B. Collection of Information Employing Statistical Methods

1. Describe the potential respondent universe and sampling method used.

Participation in the survey is mandatory for all large U.S. certificated air carriers operating scheduled passenger service. Carriers that are required to report include those which fly at least one aircraft with more than 60 seats, or operate at least one domestic-to-international route. Presently there are 48 air carriers submitting the Survey. The Department requires sampling of 1 percent for the largest domestic markets and a 10 percent sample for all other markets. All reporting carriers have voluntarily elected to submit a straight 10 percent sample. Carriers examine the flight coupons or their electronic records for flown trips for Survey reporting. The following coupons or flown trips are selected:

1. For single flight tickets, those with a serial number ending in zero.
2. All group tickets with 11 or more passengers without regard to the serial number. For group tickets with 10 or less passengers only those with a serial number ending in zero.

The selection of flight coupons ending with a zero provides a simple method for airlines to systematically achieve an unbiased sample of 10 percent of the records. Flight coupon or electronic record numbers are generally generated randomly, thus this system avoids the introduction of any systematic bias in the results.

2. Description of procedures for the collecting information, including statistical methodology for stratification and sample selection, estimation procedures, degree of accuracy needed.

Air carriers shall select a statistically valid sample of records (either flight coupons or electronic records for flown trips) for reporting purposes. The sample shall consist of 10 percent of lifted flight coupons for all large domestic markets listed in the "Instructions to Air Carriers for Collecting and Reporting Passenger Origin- Destination Survey Statistics." Flight coupons are the source of the data to be collected, and the Survey data are taken from the first reportable flight coupon in a ticket lifted by one of the reporting air carriers. (In case of ticketless travel, carriers use their internal management control system for completed trips.) The complete passenger itinerary is recorded from that flight coupon as one entry. The recording of data from the flight coupon is a matter of transcribing the information just as indicated on the ticket. The detail record for each trip reveals the complete routing from the passenger's origin to destination, including in sequence from the origin point, each point of transfer or stopover (interline or intraline), the operating carrier, the ticketed carrier, and the fare-basis code for each flight-coupon stage of the itinerary, and the total dollar value for the entire ticket. The report to the Department is submitted by electronic means.

Estimation and variance estimation procedures are as follows:

Consider a quarterly O&D survey dataset. Let n_g denote the observed number of passengers in the g-th OD pair using the O&D survey and N_g the true number of passengers. Here $g = 1, 2, ...G,$ where G is the number of possible OD pairs. Note that with the current

$$f_g = \frac{n_g}{N_g}$$

sample design of the O&D survey, could be zero, especially for smaller markets, in some quarters. Let y_{gi} denote the price paid by the i-th passenger in the sample and Y_{gi} the price paid by the i-th passenger in the quarter's population of passengers for the g-th OD pair. Then, an estimate of the average ticket price

$$\bar{Y}_g = \frac{1}{N_g} \sum^{N_g} Y_{gi}$$

11* MERGEFORMAT ()

is given by

$$\bar{y}_g = \frac{1}{n_g} \sum^{n_g} y_{gi}$$

22* MERGEFORMAT ()

It can be seen that

$$E(\bar{y}_g \mid n_g) = \bar{Y}_g$$

So, this estimator is conditionally unbiased. Also,

$$E(\bar{y}_g) = EE(\bar{y}_g \mid n_g) = \bar{Y}_g$$

and so the unconditional bias of this estimator is zero. The conditional variance of this estimator is

$$V(\bar{y}_g \mid n_g) = \left(\frac{1}{n_g} - \frac{1}{N_g} \right) S_g^2$$

where

$$S_g^2 = \frac{1}{N_g - 1} \sum_{i=1}^{N_g} (Y_{gi} - \bar{Y}_g)^2$$

The unconditional variance is obtained as follows:

$$V(\bar{y}_g) = EV(\bar{y}_g \mid n_g) + VE(\bar{y}_g \mid n_g)$$

The second term equals 0 and

$$V\left(\bar{y}_g\right) = EV\left(\bar{y}_g \mid n_g\right) = E\left(\frac{1}{n_g}\right)S_g^2 - \frac{1}{N_g}S_g^2$$

