

**SUPPORTING STATEMENT**  
**U.S. Department of Commerce**  
**International Trade Administration**  
**Survey of International Air Travelers**  
**OMB Control No. 0625-0227**

**SUPPORTING STATEMENT PART B**

Agencies are instructed to complete Supporting Statement Part B if they are using statistical methods, such as sampling, imputation, or other statistical estimation techniques; most research collections or program evaluations should also complete Part B. If an agency is planning to conduct a sample survey as part of its information collection, Part B of the ICR supporting statement must be completed, and an agency should also complete relevant portions of Part B when conducting a census survey (collections that are sent to the entire universe or population under study). For example, an agency doing a census of a small, well-defined population may not need to describe sampling procedures requested in Part B, but it should address what pretesting has taken place, what its data collection procedures are, how it will maximize response rates, and how it will deal with missing unit and item data.

Agencies conducting qualitative research studies or program evaluations, including case studies or focus groups, should also complete the relevant sections of Part B to provide a more complete description of the use of the information and the methods for collecting the information.

**B. Collections of Information Employing Statistical Methods**

1. Describe (including a numerical estimate) the potential respondent universe and any sampling or other respondent selection method to be used. Data on the number of entities (e.g., establishments, State and local government units, households, or persons) in the universe covered by the collection and in the corresponding sample are to be provided in tabular form for the universe as a whole and for each of the strata in the proposed sample. Indicate expected response rates for the collection as a whole. If the collection had been conducted previously, include the actual response rate achieved during the last collection.

NTTO Response:

The population of potential respondents consists of all international air passengers who are traveling on participating airlines whose trip originates in the United States or includes the United States in their itinerary. There are essentially two separate populations (universes): 1) non-U.S. residents originating (inbound) from foreign countries to the U.S. and 2) U.S. residents originating (outbound) travel to international destinations.

In 2023, the following were actual counts of the two populations:

<b>2023</b>	<b>Non-U.S. Resident (Inbound)</b>	<b>U.S. Resident (Outbound)</b>
Overseas	31,468,000	48,963,000
Mexico-Air	3,016,000	13,399,000
Canada-Air	10,571,000	4,608,000
Total Air Travelers	45,054,000	66,970,000
# Households (~1.6 pax per)	28,159,000	41,856,000

The sample is designed based on the geographic detail desired (all foreign countries to/from U.S. destinations or points of origin) for the resulting estimates and the specific airlines willing to participate in the survey.

The design is a **stratified**, two stage **cluster** sample, in which scheduled flights are **randomly** selected from strata defined by airline and foreign destination in the first stage. The responding travelers on each flight constitute the second stage of the sample. The selection process accounts for the anticipated passenger traffic from the various geographies. For example, we would survey more flights to Japan than flights to Russia. When the SIAT is conducted on a selected flight, whether in the boarding gate area or on-board, those passengers who respond, are considered to represent all passengers on that flight.

NTTO has access to an industry database, via third party, that contains all scheduled international flights, between the U.S. and international locations, for the next 11 months. We download a file, for the upcoming month, of flights departing the U.S. A random selection routine picks flights (clusters) representing all international regions in proportion to the amount of passenger traffic.

A flight kit (containing questionnaires, instructions, tally cards) is prepared by our contractor. These are dispatched to our field service sub-contractors at the 27 U.S. gateway airports, or to the select airlines that conduct the survey in-flight.

For the selected flights (i.e., British Airways Flight 123, from Dulles IAD to London Heathrow), field service personnel ask passengers in the gate area, starting about 2 hours before departure, if they would consider taking the survey,

In summary, the population parameters consist of all passengers, selected

- From those flying on one of 105+ international airlines (both U.S. and foreign based carriers)
- Departing from one of the top 28 U.S. gateway airports

Note: Respondent surveys are conducted for both populations (Resident originating and non-resident returning home) upon departure from the U.S. In other words, every passenger in the gate area, or on-board the flight, are available to be surveyed.

**Response Rates:** the method of survey distribution/collection has evolved since its inception in 1983. Originally, all surveys were administered on-board during flight. Consequently, the SIAT was known by the brand name “In-Flight Survey”. This approach took considerable cooperation from the airline’s airport management and the flight crew (flight attendants). Flight kits (containing questionnaires and related documents) were sent from NTTO’s contractor to the airline’s service manager at the airport. The service manager delivered the kit to the lead flight attendant. The flight crew were to distribute and collect the surveys during the flight. Some carriers, notably Mexicana Airlines, assigned an extra crew member to the flight to ensure this survey was administered!

