TRANSPORTATION COST SYSTEM (TRACS)
Statistical and Computer Documentation
(Source Code and Data on CD Rom)

I. PREFACE

A. Purpose and Content

USPS-FY08-36 documents the development of the estimated distribution keys for purchased transportation costs. It contains documentation for the TRACS subsystems used to develop these costs.

B. Predecessor Document(s)

Docket No. ACR2007: USPS-FY07-30; Docket No. ACR2008: USPS-FY08-NP24

C. Methodology

TRACS consists of four subsystems (Commercial Air, Network Air, Highway and Rail). These are all continuous, ongoing statistical sampling systems used to produce quarterly distribution key estimates. The methodology is generally the same as in Docket No. ACR2007: USPS-FY07-30 with the following exceptions.

Beginning April 1, 2008, a redesigned sampling system was implemented for the Commercial Air subsystem. The data elements and sample sizes collected did not change. The redesigned sampling system is documented in Appendix B, Section B.2.

D. Input/Output

The TRACS distribution keys are used to develop purchased transportation costs in USPS-FY08-32, CRA "B" Workpapers (Public Version) and USPS-FY08-NP2, International Cost & Revenue Analysis (ICRA) Report.

II. ORGANIZATION

The relevant source code and outputs from TRACS are provided on the accompanying CD-ROM. The 'Readme_TRACS(PublicVersion)_FY08.pdf' file describes the contents of the CD-ROM. The programs and systems used to develop the TRACS estimated distribution keys are described in the sections below. The documentation is contained in five appendixes:

- **Appendix A : Highway Subsystem**
- **Appendix B : Commercial Air Subsystem**
- **Appendix C : Network Air Subsystem**
- **Appendix D : Rail Subsystem**
- **Appendix E : Tables**

III. PROGRAM AND SYSTEM DOCUMENTATION

Since TRACS Highway, Commercial Air, Network Air and Rail subsystems are so dissimilar, each subsystem is addressed separately in its own appendix.

The following table shows the TRACS sample sizes, by quarter for FY 2008.

TRACS Sample Size by Quarter: FY 2008

	FedEx Night	**FedEx Day**	**UPS**	**Com. Air**	**Highway**	**Rail**	**Total**
PQ 1	350	875	270	540	2,200	365	4,600
PQ 2	350	875	270	540	2,202	361	4,598
PQ 3	350	875	270	540	2,200	363	4,598
PQ 4	350	875	271	540	2,201	365	4,601
FY 2008	1,400	3,500	1,081	2,160	8,803	1,454	18,397

Federal Express (FedEx) Night, FedEx Day and UPS together comprise the Network Air subsystem

Appendix A : TRACS Highway Subsystem

I. Overview

The TRACS-Highway subsystem is a continuous, ongoing statistical sampling system. On a quarterly basis, it produces separate distribution keys for four types of purchased highway contracts: Inter-BMC[1], Intra-BMC, Inter-SCF[2], and Intra-SCF.

The primary sampling unit (PSU) for all four contract types is the route–trip–stop–day, which is defined as all mail unloaded from a truck at one facility on a specific trip, on a specific day. The survey design is essentially the same for all contract types, though each has its own sampling frame. Each highway sampling frame is a list of stop–days[3].

The sample design consists of three stages. At the first stage, within each contract type, the stop–days are stratified based on the type of facility and whether the trip is inbound or outbound. A systematic random sample of stop-days is selected from each stratum. At the second stage, for each selected stop–day, a sub–sample of pallets, wheeled containers and loose items[4] off–loaded at the test facility is selected. From each container selected at the second stage, a third stage sample of items is selected. For pallets and loose items selected at the second stage, there is no third stage sample. All selected mail is recorded.

Weight and volume information by mail category is recorded for the contents of sampled items. Weights are converted to cubic-feet, based on Density Study data. For sampled pallets, the dimensions of the pallet and the percentage of mail on the pallet by mail category are recorded. Data collectors also record the facility where the item, or the pallet, was loaded onto the vehicle (to establish miles traveled) and the percentage of vehicle floor space occupied by palletized mail, containerized items, and loose items (to establish cubic–feet utilized). From the sample data, the cubic–foot–miles transported by each contract type are estimated by mail category. Distribution keys, proportions of cubic–foot–miles by mail category, are calculated for each contract type.

[1] Bulk Mail Center.
[2] Sectional Center Facility.
[3] Abbreviated name for route-trip-stop-days.
[4] Loose items include pieces, parcels, bundles, sacks, trays, or tubs. Items that are not in wheeled containers or on pallets are called loose items.

II. Statistical Study Design

The universe under study is all mail moved on contracts whose costs accrue to these highway accounts:

Inter-BMC: Account Number 53131 (regular Inter-BMC).
Intra-BMC: Account Number 53127 (regular Intra-BMC).
Inter-SCF: Account Numbers 53124 (regular Inter-SCF), 53609 (regular Inter-P&DC), 53614 (regular Inter-cluster), and 53618 (regular Inter-area);
Intra-SCF: Account Numbers 53121 (regular Intra-SCF), 53601 (regular Intra-P&DC), and 53605 (regular Intra-district).

The four contract types are described below:

- Intra-SCF – Generally these contracts involve carrying mail back and forth between an SCF or Processing & Distribution Center (P&DC) and the Associate Offices (AOs) within its service area. Outbound trips distribute mail from the SCF to the AOs. Inbound trips collect mail from AOs and bring it to the SCF.

- Inter-SCF – These contracts primarily involve carrying mail between SCFs or P&DCs. Inter-SCF trips also stop at AOs along the way.

- Intra-BMC – This category of contracts primarily involves carrying mail between a BMC and other facilities within its service area. Outbound trips distribute mail from the BMC to the SCFs and AOs, while inbound trips collect mail from the SCFs and AOs and bring it to the BMC.

- Inter-BMC – These contracts primarily involve carrying mail between BMCs, but they often include stops at SCFs and some stop at AOs. A contract that involves service between a BMC and an SCF outside of the BMC's service area may also be classified in this category.

The following table lists all the stops including destinations by facility type for each contract type in PQ4, FY2008.

Table 1. Stops by Facility Type

Contract Type	BMC	SCF/P&DC /HSP	AO/ Others	Total Stops
Intra-SCF	294	31,697	104,360	136,351
Inter-SCF	333	13,964	7,441	21,738
Intra-BMC	1,906	3,835	1,018	6,759
Inter-BMC	889	889	76	1,854

Typically, one contract route entails multiple trips, some being round–trips and others being one–way trips. One trip serves a specific route on specific days of a week. A trip consists of one or more trip segments. A segment is a movement from one facility to the next facility along the trip. There are as many stops on a trip as there are segments. Therefore, a list of trip–segment–destinations is the same as a list of trip–stops.

1. First Stage Sample

The first stage sample is a stratified random sample of stop–days. Stops are stratified based on the type of facility and whether the segment is on an inbound or an outbound trip. Within each stratum, stops are sorted by district and a systematic random sample of stops is selected with probability proportional to the trip's operating frequency. For each selected stop, all possible operating days during the quarter are listed and a test date is randomly selected from the list. The following table shows the first stage sample size by quarter for Fiscal Year (FY) 2008.

Table 2: Sample Size by PQ

Postal Quarter	INTRA SCF	INTER SCF	INTRA BMC	INTER BMC	Total Highway
PQ1	550	550	550	550	2,200
PQ2	552	550	550	550	2,202
PQ3	550	550	550	550	2,200
PQ4	551	550	550	550	2,201
FY2008	2,203	2,200	2,200	2,200	8,803

2. Second Stage Sample

For each selected stop–day, the following truck utilization data are recorded: (1) the percentage of floor space that is empty; this is further split based on the percentage of empty equipment and empty floor space; (2) the percentage of the floor space occupied by mail unloaded at the test facility; and (3) the percentage of the floor space occupied by mail that remained on the truck. The utilization data, based on square footage of floor space, are collected at the time the vehicle is opened, but before any mail is removed from or added to the vehicle.

The second stage sample is a stratified sample of off–loaded wheeled containers, pallets and loose items. Mail being off–loaded from the vehicle is stratified into five groups depending on the type of mail and the level of containerization. These five groups, or strata, are: pallets, wheeled containers, non-containerized Express Mail, loose sacks, and other loose items. The percentage of floor space occupied by each of these five strata is recorded before selecting the second stage sample. The following rules are used to select the second stage sample:

- Up to two pallets are selected.
- For wheeled containers, if two or fewer containers are unloaded, all are selected. Otherwise, a random start and a skip interval of three are used to select containers for sampling. If the combination of the random start and skip interval results in less than two containers being selected, then a second container is also selected. At most, five containers are selected.
- All loose Express Mail items (i.e., sacks, parcels, etc.) are selected.
- Up to eight loose sacks and other loose items are selected, with at least one of each item type. The item types sampled in TRACS–highway are trays, tubs, sacks, bundles, loose parcels and pieces, etc. Items are selected in proportion to their presence on the truck.

3. Third Stage Sample

The third stage sample is a stratified sample of items from each selected wheeled container. Mail in the wheeled container is stratified by item type, and one of each item type is selected. Prior to the selection of the item, the data collector records the container type and the percentage of the container occupied by each item type. These data are used to estimate the cubic–feet for each item type in the container. In some cases, the data collector counts the number of the items in the container, rather than estimating the percentage, to impute cubic–feet from the number of items.

There is no third stage sample for pallets, loose Express Mail, loose sacks or other loose items selected at the second stage. All selected mail is recorded.

For each selected pallet or item, regardless of whether it was selected at the second or third stage, two types of data are recorded: item information and mail information. For pallets, the item information includes the dimensions (height, length and width) of the pallet, and the origin facility code where the pallet was loaded onto the truck. The pallet dimensions are used for expansion; the facility code is used to determine the mileage traveled. For mail information, the percentage of pallet space taken up by mail for up to four mail categories is recorded.

For items, in addition to the origin facility code, the label content line and the item type are also required as a part of the item information. The item type determines which expansion formula applies. For mail in items, the data recorded include the number of pieces and weight by mail category. Weights are converted to cubic‑feet based on the density factors shown in Appendix E, Table 1.

III. Estimation

Quarterly distribution keys are produced for the four contract types: Inter-BMC, Intra-BMC, Inter-SCF, and Intra-SCF. Data are expanded in a three‑step process to obtain cubic‑foot‑miles by mail category for each contract type. In the first step, test data are expanded to the unloaded truck capacity. In the second step, cubic‑foot‑miles are calculated, accounting for all legs the mail has traveled on the test vehicle. In the third step, the cubic‑foot‑miles estimated from individual tests are expanded to all stop-days (the PSU) within each stratum and summed across strata to estimate cubic‑foot‑miles by mail category for the contract type. The ratios of the cubic‑foot‑miles for each mail category to the total cubic‑foot‑miles across all mail categories constitute the distribution key.

