MPA-T-5

BEFORE THE
POSTAL RATE COMMISSION
WASHINGTON, D. C. 20268-0001

RECEIVED
JUN 1 4 29 PM '00
POSTAL RATE COMMISSION
OFFICE OF THE SECRETARY

POSTAL RATE AND FEE CHANGES, 2000 : Docket No. R2000-1

DIRECT TESTIMONY OF

ANTOINETTE CROWDER

ON BEHALF OF

MAGAZINE PUBLISHERS OF AMERICA, INC.
ADVO, INC.
ALLIANCE OF NONPROFIT MAILERS
ASSOCIATION FOR POSTAL COMMERCE
ASSOCIATION OF AMERICAN PUBLISHERS
DIRECT MARKETING ASSOCIATION, INC.
DOW JONES & COMPANY, INC.
MAIL ORDER ASSOCIATION OF AMERICA
THE McGRAW-HILL COMPANIES, INC.
NATIONAL NEWSPAPER ASSOCIATION
PARCEL SHIPPERS ASSOCIATION
AND
TIME WARNER INC.

CONCERNING

CITY DELIVERY CARRIER STREET TIME COSTS

June 1, 2000

TABLE OF CONTENTS


Page

PURPOSE OF TESTIMONY ..... 1

I. SUMMARY OF RESULTS AND CONCLUSIONS ..... 2

II. THE ES WORK SAMPLING STUDY IS NOT RELIABLE FOR COSTING PURPOSES ..... 6

A. Factors That Adversely Affect The Reliability Of The Work Sample Tallies For Costing Purposes ..... 7

1. The Work Sampling Study Was Not A Central Focus Of The ES Project ..... 7

2. The Heavy Workload Of The Data Collectors ..... 9

3. The Limited Instruction And Training Of Data Collectors ..... 11

4. Turnover of Data Collectors ..... 13

5. The Lack Of Precise Definitions Of Terms ..... 13

B. The LR-I-163 Data -- The Black Box ..... 17

III. THE ES SAMPLE DESIGN IS NOT REPRESENTATIVE OF THE SYSTEM OF CARRIER ROUTES AND IS BIASED TOWARD HIGHER-LOAD ROUTES ..... 20

A. The Non-Random Sample Design And Clustering Of Sites ..... 20

B. The Exclusion Of Data From The LR-I-163 Data Set ..... 22

C. The Lack Of Statistical Support ..... 24

D. The Small Sampling Ratios And Sample Bias ..... 27

IV. THE POSTAL SERVICE'S OPERATIONAL EXPLANATIONS DO NOT EXPLAIN OR VALIDATE THE HUGE INCREASE IN LOAD TIME IMPLIED BY THE ES WORK SAMPLING RESULTS ..... 34

V. ES VIDEOTAPES AND OTHER ES DATA DEMONSTRATE THAT THE WORK SAMPLING TALLIES SUBSTANTIALLY OVERSTATE TRUE LOAD TIME ..... 41

A. The ES Videotapes and Time Studies ..... 41

B. The Postal Service and MPA Regression Analyses of ES Load Time v. Possible Deliveries Indicate A Substantial Overstatement Of True Load Time ..... 43

Page

VI. THE POSTAL SERVICE'S INCORRECT ATTRIBUTION OF THE POSSIBLE DELIVERIES EFFECT ON LOAD TIME ..... 46

Autobiographical Sketch ..... APPENDIX A

Comparison Of MPA And USPS Regressions Of ES Load Time Data ..... APPENDIX B

## PURPOSE OF TESTIMONY

I am Antoinette Crowder, a senior consultant with TRANSCOMM, Inc. in Falls Church, Virginia. My autobiographical sketch is set forth in Attachment A.

My testimony addresses the costing of city delivery carrier out-of-office (or street) costs. I focus primarily on the testimony of Postal Service witness Lloyd Raymond, USPS-T-13, which presents the results of a work sampling study of delivery carrier street activities conducted as part of the Postal Service's Engineered Standards/Delivery Redesign (ES) project. The Postal Service uses data from this study to develop delivery carrier street time proportions, in place of proportions derived from the Street Time Survey (STS) used in previous rate cases. This study implies a very large change in the time spent by carriers loading mail, and results in a substantial increase in total accrued and attributable load time costs. For 1998, use of the ES time proportions in place of those from the STS results in a $970 million increase in accrued load time, a 52% increase.

My testimony explains why these new out-of-office time proportions are unreliable, greatly overstate load time costs, and are not acceptable for postal ratemaking purposes. I evaluate Mr. Raymond's work sampling study and explain why its data are unrepresentative of the Postal Service system of routes and unreliable for postal ratemaking purposes. I also present the results of my analyses of the ES data, including: (1) ES videotape data collected at the same time as the ES work sampling data, (2) a regression analysis of the ES data, and (3) an assessment of the Postal Service's LR-I-310 regression analysis of that same data. The two regression analyses, based on the ES data, demonstrate a large overstatement of true load time. Finally, I explain why, if the ES data are used to estimate load time, then the matching ES variability model must also be used to develop load time volume variability. Otherwise, there will be a large and irrational overstatement of volume variable load time cost.

## I. SUMMARY OF RESULTS AND CONCLUSIONS

1 During this proceeding, in an effort to fully evaluate their proposals, I have
2 extensively reviewed the data and participated in discovery on USPS witnesses
3 Raymond and Baron. The intent was to acquire enough information to independently
4 (1) assess the processes by which the ES data were collected and used for
5 ratemaking purposes, (2) evaluate their reliability for postal costing purposes, (3)
6 determine whether the large increases in load time costs implied by the ES results
7 were consistent with the Postal Service's explanations of operational changes in city
8 carrier delivery, and (4) assess how well the ES load time results fit with the volume
9 variabilities developed by Witness Baron from other data sources.[1] The following are
10 my conclusions:

11 (1) The ES data are from a work sampling study conducted as part of the
12 Delivery Redesign/Engineered Standards project. Unfortunately, although they may
13 have been adequate for the purpose they were intended to serve, they are clearly
14 unreliable for ratemaking purposes. The data collectors were given no consistent
15 definitions of carrier activities and, in fact, were deliberately given latitude in
16 interpreting and recording the codes for those observed activities. Many of the study
17 codes are exceedingly vague and generalized and could be interpreted in a number of
18 ways. And, most importantly, the ES observers were unaware of the activity
19 breakpoints critical to ratemaking costing purposes. Consequently, there is no way to
20 determine what the various categories of tallies (i.e., observations of instants of
21 carrier activity) actually represent or whether they are even internally consistent. It is
22 virtually unimaginable, therefore, that the various categories of ES tallies -- collected
23 without any knowledge they would be used for ratemaking purposes -- could be

---

1 In this testimony, true load time is as defined in the Load Time Variability (LTV) study.

consistent with the precise definitions of carrier out-of-office activities that are used for ratemaking (i.e., load time, run time, street support time).

(2) My attempts to validate the ES data for purposes of accurately measuring load time have all indicated that those data are unreliable and considerably overstate the amount of true load time. From the ES data, I have developed a regression model of route-level ES load time vs. possible deliveries and delivery mode, which indicates that the ES load time estimate substantially overstates true load time (as defined for ratemaking purposes). My results are consistent with the results of a separate route-level ES model (of load time vs. volumes and possible deliveries) provided by the Postal Service in this case. Both models indicate overstatement of true load time.

(3) I have reviewed a sample of ES videotapes of carrier activities on residential park & loop routes, recorded on the same day the ES observers were collecting their work sampling data. This analysis shows that the observation codes recorded by the ES data collectors and allocated by witness Raymond to the load time category are not consistent with the load time definition used for ratemaking, and generally measure a time that is considerably greater than true load time.

(4) The sample of routes from which the ES data were collected is exceptionally small, considering the diversity of routes in the postal system. Further, the sample was not randomly selected. Indeed, it is biased toward large metropolitan areas and areas with relatively higher delivery point growth rates. Compared to the USPS system of letter carrier routes, the weighted sample routes over-represent residential curbline routes and high-load time centralized delivery points on residential park & loop routes. There is a clear bias toward geographic areas and deliveries with higher-than-system-average load time characteristics.

(5) My analysis of the Postal Service's explanations of operational changes in the delivery system since 1988 shows that they do not explain the magnitude of the large increase in load time implied by the ES data, and the corresponding large

decrease in run time. Although there have been changes, the large differences between the STS and ES load and run times are far more the result of (i) the biased ES sample, and (ii) the manner in which the ES data collectors recorded their observations and Mr. Raymond allocated them among load, run, and support categories.

(6) From the above, it is clear that the ES estimates of load, run and support time are not consistent with the definitions used for ratemaking. Because they do not match the variabilities derived from models that are consistent with those ratemaking definitions, the resulting volume variable costs are meaningless. Accordingly, witness Baron (USPS-T-12) was wrong in applying variabilities from the Load Time Variability (LTV) models, which measure the true amount of load time, to the greatly overstated load time developed from the ES data. His approach causes an extreme overstatement of variable load time and inaccuracy in all the variable out-of-office costs.

On the other hand, when the variability from the Postal Service's LR I-310 variability model (derived from the ES data) is applied to the ES estimate of total load time, the estimate of variable load costs is substantially less than that developed by witness Baron. That, of course, is because the ES load time variability is developed from a model that recognizes that a portion of the ES load time is not true load time. The ES estimate of load time and the matching ES variability produce a far more reasonable estimate of variable load time than that now offered by witness Baron.

(7) For both the stop-level LTV load time model and the route-level ES load time model, the Postal Service proposes to attribute the load time effect from the possible deliveries variables included in the models. This reflects a basic misunderstanding as to how load time varies with deliveries coverage (i.e., proportion of deliveries that receive volume). In each model, the volume variables measure the full extent of the effect of volume on load time. The effect of possible deliveries on load time is not caused by volume and should not be attributed to volume.

(8) Finally, because of the lateness of some data, I have not had time to review them, much less incorporate them into my analyses. A substantial amount of information that should have been presented at the time the Postal Service filed this case was not made available until much later. This includes information on route-days omitted from the LR I-163 data base used by witness Baron to develop the out-of-office time proportions, correct and full information on the ES observers, electronic versions of volume data, and certain recently received interrogatory responses from witness Raymond. Additionally, I have had a very short time to review the videotapes and associated written data that the Postal Service recently provided. Nevertheless, I have seen sufficient information to know, without any doubt, that the ES data are not appropriate for the use to which the Postal Service has put them.

Based on the above evaluation, I recommend that the Commission reject the use of the ES data for ratemaking purposes. However, if it decides to use these data to develop out-of-office time proportions, it should also use the ES model variability that is consistent with them. Otherwise, variable load time costs will be severely overstated.

The rest of my testimony is organized in five sections. Section II describes why the ES work sampling data are unreliable for purposes of rate making. Section III describes the biased ES sample design. Section IV addresses the Postal Service's explanations for the large difference between the STS and ES load and run time estimates. Section V explains the analyses that demonstrate that the ES work sampling data overstate load time and, accordingly, that the ES measure of load time is completely mis-matched with the load variability derived from the LTV models. Section VI explains why the effect of possible deliveries on either stop-level or route-level load time should not be considered volume-variable. An Appendix describes my ES load time model. Workpapers and a description of the videotape analysis are being filed as library references.

## II. THE ES WORK SAMPLING STUDY IS NOT RELIABLE FOR COSTING PURPOSES

1 Mr. Raymond's work sampling study was only one part of the large Engineered
2 Standards/Delivery Redesign project that had multiple but interrelated objectives,
3 including development of at least two sets of time standards for delivery carrier
4 activities, identification of preferred work methods, development of a route adjustment
5 and workload management system, and validation of the workload management
6 system. *Costing was not an objective.* The data collection portion of this project was
7 likewise large, multi-faceted and complex, requiring data collectors to employ a
8 number of simultaneous data collection techniques (time studies, videotaping, work
9 sampling, and other tasks) to gather information. This total effort is what Mr.
10 Raymond and the Postal Service call the Engineered Standards (ES) project.[2]
11 Mr. Raymond's work sampling study was not designed for ratemaking costing
12 purposes and was conducted prior to any thought that it be used in ratemaking.[3]
13 While this, standing alone, does not automatically invalidate use of the data for
14 costing, it does raise a warning flag. As MPA witness Keith Hay explains (MPA-T-4),
15 the purpose, nature, design, and execution of a study can have a huge impact on its
16 usefulness and reliability for a different purpose other that for which it was originally
17 designed.
18 A costing study should be designed to focus directly and clearly on the costing
19 question at issue. In particular, a study of carrier costs:

---

2 There has been some confusion in the use of the term "ES data" because the ES project collected more work sampling data than was presented by Mr. Raymond and provided to witness Baron for developing carrier street costs. In my testimony, I generally use the term "ES data" to refer to that portion of the work sampling data that was provided to Mr. Baron.

