THE METHODOLOGY OF THE CANADIAN AIR SCHEDULED INTERNATIONAL PASSENGER ORIGIN AND DESTINATION ESTIMATION SYSTEM[1]

Greg Hunter and Lisa DiPiétro[2]

The Air Scheduled International Passenger Origin and Destination (ASIPOD) estimation system uses the data from two air traffic surveys to produce origin-destination estimates of international passengers. The "assignment technique" is the solution to the problem caused by the non-coverage of non-interlining traffic. The assumptions of the technique are sufficiently questionable to warrant an evaluation of the bias of the estimates. However, major improvements will be made in the new system which will decrease the bias in the estimates. Also, estimates of reliability will be produced. And as a result, knowledge of the strength of the inferences made with respect to air traffic markets from these estimates will be improved in international bilateral air negotiations.

1. INTRODUCTION

In 1979, Statistics Canada embarked on a revision of the federal aviation statistics program by inviting Transport Canada and the Canadian Transport Commission (i.e. the two "user departments") to form an interdepartmental revision team. The ASIPOD estimation system is one of several projects in the revision program.

The ASIPOD estimation system uses the data from two air traffic surveys to produce estimates of the number of passengers on scheduled international flights between Canadian and foreign markets for various origin-destination combinations. The first of these two surveys, the revenue passenger origin and destination survey, provides a sample of origin-destination data on international journeys with Canadian carriers on at least one leg of the itinerary.

---

[1] Presented at the Joint Statistical Meetings of the American Statistical Association in Cincinnati, August 1982.

[2] Greg Hunter, Business Survey Methods Division, Statistics Canada. Lisa DiPiétro, Transportation and Communications Division, Statistics Canada.

A coverage problem exists, since no data are available on those international journeys with foreign carriers on all legs of the itinerary. The second survey, the airport activity survey, counts all passengers entering or leaving Canada on all Canadian or foreign scheduled carriers without consideration of the passenger's origin or destination.

This paper first outlines the requirements users have for international passenger origin and destination estimates. Then, the relevant aspects of the two air traffic surveys and the non-coverage problem are presented. And, finally, the paper describes how the ASIPOD estimation system will produce estimates of the number of passengers and the associated coefficients of variation for various international origin-destination pairs for the portions of the international market both covered and not covered by the first survey.

2. USER REQUIREMENTS

The users require estimates of international scheduled commercial air service passengers by origin and destination for bilateral air negotiations.

An international scheduled commercial air service is defined to be an operation which is between points in Canada and points in any other country, and which provides public transportation of persons, goods or mail by aircraft in accordance with a schedule and at a toll or charge per unit of traffic. Such a service is referred to as a "unit toll" service.

Before an international scheduled commercial air service can be operated into and out of Canada, some form of formal agreement must exist between the Government of Canada and the government of the second country. The formal agreement between countries may take the form of an interim diplomatic exchange of notes or of a complete negotiated Air Transport Agreement.

International bilateral air negotiations involve officials of the Canadian government from External Affairs, the Canadian Transport Commission, Transport Canada and the Ministry of Industry, Trade and Commerce. Negotiations may last several months or several years.

The routes for scheduled air services are normally the major item for negotiation, but there are many others. Some of the items, or articles written into air transport agreements, may be:

- rights to fly across, or to make stops for non-traffic purposes in, a given territory
- designation of the airline to operate each route
- compliance with laws and regulations of each country, dealing with such issues as entry, clearance, immigration, passports, customs and quarantine
- airworthiness, certificates of competency, and licences
- exchange of statistical information
- tariffs
- transfer of funds
- exemption from taxation of income

In order to negotiate these items, and particularly for an exchange of routes, the negotiating officials must know where air traffic markets are and whether they are growing. Analyses of the costs and benefits to Canada and to foreign countries of various international routes for Canadian and foreign air carriers must be available to the negotiating officials. To ensure that Canada can negotiate a fair market share, the provision of international passenger origin and destination estimates is a necessity. A crude and indirect indication of the value of such data to the Canadian economy is that the revenue[3] generated from all international air routes to and from Canada in 1980 was about 2.3 billion dollars (Canadian).

[3] Taken from tabulations internal to the Canadian Transport Commission.