The process is slightly different for Intra-SCF. Given the difficulties of maintaining a highway mileage file for the very large number of Intra-SCF routes and stops, the distribution key is based on cubic‑foot‑legs rather than cubic‑foot‑miles. If all Intra-SCF legs were the same length, this would be equivalent to cubic‑foot‑miles. Use of cubic-foot-legs is justified by the reasonable assumption that their lengths do not vary appreciably.

The program used to expand sample data and generate the distribution keys is TRACSSMN.HWYq08.EDITEXP.CNTL(ZEXP). It has three inputs:

(1) TRACSSMN.HIGHWAY.PQq08.EDITED.DATA,
(2) TRACSSMN.CUFT.DEFAULT.FLAT.TEXT.FY08, and
(3) TRACSSMN.MAILCODE.FLAT.CODE.FY08.

The first input is the final analysis file ('Z-file') that contains the clean sample data, with information necessary for expansion. The second input provides the cubic−foot value for various types of containers and items sampled in TRACS. The third file contains the three digit numeric codes for all mail categories, and the density factor associated with each.

The following notation is used in the expansion process:

h	stratum. For Inter-SCF and Inter-BMC, h=1,2,3. For Intra-SCF and Intra-BMC, h=1,2,3,4,5;
n	number of tests performed in a quarter;
i	test index within the stratum;
N	number of frame units for the quarter;
I	frame index;
Day	number of days in a week that a vehicle operates;
Capacity	vehicle capacity in cubic−feet;
%Empty	percentage of vehicle space that is empty;
%Unload	percentage of space occupied by mail unloaded;
%Remain	percentage of space occupied by mail remaining on the truck;
%Container	percentage of space occupied by unloaded wheeled containers;
%Pallet	percentage of space occupied by unloaded pallets;
%Express	percentage of space occupied by unloaded Express Mail items;
%Sack	percentage of space occupied by unloaded non−containerized sacks;
%Other	percentage of space occupied by unloaded other loose items;
S	total legs traveled on this trip, up to the test stop;
s	segment index, or leg, on the trip $\{s=1,2,\ldots,S\}$;
o	origin index − the segment of the origin facility where the item was loaded onto the vehicle $\{o \in 1,2,\ldots,S\}$.
$mile_s$	segment mileage;
r	mail category, $r \in R$;
w	net weight of mail in pounds;
d	density factor in cubic−feet/pound;
cuft	mail cubic−feet;
cfm	mail cubic−foot−mile; and
y	distribution key for the quarter.

Although the expansion process starts from the bottom and works its way up, it may be useful to first look at the formulation, conceptually, from the top down.

The distribution key for the r^{th} mail category,

$$y_r = \frac{cfm_r}{cfm} = \frac{cfm_r}{\sum_r cfm_r}$$

is the ratio of the cubic−foot−miles for the mail category, divided by the total cubic-foot-miles. The cubic−foot−miles for a mail category, from the first−stage expansion, is:

$$cfm_r = \sum_h \sum_i^{n_h} w_h cfm_{hir}$$

where w_h is the stratum weight, and cfm_{hir} is the cubic−foot−miles for a test. The stratum weight accounts for the first stage selection probability of each stratum. This is a Horvitz −Thompson type estimator. See Cochran, Sampling Techniques, 3rd edition, Theorem 9A.5 on page 260 for a proof that this type of estimator is unbiased. More details on the first-stage expansion are provided in Section VII-3. The cubic−foot−miles for a test is

$$cfm_{hir} = \sum_{s=1}^{S} cuft_{hirs} \times mile_{his}$$

where $cuft_{hirs}$ is the cubic−feet of mail that traveled on the s^{th} segment and was unloaded at the test facility, and $mile_{his}$ is the mileage of the segment. Furthermore,

$$cuft_{hirs} = \sum_{o=1}^{s} cuft_{hiro}$$

where $cuft_{hiro}$ is the cubic−feet of mail that was loaded on the vehicle at the beginning of the o^{th} segment and unloaded at the test facility. The cubic−foot−miles for a test are calculated in the second step of the expansion process.

1. Expanding to Unloaded Truck Capacity

Dropping the stratum subscript, the first step in the expansion process produces $cuft_{iro}$, which for the r^{th} mail category is the estimated cubic−feet of mail that was loaded onto the vehicle at the o^{th} facility and unloaded at the test facility, adjusted to the unloaded truck capacity.

The unloaded truck capacity, in cubic−feet, is the product of the vehicle capacity and the recorded percentage of capacity taken up by all mail unloaded at the test facility. It is also the sum of cubic−feet of unloaded mail across all mail categories:

$$Unloaded_i = Capacity_i \times \%Unloaded_i = \sum_r cuft_{irS} = \sum_r \sum_{o=1}^{S} cuft_{iro} \ .$$

All mail unloaded from the test vehicle is sampled. There are three types of second stage sampling units (SSUs): (1) pallets; (2) loose‑items; and (3) containers. The unloaded cubic‑feet of mail is developed by computing the cubic‑feet for each SSU type separately, and then summing across SSU types. Hence,

$$cuft_{iro} = cuft_{iro}^{(Pallet)} + cuft_{iro}^{(LooseItems)} + cuft_{iro}^{(Containers)}$$

1a. Palletized mail expansion formulas

At most two pallets are selected for each test. All mail on selected pallets is sampled. Data collectors measure the height (*H*), length (*L*), and width (*W*) of each selected pallet and enter the mail category and percentage (%*p*) of the pallet space taken up by mail for up to four mail categories. Let *j* denote the sampled pallet index: *j*=1 and 2. For the pallet loaded on the vehicle at the o^{th} origin facility, the actual cubic‑feet on pallet *j* occupied by mail category *r* is

$$H_{ioj} W_{ioj} L_{ioj} \times \% p_{iroj}$$

Hence the truck space taken up by mail class *r,* loaded at origin facility *o,* on unloaded pallets is

$$cuft_{iro}^{(Pallet)} = Capacity_i \times \%Pallet_i \times \frac{\sum_{j \in o} H_{ioj} W_{ioj} L_{ioj} \times \% p_{iroj}}{\sum_{o=1}^{S} \sum_{j=1}^{2} H_{ioj} W_{ioj} L_{ioj}} . \qquad (1)$$

The data set ‘PALLET’ in the 'Z‑file' contains all the sample data for pallets.

1b. Loose‑item (non‑containerized) mail expansion formulas

All the loose item mail found on the truck is stratified into: non‑containerized Express Mail, loose sacks, and other loose items. The truck utilization percentages are collected for each. All the non-containerized Express Mail is sampled. For loose sacks and other loose items, a total of eight items is sampled in each test.

Additional notation:

j	*j=1,2,...,J*	sampled Express Mail,
k	*k=1,2,...K*	sampled sack
l	*l=1,2,...L*	sampled other item (*K*+*L*<=8)
TW		tare weight of the item

The mail recorded from the sampled items is expanded to the truck space taken up by each category of loose items through a two−step process.

Step 1) The net weight of the mail in the sampled item is expanded to the cubic−feet of the item by multiplying the item's gross cubic−feet by the proportion of cubic−feet occupied by mail category *r*:

$$
\begin{aligned}
cuft_{iroj}^{(g)} &= cuft_{ioj}^{(g)} \times \frac{w_{iroj} d_r}{\sum_r w_{iroj} d_r} = cuft_{ioj}^{(g)} \times \frac{cuft_{iroj}}{cuft_{ioj}} \\
cuft_{irok}^{(g)} &= cuft_{iok}^{(g)} \times \frac{w_{irok} d_r}{\sum_r w_{irok} d_r} = cuft_{iok}^{(g)} \times \frac{cuft_{irok}}{cuft_{iok}} \\
cuft_{irol}^{(g)} &= cuft_{iol}^{(g)} \times \frac{w_{irol} d_r}{\sum_r w_{irol} d_r} = cuft_{iol}^{(g)} \times \frac{cuft_{irol}}{cuft_{rol}}
\end{aligned} \quad (2)
$$

The density factor d_r in (2) converts net weight of the mail to the cubic−feet it occupies in the item. Appendix E, Table 1, provides a complete list of the density factors by mail category. The item's gross cubic−feet varies, depending on the type of item, as follows:

a) for trays, flat tubs and con−cons, the gross cubic−feet of the item, $cuft_{iol}^{(g)}$ is as shown in Appendix E, Table 2;

b) for bundles and loose items, the gross cubic−feet of the item, $cuft_{ioj}^{(g)}$ or $cuft_{iol}^{(g)}$, is the same as the net cubic−feet of the mail in the item, or:

$$
\begin{aligned}
cuft_{ioj}^{(g)} &= \sum_r w_{iroj} d_r = \sum_r cuft_{iroj} = cuft_{ioj} \\
cuft_{iol}^{(g)} &= \sum_r w_{irol} d_r = \sum_r cuft_{irol} = cuft_{iol}
\end{aligned} \quad (3)
$$

c) for sacks, pouches, and Express Mail, the gross cubic−feet of the item is the total cubic−feet of the mail plus the tare cubic−feet for the sack/pouch:

$$cuft_{ioj}^{(g)} = \sum_r w_{iroj} d_r + TW_j \times d_j = \sum_r cuft_{iroj} + cuft_j^{(Tr)}$$
$$cuft_{iok}^{(g)} = \sum_r w_{irok} d_r + TW_k \times d_k = \sum_r cuft_{irok} + cuft_k^{(Tr)} \qquad (4)$$
$$cuft_{iol}^{(g)} = \sum_r w_{irol} d_r + TW_l \times d_l = \sum_r cuft_{irol} + cuft_l^{(Tr)}$$

The tare weight, *TW* in (4), is the difference between the gross and the net weight of the sampled item. The density factors for tare weights are shown in Appendix E, Table 3.

Step 2) The gross cubic−feet of mail in a sampled item is further expanded to the truck capacity utilized by the three sampling groups:

$$cuft_{iro}^{(LooseItem)} = Capcity_i \times \left[\begin{array}{l} (\%Express_i \times \dfrac{\sum_j cuft_{iroj}^{(g)}}{\sum_{o=1}^{S} \sum_j cuft_{iroj}^{(g)}} + \\ \%Sack_i \times \dfrac{\sum_k cuft_{irok}^{(g)}}{\sum_{o=1}^{S} \sum_k cuft_{irok}^{(g)}} + \\ \%Other_i \times \dfrac{\sum_l cuft_{irol}^{(g)}}{\sum_{o=1}^{S} \sum_l cuft_{irol}^{(g)}}) \end{array} \right] \qquad (5)$$

All loose−item data are contained in the SAS data set ‘FORM3L’ in the 'Z−file’.

