we can state that the collection was completed as described below. In addition, response rates and non-delivery rates were similar to that presented in the pilot study. As FAA is not in a position to release the specifics of the 2015 study, information from the pilot study is presented here. At this point, FAA does not know the extent of the additional collection, but expects it to no more than the 10,000 people at 20 airports discussed below.

The target population of the renewal survey is residents of households who live in proximity to an airport meeting the eligibility criteria for this study. The sampling frame of airports consists of airports with a minimum number of 100 average daily operations in 2011 that are determined by FAA to have

at least 100 jet operations per day and at least 100 households exposed to aircraft noise of 65 dB DNL or above and at least 100 households exposed to noise between 60 dB DNL and 65 dB DNL. A total of 95 airports in the United States meet these criteria.

The sampling is to be done in two stages. The first stage involves selecting a balanced probability sample of 20 airports from the frame of 95 airports. The second stage of sampling selects addresses that are within the desired ranges of noise exposure. Because this survey is intended to study the relationship between aircraft noise exposure and annoyance, persons who live at addresses with very low aircraft noise exposure would contribute little or no information toward estimating this relationship. The target population of addresses at each airport is defined to be addresses with exposures of 50 dB DNL higher.

For each airport selected, noise exposure contours will be determined using FAA's regulatory noise model. This will permit selection of potential survey participants by DNL noise exposure range, and will permit computation of specific noise exposure (DNL) for each sampled household.

The addresses on the U.S. Postal Service Computerized Delivery Sequence File (CDSF) will be used as the household sampling frame. These addresses can be geocoded to the appropriate noise strata.

B1.2 Sample Design

The study will have two phases. In the first phase, a mail survey will be sent to sampled households. The data from this survey will be used to estimate the dose-response curve. Respondents to the mail survey will be subsampled for an additional telephone interview. The purpose of this interview is to obtain further information about attitudes towards airports and airport policies. Results from the telephone interview will not be used for estimating the dose-response curve relating aircraft noise exposure and annoyance.

B.1.2.1 Selection of Airports

The population of inference consists of the 95 airports described in Section B.1.1. A balanced sampling method is used because it is desired that the sample of airports match the set of 95 airports on the factors described in Table B-1. A stratified sample, with two airports in each of ten strata, could match the population of airports on some of these factors but not on all of them. The balanced sampling procedure ensures that the sample of 20 airports is representative of the population on a much wider array of factors (Royall, 1976; Tillé, 2011).

Table B-1. Balancing factors for selection of airports

Factor	Description
1	Proportion of airports in each of the eight FAA regions[1]
2	Proportion of airports with average daily temperature above 70 degrees
3	Proportion of airports with average daily temperature below 55 degrees
4	Proportion of airports with more than 20% nighttime operations
5	Proportion of airports with more than 300 average daily operations
6	Proportion of airports with a fleet mix ratio of commuter to large jet aircraft exceeding 1
7	Proportion of airports with at least 230,000 people living within 5 miles of the airport

Restricted random sampling (Valliant et al., 2000, p. 71), with a modification to include the certainty airports, is used to select a sample that provides balance on the factors given in Table B-1. Restricted random sampling consists of the following three steps:

1. Generate a large number of random samples of size 20 from the population of airports.

2. Reject the samples that do not meet the balancing constraints.

3. Select one sample at random from the remaining samples (all of which meet the balancing constraints).

To implement this procedure, we first generate 250,000 stratified random samples using FAA Region as the stratification factor (chosen because it has the most categories).[2] Each stratified sample has the desired sample size in

---

[1] FAA has a total of 9 regions. However, the Alaska Region was not included because it has no airports that meet the selection factors.

[2] This was done purely for computational efficiency, and does not imply FAA Region is more important than other factors. By using the first factor, FAA Region, in Table B-1 for generating the candidate samples, the computational effort was substantially reduced.

each of the 8 FAA Regions, and always includes the three certainty airports and one of the New York City-area airports chosen at random. Then samples that do not meet the balancing constraints for factors 2 through 7 in Table B-1 are rejected[3], and one of the samples that meets the balancing constraints is chosen at random to be the sample of airports.

B.2.2 Sample Sizes within Airports for Mail Survey

The main purpose of the study is to provide information for constructing a national dose-response curve. The sample design for addresses selected from each airport community, therefore, is tailored for estimating a regression relationship (see Lohr, 2014). This leads to a different design than would be used if the main purpose of the survey was estimating a population mean or total for the region.

A stratified sample of addresses will be taken at each airport, with the expected national sample sizes and, based on the pilot study, with the original collection having similar results, expected response rate for both the mail and telephone survey provided in Table B-2. The response rates and match rates used in the table are based on the pilot study (similar to those found in the original collection) conducted using a similar methodology in 3 airports (see Section B.4).

This is because every generated sample was balanced for each of the eight FAA regions, and only needed to be checked for whether it was also balanced on the other six factors. The same procedure would work (and would produce similar samples) if, say, the initial samples had been stratified on Temperature, but in that case each sample would have needed to be checked for balance on 11 other criteria, so a much higher fraction of the generated samples would be rejected.

[3] The target sample size for each category of a factor was set equal to the integer closest to 20 x (proportion of airports in the sampling frame in that category). A sample met the balancing constraints if it achieved the target sample size for every category of each of factors 2 through 7 in Table B-1.