While we could record the number of completed surveys there was reliance on the flight attendant to report the response rates. For example, “x” number were handed to passengers and “y” were completed.

The in-flight approach had/have numerous logistics and personnel challenges:

The survey ‘kit’ did not make it on the selected flight for various reasons (flight did not operate, airline service manager did not board kit on flight). Also,

The flight crew fails to distribute or collect the questionnaires. This has been a particular challenge from the early 2000's due mostly to contract issues. No clauses covering the collection of federal surveys were ever agreed upon during collective bargaining between airline companies and unions. While the airlines are interested in the information collected on their passengers, airline managements have asked that NTTO arrange for survey administration to take place in the boarding gate area and not on-board. (A handful of foreign airlines still cooperate with the in-flight approach).

Since most surveys are distributed/collected in the pre-departure gate area records indicate that the passenger response rate was over 70 percent. Much better than in-flight which averaged 33 percent. **Overall, the average was 66 percent.** This is not 80 percent as mentioned in the guidance, however given the nature of the population (voluntary, no identification – therefore no means to contact pre or post flight, language/cultural challenges, limited time at the gate) an 80 percent response rate may not be practical.

NTTO is concerned with maximizing response rates, however our goal, or target, for completed surveys is contingent on available funding and the mandated goal (from the Travel Promotion Act of 2009) of one percent of inbound travelers. Consequently, we have casted a 'wider net', selecting/surveying more flights to gather enough surveys to reach the overall annual target, i.e. 110,000 surveys.

## **2. Describe the procedures for the collection of information including:**

### NTTO Responses:

- Statistical methodology for stratification and sample selection,

As mentioned in B-1, above, the sample is designed around the geographic detail desired for the resulting estimates and the specific airlines willing to participate in the survey. The design is a stratified, two stage cluster sample, where scheduled flights are randomly selected from strata defined by airline and foreign destination in the first stage. The responding travelers on each flight constitute the second stage of the sample. When the SIAT is conducted on a selected flight, those passengers who respond are considered to represent all passengers on that flight.

- Estimation procedure,

The primary data sources for computing estimates (Expanded Estimates) are the SIAT responses given that there are two separate populations of travelers, 'Inbound' and 'Outbound'.

### **Estimation and Reliability of Results for Non-Residents (Inbound)**

The survey responses are the primary data source for computing estimates. Information developed from the DHS I-94 reports is also used.

The survey responses provide information on distributions of variables and relationships among survey items as well as specifics relating country of residence and port of customs of the respondent. The DHS I-94 data provide similar information for country of residence, port of customs and mode of arrival. The Survey data are weighted based on the I-94 data by air-mode of arrival.

A weight is calculated for each survey respondent. It is defined as the number of passengers, departing (returning to country of residence) from the United States via

scheduled international air carriers, that is represented by the respondent. Calculation of the weight is a multi-step process.

1. The initial weight of a respondent is one ('1'), unless children are part of his or her travel party, in which case, the initial weight has a value greater than one, depending on the number of children and the size of the travel party.
2. Both the I-94 data and survey responses are sorted and summarized by country of residence and port of customs information.
3. The weight computed for individual survey responses is the result of directly proportioning the I-94 data to the surveys.
4. The weights determined by the limiting variables in the survey responses match the corresponding control totals from the I-94 data summarized in the same manner. The weights are then used in standard weighted ratio estimation formulas for calculating the distributions, means, and medians found in the published tables.

Because of the multistage nature of the sample design and the resulting computational burden, sampling variability has not been calculated for the estimates. Instead, the reliability of a set of related estimates is indicated by the number of respondents to the relevant questionnaire items. The more respondents, the more reliable is the estimate. Judgment must be used in deciding on the degree of confidence to place in an estimate, and in its proper use. Likewise, non-sampling (response and processing) errors have not been estimated, but may be significant, especially when combined with sampling variability. Response errors may be of some significance due to inaccuracies arising from language translations and currency conversions.

CIC Research, Inc., the contractor for the SIAT, has worked with a university statistician to estimate the effects of sampling variability and non-response errors. These issues would help accurately reflect the **reliability and validity** levels of data produced. The results are summarized in the next section.

### **Estimation and Reliability of Results for U.S. Residents (Outbound)**

The SIAT responses are the primary data source for computing estimates. Information developed from the Department of Homeland Security APIS (Advance Passenger Information System) is used in conjunction with SIAT responses.