Since N_g is unknown, an unbiased estimator of the variance is

$$\frac{1}{n_g}\left(1 - \frac{n}{N}\right)s_g^2 = (0.9)\frac{s_g^2}{n_g}$$

where

$$s_g^2 = \frac{1}{n_g - 1}\sum_{i=1}^{n_g}\left(y_{gi} - \bar{y}_g\right)^2$$

3. Describe the methods to maximize response rates, and describe how the Department deals with non-responses.

The Department contacts delinquency carriers when a report is late filed. The contact may be a telephone call or an email transmission. If no response is forthcoming, then a warning letter is sent to the carrier requesting the data be submitted within the next five business days. If the reports are not received within the 5 day period, the matter is referred to the Assistant General Counsel for Aviation Enforcement and Proceedings. DOT has the authority to fine carriers for each day that Survey report is late without just cause. However, fines and penalties are generally used as a last resort. Overall, the airline industry has an outstanding record for complying with Survey reporting obligations. Occasionally, there may be a delayed response due to a carrier strike or bankruptcy. When a delayed response does occur, the Department will place a notice on the reporting status internet page to alert users that a carrier's data are not included because of the delay.

4. Describe any tests of procedures or methods undertaken.

Carrier reports are electronically reviewed for conformance to instructions, traffic volumes and for various other relationships. Major problems discovered in this review or in later stages of processing are taken up with the carrier and resolved. Reported data are then subjected to a preliminary computerized pre-edit that detects input format problems. The pre-edit also automatically corrects certain obvious errors in the reported city/airport codes, such as the confusion of cities with the same name but in different states. In addition, the pre-edit consolidates the many detailed fare-basis codes reported into six broad categories: (1) Unrestricted First Class, (2) Restricted First Class, (3) Unrestricted Business Class, (4) Restricted Business Class, (5) Unrestricted Coach/Economy Class, and (6) Restricted Coach/Economy Class. In this consolidation, a fare-basis code that does not match in the ADP conversion matrix is automatically changed. Fare-basis codes reported on surface portions of itineraries are removed.

Next, the data are passed through other electronic edits. Each entry is tested to remove duplicates of the same ticket reported by different carriers. Surface-transportation portions at the beginning or end of ticket itineraries are removed. Airport and carrier codes on each flight-coupon stage of the itinerary are tested for validity against the Official Airline Guide electronic files. The carrier on each coupon-stage is tested to determine if it serves the airport of the flight-coupon origin and destination. If the carrier on the coupon does not serve both points, either directly or indirectly, and if the entire ticket has but one flight coupon, the carrier code is automatically changed to that of the reporting carrier when the latter serves both points. If not, the record is dropped. For itineraries with more than one flight coupon, incorrect carrier codes are changed to the code of the monopoly carrier; i.e., the only carrier serving both points. If no monopoly carrier exists, the code is changed to UK (unknown carrier). All records that have been changed by the computer are re-edited.

The computer edit also removes duplication of the same flight-coupon stages reported in juxtaposition in the itinerary, and drops itineraries that have no destination reported. Single-coupon itineraries where the origin and destination are the same are dropped. For itineraries where the carrier(s) do not serve the reported intermediate points, the city is suspected as being misreported and the record is dropped. All records that fail the editing tests (after computer corrections, if any) are dropped from the data. The passenger volume on dropped records is a fraction of one percent of the total number of sample passengers reported by each carrier.

To illustrate:

Edit for Alaska Airlines for 1st Quarter 2012

Tickets with invalid fare codes	0
Tickets with invalid point codes	66
Tickets with surface at start or end	0
Tickets with consecutive duplicate points	4
Tickets with invalid carrier on coupon	2
Tickets with invalid output format	0
Tickets requiring modification	69
Percent of tickets requiring modification	0.08 where 1.00 = 1 percent

Number of tickets in	84,740
Number of passengers in	182,806
Number of identical tickets combined	1
Number of tickets deleted	0
Number of passengers deleted	0
Number of tickets out	84,739
Number of passengers out	182,806
Percent of tickets passing edit	100.00
Percent of passengers passing edit	100.00

Average coupons per ticket   2.15
Number of coupon records   161,046

As can be seen above, less than one tenth of one percent of the coupons surveyed were incorrect. City-pair passengers from the Survey are normalized for comparison to the carrier's T-100 traffic reports as a further check. Significant discrepancies are discussed with the relevant carrier for correction.

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

Dr. Promod Chandhok is a statistician at the Bureau of Transportation Statistics, Office of the Assistant Secretary for Research and Technology, Department of Transportation. He can be reached at (202) 366-2158.

Mr. James Bouse is the contact person for the Survey. He can be reached at (202) 366-4876.