Postal Rate Commission
Submitted 9/6/2006 11:09 am
Filing ID: 53126
Accepted 9/6/2006

OCA-T-2
Docket No. R2006-1

DIRECT TESTIMONY

OF

J. EDWARD SMITH

ON BEHALF OF

THE OFFICE OF THE CONSUMER ADVOCATE

September 6, 2006

TABLE OF CONTENTS Page

I. STATEMENT OF QUALIFICATIONS 1

II. PURPOSE OF THE TESTIMONY 2

III. DISCUSSION OF DATABASE AND TESTIMONY OF WITNESS NIETO 4

A. There is No Evidence that Sampling Theory was Adequately Employed in Determining the Sampling Plan 5

B. The 7915 Transactions Observed have not been Shown to be Representative of the General Population of Transactions. 6

C. The Development of the Database Appears to Have Lacked Adequate Quality Control. 9

D. Information from Library Reference USPS-LR-L-159 May Indicate Data Collection Problems. 11

E. There Was No Analysis of Whether Enough Data Was Gathered for Each Type of Transaction in Order to Have a Statistically Adequate Sample. 14

IV. DISCUSSION OF WITNESS BRADLEY'S TESTIMONY 17

A. Witness Bradley's Recommended Model Reproduced in SAS. 18

B. Recommended Model Modified to Take into Account Walk-time 18

C. Alternative Variabilities Resulting from Adjusting for Various Outliers 20

1. Studentized Residual at 2 Variability 20
2. Studentized Residual at 2 with walk 21
3. Studentized Residual at 3 22
4. Studentized Residual at 2, Basket ID case 22

D. Choice of Model—Linear or Other 25

V. CONCLUSIONS 27

I. STATEMENT OF QUALIFICATIONS

My name is J. Edward Smith, and I am an econometrician with the Office of the Consumer Advocate of the Postal Rate Commission. I have previously worked in a variety of economic assignments in industrial, academic, consulting, and governmental positions. My economics degrees are an A.B. from Hamilton College, and an M.S. and Ph.D. from Purdue University. I have testified before the Postal Rate Commission in Docket No. R97-1, Docket No. R2000-1, and Docket MC2002-2.

II. PURPOSE OF THE TESTIMONY

Postal Service witness Bradley has presented an updated transaction supply side volume variability study for window service costs.[1] The study is based on the data collected from a transaction time study sponsored by Postal Service witness Nieto[2] and significantly revises the previously developed window service supply side variabilities presented by witness Brehm in Docket No. R97-1.

My testimony analyzes the work presented by both witnesses Nieto (USPS-T-24) and Bradley (USPS-T-17). The database developed by witness Nieto has not been shown to be adequate to serve as the basis for witness Bradley's study. Although the database appears to be deficient, given the deficiencies of the earlier study convincingly enumerated by witness Bradley, the recently compiled data may have to serve as the basis for the estimation of volume variabilities pending the development of a new database. Using the available data, with some deletions of questionable observations, I have developed alternative estimates for the variabilities presented by witness Bradley. In addition, future work in the estimation of volume variabilities should consider applying a quadratic, rather than linear, model in the estimation efforts.

I therefore recommend that the variabilities in column 7 of Table 5 of this testimony (page 17) be implemented for the purposes of this proceeding in lieu of those proposed by witness Bradley. Given the need for an improved database and additional econometric analysis, I respectfully urge the Commission to recommend the

---

[1] Direct Testimony of Michael D. Bradley, Docket No. R2006-1, USPS-T-17.

[2] Direct Testimony of Norma B. Nieto, Docket No. R2006-1, USPS-T-24.

development of an improved database and the further consideration of theoretical issues associated with the estimation process.

III. DISCUSSION OF DATABASE AND TESTIMONY OF WITNESS NIETO

Witness Nieto's window service transaction time study provides data measuring the time associated with the handling of various types of customer transactions at post office windows. Data were collected over a two week time period in April and May of 2005 at 27 randomly selected post offices in 9 of the 11 Postal Service Areas. The post offices were selected from the approximately 15,000 POS-One system capable post offices. Witness Nieto outlines the procedures for data collector training, the gathering of the data, and the subsequent processing of the data into a database comprising 7915 observations of transactions for Postal Service products. Upon review of the testimony, library references documenting the study, and cross-examination responses, I have concluded that the database appears to be deficient in several respects:

There is no evidence that sampling theory was adequately employed in determining the sampling plan.

The 7915 transactions observed have not been shown to be representative of the general population of transactions occurring on a daily basis.

There are significant quality control issues.

There may be unresolved data collection problems in the gathering of the data.

There is no analysis of whether enough data was gathered for each type of transaction in order to have a statistically adequate sample.