3. NATURE OF THE NON-COVERAGE PROBLEM

An underlying assumption of the redesign is that the same basic methodology, as implemented in the existing system, is to be used. As a result, most of the feasibility work involved identifying desirable improvements, ranking their desirability and determining how much could be done within cost and time constraints. This "same basic methodology" provided direction with respect to the calculation of estimates of the number of international passengers, but not to the calculation of estimates of reliability of the passenger estimates.

The target population of the international estimation system is the set of all tickets with an international (i.e. between a foreign country and either Canada or the United States) journey. An exchange program on passenger O & D data is maintained between Canada and the United States, whereby the United States gives Canada those records detailing the complete itineraries of the tickets collected in their survey on which:

(i) a U.S. and Canadian point is shown in the routing, or
(ii) a U.S. carrier is recorded as having flown to or from a Canadian point, or
(iii) a Canadian carrier is recorded as having flown to or from a U.S. point.

As a result of this exchange agreement, the expression "foreign" and "foreign (non-U.S.)" are both used in this paper to denote "neither Canadian nor American".

The revenue passenger origin and destination survey collects tickets issued for international journeys, but only major Canadian carriers participate in the survey. Each participating carrier selects a flight coupon on a ticket

with a serial number ending in '0', if that carrier is the first participating carrier to fly on a leg of that ticket. Hence, the survey reports a 10 percent sample of unique flight coupons on which there is at least one participating Canadian or American carrier.

Some information concerning the markets of foreign (non-U.S.) carriers is obtained from the revenue passenger origin and destination survey. For example, if a passenger travels from Ottawa to Montréal with Air Canada, then connects with Air France for Paris, the revenue passenger origin and destination survey will capture the trip because a Canadian carrier participated somewhere in the journey. The Canadian carrier would report the complete carrier and routing detail, including the Air France segment.

The revenue passenger origin and destination survey, however, does not cover coupons with foreign carriers on all legs of an itinerary. An example of such an itinerary would be that of a passenger flying on Air France from Paris to Montréal and then back to Paris on Air France. If this itinerary were the passenger's total journey, this journey would not be reported to the revenue passenger origin and destination survey. However, the itineraries of such passengers are in the target population of international journeys.

This incomplete coverage of the target population is the non-coverage problem for the ASIPOD estimation system. The coverage problem seems to be "non-coverage" as opposed to "undercoverage", since it is not even possible to include a large portion of the universe in the frame.

The existing system takes the revenue passenger origin and destination survey data and the airport activity survey data and applies a method called "the assignment technique" in order to produce total market estimates.

The airport activity survey counts passengers, on a census basis by flight, entering and leaving each Canadian airport. The survey covers all Canadian, American and foreign scheduled carriers, but it does not consider the passengers' initial origin or final destination.

Hence, the airport activity survey provides a count of the total volume of passengers for all carriers in the target population. The assignment technique is a method of estimating the non-coverage volume of passengers and assigning it to origin-destination pairs. However, in order to explain the assignment technique, a somewhat more thorough description of the two air traffic surveys is required.

4. RELEVANT ASPECTS OF THE TWO SURVEYS

The authorizing agency for the two survey programs is the Air Transport Committee of the Canadian Transport Commission, in co-operation with Transport Canada. The data are collected from the air carriers on behalf of the Air Transport Committee by the Aviation Statistics Centre (ASC) of Statistics Canada. Under the authority of the Air Carrier Regulations of the Aeronautics Act, reporting by the carriers on the ASC statements (ie. questionnaires) is compulsory.

The Revenue Passenger Origin and Destination (O & D) statistics are reported to the ASC via Statement 35. The reported data items, among others, include:

- ticket origin and ticket destination

- points of intra- and interlining (i.e. routing)

- carrier on each flight coupon stage

The revenue passenger origin and destination data are submitted monthly by major Canadian unit toll air carriers conducting scheduled passenger services. Since January 1, 1982 the seven Canadian carriers contributing information to this survey have been Air Canada, CP Air, Eastern Provincial Airways, Nordair, Pacific Western Airlines, Air Ontario and Quebecair. The American data are collected by the Civil Aeronautics Board from all certificated United States' air carriers, except helicopter operators and intra-Alaska carriers. The data for the three months of each quarter are combined, and duplicates are

eliminated; so that a file of complete itineraries, Ticket Origin and Destination (TOD) records, is obtained for the quarter.