1c. Containerized mail expansion formulas

Every third wheeled container, up to a maximum of five, is sampled. For each sampled container, the data collector records the percentage of the container taken up by each of the item types found in the container. In some cases, the data collector counts the number of the items in the container rather than estimating the percentage. For example, if only two sacks are in the sampled container, the data collector has the option of recording two sacks instead of estimating the percentage of the container space taken up by the sacks. One item is randomly selected from each of the item types. For each selected item, all the mail is counted. The pieces and weight are recorded by mail category.

Additional notations:

c	Sampled container {c=1,2,…C: C<=5};
t	Item type, *t* = Parcel, Sack, ...
p	Percentage of container space taken up by the item type
CONTCUFT	Container size in cubic−feet.

The following steps are taken to expand the containerized mail:

<u>Step 1)</u> The net weight of the mail in the sampled item is first expanded to the gross cubic−feet of the sampled item by multiplying the item's gross cubic−feet by the proportion of the item occupied by mail category *r*:

$$cuft_{iroct}^{(g)} = cuft_{ioct}^{(g)} \times \frac{w_{iroct} d_r}{\sum_r w_{iroct} d_r} = cuft_{ioct}^{(g)} \times \frac{cuft_{iroct}}{cuft_{ioct}} \qquad (6)$$

The density factor in (6) converts the net weight of mail into the cubic−feet it occupies in the item. Appendix E, Table 1, provides a complete list of density factors by mail category. The item's gross cubic−feet varies, depending on the type of item, as follows:

a) for trays, flat boxes, and con−cons, the gross cubic-feet, $cuft_{ioct}^{(g)}$ is as shown in Appendix E, Table 2;

b) for bundles and other loose pieces, the gross cubic-feet, $cuft_{ioct}^{(g)}$ is the same as the net cubic−feet of mail in the item, i.e.

$$cuft_{ioct}^{(g)} = \sum_r w_{iroct} d_r = \sum_r cuft_{iroct} = cuft_{ioct} .$$

c) for sacks, pouches, and Express Mail, the gross cubic−foot, $cuft_{ioct}^{(g)}$, is the mail cubic−feet plus the tare cubic−feet of the sack:

$$cuft_{ioct}^{(g)} = \sum_r w_{iroct} d_r + TW_t \times d_t = \sum_r cuft_{iroct} + cuft_t^{(Tr)} .$$

The tare weight, *TW* in (c) above, is the difference between the gross and net weight of the sampled item. The density factors for tare weights are shown in Appendix E, Table 3.

<u>Step 2)</u> The gross cubic−feet of mail in a sampled item are further expanded to the sampled container:

$$cuft_{iroc}^{(g)} = \frac{CONTCUFT_c}{\sum_{o=1}^{S}\sum_{t} p_{ioct}} \sum_{t}\left(p_{ioct} \times \frac{cuft_{iroct}^{(g)}}{cuft_{ioct}^{(g)}} \right) \quad (7)$$

The container size, *CONTCUFT,* for various type containers can be found in Appendix III, Table 4.

Step 3) The resulting gross cubic−feet of mail is finally expanded to the truck capacity utilized by all unloaded containers,

$$cuft_{iro}^{(Container)} = Capacity_i \times \%Container_i \times \frac{\sum_{c=1}^{C} cuft_{iroc}^{(g)}}{\sum_{c=1}^{C} CONTCUFT_c} \quad (8)$$

The SAS data set 'FORM3C' in the 'Z−file' contains all the containerized sample data.

1d. Total cubic−feet of mail unloaded

The cubic−feet of mail which was loaded on the truck at origin *o* and unloaded at the test stop is the sum of the three second stage sampling unit types:

$$cuft_{iro} = cuft_{iro}^{(Pallet)} + cuft_{iro}^{(LooseItem)} + cuft_{iro}^{(Container)}. \quad (9)$$

When added across the origin and the mail category, we obtain the reported unloaded capacity of the vehicle.

$$Unloaded_i = Capacity_i \times \%Unloaded_i = \sum_{r}\sum_{o=1}^{S} cuft_{iro} .$$

2. Cubic−Foot−Mile Calculation

Summing equation (9) across all the origin facilities up to a specific leg produces the cubic−foot estimates for the mail that traveled on the leg:

$$cuft_{irs} = \sum_{o=1}^{s} cuft_{iro} . \quad (10)$$

Each sample record contains a complete list of legs the mail item traveled on the vehicle. For each leg, the sample record also specifies the origin and destination facilities and the highway miles between the two. The cubic-foot-miles for each leg are the product of the cubic-feet estimates and the highway miles for the leg. The cubic-foot-miles is the sum of such products across all legs:

$$cfm_{ir} = \sum_{s=1}^{S} \left\{ cuft_{irs} \times mile_{is} \right\}. \quad (11)$$

The cubic-foot-mile calculation specified in equation (11) requires highway miles for all leg segments on sampled routes. The highway-mileage file contains highway miles for over 7,000 facility pairs, and covers over 90% of the Inter-BMC, Intra-BMC, and Inter-SCF routes. However, the coverage is less than 50% for Intra-SCF routes, so mileage is not used for Intra-SCF contracts.

Similar to equation (11), the test level cubic-foot-leg for the Intra-SCF contract type is calculated by:

$$cfm_{ir} = \sum_{s=1}^{S} cuft_{irs} \, . \quad (12)$$

Prior to Base Year 2000, equations (11) and (12) were adjusted to account for the unused capacity, or empty space, on the truck. The expansion process used since then adopts the 'compromise method' introduced by the Postal Service and adopted by the Commission in Docket No. R2000-1 (Op. ¶3300). Mathematically, the 'compromise method' is equivalent to removing the empty space adjustment from the expansion process as shown in equations (11) and (12).

3. Distribution Key

The test level cubic-foot-miles obtained from equations (11) and (12) in the previous section is expanded to the stratum level and summed across strata. The distribution key is a set of ratios of the expanded cubic-foot-miles for an individual mail category to the total expanded cubic-foot-miles summed across all the mail categories.

$$y_r = \frac{\sum_{h=1}^{H} \left(\sum_{i=1}^{n_h} cfm_{ir} \right) w_h}{\sum_{r \in R} \sum_{h=1}^{H} \left(\sum_{i=1}^{n_h} cfm_{ir} \right) w_h} = \frac{cfm_r}{\sum_r cfm_r} = \frac{cfm_r}{cfm} \, , \quad (13)$$

where the stratum weights are:

$$w_h = \frac{\sum_{I=1}^{N_h} Day_I \times 12}{n_h} \quad . \tag{14}$$

Appendix B : TRACS Commercial Air Subsystem

Section B.1 (Applies to PQ108 AND PQ208)

I. Overview

The TRACS Commercial Air subsystem is a continuous, ongoing statistical sampling system designed to estimate pounds for various mail categories.

The universe under study is all mail traveling pursuant to purchased transportation contracts on passenger airlines from a domestic origin to a domestic destination. It does not include mail traveling by air taxi, or via Alaska, Hawaii, HASP, Christmas, or Network Air services. The sampling frame is a list of flight−days. The sample design consists of two stages. At the first stage, flight−days are stratified and a systematic random sample of flight−days is selected within each stratum. At the second stage, a subsample of mail items is selected and information for all mail in these selected items is recorded.

On a quarterly basis, the proportion of pounds for each mail category is estimated statistically from the TRACS−Air sample data and expanded to population control totals derived from a postal administrative database named Enterprise Data Warehouse (EDW) in order to produce distribution keys for Commercial Air purchased transportation.

II. Statistical Study Design

The Commercial Air subsystem develops pound distribution keys for passenger air costs. The population of interest is all mail moved on commercial airlines traveling from a domestic origin to a domestic destination under purchased air transportation contracts. It does not include mail traveling by air taxi or via a network such as for Alaska, Hawaii, HASP, Christmas, or Network Air services.

In mail processing operations, mail is sorted into trays, tubs, sacks, or other items. Mail items for air transportation undergo a process in which each mail item is affixed with a dispatch and routing (D&R) tag. The D&R tag contains information such as item weight, airline code, flight number, and origin and destination for each flight leg. The D&R tag also shows the predominant mail class indicator that is referred to as the ACT Tag code.

Information from the D&R tagged items is consolidated into the EDW. The EDW maintains historical routing information (flight origin, destination, dispatch date and time), assigned weight and the primary mail class (as indicated on the routing label affixed to a mail item) for each mail item transported The EDW database maintains information for all scanned mail items by transaction concentrator (TC) server identifier, date, network flight and ACT tag code. Population control totals required for expansion of the TRACS‑Air sample data are obtained and distribution key estimates are developed from ACT Tag Code summary information in the EDW file.

A. First Stage Sample

The primary sampling unit (PSU) is a flight-day, which is defined as all mail being dispatched from the specified origin on a given day via a particular airline and flight with the same first‑leg destination reflected on the routing label or D&R tag. The PSUs are stratified into three strata: Inbound (if the first leg origin airport is one of the possible international airports), Outbound (if not inbound, and the first leg destination airport is one of the possible international airports), and Domestic (otherwise). The international airports are New York, Chicago, San Francisco, Miami, Los Angeles, Seattle, and Dallas TX. An equal size sample of 180 flight‑days is then systematically selected from each of the three strata.

B. Second Stage Sample

The secondary sampling unit (SSU) is an item of mail dispatched for a sampled flight‑day. Examples of mail items include sacks, letter trays, flat tubs, and loose parcels. For the second stage sample, mail items are stratified into five ACT Tag groups, based on the mail class indicator (ACT tag code): F for First Class Mail, P for Priority Mail, E for Express Mail, I for international mail and O for other mail. The ACT Tag code indicates the predominant mail class contained in the item (an item may contain one or more mail classes or categories). Two items are sampled from each ACT Tag group. If there are less than two items available for an ACT Tag group, they are replaced with items from other ACT Tag groups, with preference given first to F, then P, then E, then I and finally to O. If there are less than eight items on the flight‑day, all items are sampled.

For each selected item, the ACT Tag code and routing information are recorded directly onto the data collector's laptop computer. This information is later used to link the sample data with the records in the EDW file. The total weight of the item, as well as the count and weight of the contents of the item by mail category, are also recorded.

