

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

PERIODIC REPORTING

Docket No. RM2009-5

RESPONSES OF THE UNITED STATES POSTAL SERVICE TO CHAIRMAN'S
INFORMATION REQUEST NO. 3
(September 22, 2009)

The Postal Service hereby files its responses to questions 1-9 of Chairman's Information Request No. 3, filed on September 4, 2009. Each question is stated verbatim, and followed by the response.

Respectfully submitted,

UNITED STATES POSTAL SERVICE

By its attorneys:

Daniel J. Foucheaux, Jr.
Chief Counsel, Pricing & Product Support

Eric P. Koetting

475 L'Enfant Plaza West, S.W.
Washington, D.C. 20260-1137
(202) 268-2992, FAX: -5402
September 22, 2009

**RESPONSE OF THE UNITED STATES POSTAL SERVICE
TO CHAIRMAN'S INFORMATION REQUEST NO. 3**

1. At the technical conference, the Postal Service indicated that the ODIS-RPW subsystem of the RPW system samples volume and other data for bulk mail as well as single-piece mail and that the ZIP Code level data collected for bulk mail products are used by the Postal Service in calculating the value of the monopoly and the cost of the USO. Please indicate what other Postal Service-sponsored studies or models rely on ZIP Code-level sample data for bulk mail products collected by the ODIS-RPW system.

RESPONSE:

The service performance measurement systems for Presort First-Class Mail, Periodicals and Standard Mail currently rely on ZIP Code-level sample data collected by the ODIS-RPW system.

**RESPONSE OF THE UNITED STATES POSTAL SERVICE
TO CHAIRMAN'S INFORMATION REQUEST NO. 3**

2. Please provide coefficients of variation for the ODIS-RPW volumes for the following categories of mail:

- First-Class single-piece letters
- First-Class single-piece cards
- First-Class single-piece flats
- First-Class single-piece parcels
- Single-piece Parcel Post
- Bound Printed Matter flats
- Bound Printed Matter parcels
- Media and Library Mail
- Priority Mail

Please provide these estimates for the most recent fiscal year, both annual and by quarter. Please provide them at the national level, the District level, the 3-digit destinating ZIP Code level, and the 3-digit destinating ZIP Code pair level. Two sets of estimates should be provided—one that assumes the current sample size and one that assumes the Proposal One sample size.

RESPONSE:

CVs for the Current Sample Size:

National CV (Coefficient of Variation) estimates for FY2008 volumes were previously provided in Table 1 of Proposal One (Docket No. RM2009-5). See attachment 'ChIR.3.Q2.CV_single-piece_2008.xls' (an Excel file attached to this response electronically) for a) the quarterly FY2008 national CVs estimates ('natl' tab), b) the FY2008 quarterly and annual District level CV estimates ('csd' tab), and c) the FY2008 quarterly and annual three-digit destinating ZIP Code level current sample size CV estimates ('zip' tab).

Three-digit origin-destination pair volume CV estimates are not provided in this response. Because of the large number of 3-digit ZIP code pairs, some 3-digit ZIP code pairs are likely to have very few little to no information through the ODIS-RPW system.

**RESPONSE OF THE UNITED STATES POSTAL SERVICE
TO CHAIRMAN'S INFORMATION REQUEST NO. 3**

Missing sample information in some 3-digit ZIP code pairs would be more prominent at quarterly estimation. Therefore, model-based inference is necessary to provide the requested estimates for 3-digit ZIP code pairs with adequate precision. First, auxiliary variables at 3-digit ZIP code pair level are needed to form a model that relates volumes and auxiliary variables at 3-digit ZIP code pair level.