3 MPA/USPS-T13-1, 2, Tr. 7607-08; OCA/USPS-T13-8, Tr. 7816.

(1) Should be carefully designed to ensure a random sampling of routes that are representative of the universe of carrier routes.

(2) Should have clear systematic instruction and training of data collectors focused on the precise cost-related demarcations of the specific carrier activities to be observed, to ensure uniformity of interpretation and application by all data collectors.

(3) Should have clearly defined terms to minimize confusion or subjective interpretations by data collectors.

(4) Should employ a relatively simple, focused data collection procedure, without a lot of other data collection responsibilities to distract or interfere with the central costing objective.

By these standards, as I explain below, the ES work sampling study is unreliable for costing purposes, even though it may have been appropriate for the different purposes for which it was designed.

### A. Factors That Adversely Affect The Reliability Of The Work Sample Tallies For Costing Purposes

#### 1. The Work Sampling Study Was Not A Central Focus Of The ES Project

A work sampling project design should be geared toward the purpose it is to serve. In the case of Mr. Raymond's work sampling, its purpose was to provide ancillary support for the main focus of the Engineered Standards project: to develop engineered time standards and work methods for management. The primary tools employed to develop time standards were the time studies, a predetermined time system, and videotape analyses, not work sampling. Mr. Raymond explained that the work sampling data were intended to provide the following:[4]

- A measure of the portion of the carrier's day during which delays were experienced which could then be used as an input in developing the time standards,

---

4 USPS-T-13 at 5; ADVO/USPS-T13-1, Tr. 7368; NAA/USPS-T13-3, 7776-81; cross-examination at Tr. 8050-53.

- An estimate of the proportion of delivery receptacles by type (e.g., a #1 box vs. a one-handed slot), which could then be used to develop overall estimates of delivery workload from the standards, and

- A general indication of where carriers were spending their time, so that the Project could prioritize parts of the carrier's work day that might yield the largest productivity gains, if improvements in work methods were made.

The only information from the work sampling that was needed as a direct input in developing the time standards was a "delay factor." Tr. 8052-53. This focus on identifying a delay factor is evident in the design and layout of the barcode sheets. Nearly one-third of the barcodes for outside activities and details (21 of 66 barcodes) are related to "delay" activities. USPS-T-13, Appendix C at 23.

The other aspect of the work sampling data that Raymond was interested in was receptacle types, because the time standards vary by receptacle type and frequency of receptacle types was important for his route adjusted workload management system. Because information on receptacle types was not captured in the time studies, the work sampling was the only source for this information. Tr. 8050-51. This focus also is evident from the barcode layout. More than one-fourth of the barcodes for outside activities and details (17 of 66) are for receptacle type. The importance of receptacle type is also seen in the "details" arrow on the barcode sheet, which directs the entering of a receptacle type. See also Tr. 7581. Moreover, data collectors recorded receptacle types even when there was no delivery activity, for example, when the carrier was walking or traveling between deliveries. Data collectors were to record the receptacle type that was "near" the carrier.[5]

Indeed, it appears that the time studies, conducted at the same time as the work sampling, were a more important part of the project. The time studies were

---

5 MPA/USPS-T13-80; Tr. 7701. This focus on collecting receptacle-type information may have caused some data collectors to pay more attention to the kind of receptacle than to the precise location of the carrier when nearing the delivery point.

1 originally intended to be used to develop engineered time standards (although Mr.
2 Raymond ultimately used a "predetermined" system for time standards). The greater
3 importance placed on the time studies is evident from many factors. In determining
4 the sample size for his project, Mr. Raymond focused on the number of time studies
5 needed, rather than on the size of the work sampling. OCA/USPS-T13-1;
6 ADVO/USPS-T13-23(b); Tr. 7797-801, 7407. The training of data collectors also
7 apparently focused on the time studies. They were given specific instruction on such
8 things as how to operate the equipment and how to identify different receptacle types,
9 but were given no written instructions on how to identify the locations and activities in
10 the work sampling. Moreover, time studies interrupted and took precedence over the
11 work sampling, requiring data collectors to remember and later record intervening
12 observations.

**2. The Heavy Workload Of The Data Collectors**

13 The ambitious multiple-method scope of the ES data collection
14 effort was probably necessary for Mr. Raymond's purposes. However, even with two-
15 person teams, the data collectors were required to perform numerous, often
16 overlapping tasks for long hours. The team arrived up to an hour before the beginning
17 of the route to collect quantitative information concerning the route, such as daily
18 volumes and types of deliveries. They then followed the carrier throughout the day
19 collecting data on carrier activities and other carrier/route conditions until the carrier's
20 day was completed, following the carrier throughout the day. The team had numerous
21 overlapping and sometimes simultaneous responsibilities: the time studies, the
22 work sampling data collection, the videotaping of carrier activities, and driving/walking
23 behind the carrier. Clearly, there was a lot of work to perform:

> "Every six minutes, when the scanner beep went off, they typically performed the work [activity] sampling. They would take time studies of the various outside activities counting the appropriate items such as: number of paces walked, number of delivery points served, number of doors and gates, number of weighted or un-weighted bends made, number of trays/tubs handled, distance in tenths of miles, final odometer reading. The team also had a daily comments log for making notes about any special events, and corrections to scans. They would also videotape outside activities for approximately 1/2 hour. The video would be shot at various times throughout the street time." ADVO/USPS-T13-11, Tr. 7384-85.

In addition, the data collectors had to refer to a Postal Service Form 3999X (listing addresses, delivery points, and other important route information) and make manual notes and corrections on the form. After the carrier had finished for the day, the data collectors had to review their data, print out reports, consult their daily comments log, mark up their records, and arrange for their data to be sent to the central ES data processing location. *Id.* The typical workday for a data collector was ten to twelve hours, and some days were much longer.[6]

This heavy workload over long hours, with many responsibilities beyond just the work sampling, raises questions about the reliability of the observations for costing purposes. This is a lesser concern for some of the other data collection activities such as the time studies, videotaping, and counting of various items which

---

6 ADVO/USPS-T13-47, Tr. 7487, and cross-examination at Tr. 8046-47. On a given day, as many as 45 time study observations could have been completed, each requiring between ten and 22 scans and manual counts of various observed items related to the activity being timed. The first half of a time study required 5 scans and, after the counts were made, the second half of a time study required between 5 and 17 additional event quantities scanned. As many as 100 or more work sample tallies could occur during the day, at 6-minute intervals, each requiring 6 barcode scans. Additional scans were made to record weather conditions, satchel weight, and other items. Typically at the end of the day, there could be up to 53 additional study quantities scanned. ADVO/USPS-T13-34.

did not require much interpretation and judgment, but it is a real concern with respect to the work sampling.[7]

### 3. The Limited Instruction And Training Of Data Collectors

The most troubling aspect of the instruction and training of data collectors is the absence of any written instructions or definitions of terms about how to interpret and record the work sampling observations.

Phase 1 data collectors were given orientation classes but no written instructional material. Tr. 7411. Formal instruction appeared to concentrate on time study requirements, equipment familiarity (i.e., the video camera and the scanner), identification of receptacle types and postal forms, and role-playing situations. Id. The many subsequent data collectors were trained almost entirely on-the-job. Apparently, no individual acted as the common instructor for all new data collectors. With no written instructions, this created the clear potential for differing instructions and interpretations depending on the on-the-job training matchups.

Mr. Raymond, in fact, specifically sought out data collectors who had no prior knowledge of postal operations and intentionally simplified their training so they would "simply record what was happening":

> "We determined that the best means to collect data would be one that was easy for the data collector to learn. In this way, the data collectors could stay focused on the activities they were observing, rather than on how to record them. The pool of data collectors was selected from a non-Postal environment so as to minimize the

---

[7] Unlike the work sampling, the time studies were designed in a way to minimize the need for definitions of terms and subjective judgments about how to categorize and record observations. A good example is the time study for "Basic Delivery," which tracked a carrier during a segment of deliveries. The data collector entered the start and stop times for the study by simply scanning at a time of his choosing, and then count the number of occurrences of items or activities (such as the number of deliveries made, paces walked, and parcels delivered) during that time period. The work sampling, by contrast, required the data collector to identify, at the instant a beep sounded, where the carrier was and what he was doing, and then to properly record the observation from among numerous ill-defined barcode choices.

> potential role of preconceptions. It was our goal for the observers to simply record what was happening, not what 'should' happen. . . ."
>
> "The teams would be better off comprised of knowledgeable non-postal people in order to record what is/was and not what may look good, and to freshly evaluate all aspects of the city carrier operation. People from a wide variety of occupations, with diverse experience levels, and different educational backgrounds using technology-supported processes would contribute the most." . . . .
>
> "I made the decision that the way to document 'What happens in the day of a life of a carrier' would be best accomplished by on the job training." NAA/USPS-T13-3

Accordingly, although the data collectors were instructed on how to use the equipment and interact with the public and the observed carriers, they were given no systematic oral or written instruction on what the work sample barcodes meant, and appeared to be deliberately left to their own devices to determine and categorize (via the barcode scans) what they were observing.[8] As Mr. Raymond stated:

> "Once the observers were made familiar with the materials they would use, the observers were simply instructed to record what they saw." ADVO/USPS- T13-82, Tr. 7560.

Again, for Mr. Raymond's purposes, this may have been adequate. But for costing purposes, it is imperative to have clear and precise definitions of terms that equate to the costing distinctions intended to be measured. Simply instructing the data collectors to "record what they saw" is not adequate unless they are also given precise definitions telling them *how* to interpret and record what they saw. For familiar, well-defined activities such as counting the number of paces the carrier walked during a time study, "recording what they saw" may not be a problem, since even someone unfamiliar with postal operations or costing concepts can make those objective determinations. That is not the case with the interpretation of more specialized terms such as "point of delivery," which may mean different things to

---

8 MPA/USPS-T13-61, 62, 64, 67, 68, 70, 71 and 79; Tr. 7677-78, 7680, 7683-85, 7687-88, 7700.

1 different people, particularly if they are unfamiliar with postal operations or the precise
2 definitions used for postal costing purposes.

**4. Turnover of Data Collectors**

3 Although the total is not entirely clear, at least 52 data collectors
4 were employed at different stages over the course of the project. Tr. 7551. Eight were
5 used in Phase 1, although two left near the end of that phase. During Phase 2, at
6 least three groups of newly-hired data collectors were hired and given varied training.
7 One group included three of the original Phase 1 data collectors plus six new hires.
8 The second group consisted of 18 new data collectors. Of these 24 new data
9 collectors hired near the beginning of Phase 2, 16 left the project or accepted different
10 responsibilities. ADVO/USPS-T13-28, Tr. 7448. Approximately 20 additional data
11 collectors were brought in to fill vacancies as they occurred. Over the course of Phase
12 2, a total of 47 data collectors were employed at one time or another to fill 24
13 positions.

14 Obviously, there was difficulty retaining data collectors. Whatever the reasons,
15 this high turnover and the recurring need to train new data collectors on-the-job raises
16 concerns about the reliability and consistency of data tallies. This concern is
17 particularly pronounced because of the lack of written instructions and definitions of
18 terms, the reliance on on-the-job training of the newly-hired data collectors by other
19 data collectors, and the lack of a centralized system of instruction.

**5. The Lack Of Precise Definitions Of Terms**

20 A survey being conducted for costing purposes should start with
21 clearly defined terms that correspond to the costing categories to be measured.
22 Moreover, the meanings of key terms should be conveyed to the data collectors
23 precisely and consistently. This is not something that should be left up to the data
24 collector's interpretations, with the vague instruction to just "record what you see."

1 For the purpose of determining precise costing measures, the work sampling
2 project falls short on both scores. Many of the key barcode terms are vague and
3 overlapping, and subject to differing interpretations. The lack of precision in the
4 barcode terms might have been less of a concern had the data collectors been given
5 precise, consistent written definitions and centralized instruction on how to interpret
6 and apply the terms. But no written definitions or instructions were given to the data
7 collectors, and most of the training was decentralized on-the-job training by other data
8 collectors. This lack of precision in the barcode terms heightens the probability of
9 misrecording of carrier activities.

10 A good example is the critical term "point of delivery," which accounted for the
11 vast majority of tallies Mr. Raymond assigned to load. To the average person with no
12 familiarity with carrier operations or the costing definitions of "load time," this term can
13 mean many different things. Is the carrier at the "point of delivery" when he enters a
14 building? When he is standing by his vehicle near a centralized or NDCBU
15 receptacle? When he is approaching the doorstep in preparation for delivery? If the
16 data collector is not given precise definitions and instructions on what "point of
17 delivery" means, the resulting tallies can represent a hodge-podge of activities
18 extending beyond the precise costing definition of load time.

