REPORT TO THE CONGRESS

BY THE COMPTROLLER GENERAL
OF THE UNITED STATES

System For Measuring
Mail Delivery Performance--
Its Accuracy And Limits

United States Postal Service

The Postal Service's system for measuring mail delivery performance shows that, on a nationwide basis, 95 percent of local first-class mail receives overnight delivery. This statistic is reliable, but it includes only stamped first-class mail, which accounts for 40 percent of first-class mail volume.

The measurement system has weaknesses which make local area overnight delivery performance statistics less reliable. These weaknesses have no significant effect on nationwide statistics. However, the reliability of local performance statistics could be improved.

GGD-75-109

706585

OCT.17,1975

B-114874

C/ To the President of the Senate and
the Speaker of the House of Representatives

This report examines the accuracy and limits of the
Origin-Destination Information System which the Postal
Service uses to measure mail delivery performance. The
report suggests ways for improving both the public con-
ception of mail delivery performance statistics and the
reliability and accuracy of local delivery performance
statistics.

Since we make frequent references to Origin-Destination
Information System statistics as a measure of the quality
of mail service, it was determined that a review examining
the reliability and accuracy of the system was in order.

We made our review pursuant to the Postal Reorganiza-
tion Act of 1970 (39 U.S.C. 2008).

We are sending copies of this report to the Director,
Office of Management and Budget; the Postmaster General;
and each member of the Board of Governors of the Postal
Service.

Comptroller General
of the United States

Contents

ABBREVIATIONS

GAO   General Accounting Office

OCR   optical character reader

ODIS   Origin-Destination Information System

SCF   sectional center facility

COMPTROLLER GENERAL'S
REPORT TO THE CONGRESS

SYSTEM FOR MEASURING MAIL
DELIVERY PERFORMANCE--ITS
ACCURACY AND LIMITS
/ United States Postal Service 52

D I G E S T

The Postal Service measures the quality of mail serv-
ice by judging how well it meets standards for first-
class mail delivery. The Service's goal is 95-percent
delivery of (1) local mail in 1 day, (2) regional mail
in 2 days, and (3) cross-country mail in 3 days.

The Postmaster General's annual report for fiscal
year 1974 states that the Service is meeting the
95-percent goal for overnight delivery of local first-
class mail. However, the Service is averaging only
88 percent in other areas. These statistics are
gathered and analyzed under what is called the
"Origin-Destination Information System."

The Service's statistic for nationwide overnight
delivery performance of first-class mail is reliable,
but it is based only on stamped mail which accounts
for 40 percent of first-class mail. Since reliable
statistics are available only for this type of mail,
there is a question whether other first-class mail
moves as fast as, faster than, or slower than stamped
mail.

There are weaknesses in the Service's Origin-Destination
Information System which make local overnight delivery
performance statistics less reliable than the nationwide
statistics. Here are examples:

--The Information System can be manipulated
at the local level. (See pp. 8 and 11.)

--Samples for local areas may be too small
for reliable biweekly performance
reports. (See p. 9.)

--The system has a flaw resulting from dual
purpose measurement of mail delivery
performance and volume. (See p. 12.)

Tear Sheet. Upon removal, the report
cover date should be noted hereon.

GGD-75-109

i

--Inadequately trained clerks make many data-
recording errors. A well-trained, independ-
ent cadre would be the most effective way of
assuring that tests are properly and accu-
rately taken. (See pp. 14 and 15.)

--Some data on the characteristics of sample
mail is lost during the processing of data
collection forms at Service headquarters.
(See p. 15.)

Because of the large amount of data collected and the
relatively small amounts of erroneously recorded and lost
data, nationwide statistics are not greatly affected.
However, improving data accuracy and decreasing data
loss would improve the reliability of local performance
statistics.

RECOMMENDATIONS

To avoid possible misinterpretation, the Service should
be more specific in its public announcements about what
mail delivery performance statistics represent.

Other recommendations for improving the accuracy and
reliability of local delivery performance statistics
are on page 19.

AGENCY ACTIONS

The Postal Service concurred with GAO's recommend-
ations and has begun corrective action on many of
the problems. (See p. 19.)

CHAPTER 1

INTRODUCTION

The Postal Reorganization Act, which created the U.S. Postal Service, requires that the Service provide prompt, reliable, and efficient services to all customers. To provide quality mail delivery, the Service established the following standards for first-class mail:

--1-day (overnight) delivery within local areas (generally within sectional centers and among adjoining sectional centers).

--2-day delivery within a 600-mile radius.

--3-day delivery to all other areas.

These standards apply only to mail which has the proper address and ZIP code and which is posted by the last mail pickup time-- generally 5 p.m. The Service's goal is to meet these standards 95 percent of the time.