A. There is No Evidence that Sampling Theory was Adequately Employed in Determining the Sampling Plan.

Witness Nieto has not provided an adequate justification for the selection of sample size in terms of the total number of transactions, in terms of the number of sites sampled, in terms of the breakout between small and large sites, or in terms of the number of transactions by postal product. For example, Witness Nieto states:

> The proportion of 2 large sites to 1 small site was chosen to balance the considerations of maximizing the number of transactions observed with including small offices. Including more large offices than small is likely to increase the number of transactions observed, but small offices were also included to account for the possibility that they might have differences in transaction times despite having fewer transactions per day.[3]

The above comments relating to sample size are not sufficient support for an adequate statistical analysis. In reviewing a data collection effort based on the sampling of transactions at sites, one would expect to find an analysis of the population of sites, types of transactions and data to be collected, justification for the selection of sites, computation of sample sizes in terms of their statistical properties, and quality control procedures. The information provided in the testimony and the associated library reference USPS-LR-L-78 is inadequate for justifying the sampling plan.

Witness Nieto has essentially used a stratified sample, but has not provided adequate analysis on the statistical issues associated with the development of a

[3] Tr. 5/700 (OCA/USPS-T24-1(h)).

representative, stratified sample of data. Witness Nieto indicated:

> The goal of the study was not [to] (sic) produce an estimate of total annual transaction by type but rather to produce a dataset that permitted an update of the established transaction time econometric model.[4]

However, a database that is not based on statistical sampling theory and cannot be verified to be representative of the set of transactions studied does not provide a foundation for the development of a transaction-time econometric model. In fact, a model developed on the basis of incomplete and irrelevant data could give incorrect conclusions.

B. The 7915 Transactions Observed have not been Shown to be Representative of the General Population of Transactions.

Table 1 summarizes the POS-ONE data collection effort by size of site, numbers of transactions, revenue per site, numbers of sites sampled, and transactions:

The database consists of 7915 transactions, which have not been shown to be representative of the total number of transactions. Assuming that revenue at a site is proportional to the number of transactions at the site, Table 1 below shows that based on revenue approximately 85 percent of the 7915 transactions should have been obtained from large sites, with 15 percent from small sites. In fact, approximately 77 percent of the transactions came from large sites, with approximately 23 percent from small sites. There is no explanation of this discrepancy.

Also, the database has not been shown to be representative of total transactions in terms of the number of POS-ONE sites sampled. Approximately 50 percent of

POS-ONE sites are small, but only 33 percent of the sites in the sample are small, with 67 percent being large. Again, no adequate basis for this anomaly is offered by witness Nieto.

Table 1: Site Summary[5]

	Small POS-ONE Sites	Large POS-ONE Sites	Total POS-ONE Sites
Number of Transactions	7542	7544	15086
Percent of Total	49.99	50.01	
Revenue per Site*	245,670	1,348,940	797,013
Total Sales	1,852,843,140	10,176,403,360	12,023,738,118
Percent of Total	15.41	84.64	
Number of Sites Sampled	9	18	27
Percent of Sampled Sites	33.33	66.67	
Total Observations	1841	6074	7915
Percent of Sample	23.26	76.74	

* The number 797,013 is total average revenue per site. It is not additive in the row.

Table 2 demonstrates how a database not representative of the population could impact variability. If inaccurate proportions of single and multiple transactions are collected, the impact on variability can be significant. For instance, in his testimony, witness Bradley finds a variability of 41 percent for Bulk Stamps. An “item” in the study is defined as a unique product or service processed in a transaction, with a quantity defined as the amount of each item processed in a transaction.[6] Witness Bradley has defined three types of transactions:

---

[4] Tr. 5/706 (OCA/USPS-T24-5).

[5] Data obtained from OCA/USPS-T24-1 (Tr. 5/698) and OCA/USPS-T24-9 (Tr. 5/711).

[6] USPS-T-17, lines 15-18 at 12.

SISQ: Single Item, Single Quantity.

SIMQ: Single Item, Multiple Quantity.

MI: Multiple Items. The transaction has more than one item, and each item may have one or more units being processed.

The volume variability for Bulk Stamps of 41 percent is based on the combination of 835 SISQ and 388 SIMQ transactions. Suppose, however, that there were data collection errors and, for purposes of simplification, the SISQ and SIMQ numbers changed to other numbers shown on Table 2. Even though the same number of transactions occurred, Table 2 demonstrates volume variability could change significantly. The results from the methodology are very sensitive to the data collection process. Furthermore, if, in an effort to be helpful and efficient, the clerk asks whether the customer would like to purchase stamps or conduct any other type of business as part of the transaction, then the nature of the transaction could change entirely. The volume variability could be affected by the clerk's actions.

Table 2—Product: Bulk Stamps

	Number of Transactions		Total Transactions	Variability
	SISQ	SIMQ		
Stamps Bulk	935	288	1223	44.3%
	835	388	1223	41.0%
	735	488	1223	37.7%
	635	588	1223	34.5%

Table 2 demonstrates that if the type of transaction varies—i.e., whether a single sheet of stamps was purchased or, alternatively two sheets of stamps were purchased—then the variability is affected. Clearly, having an accurate database representative of the general population of transactions is important: a change of a few

transactions can have a major impact on computed volume variability. Furthermore, if, in the example above, the customer had also purchased a single stamped envelope as part of the transaction, the transaction would have been an MI (multiple items) transaction and would not even have entered this part of the calculations, given the formula in use.