Let $x_i = (x_{i1}, x_{i2}, \dots, x_{ip})^t$ be the vector of p auxiliary variables where i represents the index of 3-digit ZIP code pairs (approximately 810,000) and t denotes the transpose of a vector. Then we could assume, for example, a model of volume Y such as

$$Y_i = x_i' \beta + z_i v_i, \quad i = 1, \dots, 810000$$

where z_i 's are known positive constants, β is the vector of regression parameters and the v_i are independently and identically distributed (iid) random variables with $E(v_i) = 0$ and $V(v_i) = \sigma_v^2$. Writing the estimator of Y_i as

$$\hat{Y}_i = Y_i + \varepsilon_i, \quad i = 1, \dots, 810000,$$

where ε_i 's are sampling errors with $E(\varepsilon_i | Y_i) = 0$ and $V(\varepsilon_i | Y_i) = \sigma_\varepsilon^2$, that is, the estimator \hat{Y}_i , are design unbiased. So we have a mixed model

$$\hat{Y}_i = x_i' \beta + z_i v_i + \varepsilon_i, \quad i = 1, \dots, 810000$$

which includes both model-based and design-induced random terms.

Several estimation methodologies are available for such a model, e.g., empirical BLUP (best linear unbiased prediction), empirical Bayes, and hierarchical Bayes.

Theoretically, we would be able to obtain the requested estimates, but the most difficult part of this approach is to find good auxiliary information highly correlated with volume at 3-digit ZIP code pair level. As we are not aware of any census information taken at 3-digit ZIP code pair level that can be used as an auxiliary variable, there would

**RESPONSE OF THE UNITED STATES POSTAL SERVICE
TO CHAIRMAN'S INFORMATION REQUEST NO. 3**

be time needed to investigate the availability of Operations or mail processing data. It would take approximately 500 work-hours (one-fourth a work-year) to investigate and build a data base of auxiliary information and models. This could be completed within six months.

CVs for the Proposal One Sample Size:

National CV estimates were provided in Table 1 of Proposal One (Docket No. RM2009-5). The remaining Proposal One CVs can either be derived from the attached results, or are not provided in this response.

Regarding a), since the quarterly sample size is the same for each quarter, the Proposal One sample size reduction CVs can be derived by applying the Table 1 ratio of the 20% Reduction CV to the Full Sample CV to each quarter. For example, for First-Class Single-Piece Letters volume, the ratio of the 20 Percent Reduction CV of 0.43 to the Full Sample CV of 0.35 equals 1.23 (or 23 percent increase). This percent increase in the CV can reasonably be applied to each First-Class Single-Piece Letters quarterly national CV in the attachment. This same procedure would be repeated for each report category.

Regarding b), the Proposal One CVs are not provided in this response. In order to calculate the district CVs the procedures originally used for the national calculations filed with Proposal One would have to be replicated for each District. This work would require 40 work-hours, and could be completed within a month's time period.

Regarding c), the Proposal One CVs are not provided in this response as well. The difficulty with computing three-digit destinating CVs under sample size reductions is

**RESPONSE OF THE UNITED STATES POSTAL SERVICE
TO CHAIRMAN'S INFORMATION REQUEST NO. 3**

explained in response to Question 4 of this Information Request below. This work would require a replication or Monte Carlo simulation where 20 percent reductions in sample size were repetitively taken, and the resulting CVs averaged. Some data base development from the historical set of ODIS-RPW tests would probably be needed at the sample design strata level to build models of the estimates and CVs under the sample reductions. It is expected that it would take 100 work-hours to complete and could be done within two months.

Three-digit ZIP Code pair level CVs under the Proposal One reductions are not provided in this response as well. In addition to the costs for building the small area models described above, an approximate additional 200 work-hours would be needed to replication or Monte-Carlo simulations to estimate the impact of the Proposal One sample reductions.

**RESPONSE OF THE UNITED STATES POSTAL SERVICE
TO CHAIRMAN'S INFORMATION REQUEST NO. 3**

3. Please provide the information requested in Question 2 for the bulk mail products for which sample ODIS-RPW volume data are collected.

RESPONSE:

Since there was no specific request for the level of the product breakouts, class level statistics are provide in this response. Specifically, First-Class Presort, Standard Mail, Periodical, and Parcel Post totals are included.