III. Estimation

A. Estimated Parameter

The Commercial Air subsystem distribution key is the ratio of mail category pounds to total pounds across all mail categories. That is,

$$K_r = \frac{P_r}{P}, \quad (r = 1, 2, \ldots R).$$

Here, K_r is the distribution key for mail category r, P_r is the total pounds for mail category r, R is the total number of mail categories, and $P = \sum_{i=1}^{I} P_i$ is the overall total pounds across all ACT−Tag types: $i = 1, \cdots, I$. These P_i's are known control total aggregates from the EDW. The task is to estimate P_r, the pounds for mail category r, for each mail category r. The estimator of K_r is denoted as:

$$\hat{K}_r = \frac{\hat{P}_r}{P}, \quad (r = 1, 2, \ldots R).$$

B. Estimation of Mail Category Pounds

To estimate P_r, the total pounds by mail category, we note that P_r is the total pounds for mail category r, which is equal to the total pounds for mail category r across all ACT−TAG types. That is:

$$P_r = \sum_{i=1}^{I} P_{ri}, \quad (r = 1, \cdots R).$$

Here, P_{ri} is the total pounds for mail category r for ACT−TAG type i. And P_{ri} is to be estimated by the following approximate design, unbiased estimator:

$$\widehat{P}_{ri} = \sum_{h=1}^{H} \frac{P_{hi}}{\sum_{l=1}^{n_h} P_{hli}} \sum_{l=1}^{n_h} P_{hli} \frac{\sum_{k=1}^{m_{hli}} P_{rhlik}}{\sum_{k=1}^{m_{hli}} P_{hlik}} \quad (i = 1, \cdots I; \quad r = 1, \cdots R).$$

In the equation above, H is the total number of strata; n_h is the number of sampled PSUs; m_{hli} is the number of sampled SSUs (items). P_{hi} and P_{hli} are known population totals aggregated from the EDW. P_{rhlik} and P_{hlik} are calculated from recorded data.

C. Expansion Programs

The program used to expand sample data and generate the distribution keys is TRACSSMN.EXPAND.COMMAIR.PQq08.CNTL(AIRKEYS2). It has two inputs:

(1) TRACSSMN.EXPAND.COMMAIR.PQq08.SASDATA,
(2) TRACSSMN.MAILCODE.FLAT.CODE.FY08.

The first input is the final analysis file ('Z−file') that contains the clean sample data, with information necessary for expansion. The members of 'Z−file' are

SAMPLE and POP_PM_HI. The second file contains the three digit numeric codes for all mail categories.

The AIRKEYS2 program calculates $\hat{K}_r$.

Section B.2 (Applies to PQ308 AND PQ408)

I. OVERVIEW

The TRACS Commercial Air subsystem is a continuous, ongoing statistical sampling system developed to provide distribution keys for attributing purchased air transportation costs to mail categories. Weight based distribution keys are constructed for each postal quarter for this purpose. The universe under study is all mail transported by commercial air carriers. Mail not transported by commercial air carriers is excluded.

Under the TRACS Commercial Air subsystem, mail is transported from an origin city directly to a destination city. The sampling frame for an air mode is a recent historical extract of its assigned weight (lbs.) obtained from the Postal Service's Enterprise Data Warehouse (EDW). The EDW maintains historical routing data for each mail item (sack, tray, loose parcel, etc.) transported by air carrier, including the air mode, origin city, destination city, dispatch date, dispatch time, assigned weight, mail item type and primary mail class.

The TRACS Commercial Air subsystem sample design utilizes a stratified, multi-stage probability sample of mail. The primary sampling unit (PSU) at the first stage of sampling is a mail class, facility, date and time segment combination. The secondary sampling unit (SSU) is a mail item. The sampling element is a mailpiece. A sampled mail item is fully enumerated; wherein, all mailpieces contained in the mail item are measured. Data collectors capture routing information from each sampled mail item along with content weight and volume (count) by mail category. Appendix E provides a list of the mail categories measured under TRACS. The sample data are recorded into laptop computers using the Computerized On-Site Data Entry System (CODES).

The TRACS Air sample data are expanded each postal quarter sampling period using inverse sampling fractions for strata at each sampling stage along with control ratios utilizing known EDW weight totals. Annual estimates are obtained by summing the postal quarter estimates.

II. STATISTICAL STUDY DESIGN

The TRACS Commercial Air subsystem develops weight (pounds) based distribution keys for volume variable costs. The universe under study is all mail transported by commercial air carriers from a domestic origin to a domestic destination. Mail transported by Fed Ex, UPS, air taxi, or under Alaska, Hawaii, HASP, or Christmas networks is excluded.

In mail operations, mailpieces are sorted into mail items, or handling units consisting of trays, tubs and sacks. A mail item may also be a loose parcel. A mail item assigned to an air carrier gets scanned electronically and then tagged at a TC machine located at an origin postal or private mailer facility. A TC is uniquely identified in the EDW by its alpha-numeric device-id code which is normally six characters. One or more workstations may be associated with a device–id at an origin facility, and one or more device–id's in turn may be associated with an origin facility. The tagging process attaches a dispatch and routing (D&R) tag to each scanned mail item. The D&R tag identifies air mode specific information including the air carrier, origin and destination city airport codes, item weight and primary mail class. The primary mail class is the Air Contract Transportation (ACT) tag code. A mail item is assigned an ACT tag code based on the predominant mail class of the mail item's contents. The major mail class codes are: F, P, E and I, corresponding to First–Class Mail, Priority Mail, Express Mail and International mail, respectively. All other ACT tag code mail class groups are designated 'O' for sampling purposes under the TRACS Commercial Air Subystem. The EDW maintains summary weight information for all assigned (tagged) mail items, by device–id, date, network flight and ACT tag code. ACT tag weight totals from the EDW are used in the expansion of the sample data.

A. First–Stage Sample

The PSU for a network air mode is a *mailclass*facility*date*time-segment* combination.[5] The *mailclass* component is the ACT tag code. The *facility* component is the postal or private mailer site where mail items are assigned to the air mode. The *date* and *time-segment* components identify the date and sub–period when mail items are assigned. The assignment of mail items for a given mail class, facility and date combination may occur over various time segments. A time segment may be a 1–2 hour, 2–3 hour, or other period, including a 24 hour period. A time segment determines a unique PSU for a given mail class, facility and date combination.

Prior to each postal quarter, a sampling frame is constructed listing all PSUs expected during the upcoming sampling period. PSUs are stratified by weight (lb.) based size categories. Thresholds (boundaries) are dynamically formulated each quarter using a cumulative √f distribution method and a fixed number of strata.

A specified fixed total sample size is allocated among the strata proportional to each stratum's historical weight from the EDW. Prior to the draw, the PSU's are ordered by size within each stratum and a skip sample is then selected after a random starting unit is chosen. Sample dates are randomly assigned to the selected sampling units from all possible non-holiday dates of

---

[5] All mail items assigned for air transport for a PSU are subject to sampling, including mail items trucked by an air carrier.

operation (specific to each network air mode) after the dates are ordered and partitioned into a minimum number of equal size (uniform) sub-periods for workload smoothing purposes.

B. Second–Stage Sample

The second stage SSU is a mail item transported by a commercial air carrier on a first leg flight or trip. Examples of mail items include a sack, letter tray, flat tub, and loose parcel. All mail items assigned to a PSU for first-leg transport are subject to sampling.

For a sampled PSU, a sub-sample consisting of 10 mail items (minimum) if available is selected. This sample is assumed a simple random sample. The second-stage sampling frame is an implicit listing of all mail items assigned for air transport at all workstations for all device-id's at the origin facility. If there are fewer than 10 mail items available for sampling at a targeted PSU, then all mail items are selected and sampled. If no mail items are assigned for a targeted PSU, then the test is considered a zero-volume test, subject to administrative confirmation. If a test is missed, it is rescheduled to the same weekday of a future week during the sampling period (unless it is within the last week of the quarter - then rescheduled to any weekday). The mail item population counts necessary for second-stage post-stratification and expansion are obtained from the EDW.

For each sampled mail item, PSU specific routing information is recorded directly into the data collector's laptop CODES computer. The routing information links the sample and EDW frame file records during the data editing and pre-expansion processes. No tertiary stage of sampling is conducted. A census is conducted of each SSUs contents, *i.e.*, all mailpieces (sampling elements) in a selected mail item are measured. Mailpiece volume and weight totals are recorded for each mail category in addition to the sampled unit's gross weight. The data collection procedures governing TRACS Air tests are documented in Section 8, of Handbook F-65, Data Collection User's Guide for Cost Systems, and SP Letter #5 FY2008.

III. CREATING THE FIRST–STAGE SAMPLING FRAME

The frame of PSUs is constructed approximately five weeks in advance of the postal quarter from recent historical EDW information. The EDW provides the following information for each mail item transported by air mode: device-id, three-character origin and destination airport codes, route identifier, route close-out date and time, ACT tag, assigned weight, and assigned date and time. This information is cross-walked to a file containing origin facility identifier information. The sampling frame for an air mode is a listing of all PSUs expected to have non-zero activity during the upcoming postal quarter.

PSU migrations among strata, including births and deaths occurring after sample selection or as a result of inadvertent sampling of non−targeted PSUs, are addressed through post-stratification prior to first-stage expansion.

IV. FIRST–STAGE SAMPLE SELECTION

A. Selecting the Sample

The total sample size is allocated among the primary strata proportionally to each stratum's recent historical mail weight (lbs.). Prior to this step, an empirical distribution is formulated to filter small−sized units at a prescribed percentile level. The measure of size for strata assignments is average weight (lbs.) by day−of−week (DOW) among PSUs having common mail class, facility and time segment components. Dates are temporarily collapsed to DOW for this purpose. The collapsed PSUs are assigned to strata using a cumulative $\sqrt{f}$ distribution for the strata thresholds. The PSUs are then ordered within each stratum by their size measure. There are four different strata and they are classified as; 1) light, 2) medium-light, 3) medium-heavy, and 4) heavy. A starting unit is randomly selected between 1 and the computed skip interval length for each stratum. A second random assignment process provides the date component for each selected unit, from a mode−specific list of valid non−holiday operational fly dates for the sampling period.

After the random selection and date assignment processes, a workload smoothing process at the district level ensures that the randomly assigned dates for each weekday (Monday−Sunday) are distributed uniformly across equal−sized sub−periods of the full sampling period. Test facility information, including address and district code, is appended to each sample record for administration purposes.