19 This confusion over "point of delivery" is evident in tallies Mr. Raymond
20 assigned to load that show the carrier's activity as walking or traveling between
21 deliveries while supposedly at the "point of delivery."[9] These demonstrate that some
22 data collectors interpreted point of delivery broadly to include instances where the

---

[9] Another example are point of delivery tallies for dismount deliveries with an LLV (vehicle) activity code, which he assigned to load, explaining that the carrier "is delivering to a Dismount type delivery, to a residential outside delivery point *from the LLV*." (emphasis added). Aside from the impossibility of being at a dismount delivery point when inside the vehicle, this suggests that, for some ES data collectors, the delivery point began at the vehicle parking point.

1 carrier was at another location or was approaching but not physically at the mail
2 receptacle.[10]
3 Other location terms in the survey are likewise imprecise. Mr. Raymond's post-
4 survey definition of the location "on route" is when the "carrier is between the 1st
5 delivery point and the last delivery and has not deviated from his route *and is not at*
6 *another listed location*." MPA/USPS-T13-39, Tr. 7654 (emphasis added). By this
7 definition, "on route" means that the carrier is *not* at the "point of delivery," yet he
8 assigned a number of "on route" tallies to load.[11] Likewise, the location "vehicle"
9 overlaps with other locations, such as "point of delivery" on curbline deliveries.
10 Moreover, the "vehicle" location was apparently interpreted to apply when the carrier
11 was only near the vehicle, and not necessarily physically in the vehicle. This is similar
12 to the interpretation of when to enter receptacle types, which Mr. Raymond stated was
13 to represent the receptacle type "near" the carrier.
14 A similar problem occurs with the activity "delivery/collect." Mr. Raymond's
15 survey identifies two types of activities that are purportedly associated with loading
16 mail: "fingering at delivery" and "delivery/collect." The first term, fingering at delivery,
17 conveys a specific activity which even a non-postal-expert can comprehend -- the
18 activity of handling and sorting the mail in preparation for inserting into the mailbox.
19 But the second term, delivery/collect, is a generic term which, by itself, does not
20 convey a specific activity.
21 The problem with the generic "delivery/collect" term is that, like "point of
22 delivery," it could be construed to cover activities going beyond the technical Load

---

10 These ambiguities about "point of delivery" might have been resolved by a more precise term that is less subject to interpretation, such as "at the mail receptacle" which conveys a specific physical location.

11 Some of these tallies list the Activity as walking, while others list the Activity as Delivery/Collection but indicate that the carrier was walking at the time -- neither of which is consistent with the STS/LTV concept of load.

1 function. If one asked a non-expert what a carrier is doing as he approaches the
2 doorstep of a house, the answer might well be that he is "delivering the mail." After
3 all, that is the carrier's job -- delivering and collecting mail.

4 A related problem with this generic term is that for load operations which occur
5 quickly (e.g., only a few seconds), data collectors probably had difficulty discerning
6 whether a "beep" that occurred around that time actually coincided with the quick load,
7 or shortly before or after that load, which should have been access time.[12] A busy
8 data collector, even though he may be instructed to observe what is occurring the
9 instant the six-minute beep sounds, may not have been quite so focused on such a
10 difficult activity/location breakpoint -- particularly since they were not made aware of
11 the critical need for precision due to the then-unforeseen use of this data for postal
12 costing purposes. The absence of precise definitions and the other demands on the
13 data collectors' time increased the likelihood that the observations recorded by the
14 data collectors and converted to load tallies by Raymond were not observations of true
15 load time.[13] "Delivery/Collect" would be the convenient entry, particularly since the
16 data collectors were not aware at the time of the critical importance of precision for
17 costing purposes.[14]

---

[12] My review of the ES videotapes indicates that, in many single delivery situations, the physical acts of fingering mail, inserting it into a mail receptacle, and/or removing collection mail from a mail receptacle can often occur so quickly that one has to be exceptionally watchful and precise in order to determine when it begins and ends vis-a-vis run time.

[13] Based on my limited review of the ES videotapes, there do appear to be examples where the data collectors recorded activities that were not true load time activities with codes that Mr. Raymond ultimately assigned to load time.

[14] A better, more precise term for costing purposes might have been something like "depositing/collecting mail," which conveys a specific activity of placing mail into or retrieving mail from a receptacle.

**B. The LR-I-163 Data -- The Black Box**

1 I have spent much time reviewing the general categories of work
2 sampling tallies within Mr. Raymond's LR-I-163 ES data set. My first activities were to
3 review a distribution of tallies by tally type, comparable to his LR-I-281 distribution of
4 tally types, and attempt to identify the types of carrier activities described by those
5 tallies to see if they matched the STS categories to which he had assigned them.
6 After an extensive review of these data, but still having a number of outstanding
7 questions about them, I have concluded the following:
8 (1) The ES code definitions were broad, imprecise, and contained no
9 specified breakpoints (i.e., between driving a vehicle, parking it, getting out and
10 moving around it, moving to or away from it; between accessing a mail receptacle and
11 loading it; between actually loading a vehicle and moving containers to it). The data
12 collectors had no consistent guidelines as to how to apply the codes. The potential
13 for misinterpretation expands when one recognizes that two of the three codes (the
14 Activity and Activity Detail codes) were often linked.[15] It thus seems virtually
15 impossible that all the data collectors could have applied the codes consistently, or
16 that anyone could identify from the ES tallies what the carrier was actually doing when
17 he was observed.

---

15 An example is the focus on collecting information on receptacle types (needed for the important time standards and route restructuring aspects of the ES project) and the linkage in the survey between the receptacle type activity detail codes and the delivery/collect activity code. Data collectors were instructed to enter receptacle types in conjunction with delivery/collect activity tallies. Yet a receptacle tally does not necessarily mean that the carrier was delivering mail or was even at the point of delivery -- only that the carrier was "near" the receptacle. In addition to the potential that this predisposed data collectors to capture receptacle information by overusing the delivery/collect-receptacle type tally combination, it also inhibited the entry of other activity detail codes, such as walking codes, that might have more accurately described what the carrier was doing at that instant.

1 (2) Mr. Raymond cannot possibly explain precisely what carrier activities
2 were being observed when each tally was taken because the codes are too general
3 and their application too inconsistent. However, his judgmental interpretation of the
4 mix of codes in each tally, and his subsequent assignment to an STS category,
5 depend critically upon consistently applied definitions among all his ES data
6 collectors and himself. Otherwise, how can he really know what was occurring when
7 the tally was taken? In his response to POIR No. 8, he makes a number of
8 assumptions about what he thinks the codes mean and what he thinks the STS
9 definitions mean. For example, he assumes "delivery point" is defined in precisely
10 the same way for both. Thus, all Point of Delivery tallies are allocated to the STS Load
11 category. This appears to be the only way he could claim that he experienced no
12 problems in allocating the ES data to the STS categories. MPA/USPS-T13-52, Tr.
13 7667. In some cases, however, he admitted that he needed to reference the data
14 collector comments log or the USPS Form 3999X. But, in most cases, I suspect even
15 referring back to those items cannot be sufficient.
16 (3) Many of the most numerous tally groupings have the superficial
17 appearance of being indisputable and precise, "neatly falling" into STS categories
18 without any difficulty. This appearance is deceptive. The key codes to look at are the
19 Location and Activity codes, and there is little variation here. With the exception of the
20 Personal (PBL) tallies, all the tallies described in the first 3-1/2 pages of Mr.
21 Raymond's POIR No. 8 Attachment A are combinations and permutations of three
22 vague and interchangeable Location codes (Point of Delivery, Vehicle, On Route) and
23 five Activity codes (Delivery/Collection, Finger @ Delivery, Travel Between Delivery,
24 Travel Between Delivery with Sort, and Setup), which could have been interpreted by
25 the data collectors in any number of inconsistent ways.
26 (4) Despite all the effort I have made to understand these data, neither I nor
27 any other analyst could possibly identify with sufficient precision (to satisfy the

1 required STS/FAT/CAT/LTV requirements) what carrier activities are represented by
2 the various tally groupings. And, because of the potential diversity of interpretations
3 among the various ES data collectors, the general nature of the codes themselves,
4 and the brief STS definitions given to Mr. Raymond, I believe such precise
5 interpretation would be difficult even for someone who had participated in his study.

## III. THE ES SAMPLE DESIGN IS NOT REPRESENTATIVE OF THE SYSTEM OF CARRIER ROUTES AND IS BIASED TOWARD HIGHER-LOAD ROUTES

1 In a properly structured costing study, the sample of zip codes and routes
2 selected would have been chosen entirely randomly and with the assurance that all
3 route types were adequately represented (i.e., with, perhaps, some sample stratifica-
4 tion to ensure adequate sampling of rarer route types and a representative mix of
5 large and small zip codes). Mr. Raymond's sample is not representative of the entire
6 system of USPS letter routes and, in fact, is biased toward routes with a larger
7 proportion of in-office and load time.

### A. The Non-Random Sample Design And Clustering Of Sites

8 The ES database (LR-I-163) from which the Postal Service has
9 developed time proportions for its entire system of city delivery carrier routes was
10 derived from an activity sampling study of only 340 routes, most of which were clearly
11 not selected on a random basis. The approach was to choose cities from which a
12 variable number of routes would be sampled, apparently based on time and budget
13 constraints. Other than ensuring that cities were chosen in each of the Postal
14 Service's eleven regions, there was no apparent sample design for the work
15 sampling project. As the project progressed, cities and routes appear to have been
16 sampled on a relatively ad hoc basis. Although this may be appropriate for industrial
17 engineering projects, it does not meet ratemaking costing standards, where the
18 objective of the cost study is specified first and the sample is designed in a
19 representative manner to meet that objective.[16]

20 For the LR-I-163 set of routes, 53 cities were sampled. Most of the cities were
21 selected by Postal Service regional management, with Headquarters' direction that
22 the selected cities have relatively high DPS volumes, at least three delivery units (with

---

16 See, for example, the standards applied to selected Postal Service data collection systems in the A.T. Kearney Data Quality Study, Technical Report #2: "Statistical Analysis of Data Quality Issues," April 16, 1999.

1 Delivery Unit Computers), and a mix of route types (mounted, park & loop, and
2 business and residential; no rural carriers). Within each city, the routes were
3 apparently randomly selected.[17] The following table summarizes the sample in LR-I-
4 163.

SUMMARY OF USPS- AND RANDOMLY-SELECTED SITES

| | USPS Selected | Randomly Selected | Total |
|---|---|---|---|
| Regions | 11 | 8 | 11 |
| States | 22 | 9 | 26 |
| City Codes (CYs) | 43 | 10 | 53 |
| Zip Codes | 58 | 18 | 76 |
| Routes | 254 | 86 | 340 |
| Route Days | 744 | 101 | 845 |

5 The total number of 53 City Codes (CYs), however, disguises the fact that a
6 large number of these CY codes were actually geographically very close to each other,
7 with many even in the same metropolitan area, as shown in the table below:

METROPOLITAN AREAS WITH MULTIPLE OBSERVED SITES

| Metropolitan Area | USPS Selected CY Codes | Random Selected CY Codes | Total CY Codes |
|---|---|---|---|
| Cincinnati, OH | 3 | | 3 |
| Dallas/Ft. Worth, TX | 4 | 1 | 5 |
| Detroit, MI | 2 | 1 | 3 |
| Los Angeles, CA | 2 | | 2 |
| Milwaukee, WI | 3 | | 3 |
| Bay Area, CA | 2 | | 2 |
| Seattle, WA | 4 | | 2 |
| Springfield/Worcester, MA | 2 | 1 | 2 |
| Tampa, FL | 3 | | 3 |
| Trenton, NJ | 3 | | 3 |
| Washington, DC Metro | 3 | | 3 |
| Subtotal 11 Metro Areas | 31 | 3 | 33 |
| Total CY's in Database | 43 | 10 | 53 |
| **% from 11 Metro Areas** | **72.1%** | **30.0%** | **62.3%** |

17 ADVO/USPS-T13-2, Tr. 7369. However, questions remain about the true "randomness" of the random-selected sites, as discussed later.

1 Remarkably, just these 11 metropolitan areas account for 33 of the 53 City Codes in
2 LR-I-163 -- 62% of the total CY codes in the survey. Moreover, these 11 areas account
3 for 54% of both the total Zip Codes (42 of 76) and total carrier routes (182 of 340) in
4 the database. This table also shows that three of the ten randomly-selected sites
5 were in metropolitan areas that already had at least two USPS-selected sites.
6 Another 14 CY codes were in cities with populations well over 100,000, with a median
7 of 300,000.[18] Thus, the CY sites are not only highly clustered but skewed toward
8 large metropolitan areas.[19]

**B. The Exclusion Of Data From The LR-I-163 Data Set**

9 Another problem discovered shortly before Mr. Raymond's appearance
10 is that the LR-I-163 database excluded a large number of carrier routes and route-day
11 observations that were collected over the same period as the LR-I-163 data. The ES
12 project collected activity sampling data from at least six other cities, an unknown
13 number of other zip codes and routes, and 175 other route-days.[20] Based on those
14 figures, approximately 10% of the cities and 17% of the route-days were excluded
15 from the data the Postal Service used to develop the ES time proportions. Recently, in
16 response to ADVO/USPS-T13-63, Mr. Raymond briefly explained why six cities were
17 excluded from the LR-I-163 data set (including one randomly selected city). Those
18 excluded cities, however, apparently accounted for only about 14 of the missing 175
19 route-days (based on information in Raymond's response to ADVO/USPS-T13-61).
20 Moreover, for some CYs in LR-I-163, there are observed routes that are not included

---

18 New York, Chicago, San Antonio, Virginia Beach, Greensboro, Tallahassee, Little Rock, Omaha, Wichita, Ponce, PR, St. Charles, MO, Jackson, MI, Columbus, GA, and Jackson, MS.