CVs for the Current Sample Size:

Attachment 'ChIR.3.Q.3.CV_dest_bulk_tot 2008.xls' (an Excel file attached to this response electronically) provide bulk mail estimated volume CVs for a) FY2008 national annual and quarterly levels ('natl' tab), b) FY2008 district annual and quarterly levels ('csd' tab), and c) FY2008 destinating three-digit ZIP code annual and quarterly levels ('zip' tab). Three-digit ZIP code pair CV estimates are not applicable. This is because origin ZIP codes are not generally present on bulk mail entered mailpieces.¹

CVs for the Proposal One Sample Size:

Proposal One sample size CVs are either not explicitly provided, or are not provided in this response. Regarding the calculation of CVs for a) and b), it is reasonable to expect them to increase approximately 12 percent (following similar logic and assumptions as that described in response to Question 5 of Chairman's Information Request No. 1, filed Aug. 13, 2009). Specifically, the assumptions are that the national

¹ Bulk entered metered mail often contains origin ZIP codes. However, bulk metered mail is a very small proportion of all total bulk entered mail. Permit imprint indicia dominates bulk mail entry. Origin metered ZIP code distribution may not be representative of all bulk mail entered mail volume.

**RESPONSE OF THE UNITED STATES POSTAL SERVICE
TO CHAIRMAN'S INFORMATION REQUEST NO. 3**

and district sample sizes are large and the 20 percent reduction in sample size is applied to districts proportionally.

Regarding c), the CVs are not provided in this response. This work would require a replication or Monte Carlo simulation as described previously in response to Question 2 of this Information Request. It is expected that it would take 200 work-hours, and could be completed with two months.

**RESPONSE OF THE UNITED STATES POSTAL SERVICE
TO CHAIRMAN'S INFORMATION REQUEST NO. 3**

4. With respect to coefficients of variation for destinating 3-digit ZIP Codes, the Postal Service's response to Question 5 of CIR No. 1 indicates that under certain assumptions, the deterioration in precision for ZIP Codes with "moderately large" volumes would be "approximately 12 percent." Please estimate the proportion of ZIP Codes that would have too little volume to be sampled, or whose sampled volume could not yield a consistent estimate. Two sets of estimates are requested—one that assumes the current sample size, and one that assumes the reduced sample size under Proposal One.

RESPONSE:

The sample design of the ODIS-RPW system does not easily allow for the calculation of the proportion of ZIP codes where the 12 percent rule would not be applicable. The problem lies in the relationship of the Mail Exit Point to the category of mail to be estimated. The sampled Mail Exit Point (MEP)-day generally represents one three-digit destinating ZIP Code for either the letter, flat, parcel, Priority, or mixed shapes mail at the post office. The class of mail for which the CVs of interest are to be calculated (e.g., Priority Mail) may or may not be a possible mail class in that MEP. For example, Parcel Post products do not 'pass through' letter MEP. Therefore, the 'cell' for which there is no three-digit destinating volume in this case is a legitimate 'zero' cell. Other cases are less clear, such as where the occasional Priority letter appears in the letter MEP. Added difficulties exist because strata definitions change each quarter, as described in the statistical design documentation noted in response to Chairman's Information Request No. 1, Question 2 (August 13, 2009).

If we are allowed to ignore the ODIS-RPW sample design, then we could post-stratify ODIS-RPW data by destinating three-digit ZIP code and mail category. Treating post-stratified ODIS-RPW data as a simple random sample, we could count the number

**RESPONSE OF THE UNITED STATES POSTAL SERVICE
TO CHAIRMAN'S INFORMATION REQUEST NO. 3**

of tests in a cell defined by destinating three-digit ZIP codes and mail categories. We would have to decide on a threshold value for a number of tests required to obtain consistent volume estimates at destinating three-digit ZIP code level for each mail category. Then, for 'cells' with the number of tests exceeding the threshold value, a consistent estimate is obtainable; otherwise a consistent estimate is not obtainable for the cell. Applying the same process to all 3-digit ZIP codes and mail categories, we could obtain the proportion of 3-digit ZIP codes and mail categories for which consistent volume estimates are not feasible. This procedure is not obtained through the ODIS-RPW sample design. It is obtained through post-stratification under simple random sample, which does not take the ODIS-RPW design into account. We estimate it would cost one-fourth of a work-year (three months) to conduct this analysis, and it could be completed with six months.