The SIAT responses provide information on distributions of variables and relationships among survey items as well as information relating the port of debarkation to the residence of the passenger. The DHS APIS (I-92) data provide total passenger volumes by port of debarkation and the number of U.S. and non-U.S. citizens.

A weight is calculated for each survey respondent. It is defined as the number of passengers, departing from the United States via scheduled international air carriers, represented by the respondent. Calculation of the weight is a multi-step process.

1. The initial weight of a respondent is one ('1') unless children are part of his or her travel party.

In this case, the initial weight has a value greater than one, depending on the number of children and the size of the travel party.

2. Although there is non-response on each flight surveyed, the respondents are considered a random sample of the passengers, and each weight is increased to cover non-responses on the flight.

3. Each weight of a respondent in a stratum is increased to represent all travelers on all flights on the stratum.
  4. The APIS (I-92) data are incorporated into the weights by port of debarkation to represent not only the participating, but also the nonparticipating airlines in the survey.
- Degree of accuracy needed for the purpose described in the justification,

A user of statistical data is concerned with the following three questions pertaining to the 'correctness' of reported data. Are the data 1) valid, 2) accurate and 3) reliable? Applied to the SIAT estimates, we can ask those same questions:

**Are the Survey data accurate? Do the results compare favorably to known and proven outcomes? (see below, on questionnaire content and design)**

Currently there is no other known public method that produces a comparable outcome (Which is a reason for conducting the SIAT given the demand for travel data). However, the 'raw' (unweighted) Survey data is modified, through a weighting methodology, to comply with the known accurate census of non-resident arrivals to the U.S., the DHS I-94 Arrival Record and U.S resident departures via the DHS I-92. Also, based on a university study, (<https://travel.trade.gov/research/programs/ifs/Synopsis%20SIAT-DB1B%20Results.pdf>), \*\* this site is under construction \*\*international airfares reported by the SIAT strongly correlate with international airfares reported by the U.S. DOT Origin-Destination Survey.

**Are the SIAT data valid? Does the Survey measure what it is intended to measure?**

The Survey measures what it is designed to measure, namely the characteristics and perceptions of international air travelers to the U.S. (non-U.S. residents) and U.S. residents traveling abroad. Questions are focused on traveler demographics, O&D, planning, activities, spending, etc. To ensure its effectiveness the SIAT is administered in the international airport departure boarding area, a 'target rich' environment with a 'captive' audience.

**Are the Survey data reliable? Are the data consistently good in quality or performance; able to be trusted? What are the results of a Variance Analysis?**

Through general observations over time the Survey produces very stable results where we have robust sample sizes, i.e., at the national level and from the top 20 origin countries producing visitors. It is at the sublevel of country/destination, the origin-destination level, that the VOLUME estimates get 'jumpy', when compared year-over-year due to smaller sample sizes and normal sampling error.

To advance the concept of **reliability** beyond the observation stage, NTTO/CIC Research have implemented an optional task (3.5.11) 'Sample Reliability Estimates' from the current SIAT BPA contract. Initial work was done under the prior contract

(Task order 14-310) to determine the feasibility of moving forward with this effort. The endgame is to determine the 'standard error' of critical variable estimates using Variance Analysis.

CIC Research personnel worked with an Iowa State University professor of statistics to produce sample reliability estimates for statistical values appearing in published and custom SIAT reports. An industry standard methodology, known as the 'Jackknife Method', is applicable to systems involving cluster sampling. The SIAT utilizes cluster sampling in that it randomly selects flights to be sampled.

The initial production results appeared starting in the 2020 Non-resident inbound reports of visitation to U.S. destination, from Table 24 ("What U.S. Destinations did you visit (includes main destinations)). In summary, 95% confidence interval estimates were developed. For example, in the **2023 non-resident national report** :

**TABLE 24 - Q3c./Q17. What U.S. Destinations did you visit (includes the main destination)? \*\* (%)**

New York City-WP-Wayne***	28.25%
(95% confidence interval)	(26.91% - 29.59%)

95% Confidence Interval reflects two (1.96) standard deviations (standard errors) from the mean, as follows:

0.68% (one standard error, taken from tab 'Std. Error' tab of

National Report)

$\times 1.96$  (expand from one to two standard errors)  
 = +/-1.34% (95% Confidence Interval)

As applied to the point estimate of 28.25 %

$28.25\% - 1.34\% = 26.91\% \times 31,468,000$

$28.25\% + 1.34\% = 29.59\% \times 31,468,000$

Estimated # visitors	From 8,468,000
	to 9,311,000

We are 95% certain that the population estimate lies within the range from 8,468,000 to 9,311,000 with 8,890,000 as the point estimate.