However, passenger origin and destination statistics are compiled using the Directional Origin and Destination (DOD) concept. The DOD concept can be defined as "points of initial departure and ultimate destination named in the sequence which indicates the direction of travel". DOD's are pieces of itineraries which are broken up such that each component piece defines a reasonably consistent direction. To create DOD's, "open-jaw" and return itineraries, such as "symmetrical" and "circle" itineraries, must be broken into pieces which are essentially one-way trips. To obtain the DOD's, the TOD's are passed through the breakpoint routines. This breakpoint process is automated within the Passenger Origin and Destination System, and involves the calculation of various point-to-point distances within the itinerary and the comparison of these distances to the total itinerary length. As a general rule, itineraries are broken at the farthest point from the origin. Each DOD formed is recycled through breakpoint routines until no further breakpoints can be assigned.

The airport activity data are filed on Statement 32. The relevant items included for each flight are:

- the reporting carrier
- the reporting airport
- the point of origin and final scheduled destination of the flight
- the last station arrived from, for arrivals; or next station departed to, for departures
- the number of deplaned or enplaned revenue passengers

The airport activity data are submitted monthly by Canada's transcontinental

(Air Canada and CP Air) and regional air carriers (Eastern Provincial, Quebecair, Nordair and Pacific Western Airlines), by Norcanair and by all foreign carriers (including American carriers) operating scheduled international flights into and out of Canada. Since January 1, 1982, there have been 10 American and 21 other foreign carriers filing reports for each Canadian airport they served. Each new foreign carrier, granted a licence to serve Canada on a scheduled basis, is automatically included as a participant in the airport activity data collection system.

From the airport activity data, the census traffic flow data are obtained. "Traffic flow" can be defined as a count, over a certain period of time, of the number of persons who are flying on a specific carrier between a Canadian reporting airport and an adjacent point. The adjacent point is called the next stop or the last stop. For the purposes of the assignment technique only the traffic data for foreign (non-U.S.) carriers are input into the system. The data elements extracted from this survey, and used to determine the O & D international markets, are the number of revenue passengers enplaned and deplaned in Canada, the Canadian gateway carrier, and the Canadian gateway. In this survey the concept of "Canadian gateway" is defined to be that reporting airport at which a foreign (non-U.S.) carrier enters or leaves Canada.

However, in the revenue passenger O & D survey, the Canadian gateway for Canadian and U.S. carriers is the first/last Canadian point in the itinerary for a flight entering/leaving Canada. For foreign carriers the Canadian gateway refers to the point inside Canada where the passenger enters or leaves the foreign carrier. Consider the following fictitious example:

Assume that Air France flies Toronto - Montréal - Paris. Some passengers enplaned in Toronto, and some enplaned in Montréal.

Assume also that the matching single crossing DOD's are:

Winnipeg - Air Canada - Toronto - Air France - Montréal - Air France - Paris.

Toronto - Air France - Montréal - Air France - Paris - British Airways - London

The Canadian gateway in the above example would be Toronto because Toronto is the point at which the passengers enter the foreign carrier.

5. THE OBJECTIVES OF THE REVISION

The four main objectives, the fourth of which will be discussed in detail in this paper, are as follows:

(i) to eliminate problems which have been identified in the existing system.

The existing system does not impute for illegible carrier codes on flight coupons, so that "unknown carriers" becomes the third largest carrier in tabulations. Also, there is no check on the coding which indicates whether a carrier is flying to and from airports at which it actually has landing rights. As a result of even existing tabulation requirements, additional edits and imputations, to handle illegible or incorrect data on international flight coupons, are required over and above those required for domestic flight coupons alone.

Other nonsampling errors have been identified, but can not be easily corrected by an estimation system. Examples of such errors are misinterpretation by participating carriers of instructions for selecting flight coupons; systematic errors in serial numbers, used for sample selection, on ticket stock; errors in the carriers' processing systems, etc. The control of these nonsampling errors for which it is not easy to correct is not an objective of this revision.

(ii) to develop a simple computing system which is easy to use, and produces summary diagnostic information.

Some 625,000 origin and destination records, and some 820,000 airport activity records must be processed annually with minimal manual intervention. Since large volumes of passengers are dispersed across a large number of origin-destination pairs, the diagnostics at each stage of the system must summarize the processing, and still be able to point out potential problems.