Table B-2.   National sample sizes, response rates and expected completes

	Number
Initial sample	26,700
6.3% PND (Postal non-deliverables) [4]	1,682
40% response rate	10,007
Telephone	
40% match to telephone number	4,703
85.1% valid match	4,003
a. 30% complete phone interview	1,201
60% with no match to telephone number	6,004
14.9% invalid from matched group	701
Total phone number requests	6,705
35% provide phone #	2,347
b. 40% complete phone interview	939
Total telephone completes (a + b)	2,140
	%
Final mail survey response rate	40.0
Final Telephone interview response rate	8.6

The proposed allocation of sample for each airport by noise stratum is given in Table B-3. This allocation gives high precision for estimating the relationship between noise exposure and percent HA (Abdelbasit and Plackett, 1983), and allows for evaluating possible deviations from the assumed logistic model. This results in a disproportionate sample of addresses across strata, with higher percentages of addresses sampled at higher exposures, because typically the region surrounding an airport contains more addresses in low noise strata than in high noise strata.

Table B-3.   Expected number of respondents for each airport, and for the study as a whole

	Noise Exposure Stratum, dB DNL					
	50-55	55-60	60-65	65-70	70+	Total
Each airport, mail survey	100	100	100	100	100	500
Total, all airports, mail survey	2,000	2,000	2,000	2,000	2,000	10,000
Each airport, telephone survey	21.4	21.4	21.4	21.4	21.4	107
Total, all airports, telephone survey	428	428	428	428	428	2,140

---

[4] Postal non-deliverables are mailed surveys that were returned as non-deliverable by the USPS.

Table B-4 gives the anticipated margins of error for estimating percent HA from the survey for the national airport curve at three noise exposures of interest. In logistic regression, the anticipated precision depends on the parameters of the curve (Chaloner and Larntz, 1989). Therefore, the precisions are evaluated under the logistic regression model in equation (1) under four scenarios: the FICON (1992) curve, the curve reported in Fidell and Silvati (2004), a curve estimated by fitting a logistic mixed model to data in Fidell et al. (2011), and a curve that gives a "worst-case" scenario of precision for the national dose-response estimation, called "curve W." These curves are shown in Figure B-1.

Figure B-1.   Dose-response curves used for precision calculations
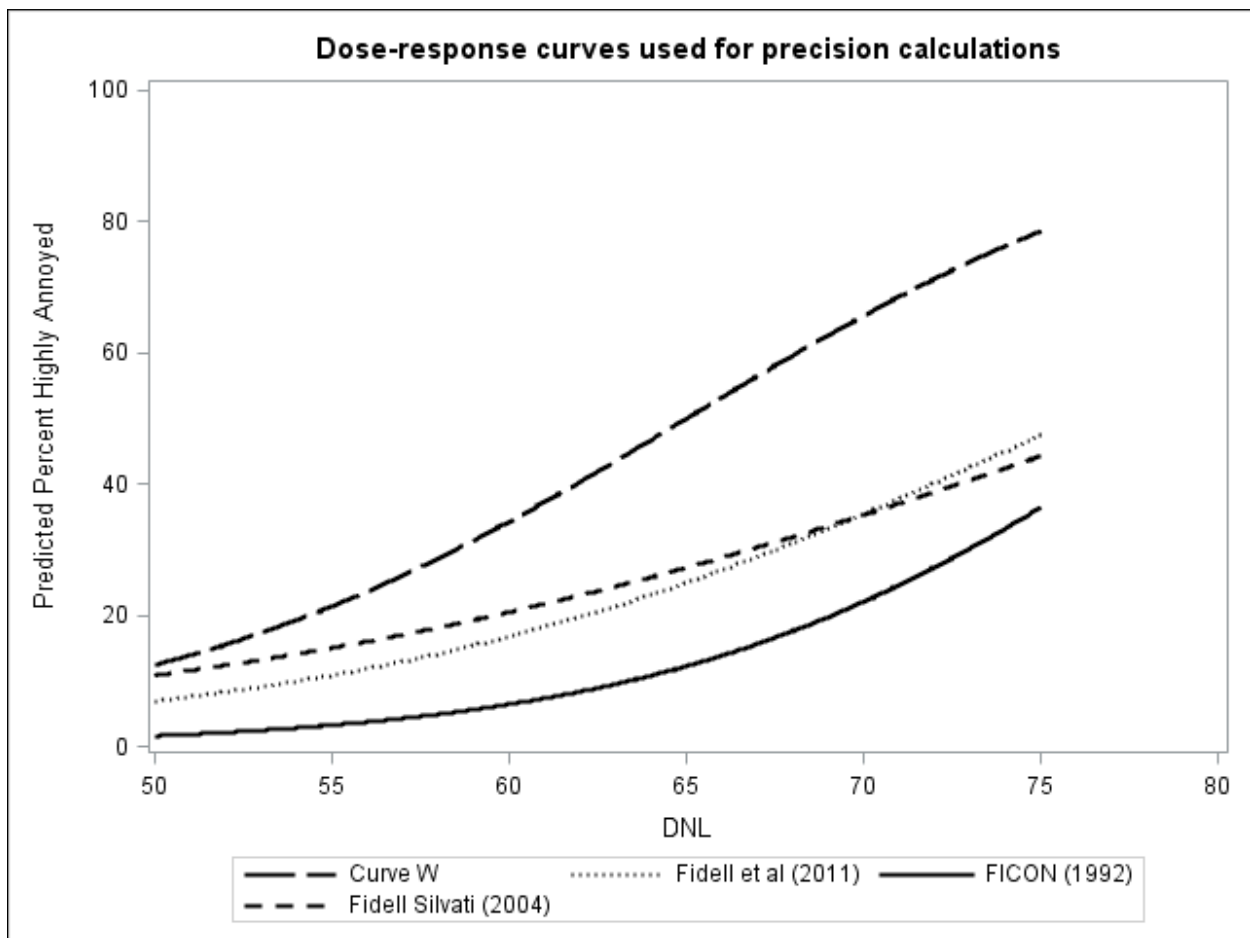