Given that the output of the model is very dependent on product mix, the elimination of over 1,500 observations during database collection is troublesome. These observations could easily have altered the computed variabilities. Moreover, as noted in Part A, above, if the sampling plan is inadequate, the database may also be deficient in terms of total number of observations collected. That deficiency, if it affected the ratio of SISQ to SIMQ transactions, could have resulted in additional errors in the calculated variability.

C. The Development of the Database Appears to Have Lacked Adequate Quality Control.

There is no substantiation that the collection of data over an April/May two week time period resulted in a database that was representative of postal transaction times or representative of the mix of postal products. Postal Service witness Hintenach, a manager with experience in Retail Operations and 40 years of service with the Postal Service, stated that, during heavy mail time periods such as holidays or tax-filing deadlines, the Postal Service manages its lobby operations differently than is the case at other times.[7] Additional information is needed to determine whether a two week data collection time period in April and May is adequate.

[7] Tr. 3/399-400.

Approximately 30 percent of the nested transactions were dropped from the database due to collection problems: 133 of 190 nested transactions were retained in the study.[8] Witness Nieto concluded that the 57 nested transactions were a small number.[9] She states:

> Therefore, dropping those 57 out of the 190 that were included does not create any unnecessary issues. It's just a transaction that happens to be separated into two parts.

In total, approximately 16 percent of all data collected was deemed unusable: Data were gathered for 9,459 transactions, but 1,535 were dropped from the database because they could not be matched.[10] Data were dropped from the database when the POS-ONE observations could not be matched with the records maintained by the data collectors. There is no analysis of the impact on the database from the elimination of the unusable observations; that is, whether the 1535 dropped observations biased the resulting sample and significantly affected computed volume variabilities. Witness Nieto concluded:

> The data just could not be matched to the product specific information, so I don't have specific concerns about the number of transactions that were dropped because we ended up with a number of transactions that were equal to or greater than the transactions in the previous study.[11]

---

[8] Nested transactions occur when a transaction is interrupted by another transaction and is then completed; for example, a customer initiates a Registered Mail transaction, steps away briefly from the window to complete some forms while the clerk continues with another transaction, and then returns to the window to complete the transaction.

[9] Tr. 5/720.

[10] Tr. 5/707 (OCA/USPS-T24-6).

[11] Tr. 5/722.

The production of a database with a number of transactions that are "equal to or greater than the transactions in the previous study" is not a statistical criterion. Witness Nieto indicated that she had not referred to the statistical literature to determine whether a significant amount of data had been dropped and was unable to cite any statistical literature covering the subject of the dropping of data.[12] In commenting on the dropping of observations, witness Bradley stated in a footnote to his response to POIR No. 7, Question 7:

> Dropping all observations with a studentized residual greater than 2.0 (in absolute value) would mean elimination of 250 data points. This is over 3% of the collected data.

Apparently witness Bradley was concerned about dropping 250 data points, in comparison to the 1535 dropped data points discussed by witness Nieto.[13] Thus, without adequate information supporting the April/May data collection timeframe and because of the large percentage of dropped nested and non-nested transactions, the quality of the database has not been substantiated and is questionable.

D. Information from Library Reference USPS-LR-L-159 May Indicate Data Collection Problems.

In Library Reference USPS-LR-L-159 witness Nieto provided walk-time data by facility. Walk-time denotes the total time period (in seconds) during which the customer approaches the window. Data collectors were instructed to gather walk-time data for each transaction in cases where the walk was long enough to be measured. Table 3 summarizes the total walk-time data (in seconds) on a facility-by-facility, daily basis.

---

[12] Tr. 5/723.

[13] Tr. 5/776 (POIR No. 7, Question 8).

Data were gathered over a period of several days at each facility during which both the amounts of total transaction times in seconds and amounts of walk-times in seconds varied at each facility. One would expect changes in total walk-time to be proportional to changes in total transaction time: the simple ratio of total recorded walk-time to total transaction time should be relatively invariant.[14] From Table 3, it appears that the collection of walk-time data was erratic from day to day at a number of sites, suggesting that either the data collection situation was very different from day to day at each facility or, alternatively, that collectors did not keep good records of walk-time, which would call into question the accuracy of other records. In either case, the differences in walk-time call into question the overall accuracy of the database.

[14] If the ratio were not invariant, one could conclude that very different types of customers were being served on different days—which seems unlikely. Alternatively, one could conclude that there were errors in the data collection process.