V. PREPARING DATA FOR EXPANSION

Data validation and edit checks are made prior to data expansion. Sample data are checked for completeness, duplicate records, and integrity of field values. Sample data are extensively validated and crosschecked with EDW (linked by D&R barcode and processing date), the Product Tracking System PTS (linked by scanned mail piece barcode) and automatically scanned barcodes (mail piece barcodes scanned by the data collector). Sample data Act Tag Codes, test dates, test facilities, time segments, air modes, and mail categories are independently checked. If the sample data are not independently vetted and confirmed with the EDW data, or if there is a conflict between PTS data or scanned barcodes, then the data are edited, if possible, or if not, then removed.

The Z−file contains unexpanded, edited sample data records for the TRACS Air subsystem. This file is produced in SAS program: TRACSSMN.FEDEX.PQq08.CNTL(ZAIR01). Each Z−file record contains population and effective sample sizes for each sampling stage and stratum within, along with the EDW frame totals required to construct a distribution key estimate for each network air mode.

In the Z−file program, updated EDW frame extract records are checked for missing values and duplicate records. A job abort is triggered if an empty field or non−unique record is encountered at the workstation and mail category level. The frame records for each subpopulation of a network air mode are then post−stratified by weight at the PSU level. The original first−stage strata thresholds determined during sample selection are updated in this program to reflect changes occurring during the sampling period. The thresholds are formulated within each subpopulation using a cumulative √f distribution. The first−stage post−stratification process adjusts the original sampling frame for PSU migrations (births and deaths) across strata occurring after the sample was selected.

In the Z−file program, the sample data at the second stage SSU (mail item) level are post−stratified by workstation to incorporate known EDW mail class (ACT tag) totals into the second−stage expansion process. Prior to this step, the program checks the sample data records for missing variable values and non−unique mail class entries. A job abort occurs if either condition is found. Ounce−pound conversion and other data record checks and edits are conducted prior to this program. Non−countable sample records are summarized and reported before being bypassed from downstream processing. Effective second−stage sample sizes are appended to the sample data records during the second−stage post−stratification process.

The sample data records at the workstation and mail category within level are merged with frame records by PSU and mail class. Strata population counts and weight (lb.) controls for each sampling stage are appended. Effective first-stage sample sizes for sampled strata are appended in this step. Extraneous frame records for unsampled PSUs and workstations are filtered. The merged data are output to the Z−file.

VI. EXPANSION

At the second stage of expansion, the sample data are post−stratified by workstation within each PSU and a mail category total is estimated for the PSU by a combined [strata] ratio estimator. In this process, the sample data for each secondary stratum (workstation) are expanded by the stratum’s inverse sampling fraction, computed as the reciprocal of the ratio of the sample item count to the known mail item count from the EDW for the stratum. An intermediate control factor, computed as the ratio of known total weight for all device‑id's in the PSU to known total weight for the sampled‑only device‑id's in the PSU, is applied to control for unsampled device‑id's. This factor is unity if a PSU has a single device‑id or if at least one mail item is sampled from each device‑id in the PSU. The combined [strata] ratio estimate of a total is computed for each mail category from the ratio of known total weight for the PSU to estimated total weight for the PSU (all mail categories combined).

PSU mail category estimates from the second stage expansion process are expanded to first‑stage post‑stratified classifications for each air mode subpopulation using a separate [stratum] ratio estimator. For the expansion phase, there are three subpopulations defined by facility type. These subpopulation groups are: 1) private mailers, 2) the JFK air mail facility, and 3) all other postal facilities. For each mail category in a primary sampling stratum, the sum of the estimated PSU weight or volume totals in the stratum is expanded by the ratio of known total weight for the stratum from the EDW to the sum of estimated PSU total weight (all mail categories combined) for the stratum. For each mail category, the sum of the estimated strata totals in a subpopulation is controlled to known total weight for the subpopulation from the EDW. The control factor is unity if at least one PSU is sampled from each primary stratum. A final factor is applied to the controlled sum for tieback to the EDW known weight for the air mode. The control factor is unity if all subpopulations in an air mode are represented by at least one non‑empty stratum.

A. Distribution Key Formulas

The TRACS Commercial Air subsystem produces distribution keys each quarter based on weight and volume.

The following variable and index notation is used in the sample data expansion formulas for the TRACS Commercial Air subsystem:

A: air mode.

c: mail class (ACT tag) component of PSU, $c \in C = \{E, P, I, F, "O"\}$.

d: device-id ($d \in D_c$*).*

D_c*: the set of device-id's for a PSU reporting non-zero activity for the mail class component (c) of the PSU.*

d'_P*,* d'_F*: reciprocal density factors (adjusted) for mail class indexes: P, F.*

h: 1st-stage index for post-stratification of PSU's into 1..L_h *strata based on known (EDW) weight.*

*i: PSU = mailclass*facility*date*time-segment (4 component combination).*

j: SSU = mail item (mail handling unit, e.g., sack, tray, loose parcel, etc.).

k: 2nd-stage index for post-stratification of SSU's into 1..L_k *wor-station-based strata.*

M, m: 2nd-stage SSU population and sample sizes (counts), respectively.

N, n: 1st-stage PSU population and sample sizes (counts), respectively.

r: mail (rate) category element ($r \in \vec{R}$*).*

$\hat{\vec{R}}$*: estimated distribution key vector (*$\vec{R}$*).*

S: subpopulation (domain) in air mode (A) .

x: sample weight (lbs) measure ($x \in y$*) for ratio estimator auxiliary variable .*

$\hat{X}$*: estimated total weight (lbs.) for ratio estimator auxiliary variable (sum of estimated total weights for all* $r \in \vec{R}$).

***X**: known (EDW based) total weight (control lbs.) from database.*

y: sample data measure of weight (lbs.), cubic feet or volume.

$\hat{Y}$*: estimated weight, cubic feet or volume total (Y).*

At the second stage of expansion, the total *(Y)* for mail category *(r)* for a sampled PSU *(i)* in stratum *(h)* is estimated by a combined [strata] ratio estimator *(Yrc)*[6], with subsequent control to known mail class *(c)* device‑id weight totals, as follows:

(1.1) $$\hat{Yrc}_{A,S,r,h,i} = \frac{1}{\delta}\left(\frac{\hat{Y}_{A,S,r,h,i}}{\hat{X}_{A,S,h,i}}\right) X_{A,S,h,i}$$, where

(1.2) $$\hat{Y}_{A,S,r,h,i} = \left(\frac{\sum_{d \in D_c} X_{A,S,h,i,d}}{\sum_{d^* \in D_c} X_{A,S,h,i,d^*}}\right) \sum_{d \in D_c} \hat{Y}_{A,S,r,h,i,d}$$, for

(1.3) $$\hat{Y}_{A,S,r,h,i,d} = \sum_{k^* \in d} \frac{M_{A,S,h,i,d,k^*}}{m_{A,S,h,i,d,k^*}} \sum_{j=1}^{m} y_{A,S,r,h,i,d,k^*,j}$$, and

(1.4) δ = d'_P , for y=ft^3 and *(c)* = P

d'_F , for y=ft^3 and *(c)* ≠ *P*

1 , otherwise ,

(1.5) $$\hat{X}_{A,S,h,i} = \sum_{r \in R} \hat{Y}_{A,S,r,h,i}$$, for y = x, and

(1.6) $$X_{A,S,h,i} = \sum_{d \in D_c} \sum_{k \in d} X_{A,S,h,i,d,k}$$.

In equation (1.2), d^* in the denominator of the ratio designates a device‑id for PSU *(i)* for which at least one workstation *(k)* is represented in the captured sample data. If all device‑id's in the PSU are represented in the sample, then this ratio is unity. In equation (1.3), k^* designates a sampled workstation stratum (post-stratified estimator) for which at least one SSU is measured. Equation (1.4) shows the adjusted density factors from equation (d.2) for the reciprocal density factor in equation (1.1) required to estimate cubic feet from the estimate of weight (not used for Commercial Air). In equation (1.6), known total weight for PSU (i) is

[6] To account for small second‑stage sample sizes during post-stratification, a combined [strata] ratio estimator of a total is formulated in lieu of a separate [stratum] ratio estimator.

obtained by summing the known total weights for all component workstations *(k)* for all device–id's *(d)* in D_c.

At the first stage of expansion for subpopulation *(S)* under air mode *(A)*, total *(Y)* for a mail category *(r)* is estimated from the estimated PSU *(i)* totals in each primary stratum *(h)*, $\hat{Y}rc_{A,S,r,h,i}$ in equation (1.1), by a separate [stratum] ratio estimator of a total, $\hat{Y}rs$, with subsequent controls to known total weight at the mail class *(c)* and subpopulation *(S)* combination levels, as follows:

(2.1) $$\hat{Y}rs_{A,S,r} = \left(\frac{\sum_{c \in S} X_{A,S,c}}{\sum_{c^* \in S} X_{A,S,c^*}} \right) \sum_{c \in C} \hat{Y}rs_{A,S,r,c} \left(\frac{\sum_{h \in c} X_{A,S,c,h}}{\sum_{h^* \in c} X_{A,S,c,h^*}} \right)$$, where

(2.2) $$\hat{Y}rs_{A,S,r,c} = \sum_{h} \left(\frac{\sum_{i=1}^{n} \hat{Y}rc_{A,S,r,c,h,i}}{\sum_{i=1}^{n} \hat{X}_{A,S,c,h,i}} \right) X_{A,S,c,h}$$, for

(2.3) $$\hat{Y}rc_{A,S,r,c,h,i} \equiv \hat{Y}rc_{A,S,r,h,i}$$ (from equation 1.1),

(2.4) $$\hat{X}_{A,S,c,h,i} = \sum_{r \in R} \hat{Y}rc_{A,S,r,c,h,i}$$, for *y*=*x*,

(2.5) $$X_{A,S,c,h} = \sum_{d \in D_c} \sum_{k \in d} X_{A,S,c,h,d,k}$$, and

(2.6) $$X_{A,S,c} = \sum_{h} X_{A,S,c,h}$$.

In equation (2.1), the expanded sample total, $\hat{Y}rs_{A,S,r,c}$, from equation (2.2) for mail class *(c)* in subpopulation *(S)* is controlled to the known total weight for all strata *(h)* comprising *(c)*, and then to the sum of known total weight for all *(c)* comprising the subpopulation *(S).* These ratios account for empty (non–sampled) strata in mail classes within each *(S)*, and empty mail classes within each *(S)*, if any. In this equation, *(h*)* and *(c*)* denote non–empty strata and mail classes, respectively, for which sample data have been recorded. If sample data are recorded for all *(h)* and *(c)* in *(S),* then both ratios are unity. Equation (2.2) is the separate [stratum] ratio estimator formula with numerator component, $\hat{Y}rc_{A,S,r,c,h,i}$, shown in equation (2.3).[7] Equation (2.3) is exactly equation (1.1), where $\hat{Y}rc$ in equation (1.1) is re−expressed to identify the mail class component *(c)* of a PSU explicitly (per its definition) instead of implicitly.