**RESPONSE OF THE UNITED STATES POSTAL SERVICE
TO CHAIRMAN'S INFORMATION REQUEST NO. 3**

5. Please apply an analysis comparable to that provided in response to Question 4 to volume data for 3-digit ZIP Code pairs.

RESPONSE:

Please see the response to Question 4 of this Information Request.

**RESPONSE OF THE UNITED STATES POSTAL SERVICE
TO CHAIRMAN'S INFORMATION REQUEST NO. 3**

6. At the technical conference it was confirmed that the External First-Class Measurement System (EXFC) relies on ODIS-RPW ZIP Code pair volumes both to allocate service performance tests at the outset, and to weight the test results once they are collected to arrive at performance scores above the ZIP Code pair level. Please provide coefficients of variation for the resulting aggregated scores. Please provide two sets of estimates--one assuming the current ODIS-RPW sample size, and one assuming the sample size under Proposal One.

RESPONSE:

Currently the coefficients of variation produced for the EXFC program do not consider the variability of the ODIS-RPW mail volume estimates. By using a Taylor Series expansion, the following formula could be used to derive the aggregate estimate and the variance of the aggregate service performance estimates.

$$\hat{p} = \frac{\sum_i w_i \hat{p}_i}{\sum_i w_i}$$

$$Var(\hat{p}) = A + B + C + D + E$$

where

**RESPONSE OF THE UNITED STATES POSTAL SERVICE
TO CHAIRMAN'S INFORMATION REQUEST NO. 3**

\hat{p}_i = performance estimate for mail originating from a single district, destinating to a single district
with a single service standard

w_i = ODIS- RPW - based volume estimate for mail originating from a single district, destinating to a district
with a single service standard

i = the origin district - destination district - service standard combination

\hat{p} = service performance on - time estimate for an aggregate level (across origin districts, destination districts, and/or service standards

$$A = \frac{1}{X^2} \sum_i w_i^2 \text{Var}(\hat{p}_i)$$

$$B = \sum_i \left(\frac{X - w_i}{X^2} \right)^2 \text{Var}(w_i) \text{Var}(\hat{p}_i)$$

$$C = \frac{1}{X^2} \sum_i (\hat{p}_i - \hat{p})^2 \text{Var}(w_i)$$

$$D = \frac{1}{X^2} \sum_i \sum_{j \neq i} w_i w_j \text{Cov}(\hat{p}_i, \hat{p}_j)$$

$$E = \frac{1}{X^2} \sum_i \sum_{j \neq i} (\hat{p}_i - \hat{p})(\hat{p}_j - \hat{p}) \text{Cov}(w_i, w_j)$$

and X is the sum of the weights derived from ODIS-RPW data

$$X = \sum_i w_i$$

The individual terms in the expansion represent

- Variance of the on time estimate (A)
- Product of the variance of the ODIS-RPW based weights and the variance of the on time performance estimates (B)
- Variance of the ODIS-RPW based weights (C)
- Covariance of the on time estimates (D)
- Covariance of the ODIS-RPW based weights (E)

In the development of the EXFC study, analyses were conducted to evaluate each of the variance components. Terms A and D were determined to be the most significant components of variance, while the remaining terms were found to be negligible and were omitted from the calculation formulae. The analyses were performed again in 1995 by the Postal Service and the estimated impact on the variance for the omission of

**RESPONSE OF THE UNITED STATES POSTAL SERVICE
TO CHAIRMAN'S INFORMATION REQUEST NO. 3**

terms B and C was estimated to be an average of 4.4 percent understatement of the range. That is, a range of +/-0.80 would be +/-0.84 if the B and C terms were included. Note that term E was not included in the analyses because it was considered to be too small to be relevant. In 1995, the ODIS volume estimates used for EXFC were based on one quarter from the same period the prior year rather than the twelve-quarter average which is used currently. Using twelve quarters of ODIS-RPW data means that the impact on the variance estimates is likely to be even lower than when only one quarter of data was used.