(iii) to tabulate the international passenger estimates, regularly and on an ad hoc basis, in ways which will simplify the analyses undertaken by users.

(iv) to produce quantifiably reliable estimates of the number of air scheduled international passengers by origin and destination.

Although the reliability of these statistics has been thought to be variable in the past, it has been, in fact, unknown to date. Estimating the reliability of these data will improve the knowledge of the strength of the inferences that can be made from analyses of these data. Inferences made without knowledge of the reliability of the data could actually be quite misleading.

6. SOLUTION OF THE NON-COVERAGE PROBLEM

6.1 Magnitude of the Problem

As in other surveys, non-coverage is a problem for the ASIPOD methodology, since no sample data are available on the origin-destination patterns of the non-coverage portion of the target population.

The following table gives an indication of the volume of passengers travelling between Canada and nine world areas. (These data are 1979 annual estimates. The world areas are not identified because these data are confidential.)

Table 1 - Estimated Number of International Passengers - 1979

Between Canada and ...	Revenue Passenger Origin and Destination	Non-coverage (percentage of Total in brackets)		Total
World Area #1	116,050	495	(0.4)	116,545
World Area #2	325,200	2,020	(0.6)	327,220
World Area #3	99,010	12,628	(11.3)	111,638
World Area #4	67,200	14,558	(17.8)	81,758
World Area #5	575,010	126,076	(18.2)	703,086
World Area #6	205,180	101,464	(33.1)	306,644
World Area #7	1,221,040	908,876	(42.7)	2,129,916
World Area #8	54,410	56,807	(51.1)	111,217
World Area #9	45,330	57,267	(55.8)	102,597
Total World	2,708,430	1,282,191	(32.1)	3,990,621

From the percentage non-coverage figures, it is evident that the non-coverage problem is a major concern.

The same table as above, but between Eastern Canada and the same nine world areas, would tend to have a higher percentage non-coverage for each world area. Hence, a lower level of geographic aggregation in origin-destination pairs generally implies a higher percentage non-coverage. For example, the non-coverage for Eastern Canada to World Area #7 is 55 percent, compared to the 42.7 percent tabulated for all of Canada to World Area #7 (as above). To clarify the idea that a higher level of geographic aggregation in origin-destination pairs generally implies a lower percentage non-coverage, consider the fact that non-coverage exists only for foreign traffic terminating at Canadian gateways. There is complete coverage of all traffic for which the Canadian end of the origin-destination pair is not a Canadian gateway. Hence, as the level of geographic aggregation becomes higher, more and more interlining traffic is included, and the percentage non-coverage becomes lower.

6.2 The Assignment Technique

The assignment technique estimates the non-coverage volume of passengers and then allocates this volume to origin-destination pairs.

From the airport activity survey, b benchmark counts of passengers entering and leaving Canada at Canadian gateway airports are tabulated by carrier. The value of b, the number of assignment groups, can be determined as follows:

$$b = 2 \times \sum_{i=1}^{g} n_i$$

where the '2' accounts for the fact that there is one count each for passengers entering and leaving Canada,

g is the number of Canadian gateway airports, and

n_i is the number of foreign (non-U.S.) carriers with landing rights at the i^{th} Canadian gateway airport.

From the revenue passenger origin and destination survey, a corresponding number of inbound and outbound passengers on international DOD's can be tabulated by crossing carrier and Canadian gateway airport. Hence, there are also b such counts from the sample survey data.

In the first stage of the assignment technique the non-coverage volume, A_i, of passengers in assignment grouping i can be estimated as follows:

$$A_i = C_i - (1/f) \times D_i \quad (i = 1,\ldots b).$$

where C_i is the airport activity census count in assignment grouping i.

f is the revenue passenger O & D survey sampling fraction (i.e. 1/10).

D_i is the sample number of international passengers in assignment grouping i.

A_i is, then, an estimate of the number of passengers, carried on the foreign carrier in the i^{th} of b assignment groupings, for which there is no origin-destination information.

The next stage allocates the non-coverage volume to origin-destination pairs. Such pairs in the non-coverage portion are called non-interlining DOD's, since they are DOD's, flown by foreign carriers, which do not interline with a participating Canadian carrier. The assignment technique imputes non-interlining DOD's in the i^{th} assignment grouping as follows:

(i) All of the DOD's contributing passengers to D_i are identified. (These would be DOD's which match on Canadian gateway, foreign carrier and direction.)