For air mode *(A)*, the estimated total *(Y)* for mail category *(r)* is the sum of the subpopulation *(S)* estimates as follows:

(3.1) $$\hat{Y}_{A,r} = \sum_{S} \hat{Y}rs_{A,S,r} \ .$$

For air mode *(A),* a distribution key element *(r)* in $(\vec{R})$ for the weight, volume or cubic feet measure *(y)* is estimated as follows:

(4.1) $$\hat{r}_{A,Y} = \frac{\hat{Y}_{A,r}}{\sum_{r \in R} \hat{Y}_{A,r}} \ .$$

B. VARIANCE ESTIMATION

The TRACS Commercial Air subsystem estimator for the distribution key element, $\hat{r}_{A,Y_r}$, in equation (4.1) is a complex type estimator which incorporates auxiliary variable information for expansion ratios at each stage of sampling. It also incorporates mail class controls for various aggregate levels upon tieback to known EDW weight (lbs.) for the sampling period. Moreover, numerous estimates are constructed each sampling period for each air mode and shape−based mail category within for the mail characteristics of interest: weight and volume.

An alternative to a direct (tractable) formula representation (or re−sampling methodology) for estimating sampling error is Taylor series (first order) approximation; wherein, sampling variation can be estimated among first−stage sampling units (ultimate clusters). To ensure that the sampling error estimates are conservative and not under−estimates (on average), the first-stage finite

[7] In equation (2.2), the sample total for stratum *(h)* in the numerator of the separate [stratum] ratio estimator is itself an estimate calculated as the sum of the combined ratio estimates, $\hat{Y}rc$, for the sampled PSU *(i)* totals in *(h)* from equation (1.1).

population correction (fpc) is ignored. A relative measure of sampling error, coefficient of variation (c.v.), is estimated, $cv(\hat{r}_{A,Y})$, for each mail category element estimate $(\hat{r}_{A,Y})$ of the distribution key vector $(\vec{R})$.

Appendix C: TRACS Network Air Subsystem

I. OVERVIEW

The TRACS–Network Air subsystem is a continuous, ongoing statistical sampling system designed to estimate distribution keys for purchased transportation costs associated with the following major air carrier networks: Federal Express (FedEx) Day Turn, FedEx Night Turn, and United Parcel Service (UPS). The day turn network normally operates six days per week, while the night turn and UPS networks normally operate up to five days a week. The TRACS–Network Air subsystem is comprised of three component subsystems: Network Air Day Turn subsystem, Network Air Night Turn subsystem, and Network Air UPS subsystem. Three distribution keys are produced: one based on cubic–feet for the day turn network, and the others based on pounds for the night turn and UPS networks.

The sample design, data collection procedures and estimation methodology for each subsystem are similar except for the inclusion of an average density factor developed to convert estimated pounds of mail to estimated cubic–feet of mail in the day turn subsystem, and an additional sampling stratum required for the UPS network. The universe under study in the TRACS–Network Air subsystem is all mail moved on the day turn, night turn and UPS networks. It does not include mail transported by commercial airlines, air taxi, Alaska, Hawaii, HASP or Christmas networks. The sampling frame for each subsystem is a list of facility–days. A facility–day is defined as all mail that is electronically scanned and tagged for air transport at a particular facility on a specific day during the quarter, and which travels on the same first–leg network flight to the carrier's hub before it is then placed on another network flight to the final destination. A network flight is a specific flight or group of flights that mail travels on from the origin city to the final destination city. Normally, a network flight's origin is a network–city inbound to the carrier's hub. For sampling purposes, the frame of facility–days for each subsystem is constructed using a recent history of the network's assigned weight obtained from a postal administrative database called the Enterprise Data Warehouse (EDW). The EDW maintains historical routing data (flight origin, destination, dispatch date and time information), assigned weight, and the primary mail class (as indicated on the routing label) for each mail item such as a sack or tray transported by the network.

The sample design consists of two stages: at the first stage, facility–days are grouped into sampling strata and a random sample of facility–days is selected within each stratum; at the second stage, for each selected facility–day, a sub–sample of mail items (e.g., sacks, trays, or loose parcels) is selected, and the information for all mail pieces associated with each selected mail item is recorded. Data collectors record routing information from each selected mail item as well as the weight and number of mail pieces by major mail category.

The data are recorded into laptop computers using the Computerized On−Site Data Entry System (CODES).

For each network air subsystem, the sample data at the mail category and primary mail class level are expanded to known population totals obtained from the EDW database. Separate quarterly TRACS−Network Air distribution keys are estimated for each subsystem. These distribution keys are used to allocate their respective network air transportation costs to the major mail categories. Annual cost estimates are obtained by summing the four quarters' costs.

II. STATISTICAL STUDY DESIGN

The TRACS−Network Air subsystem develops cubic−feet and pound based distribution keys for network air volume variable costs. Three distribution keys are developed: one for the day turn network, and one each for the night turn and UPS networks. The day turn network normally operates six days per week excluding Sundays, while the night turn network normally operates Monday through Friday five days per week. The UPS network normally operates Tuesday through Friday, with additional service on Monday or Saturday for some markets. The universe under study for the three air networks is all mail moved on each network from domestic origin to domestic destination. It does not include mail traveling on commercial airlines or by air taxi, Alaska, Hawaii, HASP, or Christmas networks.

In mail processing operations, mail pieces are sorted into trays, tubs, sacks, or other mail items. A mail item (handling unit) for air transport undergoes a process in which it is electronically scanned and then tagged with a routing label by a transaction concentrator (TC) serving the origin facility. Each TC is uniquely identified in the EDW database by an alpha−numeric code which is normally six characters. There can be one or more operation work stations associated with a particular TC at the origin facility. The scanning process generates a dispatch and routing (D&R) tag which is attached to the mail item. The D&R tag or routing label contains air carrier information, including the origin and destination cities, mail item total weight and the primary mail class code. The primary mail class code is frequently referred to as the ACT tag code. The EDW database maintains summary weight information for all scanned mail items by TC identifier, date, network flight and ACT tag code. Population control totals obtained from the EDW summary ACT tag totals (see Section VII) are used to expand the sample data.

A. First Stage Sample

The primary sampling unit (PSU) for each TRACS−Network Air subsystem is a facility−day. A facility−day is defined as a specific origin mail facility and date (during the quarter) combination, for which mail items are transported by the carrier on a first−leg flight to the carrier's hub for sort and second leg transport to the final destination city by the same carrier. For each network, the PSUs are grouped into two primary sampling strata: domestic facilities which primarily prepare mail items for domestic destinations, and international facilities which primarily prepare mail items for international exit points. The international mail stratum is comprised of origin mail facilities associated with air-stop codes: JFK, ORD, SFO, MIA and LAX. The domestic mail stratum is comprised of all other origin mail facilities.

The total sample size for each network air subsystem is allocated between the primary strata in proportion to historical pounds obtained from a recent EDW database extraction. Within a primary stratum, the PSUs are subsequently stratified by facility (secondary stratification). The primary stratum total sample size is allocated among the secondary strata proportionally to the historical pounds of mail distribution. For each subsystem, facilities that assign more than a prescribed minimum amount of mail to the network are sampled two or more times during the quarter.

The following table shows TRACS-Network Air subsystem sample sizes for each quarter of FY 2008:

Network Air Sample Size by Quarter: FY 2008

	FedEx Night	FedEx Day	UPS	Total
Q1	350	875	270	1,495
Q2	350	875	270	1,495
Q3	350	875	270	1,495
Q4	350	875	271	1,496
FY	1,400	3,500	1,081	5,981

B. Second Stage Sample

The secondary sampling unit (SSU) within a PSU is defined as a mail item transported by the network air carrier to its hub (first leg). Examples of mail items include sacks, letter trays, flat tubs, and loose parcels. For sampling purposes, a mail item is stratified into one of five ACT tag groups based on the primary mail class (ACT tag) code indicated on the routing tag, as follows: F for First−Class Mail, P for Priority Mail, E for Express Mail, I for International mail and 'O' for all other mail categories. A mail item may contain mail for one or more mail classes or mail categories; however, the ACT tag code indicates the primary mail class within the mail item. Two items are sampled from each ACT tag group, however, if fewer than two items are available for an ACT tag group, items are selected from other ACT tag groups, with preference given to P items for a day turn test, and E items for a night turn test, and then if necessary, to F, E, I and O items, and F, P, I and O items, in this order, respectively. For the UPS subsystem, if less than two mail items are available for sampling, preference is given to F, P, E, I and O mail items in this order. If there are fewer than eight mail items available for sampling, then all mail items are sampled.

For each selected mail item, the ACT tag code and routing information are recorded directly into the data collector's laptop computer. This information is used to link the sampled data to the EDW database. The total weight of the mail item, as well as the mail piece count and net weight, by mail category, are also recorded.

III. ESTIMATION

The estimation process is similar for the network air day turn, night turn and UPS subsystems, with the exception of an average density factor required for the day turn subsystem to convert estimated pounds of mail to estimated cubic−feet of mail, and a residual office stratum used in the UPS subsystem.

A. Distribution Key

The TRACS−Network Air System produces distribution keys each quarter for each network air mode based on weight, volume and cubic feet of mail. For a cubic−foot based distribution key, weight estimates are converted to cubic feet estimates using average densities (average lbs. per cubic foot) for specific mail class groups. Cubic−foot measures are unavailable in the EDW at the mail class level. As a result, a regression model is formulated to apportion EDW aggregated cubic feet to mail classes using known total weight and cubic feet for mixed city containers by mail class.

Program TRACSSMN.FEDEX.PQq08.CNTL(DENSITY) is run each quarter for the air mode serving Priority Mail to develop density factors for P and F mail classes (ACT tag groups). In this program, known EDW total cubic feet (dependent variable) for all mail classes is regressed on known total weight for each mail class (independent variables) using historical origin airport and departure date data from the sampling period. The general formula associating total cubic feet (C) and mail class P or F total weight is:

(d.1) $C = C_P + C_F$, or equivalently,

$$C = (C_P / W_P) * W_p + (C_F / W_F) * W_F \ .$$

In the above equation, C_P / W_P and C_F / W_F are the reciprocal densities for mail classes P and F, respectively.