In order to calculate the full variance of the EXFC performance estimates for the periods requested by the PRC, one would need to construct variance estimates for the twelve-quarter ODIS-RPW estimates used as weights in the performance estimates. These volume estimates are the average daily volume at the origin district-destination district-service standard level, aggregated using the origin ZIP Code-destination ZIP Code ODIS-RPW data for the prior twelve quarters (lagging by two quarters based on data availability) matched up with the service standards for the quarter being measured. Upon construction of these estimates, modifications to the software programs used to produce the EXFC service performance results would be required to include the additional variance terms. For completeness, one would also need to re-examine the assumption that the E term is negligible by calculating the covariance of the ODIS-RPW based weights and constructing the term E.

In order to assess the impact of the proposed ODIS-RPW sample reduction on the EXFC coefficients of variation, one would need to estimate the impact on the terms B, C, and E. Note that due to the use of twelve quarters of data, it would take three years

**RESPONSE OF THE UNITED STATES POSTAL SERVICE
TO CHAIRMAN'S INFORMATION REQUEST NO. 3**

following the implementation of the ODIS-RPW sample reduction for the full impact of the reduction to be realized. Furthermore, beginning in FY2009, EXFC was expanded to extend testing to virtually all 3-Digit ZIP Code areas (from the prior 463 3-Digit ZIP Code areas to 892) and test mail characteristic shifts were made to focus EXFC on single-piece mail characteristics from the prior use of all First-Class Mail characteristics. As a result of the significant changes to EXFC, the delayed impact due to the use of twelve quarters of ODIS-RPW data, as well as the prior analyses indicating that the ODIS-RPW data represents a small source of variability in the EXFC performance estimates, any estimates produced based on the FY2008 EXFC results requested in the Chairman's Information Request are unlikely to be accurate estimates for the true coefficients of variation that will result in three years.

The requested estimates could be produced, if deemed necessary, but not within the required timeframe. In accordance with the instructions, we have provided an estimate of the work-hours that would be required to apply the procedure, and an estimate of how long it would take to complete the procedure.

The estimates are 200 work-hours to produce the variance and covariance estimates for the ODIS-RPW-based weights for the four twelve-quarter periods used for the FY2008 quarterly results, along with the estimated impact of the proposed sample reductions on the variance estimates. The periods of ODIS-RPW data to be used included the following:

- Q1 was based on Q4 FY04 to Q3 FY07
- Q2 was based on Q1 FY05 to Q4 FY07
- Q3 was based on Q2 FY05 to Q1 FY08
- Q4 was based on Q3 FY05 to Q2 FY08

**RESPONSE OF THE UNITED STATES POSTAL SERVICE
TO CHAIRMAN'S INFORMATION REQUEST NO. 3**

For each time period, the volume estimates for the origin district to destination district by service standard would be needed, along with the variance of the estimates and, if appropriate, the covariance between all pairs of estimates. There are approximately 7,600 origin district-destination district-service standard combinations to produce variance estimates for, with the covariance matrix exceeding 28 million combinations. This would need to be produced for each of the four time periods noted above, and then estimates of the impact of the ODIS-RPW sample reduction would need to be applied to the data from each time period.

An additional 120 work-hours is estimated to develop and test changes to the software used to produce the variance estimates, and to reload the data from the prior year, run the programs and produce the requested estimates. In total, it is anticipated that 320 work-hours would be required to fulfill this request.

In order to maintain the schedule to deliver the Q4 and final annual reports for FY09 due in mid-October, the contractor for EXFC could not begin work on the software modifications and reloading/reprocessing the old data until approximately October 15. Based on the estimated hours required, the results could be available in mid-December.

**RESPONSE OF THE UNITED STATES POSTAL SERVICE
TO CHAIRMAN'S INFORMATION REQUEST NO. 3**

7. Please indicate whether the Postal Service plans to use sample volume data for bulk mail products from the ODIS-RPW subsystem to allocate tests or to weight test results in its planned system for measuring service performance for bulk mail products. If it does, please provide an analysis of the impact of reducing the ODIS-RPW sample size on the precision of the resulting test scores comparable to the analysis provided in response to Question 6.