Table B-4.   Anticipated precision of percent HA for national airport curve

Logistic curve	Intercept of logistic curve (β_0)	Slope of logistic curve (β_1)	Margin of Error for Percent HA at Noise Exposure (DNL)+		
			DNL=55	DNL=60	DNL=65
Curve W	-8.45	0.13	6.9	10.4	13.0

10

Fidell and Silvati (2004)	-5.854	0.075	▪6.2	▪ 6.5	▪ 6.5
Fidell et al. (2011), mixed model	-7.6	0.10	▪4.5	▪ 5.4	▪ 6.0
FICON (1992)	-11.3	0.14	▪1.1	▪ 1.7	▪ 2.5

+ Half-width of 95% confidence interval.

Curve W is derived using the following property: The variability of an estimated proportion is greatest for proportions close to 0.5, because then about half of the observations must be 0 and the other half must be 1. The variability decreases as the true proportion decreases; in the extreme case of a proportion of 0, all of the observations are the same and there is no variability.[5] We therefore use a curve with predicted value of percent HA of 50 percent at DNL 65 as a "worst case" scenario for variability, since that curve allows for the greatest variability among individual airports in terms of percent HA at DNL 65. Using the notation in equation (1), this model for "curve W" has $\beta_0=-8.45$ and $\beta_1=0.13$. These values of the intercept and slope were chosen because they give a predicted percent HA of about 12% at DNL 50, which is consistent with other dose-response curves in the literature.

The primary determinant of precision for the national airport curve is the airport-to-airport variability in the slope and intercept (Lohr, 2014). The between-airport variability estimates used for the Fidell and Silvati (2004) and Fidell et al. (2011) curves in Figure B-1 were determined from a mixed logistic regression model fit to the data given in Fidell et al. (2011). These values were multiplied by 2/3 for the between-airport variability used with the FICON (1992) curve. The between-airport variability used for curve W was again a "worst-case" scenario that allowed the individual airport curves to range between a flat curve in which the percent HA is close to zero along the range of DNL and a curve in which the percent HA is close to 100 percent at DNL 75.

One of the major purposes of the study is to determine whether the dose-response curve relating percent HA to noise exposure now differs from the FICON (1992) curve. From Figure B-1, it can be seen that the margins of error given in Table B-4 would allow for a difference to be detected between the

---

[5] This is reflected in the formula $p(1-p)/n$ that is commonly used for the variance of an estimate of a proportion p from a simple random sample.

FICON (1992) curve and each of the other three curves in the table. As an example, if curve W turned out to be the dose-response curve from the study, an approximate 95% confidence interval for the predicted percent HA at DNL 60 would be 34.3 +/- 10.4 = [23.9, 44.7]. This confidence interval does not contain the predicted percent HA of 5.2 from the FICON (1992) curve, so a significant difference would be detected. If the new airport curve is close to the FICON (1992) curve, then it would be estimated with high precision (margin of error 2.5 percentage points at DNL 65).

A second consideration is the ability of the study to be able to fit the dose-response curve at an individual airport. Table B-5 gives the anticipated margin of error for the response of percent HA at different levels for a single airport, if that airport followed the model for one of the curves displayed in Figure B-1. These margins of error are derived from the covariance matrix of the slope and intercept in logistic regression by using the delta method.

Table B-5.   Anticipated precision of percent HA for one airport

Logistic curve	Intercept of logistic curve (β_0)	Slope of logistic curve (β_1)	Margin of Error for Percent HA at Noise Exposure (DNL)+		
			DNL=55	DNL=60	DNL=65
Curve W	-8.45	0.13	5.4	5.0	5.1
Fidell and Silvati (2004)	-5.854	0.075	4.6	4.1	4.2
Fidell et al. (2011), mixed model	-7.6	0.10	4.0	3.9	4.2
FICON (1992)	-11.3	0.14	1.9	2.5	3.0

+ Half-width of 95% confidence interval.

The anticipated precisions for estimating the dose-response curve are calculated using the methods that are planned for analyzing the data from the survey. Estimates of the dose-response relationship are sensitive to the method used to combine information from the sampled airports. Combining all observations across airports in a single model (called here a combined analysis), as has been done in several studies, gives airports with higher sample sizes or certain allocation patterns higher influence in determining the national relationship. An alternative method of estimation is to fit the model to each airport separately, then average the coefficients across the airports.

We plan to use mixed models (Demidenko, 2004) and logistic regression models that account for clustering (see Lohr, 2010, Section 11.6) to estimate the overall relationship and investigate heterogeneity in the relationship among airports. Mixed models capture the best features of the combined analysis method and the averaging coefficients method. In a mixed logistic model, each airport is allowed to have its own slope and intercept, and these are estimated along with the slope and intercept that best describe the overall relationship among all airports. The models may be used to assess whether all airports have the same dose-response relationship as the overall curve. If there is heterogeneity among airports, the mixed models can estimate the degree of heterogeneity as well as investigate airport characteristics that are associated with divergence from the overall model. Mixed logistic models have been successfully used in other settings in which relationships are thought to vary across localities; see, for example, Kaufman et al. (2003).