The general formula for the multiple linear regression model, fitting total cubic feet C to F and P total weight at the combined origin airport *(o)* and departure date *(t)* level, is:

(d.2) $C_{o,t} = d_P * W_{P,o,t} + d_F * W_{F,o,t}$.

In equation (d.2), the coefficients d_P and d_F for the independent variables, $W_{P,o,t}$ and $W_{F,o,t}$, are the reciprocal densities shown in equation (d.1) at the origin-airport-day level.[8]

After density estimates are developed, they are adjusted to represent average densities relative to all mixed city and bypass container mail. The original and adjusted density ratios (d_F / d_P) and (d'_F / d'_P), respectively, are equal.

Electronic data records containing total weight and total cubic-feet measures for F and P mail classes, by origin airport and departure date, are extracted from EDW operational data in a file named: TRACSSMN.FEDEX.DAYNET.EDWCUBE.MIXED.FY08q. This file and an origin airport identifier file (TRACSSMN.FEDEX.DAYNET.EDWCUBE.MAPMC.FY08) are inputs to the density program. Records pertaining to trucked routes, offshore airports and other air modes are filtered. To account for a lag between the flight departure date and the invoice date for flights on a weekend or Monday, records for the three days are collapsed to a single 'combined weekend' day for each origin airport, resulting in summary records for a 5-day week: Tuesday through Friday and the combined weekend day.

The adjusted densities for each quarter of FY 2008 are provided in the following table:

[8] There is no intercept term in the equation since zero mail volume corresponds to zero cubic feet.

Day Turn Densities by Q for FY 2008

Q	F Density	P Density	Overall Density	Data obs.	t-value d_F	t-value d_P
1	9.69	5.02	6.05	4,147	86.90	210.32
2	9.08	5.32	6.26	4153	89.98	182.40
3	9.75	5.56	6.49	4,156	88.95	216.41
4	10.22	5.29	6.38	4,146	113.16	275.85

The TRACS−Network Air subsystem produces separate distribution keys each quarter for the three subsystems based on pounds of mail transported for the night turn and UPS networks, and cubic feet of mail transported for the day turn network. In the expansion process, the sampled pounds of mail for each test are first expanded to the known population total for the test facility and test date by ACT tag group. The expanded test−day pounds of mail for each mail category are then expanded to the total pounds of mail for the facility and quarter by ACT tag group. A final expansion is made to the total pounds for the quarter by ACT tag group for each primary sampling stratum, for each subsystem.

For the night turn and UPS subsystems, the fully expanded pounds are summed across ACT tag groups and primary strata to obtain estimated pound totals by mail category for each subsystem's distribution key. For the day turn subsystem, the fully expanded pound totals are first converted to fully expanded cubic−feet totals before summing across ACT tag groups and strata to develop the estimated distribution key for this subsystem. The conversion to cubic−feet for the day turn subsystem is explained below. The distribution keys are developed in programs: TRACSSMN.FEDEX.PQq08.CNTL(KEYDAY5, KEYNIT5, KEY5X5).

Prior to finalizing the distribution key for the day turn network, a final step is necessary to convert estimated pounds of mail to estimated cubic−feet of mail. Average densities (average pounds per cubic−foot) for ACT tag groups F and P are required for this purpose. Because cubic−feet data are available for combined ACT tag groups (not for the individual ACT tag groups F and P), regression modeling can be used to relate the known total cubic feet to known F and P ACT tag pound totals for mixed city containers, in order to construct F and P density factors.

The following notation is used in the expansion of the TRACS-Network Air subsystem sample data:

AC: ACT tag group = *F, P, E, I, O*

i: test facility index = *1,2,…,i,…,I*

j: sampled mail item index = *1,2,…,j,…,10*

t: test date index = *1,2,…,t,…,T*

h: first-stage sampling stratum = *1 (DOM), 2 (INT), 3 (RES UPS only)*

r: mail category = $1,2,\ldots,r,\ldots,r_n$

d: reciprocal density factor

x: sample weight (pounds)

X : population weight (pounds) total

$\hat{c}$: estimated cubic-feet

$\hat{x}$: estimated pounds

$\hat{R}$: estimated distribution key

For each subsystem, the test facility (*i*) and test date (*t*) mail item sample pounds (x_j) for mail category (*r*) are summed and expanded to the test facility's population (*X*) pounds total for the test date, by ACT tag group (*AC*), as follows:

$$\hat{x}_{AC,h,i,t,r} = \frac{\sum_j x_{AC,h,i,t,j,r}}{\sum_j x_{AC,h,i,t,j}} X_{AC,h,i,t} \quad .$$

The expanded facility-day sample pounds for each mail category are then summed across days and expanded to the facility's population pounds total for the quarter, by ACT tag group, as follows:

$$\hat{x}_{AC,h,i,r} = \frac{\sum_t \hat{x}_{AC,h,i,t,r}}{\sum_t \hat{x}_{AC,h,i,t}} X_{AC,h,i} \quad .$$

The expanded facility total sample pounds for the quarter for each mail category are then summed across facilities and expanded to the sampling stratum (*h)* population pounds total for the quarter, by ACT tag group (*AC*), as follows:

$$\hat{x}_{AC,h,r} = \frac{\sum_i \hat{x}_{AC,h,i,r}}{\sum_i \hat{x}_{AC,h,i}} X_{AC,h}$$

The estimated pounds of mail for each network for mail category *r* for a quarter is the sum of the $\hat{x}_{AC,h,r}$ across all strata and ACT tag groups:

$$\hat{x}_r = \sum_{AC} \sum_h \hat{x}_{AC,h,r} .$$

For the day turn network subsystem, the $\hat{x}_{AC,h,r}$ for mail category *r* are multiplied by a reciprocal density factor, d_{AC} *(AC = F, P)*, to convert estimated pounds to estimated cubic-feet by ACT tag group. The estimated cubic-feet totals are then summed across ACT tag groups and strata to obtain overall estimated cubic-feet for each mail category *r* for the quarter:

$$\hat{c}_r = \sum_{AC,h} \left(\hat{x}_{AC,h,r} * \frac{1}{d_{AC}} \right) .$$

The estimated distribution key based on pounds of mail for the night turn and UPS network subsystems is:

$$\hat{R}_{x_r} = \frac{\hat{x}_r}{\sum_r \hat{x}_r} .$$

The estimated distribution key based on cubic-feet of mail for the day turn network subsystem is:

$$\hat{R}_{c_r} = \frac{\hat{c}_r}{\sum_r \hat{c}_r} .$$

The expansion process is performed in the following programs:

Day turn network subsystem:

Program: TRACSSMN.FEDEX.PQq08.CNTL(KEYDAY5)

Inputs: TRACSSMN.DAYNET.PQq08.ZFILE

TRACSSMN.MAILCODE.FLAT.CODE.FY08

Night turn network subsystem:

Program: TRACSSMN.FEDEX.PQq08.CNTL(KEYNIT5)

Inputs: TRACSSMN.NITNET.PQq08.ZFILE

TRACSSMN.MAILCODE.FLAT.CODE.FY08

UPS network subsystem:

Program: TRACSSMN.FEDEX.PQq08.CNTL(KEY5X5)

Inputs: TRACSSMN.NET5X.PQq08.ZFILE

TRACSSMN.MAILCODE.FLAT.CODE.FY08

Appendix D : TRACS Rail Subsystem

I. Overview

The TRACS Freight Rail subsystem (TRACS−Rail) is a continuous, ongoing statistical sampling system. It produces distribution keys for Inter-BMC (Bulk Mail Center) freight rail cost (account no. 53143). The universe under study is all mail whose freight rail costs accrue to this account.

The primary sampling unit (PSU) for the freight rail subsystem is the origin−destination−day, which is defined as all mail being transported from a given origin BMC to a given destination on a given day. While the Postal Service contracts for highway transportation by route, with one route consisting of multiple trips and stops, freight rail contracts entail just one origin and one destination. A trip between the origin and the destination facility is referred to as a rail movement.

The freight rail sampling frame is a list of all origin−destination−days (PSUs) for which the destination BMC is not a mail bag depository or mailer's plant, and for which the movement is not used exclusively for empty equipment. The freight rail sampling frame is developed using 12 weeks of historical records from the Rail Management Information System (RMIS). The information extracted from RMIS includes the origin BMC, the destination facility, the date of arrival, the number of tractor trailer vans on the movement, and the cost of the movement.

The sample design consists of four stages. In the first stage, a random sample of origin−destination−days is selected from the sampling frame. In the second stage, one van is randomly selected from the vans in the selected PSU. In the third stage, a subsample of wheeled containers, pallets and loose items off−loaded from the test van is selected. From selected containers, a fourth stage sample of items is selected. For pallets and loose items selected at the third stage, there is no fourth stage sample. All selected mail is recorded. The freight rail sample design at the third and fourth stages is the same as the highway sample design at the second and third stages.

Weight and volume information by mail category and shape is recorded for the contents of sampled items. For sampled pallets, the dimensions of the pallet and the percentage of mail on the pallet by mail category and shape are recorded. Data collectors also record the percentage of van floor space occupied by palletized mail, containerized items, and loose items. The sample data are expanded, by mail category, to the cubic−foot−miles of the test van. The cost for the sampled movement is multiplied by the cubic−foot−mile proportions to estimate mail category costs for the movement. The costs for tested movements are then expanded to represent all movements in the quarter.

II. Statistical Study Design

The universe under study is all mail transported via freight rail whose costs accrue to the freight rail account (no. 53143). The mail is moved under contracts negotiated with railroads for a specified length of time. These contracts also designate rates and routing information. The freight rail sampling frame is a list of all origin−destination−days (movements) for which the destination BMC is not a mail bag depository or mailer's plant, wherein, the movement is not used exclusively for empty equipment. The frame is developed using 12 weeks of historical records from the Rail Management Information System (RMIS), thus reflecting expected activity during the sampling period. Each frame unit is a unique combination of an origin, destination and specific day. In addition to the origin and destination facility codes and the date, each frame record also specifies the average historical number of vans in the movement. The frame is updated quarterly.

1. Primary Sampling Units

The primary sampling unit is mail that travels from a given origin BMC to a given destination facility and arrives on a given day. Thus, the primary sampling unit is the origin-destination-day. Primary sampling units are selected at random from the frame. The sample is also selected in a manner which ensures that the proportion of the sample falling in each district is approximately similar to the proportion of the population falling in that district The following table shows the TRACS–Rail first stage sample size by quarter for FY 2008**.**

Sample Size by PQ

Period	Sample Size
PQ1	365
PQ2	361
PQ3	363
PQ4	365
FY 2008	1,454

2. Secondary Sampling Units

The secondary sampling unit is an individual rail van that is part of the sampled movement. A rail van is selected at random from the actual vans observed in the sampled movement.