(ii) The domestic portion of such DOD's (i.e. the portion from the Canadian point to the Canadian gateway city) is eliminated. (The domestic portion of such DOD's would be on a Canadian carrier, and, therefore, would be picked up in the revenue passenger O & D survey. The resultant "truncated DOD's" are, then, non-interlining.)

(iii) The non-coverage volume, i.e. A_i, is assigned to the resultant "sample DOD's" in proportion to their original contribution to D_i.

New DOD records consisting of "assigned passengers" are produced.

The assignment technique assumes that the truncated DOD's are representative of the non-interlining traffic. As a result, some original sample DOD's are receiving more weight than they would in the revenue passenger O & D survey alone. Hence, the estimator, $\hat{d}_j^T$, for the total market number of

passengers for the j^{th} origin-destination pair can be derived by adjusting the sampling fraction as follows:

$$\hat{d}_j^T = (1/f_j) \times d_j$$

$$\text{where } f_j = \frac{d_j}{(d_j/f) + a_j} \quad (1)$$

and where f_j is the adjusted sampling fraction associated with the j^{th} origin-destination pair, from the revenue passenger O & D sample survey.

d_j is the sample number of international passengers in the j^{th} origin-destination pair, from the revenue passenger O & D sample survey.

a_j is the number of passengers assigned to the j^{th} origin-destination pair.

Note that a_j, according to point (iii) above,

is $$a_j = \frac{d_j}{D_i} \times A_i \qquad j \epsilon i$$

where $$\sum_{j \epsilon i} d_j = D_i$$

and $$\sum_{j \epsilon i} a_j = A_i$$

6.3 An Example of the Assignment Technique

A simple example will illustrate how the assignment technique works. Assume that the airport activity data from British Airways indicated that 120 passengers enplaned at Montréal and went to London. Therefore, C_i = 120. Assume that the only two DOD's for this assignment grouping from the revenue passenger origin and destination survey are:

Sample Number of Passengers	Estimate of the Number of Passengers	DOD's
1	10	YWG-AC-YMX-BA-LON-LO-WAW
2	20	YYZ-CP-YMX-BA-LON-LH-HAM

where the codes are to be interpreted as:

Code	Denotes
YWG	Winnipeg
YMX	Montréal (Mirabel)
LON	London
WAW	Warsaw
YYZ	Toronto
HAM	Hamburg
AC	Air Canada
BA	British Airways
LO	LOT
CP	CP Air
LH	Lufthansa

Therefore, $D_i = 3$, and $A_i = 120 - (1/.1) \times 3 = 90$.

The truncated DOD's, the proportion of their contribution to D_i, the resultant number of assigned passengers and total market estimates are then:

Truncated DOD's	Proportions	Assigned Passengers(a_j)
YMX - BA - LON - LO - WAW	1/3	30
YMX - BA - LON - LH - HAM	2/3	60

6.4 Shortcomings of the Assignment Technique

The assignment technique has some recognized shortcomings.

The basic assumption of the assignment technique is that the truncated DOD's are representative of the non-interlining traffic. Consider the hypothetical example above in order to determine whether this assumption is intuitively reasonable. The assignment technique presumes that passengers flying on a particular foreign air carrier and originating in Montréal would have the same ultimate destination as passengers originating in Winnipeg or Toronto who fly through Montréal. Hence, the travel patterns of ethnic communities, for example the Polish and German communities in Toronto and Winnipeg respectively, might be used to impute for travel patterns of the more predominantly French communities in Montréal. And, in fact, the assumption that interlining and non-interlining travel patterns are the same was proven empirically to be suspect in a pilot test (Rosen and Conroy (1977)) conducted by the Canadian Transport Commission in 1977. Therefore, there is not only some intuitive but also some empirical evidence against the basic assumption of the assignment technique.

The accuracy of the estimates of the number of passengers by origin-destination pair is jeopardized by any violations of the assumption that truncated DOD's are representative of non-interlining traffic. Large volumes of passengers are allocated to origin-destination pairs, as was seen in Table 1 above, based on a small "effective" sample size. For example, the "effective" sampling fraction of the Canada - World Area #7 market is, not 10% as in the domestic survey, but 5.7% (ie. (100% - 42.7%) x 10%), because of the non-coverage of non-interlining traffic. Hence, a smaller than 10% sample of DOD's is used to allocate a large volume of passengers. Violations of the aforementioned assumption, then, would cause a potentially large bias in the estimates.