In addition, a logistic model will be fit for each individual airport. This will help identify heterogeneity among the airports, and will be useful for assessing goodness-of-fit of the models. The weights from the disproportionate sampling fractions in the noise strata will not be used in the models. The primary interest is in estimating the regression relationship between percent HA conditionally upon the noise exposure. All of the stratification information is included in the explanatory variable in the regression model, which means the stratification weights are not needed (Scott and Holt, 1982). However, weights that adjust for potential nonresponse bias may be used if necessary (see Section B-3).

B.2.3 Sample Sizes for Telephone Survey

Similar to the mail survey information above, FAA is still analyzing the data from the original collection and cannot provide information on the actual collection for the telephone survey.

Respondents to the mail survey will be subsampled for an additional telephone interview. The purpose of this interview is to obtain further information about attitudes towards airports and airport policies. Results

from the telephone interview will not be used for estimating the dose-response curve relating aircraft noise exposure and annoyance.

The expected sample sizes for the telephone survey are given in Table B-2. The method used to estimate quantities from the telephone survey will be similar to that for the mail survey. As can be seen from the telephone survey questionnaire in Attachment I, responses to many of the questions might reasonably be expected to differ for households at different noise exposure levels. Therefore, for comparing responses across airport communities, or relating responses to airport characteristics, the proposed analysis will standardize the responses across noise exposure profiles by looking at predicted values at specified noise exposures. This will be done by using logistic regression for binary responses. For non-binary responses, a linear regression model will be used.

Tables B-6 and B-7 give the anticipated margins of error for estimates of a percentage from a binary response from the telephone survey. We used the same logistic curves and between-airport variances as for the dose-response curve; as before, curve W represents a worst-case scenario for precision of a dose-response curve that is assumed to increase with increasing noise exposure.

Table B-6. Anticipated precision of estimated national percentages from telephone sample

Logistic curve	Intercept of logistic curve (β_0)	Slope of logistic curve (β_1)	Margin of Error for Percentage at Noise Exposure (DNL)+		
			DNL=55	DNL=60	DNL=65
Curve W	-8.45	0.13	7.4	10.6	13.2
Fidell and Silvati (2004)	-5.854	0.075	6.6	6.8	6.8
Fidell et al. (2011), mixed model	-7.6	0.10	4.9	5.7	6.3
FICON (1992)	-11.3	0.14	1.4	2.0	2.9

+ Half-width 95% confidence interval.

Table B-7. Anticipated precision of estimated percentages from telephone sample, for one airport

Logistic curve	Intercept of logistic curve (β_0)	Slope of logistic curve (β_1)	Margin of Error for Percentage at Noise Exposure (DNL)+		
			DNL=55	DNL=60	DNL=65
Curve W	-8.45	0.13	12.0	11.2	11.4
Fidell and Silvati (2004)	-5.854	0.075	10.3	9.2	9.4
Fidell et al. (2011), mixed model	-7.6	0.10	8.8	8.8	9.3
FICON (1992)	-11.3	0.14	4.2	5.5	6.7

+ Half-width of 95% confidence interval.

B.2 Procedures for the Collection of Information

FAA anticipated that if it does completed collection under the renewal, the FAA would continue to use the same methodology as described below. All airports are to be sampled over a 12-month period, with the sample in each noise stratum to be distributed across 6 bi-monthly sample releases. This time period is used because annoyance to aircraft noise can exhibit seasonality factors; sampling each airport over a 12-month period removes such seasonality from the estimation of the dose-response curve.

The data collection protocol includes two main components. The first component is a mail survey to collect data that will be used to model aircraft noise to annoyance, and this will be the main data collection that all sampled addresses will receive. The second component will be a telephone interview with those that complete the mail survey. The telephone numbers used to

call households that returned the mail survey will be gathered in two ways. One will be to match the address to a reverse directory that includes phone numbers. For those that a phone number cannot be found, a brief questionnaire will be mailed, asking for a phone number.

Some consideration was given to asking for a phone number on the mail survey. This number could then be used to conduct the telephone follow-up interview. The advantage of this procedure is that it reduces the need to find phone numbers when conducting the telephone survey. It was decided to not ask for the phone number in order to maximize the mail survey response rates. The literature has shown that asking for phone number suppresses mail survey response rates (Williams, 2014). The mail instrument will be used to draw the noise-annoyance curve and is the primary objective of this study. Therefore, it was judged more important to maximize the response rates at this stage and minimize any potential sources for bias than to facilitate the collection of phone numbers for the second stage.

B.2.1 Procedures for Mail Survey

The mailing protocol used for the main data collection will follow procedures outlined by Dillman, Smyth, and Christian (2008). All sampled addresses will be contacted between 2 to 4 times, depending on when the questionnaire is returned. The contacts will include: 1) an initial survey packet; 2) a thank-you/reminder postcard approximately one week after the initial survey mailing; 3) a second survey package mailing two weeks after the thank-you/reminder postcard (three weeks after initial survey mailing); and 4) a third survey package mailing three weeks after the second survey package mailing. While FAA has considered the use of a web-based survey, but conducting a web survey, rather than a mail survey, would not permit adequate coverage of those that do not have access to the web (Dillman et al, 2008; Millar and Dillman, 2011). In addition, mail surveys yield significantly higher response rates than web surveys (Manfreda, et al, 2008; Millar and Dillman, 2011; Dillman, et al., 2008)). Some consideration was given to providing the respondents a choice between a paper mail and a web

survey. This was rejected because a number of studies have found that giving respondents a choice depresses response rates (Dillman, et al., 2008).