Table 3: Walk-time: Actual and Projected

Location ID		Total Time Observed Hours	Walk Time Seconds	Transaction Time Seconds	Walk/Transaction
128644	Day 1	7.65	88	27450	0.3%
	Day 2	6.62	2	23820	0.0%
98456	Day 1	1.80	111	6379	1.7%
	Day 2	2.27	43	8132	0.5%
	Day 3	1.02	0	3681	0.0%
116806	Day 1	6.77	563	23814	2.4%
	Day 2	8.86	902	31000	2.9%
69759	Day 1	3.41	512	11751	4.4%
	Day 2	4.29	432	15023	2.9%
27500	Day 1	4.24	45	15202	0.3%
	Day 2[3]	2.43	21	8736	0.2%
30442	Day 1	7.08	529	24966	2.1%
	Day 2[3]	4.50	245	15948	1.5%
30283	Day 1	7.18	176	25659	0.7%
	Day 2	6.86	210	24501	0.9%
118483	Day 1	11.47	196	41089	0.5%
	Day 2	7.75	271	27622	1.0%
4881	Day 1	9.26	599	32745	1.8%
	Day 2[3]	5.18	340	18307	1.9%
40832	Day 1	8.36	307	29800	1.0%
	Day 2	7.69	240	27435	0.9%
21799	Day 1	7.50	0	26988	0.0%
	Day 2	7.00	0	25208	0.0%
127869	Day 1	7.99	603	28166	2.1%
	Day 2	4.55	323	16060	2.0%
126721	Day 1	7.55	266	26900	1.0%
	Day 2	6.00	208	21405	1.0%
36211	Day 1	5.01	209	17833	1.2%
	Day 2	3.71	127	13240	1.0%
4079	Day 1	9.59	691	33844	2.0%
	Day 2	10.43	1274	36273	3.5%
20171	Day 1	9.13	11	32853	0.0%
	Day 2	10.04	2	36142	0.0%
69225	Day 1	10.30	63	37026	0.2%
	Day 2	10.43	26	37529	0.1%
70364	Day 1	2.55	230	8965	2.6%
	Day 2	2.20	85	7829	1.1%
85098	Day 1	2.52	115	8972	1.3%
	Day 2	1.38	0	4954	0.0%
119685	Day 1	7.88	187	28177	0.7%
	Day 2[3]	3.84	158	13671	1.2%
107799	Day 1	1.09	0	3921	0.0%
	Day 2	1.00	0	3610	0.0%
2303	Day 1	3.23	36	11602	0.3%
	Day 2	3.90	57	14000	0.4%
120905	Day 1	7.41	409	26254	1.6%
	Day 2	8.47	337	30168	1.1%
39717	Day 1	7.08	243	25231	1.0%
	Day 2	7.83	473	27700	1.7%
119973	Day 1	8.33	262	29726	0.9%
	Day 2	5.82	74	20886	0.4%
123775	Day 1	7.03	355	24960	1.4%
	Day 2	6.15	385	21763	1.8%
84745	Day 1	4.27	11	15356	0.1%
	Day 2	2.50	5	8983	0.1%

E. There Was No Analysis of Whether Enough Data Was Gathered for Each Type of Transaction in Order to Have a Statistically Adequate Sample.

Witness Nieto testified:

> No analysis of the number of transaction observations needed for each product was performed, because the study was not designed to provide national estimates of product-specific transaction times or product volumes. The notion of "statistically representative" product observations is not well-defined in the context of this update because many transactions contain multiple products. Rather, the objective of the transaction time study was to create a database that contained sufficient transaction to allow an update of the established transaction time econometric model.[15]

Normally, in performing survey work the researcher addresses the computation of required sample size: i.e., how much data is necessary in order to have statistically reliable conclusions from the sample at some acceptable level of confidence. There does not appear to have been any consideration of this issue in this case.

I have attempted to determine whether a sufficient sample of observations was obtained for each type of transaction. The calculations for the computation of sample size are shown in Table 4.[16] The formula for the required number of observations, "n" is

$n = (z^2/H^2)* s^2$; I used s^2 as the estimator of sigma, the variance.

I determined the necessary sample size in order that the mean value for transaction time is within a given range for a 95 percent confidence level, with z at 1.96. The half

[15] Tr. 5/701 (OCA/USPS-T24-2).

[16] Gilbert A. Churchill, Jr., *Market Research Methodological Foundations*, Sixth Edition, Dryden Press, 1995, at 631.

level of precision, denoted as H, is chosen based on the desire to have the estimate be within 5 percent of the average time per transaction. Sample size was computed with H set at 4.5 seconds.[17] In some cases, the database appears to be adequate, and in other cases the required sample is substantially larger than the sample gathered. Some additional consideration of sampling theory, stratification, and other techniques would probably decrease the required sample. However, the data indicate some concern over sample size.