3. Third Stage Sampling Units

The third stage sample is a stratified sample of off-loaded wheeled containers, pallets and loose items. Mail off-loaded from the vehicle is stratified into five groups depending on the type of mail and the level of containerization. The five groups, or strata, are: pallets, wheeled containers, non-containerized Express Mail, loose sacks, and other loose items. The percentage of floor space utilized by each of these five strata is recorded before selecting the third stage sample. The following rules are used to select the third stage sample:

- Up to two pallets are selected.
- For wheeled containers, if two or fewer containers are unloaded, all are selected. Otherwise, a random start and a skip interval of three are used to select the sampled containers. If the combination of the random start and skip interval results in less than two containers being selected, then a second container is also selected. At most, five containers are selected.
- All loose Express Mail items (*i.e.*, sacks, parcels, etc.) are selected.
- Up to eight loose sacks and other loose items are selected, with at least one of each item type. The item types sampled in TRACS-Rail are trays, tubs, sacks, bundles, loose parcels and pieces, etc. Items are selected in proportion to their presence on the rail van.

4. Fourth Stage Sampling Units

The fourth stage sample is a stratified sample of items from each selected wheeled container. Mail in the container is stratified by item type, and one of each item type is selected. Prior to the selection of the item, the data collector records the container type and the percentage of the container occupied by each item type. These data are used to estimate the cubic-feet for each item type in the container. In some cases, the data collector counts the number of the items in the container rather than estimating the percentage. Then, the cubic-feet are imputed from the number of items.

There is no fourth stage sample for pallets, loose Express Mail, loose sacks or other loose items selected at the third stage. All selected mail is recorded.

For each selected pallet or item, regardless of whether it was selected at the third or fourth stage, two types of data are recorded: item information and mail information. For pallets, the item information includes the dimensions (height, length and width) of the pallet. The pallet dimensions are used for cubic–foot expansion. For mail information, the percentage of pallet space taken up by mail for up to four mail categories is recorded.

The item type is required for all sampled items. The item type determines which expansion formula is used. For mail in the items, the data recorded include the number of pieces and weight by mail category and shape. Weights are converted to cubic-feet based on the density factors shown in Appendix E.

III. Data Expansion

After the sample data are edited, the mail category distribution key is estimated. Mail category proportions of cubic–foot–miles are estimated for each sampled van using item and container sampling weights. The total PSU (Origin–Destination–Date) cost is calculated from the accrued costs found in the RMIS frame. The mail category proportions of cubic-foot-miles are then multiplied by the total PSU cost to obtain the estimated mail category costs for each PSU.

The estimated total mail category cost is the sum of the estimated PSU mail category costs weighted by the PSU design weight. The overall estimated total cost is the sum of the estimated total mail category costs across all mail categories. The distribution key is then estimated by taking the ratio of the estimated total mail category cost to the overall cost.

1. Parameter to Be Estimated

In principle, the distribution key is the ratio of mail category cost to total cost. That is:

$$K_r = \frac{C_r}{C}, \quad (r = 1, 2, \dots R),$$

Here, K_r is the distribution key for mail category r, C_r is the cost of mail category r, R is the total number of mail categories, and $C = \sum_{r=1}^{R} C_r$ is the overall total cost. Therefore, the task is to estimate C_r, the cost of mail category r, for each mail category.

2. Estimation of the Mail Category Cost C_r

The following approximate unbiased estimator is used to estimate C_r**:**

$$\hat{C}_r = \sum_{k=1}^{n} w_k C_{psu,k} \frac{\hat{Q}_{kv,r}}{\hat{Q}_{kv}}, \quad (r = 1, 2, \ldots R),$$

Here, n is the total number of sampled PSUs, w_k is the design weight of PSU k, $C_{psu,k}$ is the total cost accrued to the PSU k aggregated from RMIS, $\hat{Q}_{kv,r}$ is the estimated total cubic−foot−miles for mail category r of van v sampled from PSU k, $\hat{Q}_{kv} = \sum_{r=1}^{R} \hat{Q}_{kv,r}$. And $\hat{Q}_{kv,r}$ is calculated as follows:

$$\hat{Q}_{kv,r} = \sum_{s=1}^{S_{kv}} w_{s|kv} \sum_{i=1}^{I_{kvs}} w_{i|kvs} \hat{Q}_{kvsi,r} ,$$

Here, S_{kv} is the total number of sampled set−aside (container/pallet/loose) items from van v, $w_{s|kv}$ is the design weight of the set−aside, I_{kvs} is the total number of sampled items from set-aside s, $w_{i|kvs}$ is the design weight of item i sampled from set-aside s, and $\hat{Q}_{kvsi,r}$ is the portion of the total set−aside cubic−feet that the mail category r mail in item i takes.

Finally, the estimator of K_r is

$$\hat{K}_r = \frac{\hat{C}_r}{\sum_{r=1}^{R} \hat{C}_r}, \quad (r = 1, 2, \ldots R).$$

3. Expansion Programs

The program used to expand sample data and generate the distribution keys is TRACSSMN.EXPAND.RAIL.PQq08.CNTL(RAILEXP). It has two inputs:

(1) TRACSSMN.EXPAND.RAIL.PQq08.SASDATA(RAILDATA)
(2) TRACSSMN.MAILCODE.FLAT.CODE.FY08.

Appendix E : Tables

Table 1: Mail Codes, Shapes, Descriptions and Surface Density Factors

Mailcode	Shape	Description	Density (lbs/cu ft)
111	Letter	1C Single Piece	16.6384
111	Flat	1C Single Piece	19.6111
111	Parcel	1C Single Piece	19.6111
111	NM-Flat	1C Single Piece Nonmachinable Flat	19.6111
112	Letter	1C Nonautomation Presorted	17.9282
112	Flat	1C Nonautomation Presorted	17.9282
112	Parcel	1C Nonautomation Presorted	17.9282
113	Letter	1C Auto - Non-carrier route	17.8108
113	Flat	1C Auto - Non-carrier route	21.592
113	Parcel	1C Auto - Non-carrier route	21.592
114	Letter	1C Auto Carrier Route	17.8108
114	Flat	1C Auto Carrier Route	21.592
114	Parcel	1C Auto Carrier Route	21.592
121	Letter	1C Single Piece Cards	14.9304
122	Letter	1C Nonautomation Presorted Cards	14.9304
123	Letter	1C Auto Cards - Non-carrier route Cards	14.9304
124	Letter	1C Auto Cards - Carrier Route	14.9304
203	Letter	Foreign Origin Air - International Expres Priority	9.0965
203	Flat	Foreign Origin Air - International Expres Priority	9.0965
203	Parcel	Foreign Origin Air - International Expres Priority	9.0965
400	Letter	Periodicals	25.058
400	Flat	Periodicals	25.058
400	Parcel	Periodicals	25.058
511	Letter	Standard ECR Basic	21.0743
511	Flat	Standard ECR Basic	25.8323
511	Parcel	Standard ECR Basic	25.8323
512	Letter	Standard ECR Basic Automation	21.0743
512	Flat	Standard ECR Basic Automation	25.8323
512	Parcel	Standard ECR Basic Automation	25.8323
513	Letter	Standard ECR High Density	21.0743
513	Flat	Standard ECR High Density	25.8323
513	Parcel	Standard ECR High Density	25.8323
514	Letter	Standard ECR Saturation	21.0743
514	Flat	Standard ECR Saturation	25.8323
514	Parcel	Standard ECR Saturation	25.8323
521	Letter	Standard Non-automation	19.3716
521	Flat	Standard Non-automation	23.0379
521	Parcel	Standard Non-automation	23.0379
522	Letter	Standard Automation	21.0743
522	Flat	Standard Automation	25.325
522	Parcel	Standard Automation	25.325
523	Letter	Standard Not Flat Machinable (NFM)	21.0743

523	Flat	Standard Not Flat Machinable (NFM)	25.325
523	Parcel	Standard Not Flat Machinable (NFM)	25.325
601	Flat	Package Services Parcel Post	5.4263
601	Parcel	Package Services Parcel Post	5.4263
603	Letter	Package Services Bound Printed Matter	9.4905
603	Flat	Package Services Bound Printed Matter	9.4905
603	Parcel	Package Services Bound Printed Matter	9.4905
604	Letter	Package Services Media Mail or Library Mail	7.7639
604	Flat	Package Services Media Mail or Library Mail	7.7639
604	Parcel	Package Services Media Mail or Library Mail	7.7639
700	Letter	USPS Mail	6.5229
700	Flat	USPS Mail	6.5229
700	Parcel	USPS Mail	6.5229
800	Letter	Free Mail	11.0727
800	Flat	Free Mail	11.0727
800	Parcel	Free Mail	11.0727
914	Letter	First Class Mail International	16.7623
914	Flat	First Class Mail International	17.9922
914	Parcel	First Class Mail International	3.4286
921	Letter	Foreign Origin Surface - Letters/Cards	19.6593
921	Flat	Foreign Origin Surface – Flats	23.6509
921	Parcel	Foreign Origin Surface - Other Article	3.4843
923	Letter	Foreign Origin Air - Airmail Letters/Cards	16.9934
923	Flat	Foreign Origin Air - Airmail Flats	20.3652
923	Parcel	Foreign Origin Air - Airmail Other Article	2.8776
925	Letter	Foreign Origin Air - International Expres Priority	9.0965
925	Flat	Foreign Origin Air - International Expres Priority	9.0965
925	Parcel	Foreign Origin Air - International Expres Priority	9.0965
999	Letter	Competitive Product	n/a
999	Flat	Competitive Product	n/a
999	Parcel	Competitive Product	n/a

Table 2: Known Item Size

Item Type	*Known Cubic-Feet*
Full size envelop tray	0.749
Half size envelop tray	0.374
Flat tub or box	1.490
Small parcel tray	2.813
Con-con	8.50

Table 3: Default Tare Weight Density Factors

Item Type	*Density Factor (pound/ft^3)*
Sack with tare weight <=1.5 (lb.)	3.05
Sack with tare weight >1.5 (lb.)	6.66
Express Mail	3.05

Table 4: Known Container Size

Container	*Default Cubic-Feet*
BMC-OCR	110.61
ERMC	49.34
GPC/GPMC/APC	48.64
Hamper	30.96
Wiretainer	33.33
Short Postal Pak	42.22
Other Wheeled	48.64
Tall Postal Pak	76.67