The contents of each survey packet will include a cover letter that provides the survey purpose and sponsorship (Attachment A), and a paper questionnaire that the respondent (Attachment C) will be asked to return via the included postage-paid envelope. All materials mailed to the respondent will reference the 'Neighborhood Environment Survey.' All survey materials will be provided in English and Spanish. This follows established procedures for eliciting response from Spanish speaking households (Brick, Montaquila, Han, and Williams 2012).

A $2 cash prepaid monetary incentive will be included with the initial mail package (See Section A.9 for rationale for incentive). The initial survey and the thank-you reminder postcard (Attachment B) will be mailed to all sampled addresses. Only nonrespondents to the prior mail packages will receive subsequent survey package mailings. Mailings returned as postal non-deliverable (PND) will be excluded from subsequent mailings.

The second survey package will be sent using express delivery. This will increase the visibility of the package and maximize response at this stage (Dillman et al., 2008). The last mailing will be sent USPS first class (see Appendix B for cover letter for third and fourth mailings).

A quasi-random selection procedure will be used to select an adult to answer the mail survey. The instructions on the inside page will ask that the adult with the next birthday fill out the questionnaire.

B.2.2 Procedures for Telephone Interview

Households that complete the mail survey will be eligible for the telephone interview. Household addresses will be directory matched to a telephone number. Those that have a successful telephone match will be mailed a letter requesting participation in the telephone survey (Attachment G). If no telephone match is available the household will be mailed a request to provide a telephone number. This survey package will include a cover letter

explaining the follow-up contact procedure and study sponsorship (Attachment D). A short form for providing the household's telephone number will also be included that would be returned by the respondent in the enclosed postage-paid envelope (Attachment F). The request for telephone number will also follow the mail contact procedures outlined by Dillman, Smyth, and Christina (2008), except there will be three contacts. All households will receive a reminder postcard (Appendix E), and nonresponding households will receive a nonresponse follow-up request (Appendix D). All mailings will be done using first-class postage.

For the telephone interview, an adult will be selected using the Rizzo method (Rizzo et al., 2004). This is a probability method of selection and gives each adult in the household an equal chance of being selected.

B.3   Methods to Maximize Response Rates and Deal with Nonresponse

B.3.1 Computing Response Rates and Adjusting for Non-Response

Response rates will be calculated for both the mail and telephone survey. National estimates of the dose-response curve will be based on the mail survey. The telephone survey will primarily be used to explore the correlates of annoyance and we do not expect to publish national estimates from these data. However, since the telephone survey will be a subsample of the mail survey completes, it will be possible to assess the non-response bias of the telephone survey, at least relative to the mail survey results (see Section B-3.3).

For both the mail and telephone survey, we will use AAPOR rate RR3 to compute the response rate. For the mail survey, we will tabulate the number of postal nondeliverables for the advance letter. This will provide a way to estimate eligibility among those that don't return the package:

$$RR3 = \frac{\#\text{completed interviews}}{\#\text{completed interviews} + \#\text{refusals} + e*\left(\#\text{unknown eligibility}\right)}$$

where #unknown eligibility is the number of sampled addresses with unknown occupancy/eligibility status, and e is the estimated proportion of

these addresses that are eligible. The proportion e can be estimated from the sampled addresses where occupancy/eligibility has been established. For the National Study we expect to mail to all sampled addresses so that the delivery status (and hence occupancy) of each address should be known.

All households in the telephone sample will be a subsample of those that returned the mail survey. Since these are, by definition, occupied, the final telephone response rate will be computed assuming all of those who we try to get telephone numbers for are eligible.

B.3.2 Maximizing Response Rates

Steps to minimize nonresponse are built into the mail and telephone surveys. As mentioned earlier, the study will take proactive measures to maximize the response rate.

B.3.2.1 Mail Survey

The steps to maximize response to the mail survey include the following:

- **Household Letters.** The letters will describe the study's sponsor, goals and objectives and will give assurances of confidentiality. Letters will be sent with each survey that is mailed to the household.

- **Multiple Followups.** All sampled addresses will receive a thank you/reminder postcard approximately one week after the initial survey mailing. If a survey has not been received after the postcard, up to two additional mailings will be sent.

- **Use of $2 incentive.** As discussed in Part A, we will include a $2 incentive in the first questionnaire mailing to the household.

- **Use of Express Delivery.** The second mailing of the questionnaire will be sent express delivery. The use of special delivery methods is an accepted method for increasing response (Dillman, Smyth, and Christian 2008).

- **Sending a Spanish questionnaire at each mailing.** Including Spanish materials will help elicit response from household where

Spanish is the dominant language spoken (Brick, Montaquila, Han, and Williams 2012). This is especially important for sampled airports with a high concentration of Spanish or Hispanic households in the surrounding sampled area.