To determine the **Error % of the Estimate**,  $1.34\% / 28.25\% = 4.74\%$ , a/k/a the '**margin of error**' from the point estimate (mean).

NYC summary:

- The error % of the estimate for **two standard errors** is 4.74%. (95% confidence)

Note: Reliability estimates have been developed for 19 additional variables, including spending, for the period 2012 - 2023 and will continue throughout the contract period (2024).

Preliminary consensus: **Increased sample size (n) correlates with increased reliability.**

See attached Appendix for explanation of the Jackknife Methodology.

### **Contributing to 'Accuracy' is Questionnaire Content and Design:**

The questionnaire development was guided by the normal standards of questionnaire design to encourage the maximum response by the surveyed passengers. The questions are stated as simply and clearly as possible, and definitions of possibly confusing terms are provided on the forms. To reach most non-English speaking travelers, the questionnaire was translated into eleven additional languages (Arabic, Chinese, French, German, Italian, Japanese, Korean, Polish, Portuguese, Russian and Spanish).

For the 2012 questionnaire change, input was solicited from the travel industry and other agencies of government. NTTO received guidance and advice from the U.S. Census Bureau as to questionnaire design and question verbiage. Draft versions were field tested and foreign language versions were translated and 'back-translated' by language specialists.

In printed form, there is an English only version and eleven versions with the English version on the first half followed by the foreign language version. An announcement at the top of each form tells the respondent of the availability of the other versions. Thus, the questionnaire has been designed to minimize the language obstacle that might discourage a passenger from responding. Both resident and non-resident questions are included in the one survey instrument. The language versions added were a requirement of the foreign flag carrier's entry into the program. Without it, the airlines felt they would not obtain a representative sample of their passengers.

The questionnaire design facilitates easy distribution and collection by eliminating the necessity for the field contractor to determine the citizenship of the passengers. They need only give one form to every adult passenger in the boarding area. Likewise, the in-flight survey method is facilitated since the flight attendants are not required to determine the passenger's citizenship. Response to this survey is dependent on the flight attendants' ability to distribute and collect the questionnaires in a timely fashion.

- Unusual problems requiring specialized sampling procedures, and

The selection process described earlier was focused on 'scheduled' flights in published timetables. Scheduled flights account for about 99 percent of international air travel. Chartered flights, carrying more than 30 passengers, account for circa one percent. Many of these ad-hoc flights did serve Orlando

Sandford International Airport. Special procedures had been developed to pinpoint these flights (mostly to Europe) so they can be sampled.

- Any use of periodic (less frequent than annual) data collection cycles to reduce burden.

A bi-annual collection, every two years, would distort accurate and timely reporting

**3. Describe methods to maximize response rates and to deal with issues of non-response. The accuracy and reliability of information collected must be shown to be adequate for intended uses. For collections based on sampling, a special justification must be provided for any collection that will not yield "reliable" data that can be generalized to the universe studied.**

NTTO Response:

The SIAT survey is voluntary, both on the part of the airline and the passenger. Basically, all international airlines (99+%) serving the United States are cooperative in the process (they grant access to departure gates for our field service; and a handful of Asian carriers participate in in-flight). To ensure enhanced response rates by passengers will require a compelling reason to respond to the questionnaire.

The issue of how to deal with non-response (over 30%) is elusive.

Recall that the sampling unit (cluster) is the departing flight, a **perishable** item.

Q. (Field Service) Would you like to take the U.S. DOC survey about your travel experience?

Note: introductory paragraph on front of questionnaire, inviting participation...

R1. Yes, or

R2. No.

Hypothetical follow up:

Q. Could I ask why not?

R.2... "I don't have time", "I'm tired of surveys", "I don't speak English", "What's the point?", "I'm wary of the government", "Privacy concerns"; "Covid-19"; "Your QR code might have a computer virus in it"; "I just don't want to:", etc.

Given that the survey doesn't not ask for names, or any PII, there is no effective means of follow up or advance contact of passengers.

Consequently, we have **focused on increasing the total number of surveys** allowable by budget constraints. These are still subject to the discipline of random selection based on volume of expected passenger traffic, by carrier and U.S. gateway. For example, more surveys are expected at NY JFK (the largest U.S. international gateway) than at Baltimore (BWI). And more surveys are expected from larger airlines (American Airlines) than smaller carrier (Azul, from Brazil). And thirdly, more surveys



are expected from Europe than Oceania.