Table 4: Actual and Required Samples

Variable and Product Category	More than one item may be in be in a transaction			Cases with one item in a transaction			Total Time for One Item Transactions	Required Sample Size Based on H = 4.5	Actual Sample Size
	Transactions	Time Seconds	Std.Dev.	Transactions	Time Seconds	Std.Dev.	Col(E)*Col(F)		
CERT (Certified Mail)	391	163.8	127.2				0	na	391
FC (First Class)	1777	129.6	115.2	674	74.8	52.6	50415	525	1777
STMPSCN (Samps Scanned)	2019	97.5	94.7	1223	61.8	46.1	75581	403	2019
STMPNO (Stamps Not Scanned)	1266	108.6	102.3	627	64.1	58.1	40191	640	1266
PM (Priority Mail)	1550	155.5	135.6	602	105	75.3	63210	1076	1550
MO (Money Order)	851	139.3	150	455	118.4	111.9	53872	2375	851
PP (Parcel Post)	291	176.3	156.2	85	101.3	57.2	8611	621	291
OWR	159	188.6	167.9	57	112.6	94.6	6418	1698	159
Bounded Printed Matter	13	203.5	204.1	5	82	64	410	777	1
Library Mail	1	86	na	na	na	na	0	na	323
Media Mail	148	192.5	170.4	52	115.6	97	6011	1785	148
EM (Express Mail)	322	182.2	136.5	235	155.2	78.9	36472	1181	322
PVI (PVI Strips)	101	165.5	157	43	100.6	112.5	4326	2401	101
INS (Insurance)	314	209.9	158.1	na	na	na	0	na	314
RP (Ready Post Items)	334	140.4	114.7	57	88.9	74.6	5067	1056	334
INTERNATL	332	243.7	207	152	188.8	151.6	28698	4360	332
STMPEN (Stamped Envelopes)	163	87.8	73	77	53.8	52.6	4143	525	163
REGINS (Registered with Insurance)	15	370.4	387.5	na	na	na	0	na	15
PASS (Passport)	35	803	428	13	807	510	10491	49343	35
RETAIL	8	312.9	147.7	3	337.3	101.8	1012	1966	8
BOX (PO Box Items)	82	213.8	198	68	191.6	195.6	13029	7258	82
DOMCOD (Domestic COD)	1	823	na	na	na	na	0	na	1
FCENCL (First Class with Enclosure)	1	161	na	na	na	na	0	na	1
OSS (Other Special Services)	839	184.7	153.5	na	na	na	0	na	839
Return Receipt	319	175.9	142	na	na	na	0	na	319
Delivery Confirmation	480	200.9	191.9	na	na	na	0	na	480
Signature Confirmation	18	348.6	542.8	na	na	na	0	na	18
Certificate of Mailing	22	102.8	98.8	na	na	na	0	na	22
Postage Due	9	123.4	48.2	8	129.5	47.8	1036	433	9
SERVICES	490	110.5	95.3			82		1276	490
Hold Mail	78	97.5	63.2	74	95.5	62.4	7067	739	78
Pickup	402	111.4	98.2	345	102	83	35190	1307	402
Mailing Payments	12	158.4	146	8	161.8	157.2	1294	4688	13

Total Transactions: 4863 Total Time: 452544

Time per Transaction: 93.058524

[17] 5 percent of 93 seconds is 4.5 seconds.

Adequate data appears to have been lacking for a number of window service items. For example, “Domestic COD” and “First Class with Enclosure” each had only one observation. “Mailing Payments” had only 13 observations whereas the calculated required sample size is 4688, and other services, such as Express Mail, Parcel Post, Media Mail and others, also appear to be lacking observations. In addition, witness Bradley combined a number of the data items for estimation purposes. However, in the case of “Services” and “OSS (Other Special Services)” there was no analysis that indicated such combinations were statistically appropriate. A number of disparate data items appear to have been combined without adequate explanation of the basis for the combination.

I conclude that the database presented by witness Nieto has a number of statistical problems which render it unsuitable for a final estimation of volume variabilities. Although the database may be better than the previous database, witness Nieto has not established that the database is adequate for witness Bradley’s analysis which it supports.

IV. DISCUSSION OF WITNESS BRADLEY'S TESTIMONY

Witness Bradley presents the update of the supply side variabilities for window transactions, based on the database developed by witness Nieto. Witness Bradley used EViews in developing his projections.[18] Table 5, below, "Comparison of Cases," summarizes witness Bradley's recommended variabilities and a number of variabilities for alternative cases which I have developed.

I recommend the variabilities shown in Table 5, column 7. I have performed the analysis in the SAS programming language and present witness Bradley's recommended case and the alternatives as SAS programs, logs, and output in OCA-LR-L-3.[19]

Table 5: Comparison of Cases
(Each Case is Discussed Below in the Text)