B.3.2.2   Telephone Survey

The initial step for the telephone survey will be to match the address to a phone number. Those households that do not match will be sent a survey requesting they provide a telephone number for purposes of participating in the telephone interview (Appendix F). To maximize the response to this survey, we will promise a $10 post-paid incentive (Appendix D) and sent a thank-you/reminder postcard (Appendix E). If necessary, a follow-up mailing will be sent to those that do not initially respond (Appendix D).

Those that do match to a telephone number will be sent an advance letter that describes the study's goals and objectives and will give assurances of confidentiality. A promise of a $10 post-paid incentive will be mentioned in the letter to promote responses (Appendix G).

To maximize response to the telephone survey, interviewers with experience conducting telephone surveys will be used to the extent possible, and multiple callbacks will be made using an automated call scheduling system. Interviewer training will focus on gaining cooperation in the first minute or so of the initial contact with a potential respondent. To maximize the contact rate, we will use a calling algorithm that handles all dimensions of call scheduling, including time zone (respondent and interviewer); skill level of interviewer, appointments; call history; and priority of case handling.

The data collection design includes up to 7 call attempts for the matched phone numbers to determine whether the telephone number reaches the sampled household; if there 7 non-contacts (across a variety of times and days) then the sample record will be closed as final nonresponse. If someone is reached and the phone number is determined to be an invalid match for the sampled address, then we will send a request to the household to provide their phone number. Telephone numbers provided by the household and matched phone numbers that are determined to belong to the correct

household will be called until we obtain a completed interview or until the household refuses to participate. Those respondents deemed to be hostile to the survey request will be coded out as a final refusal. For the non-hostile respondents who refuse to participate, a specially-trained interviewer will recontact and attempt to convert the refusal to an interview. Those respondents who refuse twice will be coded as final nonresponse. Refusal conversions will be attempted at both screening and main interview stages. Telephone interviews will be conducted in English only.

B.3.3 Addressing Non-Response

In addition to efforts to maximize response rates, we will use three approaches to identify and adjust for nonresponse bias.

First, we will use logistic regression to model response propensity as a function of the airport, noise exposure, survey mode (mail vs telephone), whether the address had a matching phone number, and demographic characteristics of the block containing the sampled address, using Census 2010 data since this is the only data available for both respondents and nonrespondents. Examples of block demographic factors are percent of the population age 50 and up, percent black nonHispanic, percent Hispanic, percent male, percent renters, and household size. This was done in the ACRP pilot study and showed that the following factors are associated with having a higher response rate: living near Airport 1 or 2 (as opposed to Airport 3), having a matching telephone number, living in a census block that has a high percentage of persons age 50 and over, and living in a census block that has a low percentage of Hispanics. Similar analysis for the current collection has not been completed. The model showed that households that received the telephone survey were only one-fifth as likely to respond as households receiving the mail survey. Notably, the noise exposure level, as measured by DNL, was not significantly associated with the probability of responding to the survey.

In the ACRP study, models of propensity to be annoyed as a function of demographic covariates, noise exposure, airport, and survey mode showed that persons age 50 and up were significantly more likely to be highly

annoyed than persons below 50, but there was no evidence that the percent highly annoyed differed between the telephone and mail surveys. In the National Airport Study, we will similarly identify factors associated with both response propensity and the probability of being highly annoyed as potential causes of nonresponse bias.

Second, if there is evidence of nonresponse bias among survey respondents we will employ statistical weighting adjustments to help compensate for differential interview response rates. The base weights will be adjusted within strata for nonresponse by cells within noise strata formed from variables correlated with interview response rates, as identified by a logistic regression model, such as airport, survey mode, noise exposure, and Census block demographic characteristics (e.g. high percent age 50+ population, high percent minority population, high percent renters, percent households with 2+ adults). Nonresponse adjustment factors are designed to reduce the potential bias caused by differences between the responding and non-responding population. The adjustment factors are calculated as the reciprocal of the weighted response rates for the adjustment cells.

Poststratification of the nonresponse-adjusted weights by noise stratum will also be considered, using Census population control totals for the geographic area covered by the noise strata surrounding each airport. A difficulty is that the set of Census blocks that include all the sampled addresses do not necessarily correspond closely to the geographic area covered by the noise strata surrounding an airport. The percent of addresses in each block that are included in each noise stratum could be used to construct Census control totals that are a better geographic match, but unfortunately this data is not available from the address frame vendor. The percent of each block's geographic area that is included in each noise stratum is available, but as the population is not necessarily evenly distributed throughout the block the utility of this information will be investigated.

Third, we will compare the respondents for the National Study with the distribution of the population for the set of blocks covering the noise strata by airport, using the 2010 Census as the source of demographic data for the set of blocks and the respondents' reported characteristics from the

interview. The comparison would be done three ways: 1) unweighted, 2) using base weights, and 3) using the nonresponse-adjusted weights (see example below). Confidence intervals around the respondents' estimated proportions for demographic categories (using the nonresponse-adjusted weights) will give an indication of whether the respondents differ significantly from the population in the noise strata surrounding each airport. This was done in the ACRP pilot study and showed that the set of respondents for both the mail and telephone surveys were disproportionately likely to be white non-Hispanic and age 50+ when compared with the Census population. However, the distribution of mail survey respondents was closer to the population distribution than the telephone survey respondents.