19 Even the total of 845 route-days of observations is misleading. Just four cities accounted for 55% of those total observations (465 of 845).

20 LR-I-293, response to ADVO/USPS-T13-23(b).

1 in the data set. Also, for some routes in LR-I-163, there are some excluded route-
2 days. Unfortunately, full identification and evaluation of excluded data was apparently
3 not conducted by Mr. Raymond and could not be independently performed.[21]
4 In any type of study presented in rate proceedings or elsewhere, it is assumed
5 (unless otherwise stated) that the filed database includes the entire data set
6 encompassed by the study. If portions of the data have been purged from the data set
7 for whatever reason (e.g., deletion of outliers, data scrubbing, problems with the data
8 collection), this information is customarily divulged at the outset along with
9 explanations.[22] Such information is obviously of critical importance in assessing the
10 reliability of the data being presented, for a number of reasons. First, purged data
11 may have a bearing on the study's sample design and the representativeness of the
12 remaining data. Second, elimination of any data should be based on well-designed
13 decision rules that should be carefully explored and tested. Third, it may be that, upon
14 investigation, some or all of the omitted data should have been included in the
15 database and were improperly omitted. Fourth, and conversely, disclosure and
16 explanation concerning the purged data may, upon investigation, reveal that additional
17 portions of the filed dataset should also be purged for similar reasons. Fifth, if the
18 quantity of purged data is large (as is clearly the case here), exploration of the
19 reasons for large-scale purging may indicate that the study was poorly designed or
20 implemented, or that the results are unreliable. In short, independent review and
21 validation of a study requires an assessment of both the excluded data and the

---

21 Less than three weeks ago, we finally received a diskette containing information on the missing route-days. However, in the press of preparing testimony and reviewing many other just-received discovery responses, including ES videotapes, I have not had time to look at the missing route information, much less analyze it.

22 Mr. Raymond's written testimony and his response to an early interrogatory that no route days were purged from the data (ADVO/USPS-T13-16, Tr. 7393) led me to believe initially that LR-I-163 contained the entire data set.

1 reasons for their exclusion. This has not been possible due to the extreme lateness
2 of the disclosure of these problems and the inadequacy of Raymond's explanations.

### C. The Lack Of Statistical Support

3 Although statistical validity and representation of the universe of carrier
4 routes should be a prime concern in ratemaking, Mr. Raymond conducted no analysis
5 of the LR-I-163 data to determine whether it was representative of the USPS system of
6 routes on an annual basis.[23] In response to questions concerning the represen-
7 tativeness of his sample, he has instead provided LR-I-293. This library reference
8 includes three items: (a) a summary comparison of the unweighted results from all of
9 his USPS-selected routes versus those of his randomly selected routes, (b) a
10 comparison of age and gender of his sampled carriers to all carriers in the system,
11 and (c) a spreadsheet, with some data missing, that purports to calculate the
12 sampling requirements for the ES time studies (but not for the work sample study).
13 Despite his assertions to the contrary,[24] none of these demonstrate
14 statistically that his sample is representative of the Postal Service system of letter
15 carrier routes on an annual basis.[25] His only analysis of sampling requirements was
16 based on statistical requirements for the time studies. As Dr. Hay explains, there is a

---

23 Response to ADVO/USPS-T13-56, Tr. 7497. Apparently, the Postal Service did not either. In a response redirected from Mr. Raymond, Mr. Baron attempts to support the ES data by segmenting the data into route types and then calculating confidence intervals for the weighted sample results by route type. As Dr. Hay explains, because the data were not collected in a statistically valid random manner, Mr. Baron's attempt is simply a mechanistic exercise having no real meaning. MPA-T-4 at 7-8.

24 OCA/USPS-T13-1, MPA/USPS-T-13-13; Tr. 7797-7801, 7615.

25 His comparisons of unweighted volumes and tallies for the random v. USPS-selected routes (LR-I-293, "ADVO23lr1Tbl.xls") is particularly unpersuasive. The unweighted totals skew the results by giving too much weight to routes that were heavily sampled over multiple days. More importantly, the "random" routes are not representative of the system. They share the same skew toward zip codes with substantially-above-average numbers of carrier routes as do the USPS-selected routes.

1 considerable difference between (1) quantifying the number of sufficient time and
2 motion segments for purposes of time studies and (2) quantifying the appropriate
3 number of routes, by route type, for purposes of determining a statistically valid
4 sample of routes for purposes of ratemaking. MPA-T-4 at 13-14. Mr. Raymond has
5 not demonstrated that his sample is representative of the entire system of letter
6 carrier routes.

7 Further, his comparison of the results from routes in random and USPS-
8 selected cities satisfies no statistical requirements.[26] For each route type, the
9 number of routes from the randomly selected cities is quite small and cannot be
10 considered to be either reliable or representative of the entire system.

RANDOM ROUTES BY ROUTE TYPE

| Randomly Selected Sites | Number of Routes | Number of Regions |
|---|---|---|
| Residential Curb | 56 | 8 |
| Residential Park & Loop | 19 | 7 |
| Mixed Park & Loop | 3 | 2 |
| Foot | 3 | 2 |
| Mixed Curb | 2 | 1 |
| Business Motorized | 1 | 1 |
| Unknown | 2 | 1 |

11 Mr. Raymond does not quibble with the fact that the random routes cannot be
12 considered representative:

13 "The randomly observed routes are a respectable sample but is
14 [sic] not large enough to represent the total population of routes. It
15 does not include the demographics of: carrier classification mix,
16 route type mix, delivery point mix, age and gender mix for the ES
17 study." OCA/USPS-T13-6, Tr. 7807.

---

26 Surprisingly, although he excluded six cities and 175 route days from the LR-I-163 database, they are included in LR-I-292 which he proffers as evidence of the representativeness of the data in LR-I-163. And, although his sampling approach clearly involves a cluster sample, he did not sample weight any of his data. This is because he had no rigorous sampling plan related to route-level data..

This obviously and strangely conflicts with the response he gave to OCA/USPS-T13-1: "Based on the comparison of the data we collected from the random routes to the Postal selected routes we feel the all data [sic] should be considered as random and representative of the population." Tr. 7801. If the random routes are not representative of the system, then they obviously cannot be used as a benchmark to claim that the USPS-selected routes are representative. The number of routes from the randomly-selected cities could not possibly be large enough to validate the representativeness of the routes, by route type, which the Postal Service selected in a very deliberate manner.

Moreover, it is not clear how "random" those routes from the randomly-selected cities were. Mr. Raymond describes the selection of the random cities:

> "We used Excel® to generate a random number list for the Postal Service to use in the selection of the random sites. The Postal Service picked the sites in my presence from a listing of finance numbers." OCA/USPS-T13-6, Tr. 7807.

Yet three of the ten randomly selected sites were from metropolitan areas that had at least two USPS-selected sites. In addition, the zip codes in both the random sites and the USPS-selected sites have a substantially larger-than-system-average number of city delivery carrier routes -- an average of 26 routes per zip code for the USPS-selected sites and 24 for the random sites, compared to the system average of only 14. Tr. 7262-68.

This raises questions about whether the routes from the randomly-selected cities were truly random. Were all finance numbers in the country on the USPS list, or only finance numbers for a subset of cities or regions? Once a city was randomly selected, how were the sampled zip codes chosen from among multiple zips within the city? Did the selection of a finance number lead, in a truly random manner, to the selection of specific random zip codes and routes? These questions arise because of (a) the geographical proximity of some random cities to selected cities, (b) the high

1 routes-per-zip code skew of both the random and selected cities, and (3) differences
2 in the mix of city delivery offices and zip codes among the randomly selected cities.[27]
3 A non-random selection of delivery offices or zip codes within a randomly-selected city
4 does not produce a random result.

**D. The Small Sampling Ratios And Sample Bias**

5 To determine the representativeness of the ES routes, I have compared
6 them to available information on the total system of letter routes. Sampling ratios
7 were among the first items checked. Two of the more obvious problems are that: (1)
8 the sample is extremely small, and (2) the sample appears biased as a result of
9 Postal Service Headquarters' directives and the need of the ES project to cluster
10 sample cities and routes.

**1. Sample Size**

11 For each individual route type, all the ES unweighted sampling
12 ratios are less than 0.5% of system total routes of that route type. For the three most
13 populous route types, the ES foot and residential park & loop routes have a sampling
14 ratio of less than 0.2% while the ES residential curb routes have a sampling ratio of
15 little over 0.3%. In total, the ES routes are approximately 0.2% of total routes in the
16 system. On a regional basis, the sampling ratios range from 0.04% for the Northeast
17 region to 0.39% for the Pacific region. The business and mixed routes do not have a
18 sufficient sample size to even attempt to claim randomness. Given the size and
19 diversity of the Postal Service's system of routes, these unweighted sampling ratios
20 are extremely small.

---

27 A city finance number can encompass a large number of zip codes served from different delivery stations. Yet for some random sites with a large number of city-wide zip codes, the selected zip codes are all served out of the same delivery office, while for others the selected zip codes are served out of different delivery offices.

Although Mr. Raymond has dismissed the concern about diversity within the Postal Service system of routes,[28] it is an important and recognized consideration. In order to adequately collect a statistically valid and representative sample of overall activity time proportions on routes, by route type, one would have to design a sample that would account for the variation in, for example, the following route/carrier characteristics:

- General geographic and climate conditions;
- Residential and commercial delivery point conditions (i.e., mode of delivery, type of receptacle, type of delivery point, customer types, distances between delivery points, coverage percentages);
- Daily, weekly and seasonal volume effects (both in quantity and mix);
- In-office work space and methods;
- Out-of-office travel accommodation, traffic and walking conditions, type of vehicle used, parking point availability;
- Local supervision and union considerations.

I do not believe the ES sample of 340 routes with an unweighted sampling ratio of 0.2% of total Postal Service routes is adequate to fully represent the above conditions -- particularly given the non-random, clustered nature of the chosen sites. This non-randomness is further illustrated by the fact that 44% of the routes in LR-I-163 (148 of 340) were observed during the peak season of October, November, and December.

**2. Sample Bias**

Sample bias is indicated by the fact that there are noticeable differences in the mix of route types between (a) the weighted sample routes and (b)

---

28 In response to NAA/USPS-T13-4 (Tr. 7788), Mr. Raymond stated: "My experience with sampling populations is that when the population gets very large, no matter what criteria the company uses to pick a sample, the results are very close to a random sample." I disagree completely. To ensure adequate representation, a sampling plan becomes especially important when a population is very large and very diverse.

1 total routes in the Postal Service system, a difference which also occurs by postal
2 region. In particular, the weighted ES sample has a substantially greater proportion of
3 residential curb routes and a substantially lesser proportion of residential park & loop
4 routes than exist in the total system.

SYSTEM v. ES SAMPLE – DISTRIBUTION OF ROUTE TYPES

| Route Type | Total System Routes | Distribution of Total Routes | Un-Weighted ES Routes | Weighted ES Routes | Distrib. of Wghtd. ES Routes |
|---|---|---|---|---|---|
| Foot | 19,115 | 11.5% | 36 | 223 | 13.8% |
| Resident. Loop | 82,908 | 49.9% | 110 | 670 | 41.7% |
| Resident. Curb | 51,486 | 31.0% | 159 | 680 | 42.2% |
| Mixed Loop | 5,220 | 3.1% | 8 | 10 | 0.6% |
| Mixed Curb | 4,056 | 2.4% | 20 | 26 | 1.6% |
| Business Motor. | 3,332 | 2.0% | 3 | 3 | 0.2% |
| Total | 166,107 | 100.0% | 336 | 1,612 | 100.0% |

5 As residential loop and residential curb are the two principal route types in the
6 system, the disparity between the ES sample and the universe is particularly
7 indicative of sample bias. That bias is toward cities and zip codes with a greater-
8 than-average number of residential curbline routes and more densely populated
9 delivery areas.