The Service's annual report for fiscal year 1974 stated that 1974 was a very good year for mail service, which improved over that of the previous year. The report pointed out that the Service was meeting its 95-percent standard nationwide for overnight delivery of first-class mail but was averaging only about 88 percent on 2- and 3-day delivery. These delivery performance measures were provided by the Service's Origin-Destination Information System (ODIS).

Although ODIS' statistics indicate that delivery is good, criticism abounds. Constituents continually write their Congressmen, and newspaper articles regularly complain about poor delivery despite higher postage rates. Much of the criticism results from first-class mail being delivered late. Even at a 95-percent delivery performance, as many as 2.6 billion of the 53 billion pieces of first-class mail delivered annually would be late and, therefore, a potential cause of customer dissatisfaction.

ODIS--WHAT IT IS AND WHAT IT MEASURES

ODIS collects, analyzes, and presents mail delivery data in a variety of report formats for use by Service management.

ODIS is best known for measuring mail delivery performance, which is the time between when a letter is postmarked and when it is received at the last delivery unit before it is placed into the addressee's hands. ODIS' nationwide data collection activities started in April 1970.

ODIS collects data for all classes of mail, except second class, on such mail characteristics as physical shape; cities of origin and destination; postmark date; types of indicia (stamp, meter, etc.); and absence, presence, and correctness of ZIP code. Volumes and delivery times are measured between the 570 sectional center facility (SCF) areas, independent cities, and postal concentration centers throughout the country which are known collectively as ODIS areas. (An SCF is generally a mechanized mail handling facility for processing incoming and outgoing mail for a number of peripheral post offices.)

ODIS' sampling techniques can measure characteristics of all the mail of a particular class by examining only a fraction of that mail. Before a sample is drawn, all delivery units in any post office from which the customer gets mail are identified. These units include postal carrier routes, contract carrier routes, box sections, and third- and fourth-class post offices. There are about 500,000 individual delivery units at the 12,000 post offices where ODIS tests are conducted.

The 500,000 delivery units are divided into sample groups on the basis of their type (business, resident, etc.) and size. Sample groups of about 75,000 units and the date on which each unit will be tested are randomly selected every quarter.

On the test date a data collection clerk records the characteristics of the mail being sampled on specially designed optical character machine-readable forms. The number of pieces of mail selected is determined by the estimated mail volume of the delivery unit being tested. For example, if the estimated volume for that day is 2,001 to 4,000 pieces, every 10th piece is recorded. After the forms are completed and edited for neatness, completeness, and accuracy, they are sent to headquarters in Washington, D.C., where the management reports are prepared. ODIS' annual operating costs are estimated at $6 million to $10 million.

MAIL DELIVERY STATISTICS

To understand and appreciate ODIS' mail delivery statistics, it is necessary to become familiar with the composition of first-class mail and the percentage of that mail included in the statistics. First, the Service's announcements on first-class mail delivery include only stamped mail which, for the year ending October 11, 1974, accounted for 40 percent of all first-class mail, as shown below.

Indicia	Estimated annual volume (billions of pieces)	Percent
Stamped	21	40
Metered	23	43
Permit, government, and other (note a)	9	17
Total	53	100

a/ Includes mail with unknown indicia.

No delivery performance statistics are possible for permit and government mail because these are not postmarked and, therefore, do not bear the mailing date needed to determine delivery time. Further, although the Service maintains delivery statistics for metered mail, it considers them less reliable than those for stamped mail because the postmark date is affixed by the mailer before the mail is delivered to the Service. Thus, a delay may occur between when a mailer postmarks a letter and when he deposits it in the post office.

Second, most first-class mail orginates and is ultimately received locally. However, when ODIS' statistics show that the Service is consistently meeting its 95-percent standard for overnight delivery of first-class mail nationwide, only about 41 percent of stamped first-class mail--or about 8.7 billion pieces of the total volume of 21 billion--is included in these statistics.

Third, the combined volume of first-class mail in ODIS' overnight, 2-day-area, and 3-day-area statistics represents only 33 percent of all first-class mail.

- - - -

We primarily reviewed mail-processing operations and ODIS procedures related to stamped first-class mail within

overnight delivery areas because the majority of all first-class mail is generated and delivered within these areas. Our observations concerning the reliability of overnight performance statistics also generally apply to 2- and 3-day statistics since ODIS' data collection and reporting procedures are the same for all first-class mail.

CHAPTER 2

NATIONWIDE OVERNIGHT DELIVERY PERFORMANCE STATISTICS

FOR STAMPED FIRST-CLASS MAIL ARE RELIABLE

Our evaluation of the reliability of ODIS' nation-wide overnight delivery performance statistics for stamped first-class mail showed that:

--Overnight delivery areas are tailored to meet a 95-percent performance.

--Tests conducted by an independent task force support the Service's claim that nationwide overnight delivery standards are being met 95 percent of the time.