Table B-8.   Example table for comparing National Study respondents with 2010 Census population

Mail Survey	Unweighted	Base Weights	Nonresponse-Adjusted Weights	95% CI	2010 Census Percentage
Airport 1					
Percent Male					
Percent Hispanic					
Percent Black, nonHispanic					
Percent Age 50+					
Percent HH with 2+ adults					
......					
All					
Percent Male					
Percent Hispanic					
Percent Black, nonHispanic					
Percent Age 50+					
Percent HH with 2+ adults					

Non-response for the telephone survey will be evaluated using the results of the mail survey. The fact that the telephone sample is a subset of the mail sample means that results from the mail survey can be used to evaluate potential nonresponse bias from the telephone survey. If there is evidence of nonresponse bias, an additional set of weights will be developed for the telephone respondents, using a two-phase procedure that starts with the mail respondents' weights and then adjusts for the additional second-phase nonresponse from the telephone survey.

B.4   Test of Procedures or Methods to be Undertaken

The questionnaires and data collection procedures used in this study are based on a pilot study conducted through the National Academy of Sciences (ACRP Project 02-35; Research Methods for Understanding Aircraft Annoyance and Sleep Disturbance, Task 15 – Final Report; publication expected July 2014). The purpose of this project was to develop a methodology for collecting and analyzing data for estimating the dose-response curve relating noise to annoyance. This study was conducted in

communities near 3 airports and compared two different strategies to collect the data:

1. A 20 minute telephone survey using an address-based sample frame (ABS). An ABS frame was used because of the need to map households within highly specific noise strata surrounding the airport. Households where a telephone number could not be found were sent a mail request to provide one.

2. A short mail survey sent to a sample of households. This survey consisted of approximately 25 questions, including the annoyance question used in estimating the dose response curve.

The questionnaires were developed after an extensive review of the literature on airport noise and its relationship to annoyance. The final questionnaires were reviewed and modified based on comments from the NAS panel.

The study found the mail survey to have a significantly higher response rate than the telephone survey (35% vs. 12%). Interestingly, there were no significant differences in estimates of the dose-response relationship between the two methods of data collection, even though their response rates were much different.

Based on these results, a national study was designed that relied on a mail survey to collect the data necessary to estimate the dose-response relationship. We expect the response rate for the national mail survey will be enhanced by several factors discussed above in Section B.3, including use of express delivery for the second mailing of the questionnaire, an additional mailing to households and mailing a Spanish questionnaire. The pilot study also did not include any distinct government sponsor, as will be the case for the national study. The national study will reference the DOT as the sponsor, which should increase the overall response rate (National Research Council, 1979).

The national design includes following up with households that respond to the mail component to complete a telephone interview. This interview asks additional questions about the respondent's personal and environmental issues that have been hypothesized to be related to annoyance. These data

will be used to explore the correlates of annoyance. There are several advantages to using the mail respondents for the telephone sample, rather than having two separate samples. First, and most importantly, in some airport communities there may be relatively few addresses in the population at high noise exposures. It is important for the precision of the estimated dose response curve to be able to have sufficient sample size for the residents at high noise exposures. Using the respondents from the mail survey for the telephone survey allows the telephone survey to have representation from the high noise strata which otherwise might not be possible (because those addresses would be "used up" for the mail survey). Second, it provides a way of evaluating nonresponse bias in the telephone survey because the mail responses are known for telephone nonrespondents as well as respondents. And third, the mail survey can be used to leverage and augment the precision obtained from the telephone survey because of the additional information available for the full set of mail respondents.

B.5 Individuals Consulted on Statistical Aspects and Individuals Collecting and/or Analyzing Data

FAA project lead and other technical lead
Natalia Sizov
Physical Scientist
Federal Aviation Administration

Nick Miller
Principal Investigator
Senior Vice President
Harris Miller Miller & Hanson Inc.

James Fields
Review of Annoyance Survey Literature
Development of Survey Instruments
Consultant

Sharon Lohr
Statistician
Westat

David Cantor
Westat
Survey Methodogist

Eric Jodts
Westat
Survey Operations

References

Abdelbasit, K.M. and Plackett, R.L. (1983). Experimental Design for Binary Data. *Journal of the American Statistical Association,* 78, 90-98.

Brink, M., Wirth, K., Schierz, C., Thomann, G., and Bauer, G. (2008). Annoyance responses to stable and changing aircraft noise exposure. J. Acoust. Soc. Am. 124, 2930-2941.

Brick, J.M., Montaquila, J.M., Han, D. and D. Williams (2012) "Improving Response Rates for Spanish-Speakers in Two-Phase Mail Surveys" *Public Opinion Quarterly* 76(4): 721 – 732.

Chaloner, K. and Larntz, K. (1989). Optimal Bayesian Design Applied to Logistic Regression Experiments. *Journal of Statistical Planning and Inference,* 21, 191-208.

Church, A. (1993) "Estimating the effect of Incentives on mail survey response rates: A Meta-Analysis." *Public Opinion Quarterly* 57: 62-79.

Connor, W., and Patterson, H. (1976). Analysis of the effect of numbers of aircraft operations on community annoyance, NASA CR-2741, National Aeronautics and Space Administration, Washington, D.C., pp. 1–93.

Demidenko, E. (2004). *Mixed Models.* New York: Wiley.

Dillman, D., Smyth, J.D. and L.M. Christian (2008) *Internet, Mail, and Mixed-Mode Surveys: The Tailored Design Method*) John Wiley & Sons: Hoboken, NJ.