10 In addition, the sampled zip codes differ significantly from the system in terms
11 of the average number of routes per zip code. The ES sampled zip codes averaged
12 26 routes while the national average for zip codes with letter carrier routes is only 14.
13 Of the 76 zip codes sampled, only 11 (14.5%) had less than the national average
14 number of routes, 65 (84.5%) had more than average, and 28 (36.9%) had more than
15 twice the national average number of routes. This disparity extends to the regional
16 level as well. In each Postal Service region, the sampled zips averaged more routes
17 per zip code than the average for that region. For example, in the Allegheny region, the

1 sampled zips averaged 30 routes per zip compared to the region average of only 13
2 routes.[29]
3 This bias toward larger zip codes (i.e., more densely populated areas) is not
4 surprising since the Postal Service Headquarters' directive to its Regions was to
5 identify cities with multiple delivery units having a mix of both business and residential
6 curbline, loop, and foot routes. Given the increased Postal Service reliance on
7 curbline and dismount delivery modes over the past 20 years, this likely creates bias
8 toward cities in high-delivery point growth areas.[30] And, in order to sample foot and
9 business/mixed routes, relatively dense delivery areas would have likely been
10 selected.
11 In addition, the sampled distribution of routes by route type among the regions
12 differs from the distribution of the total. For example, regions in the north and east
13 (the Northeast, New York Metro, Allegheny, Capital Metro, Mid-Atlantic, and Great
14 Lakes regions) account for 51.7% of the systemwide residential park & loop routes,
15 but only 42.6% of the ES sample. Conversely, regions in the south and west (the
16 Southeast, Midwest, Southwest, Western, and Pacific regions) account for only 48.8%
17 of the systemwide residential park & loop routes, but 57.4% of the ES sample A
18 representative sample would be expected to have a distribution among regions more
19 in line with the distribution of total system residential park & loop routes. Instead, it
20 appears that residential park & loop routes in the north and east are under-
21 represented while those in the south and west are over-represented. The under-

---

29 All calculations are presented in my workpapers.

30 One way to see this is to consider the four high-growth regions of the country: the Southwest, Southeast, Pacific and West. About 60% of the ES weighted residential curbline routes were from those four regions while only about 40% of the ES weighted residential park & loop routes were from those regions. Obviously, the older, lower-growth regions of the country have a greater concentration of park & loop routes.

represented north and east likely have a greater proportion of lower-population growth areas or older delivery areas that are more likely to be served by park & loop routes., whereas the south and west likely have a greater proportion of higher-population growth areas or newer delivery areas. Similar situations occur for the other route types as well.

Results from the ES foot routes provide a clear demonstration of how unrepresentative (biased) sampling can affect results. Almost 85% of the ES foot routes come from the New York Metro, Pacific and Great Lakes regions while only 48% of total system foot routes are in those regions. The results of a single, extreme ES foot route in the New York Metro region (out of only four foot routes sampled in that region), after sample weighting, accounts for approximately half of the weighted load time proportion calculated for foot routes. Tr. 7239-55. Even Mr. Baron was "somewhat troubled" by that result. Id. at 7255.

Another indication of non-representativeness and bias of the ES sample compared to the Postal Service system is the disparity in the number and mix of delivery types for residential park & loop routes, the most numerous route type. The ES weighted sample clearly has more residential NDCBU and central deliveries and fewer residential other/loop deliveries than do all such routes in the system. The ES sample also has more possible deliveries.

RESIDENTIAL PARK & LOOP ROUTES -- DELIVERY TYPES

| Residential Park & Loop Route Delivery Types | ES Weighted Sample Average Per Route | ES Weighted Sample Distribution | Total System Average Per Route | Total System Distribution |
|---|---|---|---|---|
| Resident. Curbline | 31 | 6.3% | 38 | 7.9% |
| Resident. NDCBU | 35 | 7.1% | 29 | 6.0% |
| Resident. Central | 131 | 26.2% | 88 | 18.3% |
| Resident. Other/Loop | 274 | 54.8% | 298 | 61.8% |
| Business | 28 | 5.7% | 29 | 6.0% |
| Total | 500 | 100.0% | 482 | 100.0% |

NDCBU/central deliveries constitute one-third of the total ES park & loop route deliveries, compared to less than one-fourth in the system. Moreover, the average ES park & loop route has 42% more NDCBU/centralized deliveries than the average system park & loop route.

This bias toward NDCBU and centralized deliveries results in an overstatement of average load time on park & loop routes. When compared on a regional basis, the ES sample characteristics (mix of delivery types and total number of possible deliveries) differ considerably from those for the total residential park & loop routes in the same region. In particular, residential park & loop routes in the Allegheny region are severely under-represented and those routes, on average, have a low proportion of NDCBU/centralized delivery points. This is further evidence that the ES sampled zips are biased toward more densely-populated and/or higher delivery-point growth areas. In any case, greater load time, as a proportion of total out-of-office time, is expected with NDCBU/centralized delivery points, and the ES park & loop sample is biased toward delivery points requiring high load times.

* * *

The kinds of problems I have found with the ES data collection, described above and in Section II, are the very kinds of survey design and execution problems that the Data Quality Study, prepared by A.T. Kearney, Inc., cautions about:

"Bias includes three components:

1. Bias due to survey design errors (e.g., incomplete lists from which to sample) includes data collection errors and errors in analysis. In each case, the appropriate quantity was estimated, but it was estimated incorrectly or incompletely.

2. Bias due to failures during implementation (e.g., data recording errors, missing data, failure to follow sampling procedures).

3. 'Definitional bias' is estimating the wrong quantity or using the wrong data source for the relevant economic issue (e.g., using the wrong distribution key)."[31]

As the Data Quality Report observes, "[s]implicity is important for implementation by geographically dispersed data collectors." Id. at 13. It also notes that "[i]ncreasing the sample size of a survey may not increase accuracy appreciably if the survey is subject to bias." Id. at 11. For these and the other reasons I have described, the ES work sampling, while perhaps appropriate for its original purposes, is not appropriate for postal costing purposes.

---

[31] Technical Report #2: "Statistical Analysis of Data Quality Issues" at page 11, A.T. Kearney Data Quality Study, April 16, 1999.

## IV. THE POSTAL SERVICE'S OPERATIONAL EXPLANATIONS DO NOT EXPLAIN OR VALIDATE THE HUGE INCREASE IN LOAD TIME IMPLIED BY THE ES WORK SAMPLING RESULTS

The ES time proportions for letter carrier out-of-office time are dramatically different from the 1986 STS proportions. System-wide, the changes for the major categories are as follows:[32]

| | STS | ES |
|---|---|---|
| Load Time | 25.15% | 38.15% |
| Foot Run Time (FAT) | 41.59% | 29.49% |
| Curbline Run Time (CAT) | 9.14% | 3.92% |
| Drive Time | 7.20% | 11.01% |
| Sum of Load, FAT, CAT, and Drive | 83.08% | 82.57% |

The ES data imply a large, unexplained shift away from both CAT and FAT run time and toward load and drive time, as shown by the following load and run time per stop comparisons.[33]

**Load Time**

| | 1986 Load Time Per Actual Stop (STS) | 1998 Load Time Per Actual Stop (ES) | Change |
|---|---|---|---|
| SDR | 11.79 sec. | 17.04 sec. | 44.6% |
| MDR | 75.56 | 114.35 | 51.3% |
| B&M | 21.67 | 36.21 | 67.1% |
| Wtd. Avg. | 17.37 | 26.01 | 49.7% |

**Run Time**

| | 1986 Run Time Per Actual Stop (STS) | 1998 Run Time Per Actual Stop (ES) | Change |
|---|---|---|---|
| CAT runtime | 5.47 sec. | 2.66 sec. | - 51.4% |
| FAT runtime | 25.09 | 20.11 | - 19.8% |
| CAT + FAT | 30.56 | 22.77 | - 25.5% |

32 STS is based on 1998 accrued costs by route type. This table excludes street support which, on a system-wide basis, has remained relatively unchanged. All data and calculations are in my workpapers.

33 Averages represent total accrued load, CAT, and FAT times converted to minutes and divided by CCS total actual stops.

These are enormous changes. Even if the ES proportions were assumed to be representative and reliable (which they are not), average accrued load time per actual stop would have had to increase 50% since 1986 while average run time would have had to decline by 26%. This implies that the proportion of route time (excluding street support) spent by carriers loading mail has increased from 30% to 50%. The obvious question is how could the system have changed so much?

Postal Service witness Kingsley offered several explanations why load costs have increased since 1988. She cited increases in cluster boxes and other centralized delivery points, motorization and curbline, delivery point coverage, and volume per delivery. She also cited changes in mail mix and DPS implementation. USPS-T-10 at 27-28.

Upon closer examination, however, her explanations fall far short of explaining the magnitude of the changes implied by the ES data. I believe this is principally because the ES data, for the reasons I have described earlier, do not provide a representative or reliable indication of true cost proportions conforming to ratemaking costing criteria or standards. Ms. Kingsley's specific points are addressed below.

**Growth in Cluster Boxes**. While there has been a growth in centralized and NDCBU receptacles since 1988, this growth has undoubtedly been concentrated in new housing developments and outlying suburban areas. It should not have had much impact on foot and park & loop routes in established neighborhoods.[34] Moreover, data from the USPS Carrier Cost System (CCS) indicate that this growth in cluster boxes has not had as much impact on the system as implied:

- The proportion of actual Multiple Delivery Residential (MDR) stops (which include all residential NDCBU and centralized stops) has

[34] The vast majority of the possible stops in existence in 1998 were in existence and embedded in the system in 1986 (e.g., 16.4 billion CCS stops in 1998 compared to 14.7 billion in 1986).

remained virtually unchanged since 1986 -- from 7.49% of total system stops in 1986 to 7.55% in 1998.

- The proportion of actual Single Residential Delivery (SDR) stops has declined by less than 0.5% (from 84.44% to 83.99%) while the proportion of actual Business and Mixed (B&M) stops has increased by less than 0.4% (from 8.07% to 8.46%).

- Coverage ratios (actual stops to possible stops) for SDR and MDR stops have increased only slightly, while that for B&M has actually decreased slightly between 1986 and 1998.

| | 1986 CCS Coverage | 1998 CCS Coverage |
|---|---|---|
| SDR stops | 92.1% | 92.9% |
| MDR stops | 97.3% | 97.4% |
| B&M stops | 91.1% | 89.9% |

Since MDR stops represent only 7.6% of the total CCS stops in the system, this change cannot possibly explain the huge differences between the STS and ES time proportions.

**Replacement of Foot Routes and Increased Motorization.** Foot route costs have declined from 20% to 11% of total letter carrier costs since 1986. However, the proportion of foot *plus* park & loop accrued costs has actually increased from 82.1% to 82.4%, indicating that most replaced foot routes (which are generally in more urbanized areas) have been converted to park & loop routes. Such conversions should save some walking-related "travel time," but would not likely change the actual geographic delivery conditions of those areas (i.e., location of delivery unit, delivery density, and walking conditions). Accordingly, the motorization of foot routes would likely save some FAT walking time related to traveling between groupings of deliveries in urbanized areas, but would be partially offset by added drive time traveling between parking points.

The motorization of some foot routes, however, cannot explain the extremely large decrease in the systemwide proportion of FAT walking time -- from 42% to 29% -- implied by the ES data. First, converted foot routes appear to account for only

1 roughly 10% of all routes. Second, the total of foot plus park & loop routes, which
2 account for the great majority of FAT walking time, has not changed since 1986. Third,
3 because foot routes are concentrated in relatively dense areas with shorter distances
4 between delivery points, the amount of FAT walking time saved by conversion to park
5 & loop routes would not be large.[35]

6 Changes in curbline routes likewise cannot explain the ES results. The
7 proportion of curbline routes in the system appears to have actually declined slightly
8 since 1986 -- from 17.0% to 16.1%. While this does not necessarily mean that the
9 proportion of curbline stops in the system has declined, it clearly does not support the
10 Postal Service's suggestion that curbline stops have increased substantially. In fact,
11 the ES data appear to be inconsistent with such an increase. Those data imply not
12 only that total run time (CAT curbline plus FAT walking time) has declined sharply
13 since 1986, but that the CAT proportion of total run time has also declined sharply,
14 from 18% to less than 12%.[36] For this to be believed, there would have to be dramatic
15 declines either in the proportion of curbline stops or the average distances between
16 curbline stops. Yet if, as the Postal Service implies, curbline stops have increased

---

[35] Average drive time likely has increased somewhat due to an increase in dismount stops to serve centralized and NDCBU delivery points. MPA/USPS-T10-8, Tr. 1710. However, dismounts are operationally comparable to small loops and should also have FAT run time associated with them. Although it is not surprising that the ES data show an increase in drive time relative to the STS, that increase is likely overstated because the sampled ES routes have a substantially greater-than-system-average proportion of centralized and NDCBU deliveries. Further, the under-representation of residential single delivery points and over-representation of multiple delivery points in the ES sample, especially in more urban/suburban areas, may help explain why the ES FAT and CAT run time proportions are substantially less than those from the STS. Another contributing explanation is the misrecording of some FAT access time as load time.