Since airport activity census counts are used as benchmark figures for traffic volumes, their accuracy is very important. Although no evaluations have been

undertaken to investigate the magnitude of bias from nonsampling errors in the airport activity census counts, aviation statistics economists feel that this is a survey in which such errors would be small. There are currently, however, ongoing discussions with the major Canadian air carriers on how the reporting requirements of government agencies can be minimized. As a result of these discussions, the airport activity survey could become a sample survey. If a sample is to be designed, the accuracy of the activity counts of gateway airport passengers would be an important design consideration.

The assignment technique also assumes that all of the non-coverage is accounted for by the non-interlining traffic. An interesting way of validating this assumption would be to compare, for the same reference period, the airport activity census count for a Canadian air carrier to the analogous estimate from the origin-destination sample survey. This analogous estimate would be the sum of all passengers on DOD's with the same Canadian gateway airport and crossing carrier. It would be necessary to be able to determine whether differences in the estimates were ascribable to differences in the concepts of the two surveys, and if so, and whether these differences have been taken into account in the ASIPOD estimation system. If such differences have not been accounted for, then it could be that there is a problem with the assumption that all of the non-coverage is accounted for by non-interlining traffic.

Many of these shortcomings should be investigated. However, there are major problems in the existing system, and no alternative solutions, which are superior to the same basic methodology of the assignment technique and which can be implemented within time and cost constraints, have been found. Furthermore, the improvements in the estimates of the number of passengers by air traffic market in the new ASIPOD system will be substantial.

The use of the assignment technique to estimate for international markets is an innovative solution to a large problem. It does not completely solve the

non-coverage problem, but it is a major step in the right direction, as will be the production of estimates of the variance of the estimates. These variance estimates should take account of the assignment technique and its assumptions, and, at the same time, give a meaningful measure of the reliability of the DOD estimates.

7. ESTIMATION OF VARIANCE

The estimator of the variance of the international origin-destination estimates is a simple extension of the variance estimator for the revenue passenger origin and destination survey.

7.1 Variance of the Estimate of Interlining Traffic

The development of the estimator of the variance of the revenue passenger origin-destination estimates is dependent upon the way in which tickets contribute passengers to origin-destination pairs (i.e. the domains of interest). Recall that each ticket is selected with probability 0.1, and that each ticket may be broken up into several segments or DOD's. Each ticket may contribute 0,1,2, etc. passengers to a given domain of interest. For example, the itinerary

YWG - AC - LON - BA - YYZ

would be broken, via the breakpoint routines, into the two DOD's

YWG - AC - LON
and
LON - BA - YYZ.

Consider the inbound plus outbound estimates which are total passenger figures, independent of direction. For such estimates this ticket would add passengers to, among others, the following domains of interest:

Domains of Interest	Passenger Count
Winnipeg - London	1
Toronto - London	1
Canada - London	2
Canada - Europe	2
Eastern Canada - Europe	1
Western Canada - Europe	1

Note that the number of passengers per ticket depends on the geographic level of aggregation, and, therefore, on the particular origin-destination estimate (i.e. domain of interest).

The estimate, $\hat{d}$, of the number of passengers from the revenue passenger origin and destination survey can be developed for a particular domain of interest as follows:

$$\hat{d} = \sum_{i=1}^{n} \frac{x_i}{f}$$

where x_i is the number of DOD's belonging to the domain of interest on the i^{th} ticket.

n is the number of sample tickets selected, and

N is the population number of tickets in the revenue passenger origin and destination survey.

f is the sampling fraction, i.e. $n/N = .1$

The revenue passenger origin and destination survey sample is effectively a 10% simple random sample because

(i) the selection of coupons with serial numbers ending in the digit '0' produces a systematic sample, and

(ii) there is no cycle associated with the distribution of tickets which would cause a relationship between the survey estimates and the last digit of the serial number.