As described in the above section on 'Reliability' we are now able to quantify the standard error(s) associated with a point estimate and provide assurance (95%) regarding the estimate. We have tentatively **concluded that an increased sample is associated with a lower error rate.**

Note the following example regarding estimated overseas visitation to New York City:

Table-24	2023	2020 (Covid-19 in late March)
Total overseas travel to U.S.	31,469,000	7,594,000
Total sample (n) to U.S.	33,895	14,273
Visitation to NYC - Point Estimate	28.25% (8,890,000)	21.12% (1,603,852)
95% confidence interval	26.91% (8,468,000) to 29.59% (9,312,000)	18.75% (1,423,875) to 23.49% (1,783,830)
95% confidence interval +/-	+/- 1.34%	+/- 2.37%
Error % of Estimate	1.34% / 28.25% = <b>4.74</b>	2.37% / 21.12 = <b>11.22</b>

Note **the error rate was reduced by 58% by an increase in sample size** from 2020 to 2023. During Covid-19, even with the international travel restrictions, reduction in flights and closed airports, NYC was still in the top two cities for the most visited cities.

With the new variation assessment process, we are now able to report on numerous variables with a quantified degree of reliability.

Back to the question on maximizing response rates, see the next section.

**4. Describe any tests of procedures or methods to be undertaken. Testing is encouraged as an effective means of refining collections of information to minimize burden and improve utility. Tests must be approved if they call for answers to identical questions from 10 or more respondents. A proposed test or set of tests may be submitted for approval separately or in combination with the main collection of information.**

NTTO Response:

NTTO continues to research the possible use of **incentives** to encourage not only an increase in the number of responses, but also an increase in the response rates.

As cited in Part A #9 the use of \$10 Starbucks cards, as an incentive to potential respondents, yielded negligible results.

Given the fact that the SIAT contract is fixed price, is funded by a fixed level of appropriations, and the target number of respondents is always reached, there does not

appear to be room for additional payouts. In other words, if the target of 105,000 (2024) responses are collected what would be the point of incentives, just adding to unit cost.

We invite the participation of OMB desk officers to discuss options that are both feasible and can be justified, as mentioned in Part A #9.

Attempt to 'minimize' burden and improve utility:

NTTO/CIC Research are testing a shorter version of the same questionnaire. (Note: the short version has 23 of 33 questions. The 23 questions are **exactly** the same as in the 'legacy' 33 question version). The hypothesis was that passengers might be more inclined to respond to a short questionnaire. That doesn't appear to be the result. However, respondents might be inclined to respond to more of the questions in the instrument. This is being evaluated.

A quick look at results in 2023. Responses from passengers using the:

- Short (23) questionnaire = average of 10.3 completed surveys per flight
- Long (33) questionnaire = average of 9.6 completed surveys per flight, for example:

British Airways	12.0
Jet Blue	11.2
American Airlines	9.4

It appears that the short form (23 questions) doesn't offer any advantages, per ser.

**5. Provide the name and telephone number of individuals consulted on statistical aspects of the design and the name of the agency unit, contractor(s), grantee(s), or other person(s) who will actually collect and/or analyze the information for the agency.**

NTTO Response:

In addition to public comments reported in Part A, #8, the following gives an anthology of significant contributors to the SIAT program.

- Judith Schwenk of the Transportation Systems Center, U.S. Department of Transportation (Volpe Center) was the original mathematical statistician who developed the survey sampling design and analysis procedures. The Volpe Center can be reached at (617) 494-2000.
- Dr. Reuben Cohen, Senior Vice President of Response Analysis was the statistician responsible for the technical direction of the program from April 1984 through June 1985. Response Analysis is now part of GfK Custom Research <http://www.gfk.com>.
- **Dr. Gordon Kubota**, President of CIC Research, Inc. is the statistician that is responsible for the NTTO program beginning from July 1985 and he may be reached at (858) 637-4000.

Dr. Kubota can also provide contact information for **Dr. Jae-Kwang Kim**, PhD, Professor of Statistics, Iowa State University, Ames, Iowa.

- Jonathan W. Williams, PhD, University North Carolina – Chapel Hill, Associate professor of Economics; conducted studies on correlation between SIAT international airfares and U.S. DOT O&D Survey international air fares. (919-966-5375; [jonwms@unc.edu](mailto:jonwms@unc.edu) )