--Sample sizes are more than sufficient to reliably measure nationwide delivery performance.

--ODIS can be manipulated at the local level.

As previously stated, only stamped first-class mail is included in ODIS' overnight delivery performance statistics. Therefore, the question remains whether other first-class mail moves as fast as, faster than, or slower than stamped mail.

OVERNIGHT DELIVERY AREAS ARE TAILORED
TO MEET A 95-PERCENT PERFORMANCE

The geographic areas committed for overnight delivery are determined by each SCF. Generally, a commitment is made to only those areas where SCF postal management expects to meet its goal 95 percent of the time. Overnight delivery areas are continually being expanded and now include over 50 percent of first-class mail volume.

Although there are locations where the 95-percent standard is not consistently met because of transportation or mail processing problems, mail processing is normally geared to the overnight delivery of local mail. Stamped first-class mail destined for delivery overnight and mailed by 5 p.m. is collected and canceled with an a.m.

or p.m. postmark of the date on which it is mailed. Mail collected after 5 p.m. receives a minus-p.m. cancellation. The minus-p.m. cancellation is recorded on an ODIS test as not qualifying for overnight delivery and that mail is not included in the Service's overnight delivery performance statistics. About 20 percent of all overnight area mail receives a minus-p.m. postmark and is included in the Service's 2-day-area delivery statistics.

Within the overnight delivery area, qualified stamped mail is generally given priority over nonqualified mail. After sorting, mail is dispatched by truck to the various post offices in the overnight area in time for delivery on that day. Our observations of mail processing confirmed that the 95-percent standard is realistic in these circumstances.

TASK FORCE TESTS SUPPORT ODIS STATISTICS

To evaluate the accuracy of ODIS' 95-percent delivery statistics for overnight areas, a task force of GAO and Postal Inspection Service auditors conducted independent tests at randomly selected locations throughout the continental United States during the 2 weeks ended October 4, 1974. The results of the task force's tests supported the Service's 95-percent overnight delivery achievement claims on a nationwide basis.

In conducting the tests, the task force:

1. Limited the universe to the 100 largest ODIS areas, which process over 60 percent of all first-class mail.

2. Tested a random sample of about 400 delivery units --selected by the Service--1 week after the units were tested by ODIS clerks.

3. Limited advance notice to the delivery units selected to 24 hours or less, when possible.

4. Used ODIS procedures and data collection forms.

5. Analyzed the results and developed conclusions on the basis of statistical inference.

The following table compares task force and Service test results.

	Postal Service	Task force
Estimated volume based on test results (note a)	691,199	1,038,397
Estimated volume delivered within standard	665,004	964,877
Estimated volume not delivered within standard	26,195	73,520
Estimated percent delivered within standard	96.2	92.9
Sampling error (95-percent confidence level) (note b)	±1.1%	±2.3%

a/ Of the 400 delivery units selected for the tests, many
   did not have any qualified stamped first-class mail for
   overnight delivery and were dropped from the analysis.
   Certain other units were discontinued and some tests were
   missed.

b/ The range around the estimated percentage within which the
   "true" delivery percentage would be expected to fall 95 per-
   cent of the time.

The 3.3-percent difference between the two performance
scores--96.2 and 92.9--is statistically significant, but
neither score is inconsistent with the Service's nationwide
performance level of 95 percent. The difference between the
Service's test results and the task force's could be due to
the fact that the task force tests were taken 1 week later,
when mail volume was 50-percent greater.

SAMPLE FOR NATIONWIDE STATISTICS
IS MORE THAN ADEQUATE

A sample as small as about 500 delivery units may pro-
vide a reasonable estimate 1/ of nationwide delivery performance
for a 1-year period. The more frequent the reporting re-
quirement, however, the larger the sample should be, as
shown in the following table.

_____

1/ Assuming a 95-percent confidence level and a sampling error
   of ± 1 percent.

7

Report frequency	Estimated sample size
Annual	500
Semiannual	1,000
Quarterly	2,000
Biweekly	13,000

The Service currently samples about 286,000 delivery units annually to obtain data on local ODIS performance. We believe that this sample size is more than sufficient for reliable nationwide performance statistics.

CAVEAT ON MANIPULATION

Generally, mail processing personnel both move the mail and measure their own proficiency in doing so. ODIS' reliance on these personnel coupled with the pressure from postal management for high scores leaves ODIS vulnerable to manipulation. In fact, ODIS can be manipulated at all points. Test scores can be influenced by (1) controlling the mail available to be sampled and the information recorded on the data sheets and (2) making changes on data sheets during editing.