1	2	3	4	5	6	7
Product	**Bradley Recommended Variability**	**Recommended with Walk Variability**	**Studentized Residual at 3 Variability**	**Studentized Residual at 2 Variability**	**Studentized Residual 2 with walk time Variability**	**Smith Recommended Studentized Residual 2 with Walk time, BasketID Case**
Stamps Bulk	41.0%	40.8%	42.6%	40.28%	40.07%	40.0%
Stamps Non-Bulk	68.0%	68.1%	80.9%	72.55%	72.52%	72.5%
Priority	70.2%	69.9%	69.5%	69.54%	69.18%	68.9%
First Class	64.2%	64.0%	66.4%	67.35%	67.02%	66.9%
Parcel Post	75.3%	74.9%	72.2%	73.29%	72.85%	72.8%
Other W& R	67.5%	67.1%	67.7%	68.29%	67.94%	67.9%
Express Mail	66.4%	66.0%	67.6%	68.72%	68.30%	67.8%
Money Order	64.7%	64.4%	61.2%	62.26%	61.88%	61.6%
Certified	100.0%	100.0%	100.0%	100.00%	100.00%	100.0%
Insturance	100.0%	100.0%	100.0%	100.00%	100.00%	100.0%
Registered	100.0%	100.0%	100.0%	100.00%	100.00%	100.0%
International	78.5%	78.2%	77.5%	78.39%	78.06%	77.9%
PO Box	72.5%	72.2%	68.4%	67.32%	66.92%	66.4%
COD	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Other SS1	95.2%	95.1%	95.3%	95.35%	95.28%	95.6%
PVI	59.6%	59.2%	58.6%	56.10%	57.51%	58.0%

[18] USPS-LR-L-80 at 7 and the corresponding Excel file "Calculating Variabilities Addendum_49292.xls."

[19] For purposes of convenience, I have reproduced both witness Bradley's original Excel file as well as a corrected version based on his response to POIR No. 11, Question 1, in my library reference OCA-LR-L-3 along with the cases which I have developed.

A. Witness Bradley's Recommended Model Reproduced in SAS.

The SAS based Recommended Model contained in my library reference was prepared by me to reproduce witness Bradley's results using the database he provided in response to POIR No. 3, Question 10.[20]

In determining the number of SISQ and SIMQ transactions, I obtained identical results as witness Bradley, except for "Stamps Non-Bulk" (370 instead of 372 SIMQ transactions), "First Class" (169 rather than 170 SIMQ transactions), and "Other SS1" (9 SISQ instead of 1 SISQ transaction). In his response to POIR No. 11, Question 1, witness Bradley acknowledges that the SISQ value should be 9 for "Other SS1." In the case of "Other SS1," where the volume variability became 95.2 percent, based on the switching of 8 transactions, instead of witness Bradley's original calculation of 99.4 percent variability. This case shows that the impact of failing to account for a few transactions became very important: Witness Bradley's model is sensitive to changes in data.

B. Recommended Model Modified to Take into Account Walk-Time

Witness Bradley did not include walk-time in his model, but walk-time should have been included. If it had been included in the model, the resulting variabilities would be as shown in Table 5, column 3.[21]

Walk-time is defined as the time required by a customer to reach the service window. In some cases—e.g., small post offices with minimal walk-times—it is reported

[20] Tr. 5/820.

[21] The corresponding files are the Excel file "RMwithWalk.xls" and the SAS file "RMwithWalk.sas."

that the data collectors did not record walk-time. However, walk-time was recorded for larger offices or where the time was considered to be significant. This case incorporates walk-time as being a part of the window service time where such time was recorded as part of the transaction.

Witness Bradley summarized the definition of a window transaction, indicating:

> The "transaction time" recorded in POS is simply the time from when the window clerk presses the button to enter the first item to be sold until the time when the receipt is printed. This amount of time is short of the total transaction time and the gap between the actual time and the POS ONE recorded time will vary with the type of transactions.[22]

Witness Nieto explains that on-site data collectors recorded information on the time associated with the customer approaching the window (if applicable), the time the transaction began, and the time the transaction ended.[23] She stated that customer walk-time is not included in the transaction time:

> d. No. The purpose of the transaction time study was to construct a database permitting an update of the econometric model of transaction time, thus any non-transactional time was not relevant.

> e. The "walk" part of the transaction was recorded for those transactions in which the walk was long enough to allow a data collector to record a separate measurement. However, the time associated with the "walk" part of the transaction was not included in the calculation of transaction time for any transactions….[24]

[22] USPS-T-17 at 5, lines 6-10.

[23] USPS-T-24 at 6, lines 13-17.

[24] Tr. 5/703 (OCA/USPS-T24-3(e)).

During walk-time, the clerk may be indicating through body language a readiness to serve, may be actually greeting the customer, or may simply be standing and waiting. The approach of the customer requires the clerk to be available and prepared to serve; whether the clerk is actually doing any other meaningful work is irrelevant, for the walk-time procedure requires that the clerk devote time to being present and prepared to serve. Accordingly, customer walk-time should have been included as part of the window service transactions.

C. Alternative Variabilities Resulting from Adjusting for Various Outliers

I have further calculated variabilities for three different cases adjusting for the elimination of various outliers and walk-time as shown in columns 4-6 of Table 5. I do not recommend these variabilities, but they are exhibited for informational purposes to demonstrate the impact of various adjustments that might have been made by witness Bradley. I also calculated the variabilities for a fourth case, discussed below, which I do recommend, the "studentized residual 2 with walk time, basketID" case appearing in column 7 of Table 5.