Edwards, P., R. Cooper, I. Roberts, and C. Frost, 2005. "Meta-analysis of randomised trials of monetary incentives and response to mailed questionnaires." Journal of Epidemiology & Community Health 58: 987-999.

Edwards, Phillip J., Ian Roberts, Mike J. Clarke, Carolyn DiGuiseppi, Reinhard Wentz, Irene Kwan, Rachel Cooper, Lambert M. Felix, and Sarah Pratap. 2007. "Methods to Increase Response to Postal and Electronic Questionnaires (Review)." *Cochrane Database of Systematic Reviews* 3: 1-12.

Federal Aviation Administration (FAA) (2007). *Environmental Desk Reference for Airport Actions.* Washington DC, Federal Aviation Administration, available at http://www.faa.gov/airports/environmental/environmental_desk_ref/.

Federal Interagency Committee on Noise (FICON) (1992). Federal Agency Review of Selected Airport Noise Analysis Issues. In *Final Report: Airport Noise Assessment Methodologies and Metrics*, Washington, D.C.

Fidell, S., Horonjeff, R., Mills, J., Baldwin, E., Teffeteller, S., and Pearsons, K. (1985). Aircraft noise annoyance at three joint air carrier and general aviation airports. *J. Acoust. Soc. Am*. 77(3), 1054–1068.

Fidell, S. and Silvati, L. (2004). Parsimonious Alternatives to Regression Analysis for Characterizing Prevalence Ratings of Aircraft Noise Annoyance. *Noise Control Eng. J.,* 52, 56-68.

Fidell, S., Mestre, V., Schomer, P., Berry, B., Gjestland, T., Vallet, M., and Reid, T. (2011). A first-principles model for estimating the prevalence of annoyance with aircraft noise exposure. *The Journal of the Acoustical Society of America*, 130, 791-806.

Fields, J. M. (1991). An updated catalog of 318 social surveys of residents' reactions to environmental noise (1943-1989), NASA Contractor Report No. 187553, Contract NAS1-19061, pp. 1–169.

Groothuis-Oudshoorn, C. and Miedema, H. (2006). Multilevel grouped regression for analyzing self-reported health in relation to environmental factors: the model and its application. *Biometrical Journal*, 48, 67-82.

Janssen, S., Vos, H., van Kempen, E., Breugelmans, O., and Miedema, H. (2011). Trends in aircraft noise annoyance: The role of study and sample characteristics. *J. Acoust. Soc. Am*. 129, 1953-1962.

Kaufman, J.S., Dole, N., Savitz, D.A., Herring, A.S. (2003). Modeling Community-level Effects on Preterm Birth. *Annals of Epidemiology,* 13, 377-384.

Lohr, S. (2014). Design effects for a regression slope in a cluster sample. Accepted for publication in *Journal of Survey Statistics and Methodology.*

Manfreda, K. L., Bosnjak, M., Berzelak, J., Haas, I., and Vehovar, V. (2008). Web surveys versus other survey modes: A meta-analysis comparing response rates. *International Journal of Market Research,* 50, 79-104.

Mercer, A., Caporaso, A., Cantor, D. and R. Townsend (2014) "Monetary incentives and response rates in household surveys." Paper to be presented at the annual meeting of the American Association for Public Opinion Research, Anaheim, CA, May 15-18, 2014.

Miedema, H. and Vos, H. (1998). Exposure-response relationships for transportation noise. *J. Acoust. Soc. Am.* 104, 3432-3445

Millar, M.M. and Dillman, D.A. (2011). Improving response rates to web and mixed-mode surveys. *Public Opinion Quarterly* 75(2): 249 – 269.

National Research Council (1979) *Privacy and Confidentiality as Factors in Survey Response.* Washington, DC: National Academy Press.

Rizzo, L., Brick, J.M., and Park, I. (2004). "A Minimally Intrusive Method for Sampling Persons in Random Digit Dial Surveys," *Public Opinion Quarterly,* 68, 267-274.

Royall, R. (1976). Current Advances in Sampling Theory: Implications for Human Observational Studies. American Journal of Epidemiology, 104, 463-474.

Schultz, T.J. (1978). Synthesis of social surveys on noise annoyance, *J.Acoust.Soc.Am*, vol. 64, No. 2, pp. 377-405.

Scott, A.J. and Holt, D. (1982). The effect of two-stage sampling on ordinary least squares methods. *Journal of the American Statistical Association,* 77, 848-854.

Singer, E., Van Hoewyk, J., Gebler, N., Raghunathan, T., and McGonagle, K. (1999). The effect of incentives on response rates in interviewer-mediated surveys. *Journal of Official Statistics* 15: 217-230.

Tillé, Y. (2011). Ten Years of Balanced Sampling with the Cube Method: An Appraisal. *Survey Methodology,* 37, 215-226, available at http://www.statcan.gc.ca/pub/12-001-x/2011002/article/11609-eng.pdf.

Williams, D., Brick, J.M., Montaquila J.M., and Han, D. (2014): Effects of screening questionnaires on response in a two-phase postal survey, *International Journal of Social Research Methodology*, DOI: 10.1080/13645579.2014.950786

Valliant, R., Dorfman, A.H. and Royall, R. (2000). *Finite Population Sampling and Inference: A Prediction Approach*. New York: John Wiley & Sons.

Yates (1946). A Review of Recent Statistical Developments in Sampling and Sampling Surveys. *Journal of the Royal Statistical Society*, A, 109, 12-43.