[36] Total system CAT and FAT proportions. This ES-implied reduction in the CAT proportion of total run time occurs also within the two major categories of routes, residential curbline and residential park & loop.

substantially since 1986, then one would expect to see an increase in the proportion of CAT run time and a reduction in FAT run time. The ES data imply just the opposite.

Ms. Kingsley could not explain these seemingly anomalous FAT/CAT run time results, and redirected questions to Mr. Baron. See MPA/USPS-T10-21, Tr. 7181-83. Baron's attempt at explanation was not illuminating. When asked to provide an "opinion and rationale" why average access times have declined so dramatically since 1988, he simply said the reason was the "substitution into the BY98 worksheets of the new street-time percentages presented" in Mr. Raymond's testimony. MPA/USPS-T10-21 (redirected from Kingsley), Tr. 7183. In follow-up to that response, when asked to "provide your explanation or opinion of why" access time "has declined so much in nine years," he again merely repeated that the reason was the "substitution of the new street-time proportions derived from the ES database for the proportions derived from the 1986 STS study." ADVO/USPS-T12-12, Tr. 7095. *The Postal Service's inability to explain these results suggests that the "substitution of the ES proportions" is the problem, not the explanation.*

**Introduction of DPS.** DPS letters are obviously new since 1986. However, their effect on load costs is not fully apparent. Kingsley herself stated that carriers do not spend more time at the delivery point "fingering" DPS mail because the fingering activity occurs as they walk between deliveries. Tr. 1774, 2110-11. Further, using the USPS estimate of the total increase in carrier time due to DPS, and assuming that all letters are DPS and that all additional cost is load cost (clearly extreme assumptions), the increased time would not come close to explaining the large ES-estimated accrued load time per stop.[37]

---

[37] See USPS-T-10 at 27, where Ms. Kingsley states that the additional handling cost for DPS is estimated at approximately one hour per 5000 pieces, including all in-office and out-of-office additional time associated with DPS.

1 **Increases in Deliveries Coverage, Volume Per Delivery, and Mail Mix**. Ms.
2 Kingsley implied that the large change in average load time is caused by large
3 increases in volume per delivery and changes in mail mix since 1988. However, she
4 overstated the effect of volume growth on load costs by focusing on pieces per
5 *possible* delivery, rather than pieces per *actual* delivery which have likely grown by only
6 roughly 3% since 1988. Tr. 1948-52, 2107. Even compared to 1986, the changes
7 have not been large. For Single Delivery Residential stops which represent 84% of
8 total stops, volume per stop has increased only 8.2% and the mix of volume has
9 changed only slightly since 1986.[38]

| SDR Stops | 1986 | 1998 |
|---|---|---|
| Total Volume per Stop | 5.03 | 5.44 |
| Proportion of Letters | 75.4% | 73.2% |
| Proportion of Flats | 23.5% | 25.5% |
| Proportion of Accountables | 0.2% | 0.1% |
| Proportion of Parcels | 0.9% | 1.2% |

10 In contrast to these relatively small changes in volume and mix since 1986, the
11 ES time proportions imply that accrued load time per SDR stop has increased 44.5%
12 -- from 11.8 to 17.0 seconds! By comparison, the SDR Load Time Variability (LTV)
13 model indicates that these volume changes since 1986 should have caused only a
14 6.6% increase in stop time.[39] The extremely large increase in accrued SDR load time
15 implied by the ES data is implausible and unexplained.
16 Such large disparities raise substantial questions. Although the LTV models
17 do not reflect DPS handling, they have been the best available means for determining
18 how system-wide changes in volume, mail mix, and possible deliveries per stop

---

38 Following Ms. Kingsley's appearance, the Postal Service acknowledged that "there have not been any substantial changes in the proportion of residential delivery volumes which were flats." ADVO/USPS-9, Tr. ___.

39 This disparity between ES implied load time and LTV modeled time occurs for MDR and B&M stops as well.

1 impact load time per stop. They also were constructed using a clear and reliable
2 definition of load time, while the same cannot be said for the ES data. Certainly they
3 provide a reasonable general indication of how much true load time should have
4 changed since 1986, and cast serious doubt on the validity of the ES load time
5 estimates.

6 In short, while there have been operational changes since 1988, the Postal
7 Service's explanations could not account for the enormity of the increased load time
8 implied by Raymond's data and analyses. To the contrary, some of the explanations
9 seem inconsistent with the ES results, casting further doubt on their reliability for
10 postal costing purposes.

## V. ES VIDEOTAPES AND OTHER ES DATA DEMONSTRATE THAT THE WORK SAMPLING TALLIES SUBSTANTIALLY OVERSTATE TRUE LOAD TIME

1 For the reasons discussed in previous sections, it became apparent to me that
2 Mr. Raymond's results substantially overstate the proportion of true load time costs in
3 the Postal Service's city delivery carrier system. This overstatement is due to a
4 combination of two factors: sampling errors due to a non-representative sample
5 design, and misrecording errors due to data collectors not always interpreting and
6 recording observations of carrier activities in accordance with the costing definitions
7 used for rate case purposes. To test this second factor, I analyzed some ES
8 videotapes and regression models of ES data as described below. These analyses
9 confirm that the ES load tallies include a portion of time that does not represent true
10 load time, as defined for postal costing purposes.

### A. The ES Videotapes and Time Studies

11 ES data collectors following a particular carrier not only recorded work
12 sample observations of the carrier's activities at six-minute intervals, but also
13 videotaped portions of the carrier's route and conducted time studies of delivery
14 segments, recording information on when the time study began and ended, the
15 number of deliveries the carrier made, the total paces the carrier walked during that
16 time, and other information. From my earlier investigation of the ES data, it had
17 become apparent that (1) the proportion of load time tallies was overstated, and (2) at
18 least part of this overstatement was likely due to ES data collectors recording non-
19 load activities with codes that Raymond ultimately allocated to load time. Moreover,
20 the extremely large increase in the load time proportion on park & loop routes led me
21 to suspect that these routes were one source of overstatement. In April, when we
22 finally were allowed to obtain restricted access to confidential materials, I decided to
23 review some of the videotapes to see if they might be useful in analyzing load time
24 proportions. The results of my review are provided below.

1 Videotape Analysis. My analysis of ES videotapes focused on a selection of
2 park & loop routes that have relatively consistent delivery characteristics throughout
3 the day, and sufficient videotape data to obtain a reasonable understanding of route
4 conditions and load-run time characteristics.[40] Due to time constraints and tape
5 quality limitations, I have been able to fully review and process the videotape data for
6 11 the 29 routes initially viewed at the Postal Service's Merrifield office.[41] The
7 videotapes for each of these routes were carefully reviewed and timed to determine
8 the amount of load time for each delivery and the elapsed loop times for the segment,
9 as described in more detail in MPA-LR-7. The following table summarizes the results
10 of my analyses of these routes:[42]
11

---

[40] I focused on park & loop routes with consistent delivery-type characteristics for several reasons. First, because of the limited time provided for initial review of videotapes at Merrifield, the lateness in receiving tapes for further analysis, and the substantial time required for proper analysis of the tapes, I simply did not have time to analyze all of the tapes. I had to focus on those that were easiest to analyze. Second, it quickly became apparent that many of the tapes (especially for non-park & loop routes) were not particularly useful for my analysis, because they either show choppy, incomplete segments of activities or involve complex routes with a mix of different delivery types that are difficult to analyze. Third, the videos of park & loop routes tended to have longer segments of multiple load-run time cycles that were easier to interpret (e.g., where the breaks between load and run time activities could be more easily identified). These aspects are discussed more fully in MPA-LR-7.

[41] A twelfth route was fully analyzed but is not included in my summary of results because of the limited time span of the videotaped segments. Although the taped portions showed a low percentage of load time (about 20% of loop time), the tape spanned only about a one hour period in the morning.

[42] In all cases, the tally data and the videotape data are from a day included within LR-I-163. (Routes 1913, 8212, 611, and 6410 had tape for two days; routes 415 and 820 had tape for three days.) The ES ratio has been adjusted to match the same time period as observed in the videotapes.

| CY and Route | Number of Stops Observed in Videotape | Amount of Videotape Time Observed (Minutes) | Average Stop Time (Seconds) | Comparable ES Proportion of Load vs. Load plus FAT Run Time | MPA Proportion of Load vs. Load plus FAT Run Time |
|---|---|---|---|---|---|
| CY04 4233 | 68 | 27.1 | 3.6 | 35.5% | 15.2% |
| CY05 2822 | 49 | 23.4 | 3.2 | 30.0 | 11.3 |
| CY27 1428 | 23 | 16.2 | 10.6 | 46.2 | 25.1 |
| CY37 0321 | 28 | 14.8 | 5.4 | 40.7 | 17.1 |
| CY51 6410 | 128 | 53.5 | 7.3 | 57.5 | 29.2 |
| CY55 0611 | 147 | 104.1 | 4.2 | 32.3 | 9.8 |
| CY58 8212 | 50 | 25.1 | 4.6 | 56.0 | 15.2 |
| CY60 1913 | 51 | 23.5 | 3.6 | 48.7 | 13.1 |
| CY62 0415 | 172 | 86.1 | 3.6 | 33.3 | 11.9 |
| CY63 0806 | 27 | 11.9 | 6.5 | 45.0 | 24.7 |
| CY63 0820 | 90 | 45.5 | 6.6 | 30.5 | 21.7 |

The videotape results for these routes show a significantly lower proportion of load-time-to-loop-time than indicated by the ES work sample tallies. This strongly suggests that on walking sections of park & loop routes, the ES tallies overstate true load time by a significant amount. This overstatement is most likely due to ES data collectors recording non-load activities with codes that Raymond ultimately allocated to load time.

While my analysis was necessarily limited due to time constraints, it does provide clear evidence of overstatement of true load time proportions. Although it is not possible to extrapolate these results to quantitatively estimate the amount of load time overstatement in the ES database, this analysis does demonstrate why the ES results cannot be deemed to be a reliable indication of true load time costs in the USPS city delivery carrier system.

### B. The Postal Service and MPA Regression Analyses of ES Load Time v. Possible Deliveries Indicate A Substantial Overstatement Of True Load Time

To investigate the relationship of the ES load time data to the expected structure of load time activities (using the standard ratemaking-costing definition) and

1 to the results of the LTV models, MPA developed a route-level regression model of ES
2 load tallies with possible deliveries and other variables.[43] Separately, the Postal
3 Service also developed a route level regression model of ES load time with volumes,
4 possible deliveries, and other variables (USPS-LR-I-310). Not surprisingly, the
5 results of both analyses are exceptionally consistent and demonstrate that the ES
6 data overstate true load time.

7 The MPA ES regression model explains route-level ES load tallies as a function
8 of (a) number of possible deliveries by delivery type and (b) dummy variables
9 identifying delivery mode of the route. Based on a qualitative analysis of the ES data, I
10 hypothesized that if there were any overestimates of true load time in the ES data, they
11 would likely vary by the delivery mode employed on the route. (There are five delivery
12 modes: curb, park & loop, dismount, central, and other.) The delivery mode dummy
13 variables were included to test this hypothesis.

14 Separately, the USPS ES regression model explains ES load time by route-day
15 as a function of (a) volume by shape, (b) proportion of possible deliveries by delivery
16 type, (c) number of total possible deliveries, (d) volume/possible deliveries interaction
17 terms, and (d) shape related dummy variables related to high per-piece ES load time
18 observations.

19 In both the MPA and USPS ES models, the intercept and intercept-related terms
20 are large and statistically significant. As explained below, these values demonstrate
21 that the ES load tallies include activities which are not true load time. In the MPA ES
22 model, the intercept applies to the park & loop mode of delivery. The related terms
23 are the dismount and "other" delivery mode dummy variables which add to or subtract
24 from the load time effect included in the park & loop intercept value. In the USPS ES

---

43 The possible deliveries variable was used as a proxy for volume and actual deliveries. Apparently, actual delivery data are not available in summary form and volume data were provided only recently. Appendix B provides a more detailed examination and explanation of both the MPA and USPS ES models.

model, the proportion of possible deliveries by delivery type and the high-load shape-related dummy variables are the intercept-related terms. Coefficient values for these terms are also large and statistically significant.

The key point is that the intercept and intercept-related terms in both models represent relatively "fixed" time in the ES load time data (i.e., time that does not vary directly with number of possible deliveries). If the activities encompassed by the ES load time data only included true load time, then the intercept value and the coefficients for the other related terms would be close to zero and statistically insignificant. This result is expected for true load time, since zero possible deliveries should produce zero load time. Thus, this fixed time identified in the regressions should be considered non-load time that belongs in another out-of-office time component. This non-load time is the result of data collectors recording non-load carrier activities with codes which Mr. Raymond allocated to load time.