1. Studentized Residual at 2 Variability

Witness Bradley's study removed a number of outlying observations, as mentioned on pages 22-24 of USPS-T-17. POIR No. 7 requested that witness Bradley supply a test for outliers. In response, witness Bradley noted that he had removed a number of outliers, but that he had not used a formal statistical test.[25]

[25] POIR No. 7, Question 7. (The response to this POIR question was designated and commences at Tr. 5/779, but most of the response, including the portion pertinent here, was not transcribed.)

Witness Bradley's response then discussed the use of studentized residuals:

> In general, a studentized residual is a residual divided by its standard error, but in outlier analysis it is important to have an "externally' studentized residual. Externally studentized residuals have the standard error calculated with the observation removed, thus eliminating the possibility that a large outlier could contribute to a large standard error and thus make it harder to identify the outlier itself.[26]

Witness Bradley's response to Question 7 of POIR No. 7 provided a list of studentized residuals with an absolute value above 2. He concluded that observations with a studentized residual above absolute value 3 are likely outliers, and that those with a studentized residual above 2 in absolute value bear investigation. He furnished a list of 250 observations with a studentized residual above 2 in absolute value and 117 observations with a studentized residual above 3.

Using the list of 250 observations with a studentized residual above 2, I have calculated the variabilities which appear in column 5 of Table 5.[27]

2. Studentized Residual at 2 with walk

I also calculated varibilities including walk-time as part of the overall time along with the studentized residual at 2, as in the analysis above.[28] The resulting variabilities are shown in Table 5, column 6.

---

[26] *Ibid.* at 2.

[27] The Studentized2 SAS program in my library reference OCA-LR-L-3 provides a regression with the 250 observations deleted. The associated Excel file is "Studentized2.xls."

[28] The relevant computer files in my library reference OCA-LR-L-3 are "Studentized2withtime," either in Excel or in SAS.

3. Studentized Residual at 3

I further calculated variabilities in the case of studentized residual at 3 which eliminates 117 observations.[29] The variabilities appear in column 4 of Table 5.

4. Studentized Residual at 2, Basket ID case

The specific case and consequent variabilities that I recommend, based on the above paragraphs and other information developed, are listed in column 7 of Table 5 and take into account the various data issues previously discussed.[30]

Question 3 of POIR No. 7[31] notes that the input data file contains positive values of the variable "items" and zero values for the variable "quantity" in a number of cases. Witness Bradley noted in his response to Question 4 of POIR 7[32] that a zero value for a window service item means that there was a transactional activity for an item, although no quantity was ultimately purchased. He cited examples, including an inquiry about a product or the acceptance of a previously stamped product. That is, there was a transactional activity without, however, the purchase of a quantity. The regression and variability analysis were performed with the value for "item" set to zero whenever "quantity" was equal to zero.

The regression underlying my preferred and recommended model is summarized in Table 6. The output from the SAS program provides a standard error and t statistic

---

[29] The relevant files in my library reference OCA-LR-L-3 are "studentized3.xls" and "studentized3.sas."

[30] The associated Excel and SAS files in my library reference OCA-LR-L-3 are denoted as "Studentized2withtimebasketid."

[31] Tr. 5/809.

[32] Tr. 5/808.

on the assumption of normality in the residuals. Since the residuals in this case are not normally distributed, the relevant statistics are the HC Standard Error and HC t-statistic, as discussed by witness Bradley in response to an interrogatory in Docket No. R2005-1:

> The HC standard error is calculated by taking the square route (sic) of the variances along the main diagonal of the consistent variance/covariance matrix from the regression. This variance/covariance matrix is provided in Library Reference USPS-LR-K-81. The HC T-statistic is calculated by the ratio of the estimated coefficient to the HC standard error.[33]

SAS does not provide the HC statistics, so they must be computed separately, as presented in the Excel workfile "Regstats.xls." In general, I have not computed the HC statistics for the various regressions. The procedure is time consuming, and I have not used the HC statistics in terms of decision making. One of the advantages of EViews is the computation of the HC statistics within the program.

[33] Docket No. R2005-1, OCA/USPS-T-14-17.