RESPONSE:

Yes, the Postal Service is currently using the sample volume data from the ODIS-RPW subsystem to weight service performance results for Presort First-Class Mail, Periodicals, and Standard Mail. For Periodicals and Standard Mail, only the destination volumes from ODIS-RPW are used because the ODIS-RPW sampling process cannot determine the origin of most pieces which were dropped in bulk. Due to the lack of origin-destination mail volume data, alternate data sources are being explored for use in the future for bulk mail products. As the mail volume and the number of mailers participating in Full Service Intelligent Mail® grows, it is likely that mail flow data from this population will be more advantageous to use because the data will have both origin and destination information, as well as presort level. The plan is to monitor the coverage of Full Service Intelligent Mail® and assess the data against the estimates from ODIS-RPW sampling data as well as using the overall national level estimates from the ODIS-RPW reports for bulk products, which are derived from postage statement data.

We believe that it is likely that the proposed reduction to the ODIS-RPW sampling data will have very limited impact on the service performance estimates for these bulk mail products. Similar to EXFC, a twelve quarter average of the volume estimates are being used. As noted in the response to Question 6 of this Information Request, it will be three years from the time of implementation before the full impact of the reduction would be realized. Given the plan and need to revise the methodology for

**RESPONSE OF THE UNITED STATES POSTAL SERVICE
TO CHAIRMAN'S INFORMATION REQUEST NO. 3**

weighting the service performance results, as well as the current very limited data available in the pilot systems being used for service measurement in FY09, it is not possible to calculate the impact of the proposed ODIS-RPW sample reduction.

**RESPONSE OF THE UNITED STATES POSTAL SERVICE
TO CHAIRMAN'S INFORMATION REQUEST NO. 3**

8. The Postal Service's response to Question 8 from CIR No. 1 acknowledges that transportation optimization modeling relies on 3-digit ZIP Code pair volume flows taken from the ODIS-RPW system but asserts that "the corresponding dispatch plans and routing instruction[s] for the procured transportation" do not.

a. Does optimization modeling play no role in the dispatch plans and routing instructions developed for purchased transportation?

b. What volume flow data are used to develop dispatch plans and routing instructions for purchased transportation?

RESPONSE:

a. The optimization model plays an indirect role in the dispatch plans and routing instructions. For example, the ODIS data are being used in TOPS for long term planning.

b. For dispatch plans and routing instruction, SPLY, plan vs. actual data are utilized, and for long term planning, SPLY, EOR and ODIS data are being utilized.

**RESPONSE OF THE UNITED STATES POSTAL SERVICE
TO CHAIRMAN'S INFORMATION REQUEST NO. 3**

9. For the nine mail categories indicated in Question 2, please provide the current ODIS-RPW target CVs for the month by sample area estimates. Please provide the corresponding target CVs associated with the proposed sample size reduction.

RESPONSE:

The following table provides the monthly coefficient of variation (CV) targets for the approximate 196 sample areas. These targets are from the most recent quarterly sample selection: Quarter 4 FY2009.

Table: ODIS-RPW Sample Selection Monthly Targeted CVs

	-----Sample Areas -----					
Stratification Variable	All Sample Areas Except those provided in Columns to the Right	Portland, Maine	Mid-Island, N Y	Merrifield, VA	Little Rock, Arkansas	Memphis, Tenn.
Predicted First Class Letters / Cards	0.20	0.10	0.15	0.20	0.10	0.25
Priority Mail	0.40	0.15	0.40	0.40	0.20	0.40
Mixed Parcels	0.35	0.35	0.35	0.35	0.35	0.35
Predicted Standard Mail Letters / Predicted First Class Flats	0.75	0.60	0.75	0.75	0.35	0.35
Predicted Standard Mail and Periodical Flats	0.35	0.35	0.35	0.35	0.35	0.35
Accountable / Business Reply Mail	3.00	3.00	3.00	3.00	12.00	20.00

CERTIFICATE OF SERVICE

I hereby certify that I have this date served the foregoing document in accordance with Section 12 of the Rules of Practice and Procedure.

Eric P. Koetting

475 L'Enfant Plaza West, S.W.
Washington, D.C. 20260-1137
(202) 268-2992, FAX: -5402
September 22, 2009