Test results for each of the 570 ODIS areas are disseminated to the regions, districts, and individual offices. The results are used at each level (1) to judge the performance of each lower level--the region judges district performance and so on down to the SCF, which judges the performance of its mail-processing personnel--and (2) as a basis for taking action against supervisory personnel. For example, one postmaster-SCF manager said he was denied a quality pay increase because his overnight ODIS scores averaged less than 95 percent for the previous year. The district manager at another location said he received considerable pressure from the regional postmaster general when ODIS scores were low. He said he used ODIS in turn as a disciplinary tool in his district by threatening SCF management officials with replacement if their ODIS scores did not improve.

As we point out in the following chapter, manipulation to achieve higher ODIS scores was a problem at 1 of the 10 SCFs we visited. This problem should not be widespread since overnight delivery areas are generally tailored to mail-processing capabilities. However, it may occur at other offices where pressure for high ODIS scores is great and mail-processing resources are inadequate to handle peak mail volumes.

CHAPTER 3

PROBLEMS OBSERVED MAY AFFECT

LOCAL DELIVERY PERFORMANCE STATISTICS

The following weaknesses in ODIS make local overnight delivery performance statistics less reliable than nationwide statistics:

--Samples for determining local overnight delivery performance may be too small.

--Advance notice of ODIS tests can affect their representativeness.

--Late mail was removed from ODIS tests at one major post office to improve ODIS test results.

--ODIS has a flaw resulting from dual purposes being served.

--Inadequately trained personnel make many data-recording errors.

--Some data on the characteristics of sample mail is lost during the processing of data collection forms at Service headquarters.

Because of the large amount of data collected and the relatively small amounts of erroneously recorded and lost data, nationwide statistics are not greatly affected. However, improving data accuracy and decreasing data loss would improve the reliability of local statistics.

The Postal Inspection Service also made some of these same observations.

SAMPLES FOR DETERMINING LOCAL STATISTICS MAY BE TOO SMALL

Because of the small number of delivery units and letters sampled, biweekly statistics for individual ODIS areas may not be sufficiently reliable. Clerks in the largest areas conduct biweekly tests at as many as 34 delivery units, but those in the smallest areas conduct as few as 4. About 45

percent of all ODIS areas have 15 or fewer biweekly ODIS tests. The number of letters actually counted at a delivery unit is also very small--over 60 percent of the delivery units in the task force's tests had 15 or fewer letters.

The combination of a small number of delivery units and letters would, for the units involved, hold out the possibility of large sampling errors. 1/ That is, the actual aggregate characteristics of all letters could vary considerably from the characteristics of those letters included in the sample. To determine the extent of sampling errors, we obtained test results on eight ODIS areas for a 1-month period. As shown below, the sampling errors varied widely among the areas.

ODIS area	Delivery units in sample	Performance (percentage)	Sampling error at 95-percent confidence
New York, N.Y.	88	95.5	+ 2.7%
Lehigh Valley, Pa.	47	99.6	+ 0.4%
Philadelphia, Pa.	118	99.1	+ 0.4%
Houston, Tex.	83	99.0	+ 0.8%
Greenville, S.C.	33	90.9	+ 4.3%
Gary, Ind.	41	96.2	+ 1.7%
South Bend, Ind.	36	96.9	+ 2.1%
Portland, Oreg.	60	96.2	+ 1.7%

The eight areas above cannot be considered typical because, for the most part, the number of delivery units selected within each area for a biweekly period was larger than the national average. Therefore, the sampling errors shown are likely to be smaller than those which would be found in a typical ODIS area biweekly report.

Because of technical problems, we could not estimate sampling errors for biweekly reports. Nevertheless, we believe the biweekly sampling errors would probably be even larger than those on a monthly basis.

ADVANCE NOTICE AFFECTS
REPRESENTATIVENESS OF ODIS TESTS

Notice of ODIS test dates and units to be tested is given to mail-processing personnel at most SCFs well in advance.

1/ The range around the estimated percentage within which the "true" delivery percentage would be expected to fall 95 percent of the time.

Advance notice affords an opportunity to expedite mail proc-
essed for delivery units scheduled for testing. At two SCFs
mail-processing officials readily admitted that this notice
enabled them to divert extra staffpower to clean up delayed mail
before the test.

About a month before the beginning of each postal quarter,
headquarters sends a printout to every participating post office
showing the delivery units to be sampled and the dates of the
sampling. Advance notice is then given to various officials
within each post office.

Advance notice is necessary to insure that (1) all mail
is available for sampling and (2) the data collection clerk
takes the test. It is generally not necessary, in our opinion,
to notify mail-processing personnel of an impending ODIS test
more than a day in advance. Notice of tests was given a week
or more in advance at most of the offices we visited.

At the Houston SCF, mail-processing and quality control
officials said they were "beefing up" staffpower at units to be
tested to avoid possible repercussions from management about
low test scores. This special attention insures that no mail
is left unprocessed in those units scheduled for a test and
has a positive influence on overall test results to the extent
that the unit is not representative of mail conditions in other
units. At the Cleveland SCF, mail-processing officials said
that advance notice enabled them to give mail-processing pri-
ority to such units, if necessary, for several days before the
test.