The estimate of the variance can be written, then, as

$$\hat{var}(\hat{d}) = N^2 \frac{1}{n} (1 - f) v_s$$

$$\text{where } v_s = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

$$\text{and } \bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \quad ,$$

and the coefficient of variation, as

$$\hat{cv}(\hat{d}) = \sqrt{\hat{var}(\hat{d})}/\hat{d} \quad .$$

Note that $\hat{var}(\hat{d})$ can also be written as

$$\hat{var}(\hat{d}) = \frac{(1-f)}{f^2} \sum_{i=1}^{n} (x_i - \bar{x})^2 \qquad (2)$$

where n is assumed to be sufficiently large for n/(n-1) to be approximately equal to 1.

7.2 Variance of Total Market Estimates

The method for calculating total market coefficients of variation has to recognize that the assigned data are a function of the sample data. In other words different samples will produce different assigned data and, thereby, different values of the total market estimate. The re-use of certain portions

of the sample has an effect on the sampling distribution of the estimates.

The method which will be used adjusts the sampling fraction from 10% to be the percentage for which sampled records for a particular domain of interest are actually accounting. Since the use of the assignment technique is a given, it has to be assumed that truncated DOD's from the revenue passenger origin and destination survey are representative of the non-interlining traffic. The measure of reliability will be a measure of the precision of the DOD estimates, only to the extent to which this assumption is valid. As a result, it is reasonable to adjust the weights of the sample DOD's to take into account the non-interlining DOD's. The sampling fraction, then, for the estimation for the j^{th} domain of interest would be f_j as developed in equation (1) above. f_j would replace f in the formula for $\hat{var}(\hat{d})$ in equation (2) above in order to yield the formula for the variance of the total market estimates:

$$\hat{var}(\hat{d}_j^T) = \frac{(1-f_j)}{f_j^2} \sum_{i=1}^{n} (x_i - \bar{x})^2.$$

And, the coefficient of variation for the total market estimate would be

$$\hat{cv}(\hat{d}_j^T) = \hat{var}(\hat{d}_j^T) / \hat{d}_j^T.$$

Note that $f_j < .10$ for $a_j > 0$. This means that less than a 10% sample is achieved when it is necessary to include assigned data in a total market estimate. Hence, by using f_j instead of .10 in the expression for the variance of a DOD estimate, the coefficient of variation is adjusted to relate to the total market estimate.

This method gives credit to the use of sample data in the assignment technique; but it is dependent, as is the determination of the estimates themselves, upon the assumption that truncated DOD's are representative of non-interlining traffic.

8. FUTURE CONSIDERATIONS

Earlier, the need for data on air traffic markets in international bilateral air negotiations was explained. The exchange of statistical information is one of the negotiable articles in air transport agreements. Currently, agreements for the exchange of statistical information exist with several countries. The concepts and quality of the data from some of these countries indicate that these data could be used in the ASIPOD estimation system. Such data would provide sample information on the non-coverage portion of the international universe of tickets. Feasibility work is currently underway to determine whether the number of countries for which these data can be used would improve the accuracy of a sufficient number of estimates to justify the expansion of the ASIPOD system to use "exchange data".

9. ACKNOWLEDGEMENTS

The authors would like to thank their colleagues in the Transportation and Communications Division and the Business Survey Methods Division of Statistics Canada, and in particular Robin Dunn, for their many helpful comments during the development of this paper.

REFERENCES

[1] Burchell, J.M.: "International Air Passenger Origin/Destination Project"; presentation to the Transportation and Communications Division of Statistics Canada, March 1981.

[2] "International Air Passenger O & D Statistics System - User Requirements Specifications"; prepared in the Aviation Statistics Centre (ASC) of the Transportation and Communications Division of Statistics Canada, revised March 1981.

[3] "International Air Passenger O & D Statistics System - Feasibility Study Report"; prepared in the ASC, January 1981.

[4] "International Air Passenger O & D Statistics System - Requirements for Functional Specifications (Concepts)"; prepared in the ASC, revised September 1981.

[5] Rosen, F.G. and Conroy, P.F.: "A Test of the Assignment Technique Based on a Survey of International Air Travellers at Montreal and Toronto"; a confidential report prepared for the Passenger and Aviation Economics Directorate of the Research Branch of the Canadian Transport Commission, February 1977.

[6] Carpenter, R.: "Specifications for Estimates of the Reliability for the Air Scheduled International Passenger Origin and Destination Estimates"; prepared in the Business Survey Methods Division of Statistics Canada, revised January 1982.