Table 6: Summary of Model

	Paramter Estimate	HC Standard Error	t value	Variances
CERT	6.16336	2.901947622	2.1239	8.4213
FC	17.10971	1.09795264	15.583	1.2055
STMPSCN	3.12274	0.512347538	6.095	0.2625
STMPNO	0.22611	0.076811457	2.9437	0.0059
PM	27.61926	1.471937499	18.764	2.1666
MO	30.67455	1.228356626	24.972	1.50886
PP	36.92101	2.851894107	12.946	8.1333
OWR	27.87613	2.833354902	9.8386	8.0279
EM	81.04351	3.321309983	24.401	11.0311
PVI	28.79125	6.494905696	4.4329	42.1838
INS	27.1165	3.153632826	8.5985	9.9454
RP	8.22873	2.817392411	2.9207	7.9377
INTERNATL	60.69035	3.466958321	17.505	12.0198
STMPEN	0.17368	0.024494897	7.0905	0.0006
REGINS	148.9322	30.13016263	4.943	907.8267
PASS	422.2888	28.57551749	14.778	816.5602
RETAIL	86.03752	29.09050876	2.9576	846.2577
BOX	88.79309	10.51655837	8.4432	110.598
OSS	8.63649	2.16981566	3.9803	4.7081
SERVICES	35.6945	2.567605889	13.902	6.5926
CHECK	29.2073	3.162736157	9.2348	10.0029
CREDIT	20.24144	2.413420809	8.387	5.8246
DEBIT	14.59669	3.183771349	4.5847	10.1364
ITEMS	15.4473	1.351961538	11.426	1.8278
D1	17.32713	3.299590884	5.2513	10.8873
D2	21.00086	3.934094559	5.3382	15.4771
D3	28.39567	3.3245902	8.5411	11.0529
D4	30.45268	3.939352739	7.7304	15.5185
D5	30.64936	2.577130187	11.893	6.6416
D6	34.5585	2.822073706	12.246	7.9641
D7	29.84497	2.416691954	12.35	5.8404
D8	31.06525	3.471397413	8.9489	12.0506
D9	35.69041	3.497441922	10.205	12.2321
D10	38.16129	6.931204801	5.5057	48.0416
D11	28.06793	2.967237773	9.4593	8.8045
D12	38.80215	2.934160868	13.224	8.6093
D13	47.01444	4.637283688	10.138	21.5044
D14	91.77532	5.899542355	15.556	34.8046
D15	66.87998	3.901820088	17.141	15.2242
D16	72.08817	4.406143438	16.361	19.4141
D17	68.60539	3.889408695	17.639	15.1275
D18	24.85237	4.609186479	5.3919	21.2446
D19	47.30878	3.916260972	12.08	15.3371
D20	47.03312	4.914539653	9.5702	24.1527
D21	47.11763	3.788799282	12.436	14.355
D22	58.25412	7.42806839	7.8424	55.1762
D23	41.37101	3.817603961	10.837	14.5741
D24	48.53171	3.003647782	16.158	9.0219
D25	43.61772	3.122034593	13.971	9.7471
D26	37.48149	3.178411553	11.793	10.1023
D27	45.64001	3.809685026	11.98	14.5137

D. Choice of Model—Linear or Other

Witness Bradley used a linear model as the basis for his analysis. He noted that the model had been used in Docket Nos. R2001-1 and R2005-1, with no party criticizing or objecting to the analysis, concluding that it seemed appropriate once again to adopt a linear specification.[34] During the hearings, he reiterated his position and noted further that the linear model provides variabilities which are consistent with the entire structure of the window service costing model.[35] It should be noted, however, that the use of a model other than a linear model for this phase of the window operation would not be theoretically inappropriate: no advantage of a linear model has been established for this purpose, nor would use of a different type of model in any way conflict with the analysis. A quadratic model or other type of suitable flexible functional form could be used; in other cases witness Bradley has strongly advocated the use of flexible functional forms.

A deficiency of the linear model is that it effectively sets volume variabilities for a number of products sold with other products, such as insurance and return receipt, at 100 percent because of the selection of a linear model. The products are always sold in conjunction with another product. Witness Bradley explained the reason for the assumption of 100 percent variability in those cases:

> The assumption really comes from the fact that for those special services they're never sold by themselves. They're always sold with another product and so they have no transaction core related time. All they have is the time associated with themselves in the transaction, and that's why you get the 100 percent.[36]

[34] Tr. 5/740 (OCA/USPS-T17-3(b)).

[35] Tr. 5/848.

[36] Tr. 5/853.

Witness Bradley further states:

> In a linear model, like the established model, a product never sold by itself would have no "fixed" or "common" time. It is easy to show that a in (*sic)* linear model, a product with no "fixed" or "common" time has a variability that is 100 percent.[37]

Accordingly, I advocate consideration of a flexible functional form in estimating volume variabilities of window services.

[37] USPS-T-17 at 10.

V. CONCLUSIONS

A database that is not representative of the population of data cannot yield meaningful results. Witness Bradley's model seems to be very susceptible to potential errors in the database, as indicated in Table 2. Witness Nieto has not provided affirmation that the database meets the quality standards that are needed; in fact, my analysis of walk-time concludes that there may have been some problems with the quality of the data collection, a concept further reinforced by the substantial number of data items dropped during data collection.

My analysis of witness Bradley's work results in minor changes as proposed in Table 5 to his conclusions and recommended variabilities. However, it should be noted that major changes in variabilities might result from the development of an adequate database. This is, however, a matter about which we have insufficient information.

Furthermore, witness Bradley has endorsed the use of a flexible functional form for estimation purposes in other cases. It is speculative as to whether the use of a flexible functional form in this analysis would have resulted in substantially different conclusions. It is clear, however, that for some products delineated in this testimony, such as insurance and return receipt, the use of a linear form is essentially equivalent to assuming 100 percent volume variability.