MANIPULATION OF MAIL AT ONE MAJOR POST OFFICE

At the Detroit SCF late mail was deliberately removed
from delivery units before an ODIS test to improve overnight
delivery performance statistics. Before the completion of
our review, however, Detroit postal officials began giving
priority to processing overnight stamped first-class mail
to meet delivery standards and we found no further indication
that late mail was being removed prior to ODIS tests.

Our observations of mail processing at Detroit, during
the night before certain ODIS tests, indicated that special
emphasis was being given to the mail at the delivery units to
be tested the following morning. Consequently, we designed a
test of mail processing operations to determine if late mail
was being removed from a delivery unit before an ODIS test.

On 5 separate days, letters which were canceled several days in advance were planted with other mail addressed to the delivery units to be tested. Mail for these delivery units was sorted at the Detroit SCF and sent to another office where the ODIS tests were conducted. On 4 of the days, most of our late letters were removed from the delivery units before the ODIS tests. On 3 occasions we located our missing late letters along with as many as 10 other late letters for the same delivery unit back at the Detroit SCF.

Three Detroit mail-processing foremen said they examined mail destined for delivery units to be ODIS tested and removed late letters. They then reinserted these letters in the mail-processing operation at a point when they could not be delivered in time for the ODIS test. One foreman said he held out as many as 300 late letters in a single evening. Although these foremen said they had been instructed to withhold late mail, their superiors denied knowledge of their activities.

Similar tests were conducted at the other nine offices we reviewed, but none of their late letters were tampered with. Nevertheless, while the manipulation found was isolated, the entire system is vulnerable to it.

ODIS HAS A FLAW RESULTING FROM
SERVING DUAL PURPOSES

ODIS was designed to measure mail delivery performance and volume. However, an accurate measure of both is precluded by ODIS' data collection procedures for recording misthrown mail.

Misthrown mail--mail at the right office but wrong delivery unit--is recorded as "delivered" during a test. The presumption is that mail is misthrown in equal amounts among all delivery units and will be sorted to the proper unit for delivery that same day. This presumption is not entirely correct, however, because misthrown mail is not always delivered the same day. Also, offices cannot agree on how to treat such letters and headquarters has no written policy.

Recording misthrown mail remaining at a delivery unit from the previous day overstates mail volume for that day and unit. It does, however, measure delivery performance more accurately because the majority of this mail missed delivery the previous day and is, therefore, late mail. The Service should determine whether an accurate volume measure or an

accurate delivery performance measure is more important at this level and revise its instructions accordingly.

In preparing for an ODIS test, all mail received for a unit after the cutoff time of the previous day must be included in a test. For example, if a carrier's unit is to be tested on April 9 and the carrier's departure time is 8 a.m., all mail received between 8 a.m. on April 8th and 8 a.m. on April 9th is subject to sampling. All mail received during this period is recorded as delivered on April 9 on the presumption that the carrier takes all mail with him when he leaves to make his rounds.

As a carrier sorts the mail at his unit in the order in which he intends to deliver it, he sets aside misthrown letters. These letters are then re-sorted to the proper carrier. If the misthrows are not all sorted to the proper carrier before he leaves, the ODIS test results will overstate delivery performance since this mail will be recorded as delivered when, in fact, a 1-day delay has occured.

Except in Houston, misthrown mail was not a major problem at the offices visited. We visited eight offices in Houston, all of which were receiving large amounts of misthrown mail from the Houston SCF, which does the sorting. Because of reduced staffpower, only two of the eight offices were able to re-sort the misthrows for delivery the same day.

In the event of a test, Houston's policy was to exclude misthrows remaining at a carrier's unit from the previous day because otherwise volume would be overstated. Excluding these misthrows, however, resulted in an overstatement of delivery performance since most misthrows were late mail.

To illustrate the effect of including the previous day's misthrows, we randomly selected 11 carrier routes from 4 Houston offices and calculated the overnight delivery performance rate for each route. As shown in the following table, including misthrows lowered overnight delivery performance an average of 5.9 percent.

Unit and route	Performance rate excluding misthrows	Performance rate including misthrows	Difference between rates
Julius Melcher:			
Route 2743	100.0	97.2	2.8
Route 2751	99.0	96.4	2.6
Route 2724	95.5	93.2	2.3
Route 2723	88.4	85.4 ·	3.0
East Houston:			
Route 2802	98.2	96.6	1.6
Sharpstown:			
Route 3622	85.8	77.8	8.0
Route 3626	90.6	73.2	17.4
Route 3636	91.2	80.8	10.4
Route 3635	90.8	78.8	12.0
Sam Houston:			
Route 240	94.2	91.3	2.9
Route 254	94.0	93.4	.6
Total	93.3	87.4	5.9

Most offices we visited occasionally had some misthrown mail which did not receive same-day delivery. Officials at these offices were about evenly split on the question of including a prior day's misthrows in an ODIS test.