Both the MPA and USPS ES models demonstrate why the ES estimate of accrued load time is 92.3% greater than the LTV estimate of load time, but the USPS ES model variability generates a variable cost that is only 35.8% greater than that for LTV. The lower LTV estimate of accrued load time is associated with 64.0% variability while the substantially higher ES estimate is associated with a 45.2% variability. This is precisely what is expected when the load time estimate being used contains a high proportion of fixed time that cannot be true load time. If the ES time proportions are used to disaggregate out-of-office costs, then there is no question that the variability estimate from the USPS ES model produces a more reasonable estimate of variable load time.

This latter point -- the necessity to match accrued costs and variability models in order to avoid severe errors and distortions in estimation of volume variable costs -- is extremely important. The STS estimates of accrued load time, themselves, are 47.4% greater than the Load Time Variability (LTV) model estimates. ES-based

1 estimates, however, show an even much greater disparity -- 92.3% greater than the
2 LTV modeled estimates. The disparity between modeled and STS accrued estimates
3 was troubling enough, but now it is enormously greater with the ES accrued
4 estimates. This indicates at least one of two problems: (a) inaccuracy or unreliability
5 in either the LTV or the STS/ES results, or (b) a severe mismatch between the
6 functions represented by the LTV models and those reflected in the STS/ ES load time
7 proportions.

8 In either case, this presents the Commission with a serious ratemaking
9 problem. When variability is derived from a functional model that only explains
10 approximately half of the estimated accrued time for that function, there is an
11 extremely high probability that the estimated accrued time includes much more than
12 what was modeled. In that event, applying the modeled variability to the estimated
13 accrued time would produce a result that not only cannot be explained but has no
14 meaning. Moreover, doing so would clearly produce a greatly overstated estimate of
15 variable functional costs.

## VI. THE POSTAL SERVICE'S INCORRECT ATTRIBUTION OF THE POSSIBLE DELIVERIES EFFECT ON LOAD TIME

16 In USPS-T-12, Mr. Baron continues to assert that the effect of possible
17 deliveries on MDR and B&M stop load time, as measured in the LTV stop-level load
18 models, should be attributed to the subclasses.[44] He believes that, if he replaces the
19 number of possible deliveries in the LTV stop-level models with the number of actual
20 deliveries, then he can measure the extent to which average stop load time varies with
21 actual deliveries, which in turn vary with volume. He attributes this "volume-variable"
22 actual deliveries load time among the subclasses. However, that attribution is

---

[44] Multiple Delivery Residential (MDR) and Business and Mixed (B&M) stops typically have more than one delivery per stop. Thus, the LTV MDR and B&M models reflect the fact that volume may be distributed among multiple deliveries at the average stop.

1 incorrect. All volume-related stop load effects are already captured by the volume
2 coefficients in the models. Thus, his approach amounts to attributing some of the
3 stop load time twice and is clearly excessive and inappropriate.

4 What Mr. Baron ignores is that volume affects average stop load time in two
5 ways. Greater stop volume increases average stop load time directly through an
6 increase in volume per delivery, when actual deliveries stay constant. And, when
7 there are scale economies at the delivery level, greater stop volume also indirectly
8 increases stop load time. The indirect effect may include the increase in actual
9 deliveries per stop (e.g., increases in delivery-related fixed costs) and the increase in
10 average unit time (from a reduction in pieces per delivery). Both the direct and indirect
11 effects from a volume change are necessarily captured by the statistically significant
12 volume coefficients in the LTV models. The possible deliveries variable in the model
13 does not reflect any volume effects, instead it is a control variable so that the LTV
14 volume coefficients properly measure the entire volume effect on stop-level load
15 time.[45]

16 Comparable to the MDR and B&M stop-level load time models, the USPS ES
17 model discussed in LR I-310 is a route-level model with multiple possible deliveries.
18 And that discussion makes a suggestion similar to the one made by Mr. Baron in
19 USPS-T-12: the possible deliveries effect in the route-level USPS ES model should
20 be attributed among the subclasses. This, of course, is also incorrect and for the
21 same reason. The possible deliveries variable in the route-level load time model

---

[45] The number of possible deliveries affects stop-level load time by affecting the number of actual deliveries, independently of volume. Leaving volume constant, an increase in possible deliveries increases the number of actual deliveries. This is because the volume-coverage function will distribute the constant level of volume among more actual deliveries when there are more possible deliveries. Further, the presence of possible deliveries variables in the models is appropriate. If the possible deliveries variables are excluded from the models, it would artificially inflate the coefficient values for the volume variables and lead to an overstatement of load time variability.

1 serves the same control purpose as in the stop-level load time models. Just as in the
2 stop-level models, the volume coefficients reflect the entire volume effect on route-
3 level load time model. When volume on a route increases and there is less than
4 100% delivery coverage on the stop, then some of the volume goes to newly covered
5 stops/deliveries (causing whatever fixed stop/delivery time is appropriate) and
6 average volume for all covered stops/deliveries on the route decreases. [46] This entire
7 effect is measured in the volume coefficients. Just as in the stop-level B&M and MDR
8 models, possible deliveries in the ES route-level model is simply a control variable.

---

46 Apparently the USPS is confused about how coverage changes average stop/delivery and average per piece load time. (See, e.g. the discussion in LR-I-310, page 20.) When there is less than 100% coverage, a volume increase causes an increase in coverage which reduces average volume per stop on the route. If there are stop/delivery-level load time scale economies (i.e., elemental load time variability is less than 100%), then average per piece load time actually increases (coverage-related load time is positive). On the other hand, if there are no such scale economies (i.e., elemental load time variability is 100% and there is no fixed stop/delivery time), then average load time per piece does not change and changes in coverage have no effect on per piece load time (i.e., coverage-related load time is zero).

## AUTOBIOGRAPHICAL SKETCH

My name is Antoinette Crowder and I am a senior consultant with TRANSCOMM, Inc., an engineering and economic consulting firm located in Falls Church, Virginia. I have been associated with TRANSCOMM for twenty-seven years and, during that time, have been involved in a variety of projects dealing with costing, pricing, market and demand studies, economic and financial analyses, survey design, and research on numerous regulatory and policy issues. These activities have concerned the electric power, gas, communications, and postal/publishing industries. I have prepared or assisted in preparing numerous filings at various federal and state regulatory agencies on behalf of numerous clients. In addition, I am involved in the firm's overseas consulting activities, providing financial, economic and regulatory assistance to multi-national organizations, international firms, and national governments.

I have been involved in analyses of postal ratemaking and policy issues since the beginning of the R77-1 rate case. My work has involved revenue requirement, cost attribution and distribution, subclass rate structure and discounts, institutional cost allocations, service-quality measurement, demand and market assessment, and mail classification issues. I am part of the TRANSCOMM team that provides economic/financial advice on postal matters and monitors costs, financial statements, volumes, service levels, and other aspects of Postal Service operations on behalf of several clients.

I have testified before the Postal Rate Commission in seven proceedings and have contributed to development of other testimony presented to the Commission. In Docket R84-1, I contributed to mail processing peak-load and second-class intra-SCF discount testimony. In Docket R87-1, I contributed to carrier-out-of-office and third-class/fourth-class Bound Printed Matter drop-ship discount testimony, and I also prepared and presented rebuttal testimony on third-class presort discounts. In

Dockets C89-3/MC89-1, I helped prepare and presented direct testimony on the proposed local saturation subclass. In Docket R90-1, I assisted in preparation of carrier-out-of-office cost and institutional cost coverage testimony and prepared and presented rebuttal testimony on third-class rates. In the R90-1 Remand, on behalf of a third-class mailer's group, I presented testimony concerning the attribution of city carrier coverage-related costs. I presented two pieces of rebuttal testimony in Docket R94-1 and a rebuttal testimony in MC95-1. In Docket R97-1, I presented testimony in response to Presiding Officer's Notice of Inquiry No. 3 on city delivery carrier load time costs, and rebuttal testimony on carrier costs and rate design issues.

Over the course of my 23-year involvement in postal ratemaking matters, I have had numerous opportunities to observe postal operations and have analyzed the cost aspects of those operations. I have also become familiar with economic costing and pricing concepts, both generally and as applied to postal ratemaking.

My education includes a B.S. in Biology from the University of Virginia, an M.S. in Biology from George Mason University, and additional course work in economics, mathematics and statistics.

# COMPARISON OF MPA REGRESSION OF ES LOAD TIME TALLIES TO POSSIBLE DELIVERIES TO USPS REGRESSION OF ES LOAD TIME TO VOLUME AND POSSIBLE DELIVERIES

This appendix describes the MPA and USPS regressions of ES load time, and then compares and assesses results from the two models. It also demonstrates that both models exhibit the same ES load time attributes: the ES load time data, when investigated formally through regression technique, do not generate the expected route-level structure for load time. Instead, they both generate results that demonstrate that the ES data overstate true load time, as defined for the Load Time Variability (LTV) model, because they include observations of carrier activities that are not true load time.

## A. MPA Model of ES Load Tallies vs. Possible Deliveries

### 1. MPA Model Description

An MPA route-level load model was developed to investigate the relationship of the ES load time data to: (a) the expected structure of load time activities (using the ratemaking-costing definition) and (b) the results of the LTV models. The model assumes route load time is explained by route possible deliveries by type and delivery modes. Coefficient estimates for the two sets of explanatory variables were developed by regressing route load tallies against the number of possible deliveries by delivery type by route, and a series of delivery mode related dummy variables.[1] The delivery mode variables were included in the analysis

---

1 Conceptually, true route-level load time is affected by two workload load variables - route volumes and number of actual deliveries. Ideally, regressing the ES load tallies against these two variables would allow the true load time contributions from each variable to be separately identified and assessed. However since these data were lacking, route-level possible deliveries by delivery type, provided by the USPS for ES sampled routes, were used in the regression model as proxy workload variables to explain the ES load tallies. The ES project apparently collected

(footnote continued)

based on a separate qualitative assessment of the ES data. It was clear from an examination of such data that any over-estimates of load time might be partially explained by the delivery mode employed on the route.

The model developed coefficient estimates for the following eight possible delivery types and five delivery modes:

| Delivery Types | Delivery Modes |
|---|---|
| Residential and Business Other | Park& Loop |
| Residential and Business Curb | Dismount |
| Residential and Business NDCBU | Curb |
| Residential and Business Central | Foot |
| | Other |

The residential and business other delivery type includes park & loop and foot deliveries.

Data from the 336 ES routes used to develop ES out-of-office time proportions were used to construct the model. The ES load tallies from the multiple-day routes represent averages of daily observations and each tally was assumed to represent six minutes of out-of-office time. The number of possible deliveries by delivery type and the delivery mode employed by route were provided by the USPS.[2]

The load time effect from the possible delivery and delivery mode related variables was tested through the following regression form:

---

(footnote continued from prior page)

information on actual deliveries but did not maintain summary records on them. (ADVO/USPS-T13-18) It also collected volume information and made a hard copy of such data available a few weeks ago and an electronic version of the volume data was made available within the past several weeks. Route volumes and actual deliveries can be expected to vary directly with route possible deliveries. Thus, possible deliveries should indirectly affect route-level load tallies through its relationship with volume and actual deliveries.

2 USPS LR I-219.

$$LT_i = \alpha + c_1{*}RCURB_i + c_2{*}RNDCBU_i + c_3{*}RCENT_i + c_4{*}ROTHER_i + c_5{*}BCURB_i + C_6{*}BNDCBU_i + C_7{*}BCENT_i + C_8{*}BOTHER_i + \Sigma\, d_j{*}DMODE_{ji} + \varepsilon_i,$$

where: $LT_i$ = route (i) load tallies,

$RCURB_i$ = route (i) residential curb possible deliveries,

$RNDCBU_i$ = route (i) residential NDCBU possible deliveries,

$RCENT_i$ = route (i) residential central possible deliveries,

$ROTHER_i$ = route (i) residential other possible deliveries,

$BCURB_i$ = route (i) business curb possible deliveries,

$BNDCBU_i$ = route (i) business NDCBU possible deliveries,

$BCENT_i$ = route (i) business central possible deliveries,

$BOTHER_i$ = route (i) business other possible deliveries,

$\Sigma\, d_j{*}DMODE_{ji}$ = sum of load effects from delivery modes j = 2 through 5.

The DMODE variables are binary (dummy) variables, sequenced according to their order in the last table, beginning with the dismount mode (j=2). They are assigned a value of one (zero) if the route is (is not) of the specified mode. They represent the delivery modes in the sequence indicated in the last table starting with the dismount mode (j=2). The (c) coefficients represent the load tallies per possible delivery for the different delivery types. The intercept coefficient ( ) represents the quantity of ES load tallies that are relatively "fixed" on a per route basis on park&loop delivery mode routes (i.e., are not true load time). For the other delivery modes, the relatively "fixed" ES load is represented by the (d) coefficient values acting on the binary variables.