MANY DATA-RECORDING ERRORS
MADE BY DATA COLLECTION CLERKS

All delivery units at which ODIS tests are conducted are in the 12,000 largest post offices. The largest of these, usually the SCFs, have the most delivery units and, consequently, may have several ODIS tests each workday. The smaller offices conduct very few tests.

At the SCFs the data collection clerks are generally more familiar with ODIS testing instructions because they conduct many tests. The clerks at the smaller offices, however, are less knowledgeable and make many recording errors because they conduct tests infrequently.

At the Los Angeles Post Office a cadre of clerks took all ODIS tests, whether in the SCF or another post office within the ODIS area. We believe that a well-trained cadre would be the most effective way of assuring that ODIS tests are properly and accurately taken.

Training of data collection clerks

Eight of the ODIS areas we visited formally instructed all clerks at the SCFs. This was often coupled with on-the-job training. One SCF provided on-the-job training only and one gave no training--clerks were required to familiarize themselves with ODIS.

In three ODIS areas, however, no training was given to the clerks in other offices. We believe that continuing training, especially for clerks at offices outside the SCFs, is necessary to maintain the integrity of local ODIS test results.

The Houston SCF has developed an effective technique for assuring that data collection clerks at its offices who conduct infrequent ODIS tests are familiar with ODIS instructions. A questionnaire on ODIS is sent to each office. Only those clerks who accurately complete the questionnaire can conduct an ODIS test. A questionnaire of this type could be used periodically to identify those clerks needing additional training.

NEED TO ANALYZE CAUSES OF DATA LOSS DURING PROCESSING

ODIS data collection forms are scanned at Service headquarters by an optical character reader (OCR) which transfers the coded data to computer tapes for final report preparation. Some data, however, is not accepted by the OCR. Because of the small samples at many ODIS areas, losing certain types of data could affect local performance scores.

Although a categorization of the reasons for data loss from each form is available on a special error list, the Service has not analyzed these lists to identify ways to minimize the losses. A Service official said the Service did not have sufficient staff time to perform such an analysis.

Each data collection form is divided into two basic information areas--test identification and mail characteristics of up to eight pieces of mail. (See app. II.) Recorded mail characteristics include postmark date, origin, class and type, indicia, and ZIP code.

Clerks use special lead pencils and a writing font recognizable to an OCR to complete the forms. The forms are then edited for neatness, completeness, and accuracy and forwarded to Washington, D.C., for processing and report preparation.

Arriving in Washington at a rate of about 20,000 a day, the forms are batched and fed through the OCR which either rejects or accepts them. If the forms are accepted, the characteristics of the sample mail are transferred to computer data tapes for further processing into the final report formats.

Data forms are rejected only if their test identification data is incorrect or unreadable to the OCR. Rejected forms are reintroduced after the incorrect or unreadable items have been erased and rewritten by clerks at head-quarters.

Whether the mail characteristics on each form are accepted by the OCR depends on whether they are complete and readable. However, forms are not rejected for unacceptable mail characteristics. If the data on the form is not accepted by the OCR, the data is earmarked by a special error code and excluded from the file for report purposes.

After all forms for each batch are fed through the OCR, an error list is printed out showing the total number of records, forms, errors, and forms with errors and the per-centage of forms with errors. The printout also shows a breakdown of types of errors, indicated by alphabetic codes.

For the first quarter of fiscal year 1975, the error list showed that an average of 8 percent of the 1.4 million data forms processed were rejected by the OCR because of unreadable test identification data, erased and corrected by clerks, and rerun through the OCR. In addition, about 16 percent of the forms had mail characteristics data which was not accepted by the OCR and, therefore, was lost. The Service has not evaluated error rates or codes to determine causes of data loss or ways to minimize them.

OBSERVATIONS BY POSTAL INSPECTION SERVICE

At the request of the Postmaster General, the Inspection Service visited eight post offices to determine, among other things, the integrity of ODIS data and related reports. In

its September 1974 report, the Inspection Service concluded that "the procedure used in conducting ODIS tests at most offices wherein units to be tested are notified in advance * * * tend[s] to weaken the integrity of the ODIS system."

The report further noted that some management decisions made to improve ODIS test results were detrimental to mail service. That is, achieving a high ODIS score took priority over moving mail which did not affect the test results.

CHAPTER 4

CONCLUSIONS AND RECOMMENDATIONS

CONCLUSIONS

The Postal Service claims that it is achieving 95-percent delivery of first-class mail in overnight delivery areas on a nationwide basis. This statistic, however, is based only on stamped first-class mail.