This MPA ES model tests the ES load time estimates in two ways. First, inclusion of the intercept and the modifying dummy variables determines whether the ES load tallies systematically include some relatively "fixed" time per route. If the ES load tallies have been classified correctly, then the regression intercept value and all

(d) values should be close to zero and statistically insignificant. This just validates the common sense expectation that zero possible deliveries should produce zero load time tallies. Second, the per delivery load times from the ES model (converted from the per delivery tally estimates) can be compared to per delivery estimates from the LTV models after the LTV estimates are adjusted to represent average possible delivery time. If the coefficient values for the possible delivery variables are significantly different from the LTV figures, then the comparison may suggest that either (a) the ES load tallies mis-state the true amount of load time at delivery points or (b) the LTV models are no longer representative of the current structure of load activities.

## 2. MPA Model Results

Coefficient values and the accompanying t statistics for the model are indicated below:

$$
\begin{aligned}
Lt_i = \; & \underset{(9.0575)}{10.4640} + \underset{(7.6483)}{.0233}*RCURB_i + \underset{(5.8359)}{.0168}*RNDCBU_i + \underset{(8.1055)}{.0192}*RCENT_i + \\
& \underset{(2.5421)}{.0074}*ROTHER_i + \underset{(.8080)}{.0289}*BCURB_i + \underset{(.7998)}{.0217}*BNDCBU_i + \\
& \underset{(1.9406)}{.0393}*BCENT_i - \underset{(-1.8094)}{.0111}*BOTHER_i + \underset{(3.0600)}{3.0588}*DMODE_{2i} + \\
& \underset{(.5809)}{.6920}*DMODE_{3i} + \underset{(.4327)}{.4830}*DMODE_{4i} - \underset{(-2.2941)}{9.4355}*DMODE_{5i}
\end{aligned}
$$

$R^2 = .4057.$

All residential possible delivery and the business central delivery coefficients are significant at the 5% level.

**Intercept and Dummy Variables**

The intercept and the coefficient for the $DMODE_2$ variable for dismount routes are also significant at the 5% level. The positive intercept value suggests that some portion of the ES load tallies is not true load time but, instead, represent some relatively "fixed" ES load time on park & loop routes. The positive coefficient for $DMODE_2$ indicates even more relatively "fixed" ES load time on dismount routes - about 33% more than compared to park&loop routes.[3] The coefficients for the curb and foot delivery mode dummy variables are statistically insignificant, suggesting about the same level of relatively "fixed" ES load time for those types of routes as on park & loop mode routes.[4]

**Possible Delivery Coefficients**

In order to compare the load time results from this MPA ES model to those of the LTV models, the coefficients for each of the possible delivery types were converted to seconds per delivery estimates (assuming each ES tally is six minutes). The estimates for residential curb and other (foot/loop) deliveries were then averaged to develop a single delivery residential (SDR) delivery estimate for comparison with the LTV SDR model estimate (using current CCS volume data). Similarly, the residential

---

3 This could occur, for instance, if the some ES observers assumed that the carrier was at the point of delivery (and in the process of delivery/collect) after he had parked his vehicle for a dismount but had not yet reached the receptacle point. It is very easy to imagine an observer making this interpretation and, as discussed in the text, there are tallies which suggest that this interpretation did occur.

4 The fixed time in this model may not be strictly fixed in the same way as travel time to and from a route. One possible explanation is that ES observers systematically included some run or support time in their load time tallies and there are small variations in the number of actual stops (i.e., accesses) within one delivery mode type. Or, it may due to the inconsistency in the amount of run or support time the ES data collectors included in the observations that were ultimately allocated to load time.

NDCBU/central delivery and business delivery time estimates were averaged to develop multiple delivery residential (MDR) and business and mixed (B&M) delivery estimates for comparison with LTV MDR and B&M model estimates, respectively. The results are indicated below:

| **Delivery Type**[5] | **Seconds/Delivery** |
|---|---|
| RCURB | 8.37 |
| ROTHER | 2.66 |
| **Weighted Average SDR** | 5.33 |
| | |
| RNDCBU | 6.04 |
| RCENT | 6.91 |
| **Weighted Average MDR** | 6.50 |
| | |
| **BCENT/Weighted Average B&M** | 14.13 |

The seconds per delivery for the two multiple delivery categories are relatively close as might be expected. There should be little difference in carrier loading activities between these two delivery types. However, residential curb deliveries generally would be expected to have a greater load time than residential "other" deliveries (e.g., loop deliveries) because, unlike loading "other" deliveries, the carrier must stop and start his vehicle at each delivery point before preparing mail for subsequent delivery.

Separately, LTV based estimates for seconds per actual stop were converted to seconds per possible delivery for comparison with the MPA model derived estimates. Witness Baron has provided load time per actual stop figures for the three stop types using the LTV model in combination with CCS volumes per actual stop.[6] The figures derived by Witness Baron under the mean volume and mean load time approaches

---

5 The SDR and MDR equivalent values represent possible delivery weighted averages of the seconds per delivery for the corresponding delivery types. Weights were calculated from the numbers of deliveries by type from the 336 ES routes.

6 USPS-T-12, page 21.

were converted to estimates per possible delivery by multiplying the actual stop figures by the actual stop coverages and then dividing these values by the average number of possible deliveries per possible stop.[7] A build-up and comparison of both sets of data are provided below:

| | LTV Seconds Per Actual Stop | LTV Seconds Per Possible Stop | LTV Seconds Per Possible Delivery | MPA ES Model Seconds Per Possible Delivery |
|---|---|---|---|---|
| SDR | | | | |
| Avg. Volume | 8.63 | 8.02 | 8.02 | 5.33 |
| Avg. Load Time | 8.41 | 7.82 | 7.82 | 5.33 |
| MDR | | | | |
| Avg. Volume | 61.71 | 60.09 | 7.74 | 6.50 |
| Avg. Load Time | 73.44 | 71.50 | 9.21 | 6.50 |
| B&M | | | | |
| Avg. Volume | 19.08 | 17.16 | 13.32 | 14.13 |
| Avg. Load Time | 8.35 | 7.51 | 5.83 | 14.13 |

The SDR and MDR load seconds per delivery are somewhat lower for the MPA ES model. However, in general, the MPA ES model appears to depict load times per possible delivery in a manner comparable to the LTV model. Thus, the significantly higher load time percentages calculated from the ES data compared to the STS percentages, used in the previous rate cases, appear to relate to the relatively "fixed" time identified in the model.

7 Coverage data are from Meehan base year workpapers (WS 7.0.4.1) and possible delivery per possible stop estimates are CCS data averages by stop type from LR-157.

## B. USPS Regression of ES Load Time with Volumes and Possible Stops

### 1. Summary of Approach

The USPS recently submitted a draft report containing route-level regression analyses of the ES/USPS accrued load time as a function of shape volumes and possible deliveries.[8] The variables were the load tally data contained in USPS LR I-163 for 750 route days plus volume and possible deliveries information for those same route days but derived from other sources. Before conducting regressions, the route tally data were converted to load time estimates by separately identifying total out of office time by ES route day and then applying load tally ratios to those times.

The USPS conducted several regressions. The initial regression specification included several variables designed to test the separate effects of route volumes and possible deliveries on route load times. As indicated in page 7 of the report, these included: route-level shape volumes, the square of these volumes, the product of the paired combinations of the shape volumes (volume interaction terms), route-level possible deliveries, the square of possible deliveries, the product of possible deliveries and each of the shape volumes, and the fraction of total possible deliveries for each delivery type. Although results for this specification were not presented, the report indicated that the regression resulted in low (t) statistics for most variables and low coefficient values for the volume variables.

The problem was allegedly corrected by adding shape-related dummy variables related to high per piece ES load time observations. The report stated that "each such dummy variable was set equal to one for all observations for which the ES load time per piece (by shape, and for accountables) fell within the upper 10% of the

---

8 USPS LR I-310, "Draft Report on Load Time Variability Provided in Response to ADVO/USPS-T12-11."

distribution of all observations of ES load time per piece." Statistics for this corrected regression, as presented in Table 3 of the report, show statistically significant (t) values for many of its volume and possible delivery variables. Table 4 presents elasticity estimates for the volume variables that the USPS considered within the range of acceptability. Although the USPS ran other regressions using different ceiling points for the high per piece load times, the USPS preferred regression is the one containing the 10% shape-related dummy variable specification.

### 2. Assessment of Regression Results and Evidence of Load Time Mis-Classification

Close examination of the Table 3 data for the preferred regression reveals a general consistency with the results achieved from the MPA ES load regression presented in the previous subsection and supports my conclusions derived from the latter. In the USPS model, the intercept variable and the variables that shift the intercept (the shape-related dummy variables and the fractions of total possible deliveries represented by each delivery type) collectively indicate the presence of relatively "fixed" load time per route and therefore load time overstatement. This is true for both normal and high load time observations. In both instances, the estimated load time is positive when all shape volumes are zero because the sum of the intercept and shift variables acting on their respective coefficients is positive. This result is clearly nonsensical at the route level because without volumes to deliver there are no delivery points to be accessed, and therefore there is no load time to be incurred anywhere along the route.

Consider what happens when route volume decreases. A loss of route volume decreases load time because (a) previously covered stops are now uncovered and (b) volume on the remaining covered stops is reduced. For the newly uncovered stops, the previously observed fixed and variable costs are gone since volume on those stops has been entirely eliminated. For the remaining stops, variable costs are

reduced but fixed costs are still incurred. Obviously, if a route's entire volume is eliminated, there is no fixed or variable load time since zero volume means zero covered/actual stops. Thus, it is clear that with per stop scale economies, the route load time-route volume relationship would be described by a positively sloped curve, passing through the origin and exhibiting declining marginal load times.[9]

Unfortunately, the non-volume related coefficient values from the USPS regression do not produce the required zero load time - zero volume result. And these results are not limited to the high per piece load time route days. The USPS's own regression results shows that the incidence of load time overstatement is pervasive to most if not all route days used in their analysis. The relevant coefficient values are reproduced below from Table 3 of LR I-310 for easy reference.

| Variable | Coefficient Value |
|---|---|
| Intercept | -5.597.31 |
| % of PD - Residential Other | 6.132.81 |
| % of PD - Residential Curb | 9,152.67 |
| % of PD - Residential Central | 7,979.13 |
| % of PD - Residential NDCBU | 7,430.93 |
| % of PD - Business Other | 4,828.31 |
| % of PD - Business Curb | 1,692.19 |
| % of PD - Business Curb | 10,486.00 |
| Load Time/Letters Dummy | 2,777.49 |
| Load Time/Flats Dummy | 2,161.75 |
| Load Time/Accountables Dummy | 2,292.01 |
| Load Time/Parcels Dummy | 1,164.32 |

9 Note that there is a clear difference between this result and the volume- load time relationship observed at the stop level. At the stop level, the cost-volume curve does have a positive intercept, indicating fixed stop time. Assuming a covered stop remains covered after a drop in route volume, the related drop in stop volume would plot as a less than proportionate decrease in stop load time. Extending the plot to zero volume would indicate a positive intercept value, revealing the fixed stop load time. However, the stop level plot does not capture the fixed load time changes occurring at the same time on other stops that are also uncovered. The changes in all stop fixed load time associated with covering or uncovering stops should be correctly captured in a route level view.

For any normal (non-high load time) route day, zero volume produces a load time explained by the negative intercept value plus a weighted average of the positive percent of P.D. coefficient values for the different delivery types. The weights applied are the fractions of total deliveries by delivery type for that route. Note that all the positive P.D.-related coefficient values for residential delivery types are higher than the intercept in absolute value. Since these delivery types comprise the vast majority of all possible deliveries along routes, zero volumes for any normal route will produce positive load times at zero volume for most if not all of these routes. Of course, for high load time routes the overstatement is accentuated. One, several or all of the positive coefficient values for the shape related dummy variables must be added to the intercept and P.D.-related values to determine the extent of mis-classification for these routes.

### 3. The Effect of Possible Deliveries

In the report presented in LR I-310, the USPS continued to insist on a possible deliveries effect to be included as part of the load-related volume variable cost estimate. Prior to the ES load time analysis in LR I-310, the deliveries effect had been described as a stop level effect pertaining to MDR and B&M stops described by the LTV models. Now the effect has been elevated to the route level and, by implication, now applies to all stop types including SDR.

Regardless of which models are used, possible delivery elasticities calculated by the USPS should not be included as part of the volume variability calculation. These values are meaningless for use in calculating volume variable costs. Possible deliveries affect system level load time through effects on actual deliveries, independently of volume. Thus, possible deliveries must be recognized in any load time model as a control variable to ensure that the volume-related coefficients that subsequently become part of the volume variability calculation are not biased.

## CERTIFICATE OF SERVICE

I hereby certify that I have on this date served the foregoing document upon all participants of record in this proceeding in accordance with section 12 of the Rules of Practice.

Thomas W. McLaughlin

June 1, 2000