A reliable delivery time performance measure is only available for first-class stamped mail. Therefore, the question remains whether other first-class mail moves as fast as, faster than, or slower than stamped mail.

Our review supports the Service's claim that it is achieving 95-percent overnight delivery of stamped first-class mail. ODIS' measurement of this delivery, however, can be manipulated by those who have a vested interest in the mail delivery performance it measures and has other weaknesses which could distort local delivery performance.

ODIS' sample sizes are more than adequate to reliably measure nationwide delivery performance for stamped first-class mail but may be too small for reliable biweekly results for each local ODIS area.

Ideally, little or no advance notice of tests should be given to mail-processing personnel, and tests should be conducted by a group independent of the mail-processing function. We believe that a well-trained cadre would be the most effective way of assuring that ODIS tests are properly and accurately taken. In those areas where a cadre is impractical, continuing training should be provided to data collection clerks to minimize recording errors.

The accuracy of ODIS performance statistics could be improved, particularly for local areas, by providing for uniform treatment of a prior day's misthrown mail in a test. Also, analyzing the causes of OCR data losses may help find ways to reduce such losses.

RECOMMENDATIONS

We recommend that, to avoid possible misinterpretation of the mail delivery performance statistics, the Postmaster General have the Service be more specific in its public announcements about what the statistics represent.

We also recommend that the Postmaster General improve the accuracy and reliability of local delivery performance statistics by

--not giving routine advance notice of tests to mail-processing personnel,

--determining whether an accurate measure of delivery performance or volume is more important at the local level and issuing instructions accordingly,

--exploring the feasibility of having an independent cadre conduct tests in each ODIS area,

--providing initial and continuing training to all data collection clerks, and

--evaluating OCR error lists to determine the causes of data loss and ways to minimize such loss.

AGENCY ACTIONS

The Postmaster General concurred with our recommendations and has initiated corrective actions in response to many of the recommendations. (See app. I.)

Specifically, the Service

--agreed that its public announcements about what mail delivery performance statistics represent should be as clear and specific as possible, and will try to make them so; .

--will reemphasize its policy of limiting notice of ODIS tests to those personnel who need to be informed in order to conduct the tests properly;

--is reviewing the uses made of ODIS information, the degree of accuracy and reliability required, and the relative priority of these various uses;

--is actively exploring the extension of the cadre system to additional offices;

--has developed an audio-visual program to provide better
  training for ODIS data collection clerks; and

--is initiating a study to evaluate the OCR's error
  message printouts to determine the types of data
  lost and possible corrective action to minimize
  such losses.

CHAPTER 5

SCOPE OF REVIEW

Our review had two phases. During phase I we spent 1 month in each of the following major post offices:

1. Continental Facilities SCF, King of Prussia, Pennsylvania
2. Philadelphia SCF, Philadelphia, Pennsylvania
3. Detroit SCF, Detroit, Michigan
4. Cleveland SCF, Cleveland, Ohio
5. North Suburban SCF, Chicago, Illinois
6. Rockford SCF, Rockford, Illinois
7. San Antonio SCF, San Antonio, Texas
8. Houston SCF, Houston, Texas
9. Oakland SCF, Oakland, California
10. Los Angeles Post Office, Los Angeles, California

During this phase, we primarily reviewed the processing of stamped first-class mail for overnight delivery areas. Our work included:

1. Counting all qualified first-class mail at delivery units on the day before, the day of, and the day after selected ODIS tests and taking other steps to ascertain whether delayed mail was being extracted during processing.

2. Interviewing employees to determine whether postal management exerted excessive pressure to achieve better ODIS results.

3. Reviewing cancellation die changing procedures to assure that mail was properly postmarked.

4. Reviewing transportation delays to determine if they affected ODIS scores.

5. Reviewing staff-hours expended in mail-processing operations to determine if extra staffpower was used to process mail destined for delivery units to be tested.

6. Searching for delayed mail at various mail-processing locations.

21

ODIS' testing procedures were evaluated to determine whether:

1. All mail was tested and the data was properly recorded.

2. Data collection forms were properly edited.

3. Data collection clerks knew the proper procedures.

4. Tests were made on schedule.

5. Delivery unit lists were current.

During phase II, in conjunction with the Postal Inspection Service, we conducted tests at about 400 delivery units throughout 35 States and the District of Columbia. (See app. III.) The tests duplicated ODIS tests conducted by the Service 1 week previously and were randomly selected from all tests conducted during a 2-week period. Our test results were compared with those of the Service to determine the reliability of ODIS' overnight performance statistics for stamped first-class mail.

The data from our tests was recorded on ODIS data collection forms and processed on the Service's OCR; final results were tabulated using the Service's programs and computer. Data printouts were obtained at the completion of each processing step to determine whether the integrity of the data was being maintained and to evaluate the causes of any data losses.

THE POSTMASTER GENERAL
Washington, DC 20260

August 11, 1975

Mr. Victor L. Lowe
Director, General
  Government Division
U. S. General Accounting Office
Washington, D. C. 20548

Dear Mr. Lowe:

Thank you for the opportunity to comment on your proposed report on
the Postal Service's system for measuring mail delivery performance.

We are pleased with the report's principal findings:

(1)   The Postal Service's overnight delivery performance
       statistics can be believed.

(2)   Although the system has some weaknesses affecting local
       statistics, the relatively small amounts of erroneously
       recorded or lost data involved have no significant effect
       on national statistics because the total amount of data
       recorded nationally is so large.

The report concentrates on overnight delivery performance on first-class
stamped mail, but we agree with your observation that these findings
would also hold true for our national statistics on second and third day
delivery, since they are derived from the same system as our overnight
statistics and the amounts of data collected nationally are also very large.

As the report notes, the Origin Destination Information System (ODIS)
which produces our performance statistics on stamped mail, based on
cancellation time, cannot produce the same kind of data on metered,
permit or government mail, since such mail is not cancelled. However,
we strongly believe that such mail does move at least as fast as stamped
mail, since it by-passes the cancellation process and much of it is pre-
sorted.
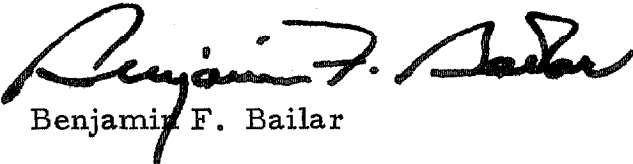
We agree that our public announcements about what our mail delivery performance statistics represent should be as clear and specific as possible, and we will certainly try to make them so.

As to the report's other recommendations:

(1)   We will re-emphasize our policy of limiting notice of ODIS tests to those personnel who need to be informed in order to conduct the tests properly. It is not possible to restrict all mail processing people from advance notice since some of them must provide information, and sometimes manpower, necessary to conduct the tests.

(2)   We are reviewing the uses made of ODIS information, the degree of accuracy and reliability required, and the relative priority of these various uses. After we have completed this study, we will be in a better position to act on your recommendation about determining whether performance or volume accuracy is more important in regard to counting misthrown mail. As the report indicates, the amount of data involved is not significant in terms of national performance statistics, even though it can affect local statistics.

(3)   We are actively exploring the extension of the cadre system to additional offices.

(4)   We have developed an audio-visual program to provide better training for ODIS data collection clerks.

(5)   We are initiating a study which will evaluate the optical character reader's error message printouts to determine the types of data lost and possible corrective action to minimize such losses.

We greatly appreciate the thoroughness of your study and the resources and technical expertise your organization brought to bear in accomplishing it. We have benefited from our technical discussions with your staff and we feel sure the Congress and the public will benefit from your report through an improved understanding of how the quality of mail service is measured and the soundness of our performance statistics.

Sincerely,

Benjamin F. Bailar

BEST DOCUMENT AVAILABLE

THIS FORM HAS SPACE
FOR RECORDING DATA
FOR EIGHT PIECES OF
MAIL.

INSTRUCTIONS:
USE A NO. 2 PENCIL.
DO NOT FOLD, STAPLE, TEAR OR SMUDGE.
DO NOT WRITE IN BLUE AREA.
READ INSTRUCTIONS BEFORE COMPLETING FORM.

MAKE YOUR NUMBERS
AND LETTERS LIKE THIS   1234567890X

TEST
IDENTIFICATION NUMBER

DATE OF TEST
MONTH   DAY

SKIP INTERVAL

PAGE NUMBER

START
(WITH "01")

POSTMASTER'S USE ONLY

A. POSTMARK DATE — MONTH — DAY — CANNOT BE READ
B. POSTMARK OF ORIGIN — LOCAL — NON LOCAL — FOREIGN — CANNOT BE READ
C. MAIL CLASS — D. MAIL TYPE
IF NON-LOCAL
3 DIGIT ZIP CODE
IF ZIP UNKNOWN
POST OFFICE AND STATE (IF FOREIGN, SHOW COUNTRY)

FIRST CLASS — AIR MAIL — PRIORITY MAIL — PARCEL POST — OTHER FOURTH — THIRD CLASS — LETTER — CARD — SPR — FLAT — PARCEL — CAT. AUG

E. INDICIA — STAMP | METER | PERMIT | GOV'T
F. ZIP IN MAILING ADDRESS — CORRECT — INCORRECT — NO ZIP
IF INCORRECT FILL IN ZIP SHOWN IN ADDRESS

USE ONE COLUMN FOR EACH PIECE

PS FORM 1300
Feb. 1973

ORIGIN-DESTINATION INFORMATION SYSTEM

☆GPO 1973—496. 619

25