# AN EFFICIENT PROCEDURE FOR AUDIT OF ACCOUNTS RECEIVABLE

By

W. Edwards Deming
Consultant in Statistical Surveys
Washington

and

T. Nelson Grice
Ernst & Ernst
Winston-Salem

Reprinted from
Management Accounting
March 1970

# An Efficient Procedure for Audit of Accounts Receivable

ADDITIONAL GAINS, OVER AND ABOVE THOSE ATTRIBUTABLE TO THE MORE CONVENTIONAL APPROACHES

By W. Edwards Deming and
T. Nelson Grice

The procedure described here illustrates how stratified sampling and ratio-estimates may increase the efficiency of a sampling procedure under appropriate conditions.

"It was not practicable to confirm accounts receivable by direct communication with the customer, wherefore we employed other auditing procedures." This statement or similar wording appears in the accountants' reports of a substantial majority of Class I motor carriers in the United States. The statement poses three questions: Why was it not practicable? What were the other possible auditing procedures? How effective were they?

This article explores these questions and describes a statistically designed audit. The statistical procedures used here are in textbooks and require only a reference. What is new here is the application. The purpose of the paper is thus exposition of a particular application.

## The Sampling Procedure

Any statistical method has the advantage of providing valid estimates of: (1) the proportions of accounts for which the charges recorded by the motor carrier and the charges computed by the shipper were different, (2) the number of dollars involved in these differences, and (3) the average difference per dollar in favor of the shipper, and of the carrier.

The sampling procedure described here with a ratio-estimate for the proportion of accounts in error turned out to be, for this characteristic, five and one half times as efficient as simple random sampling. Put another way, the sample of 221 accounts, drawn and processed by the method prescribed here, for an estimate of the proportion of accounts in error, was equivalent to a sample of 1,220 accounts drawn by simple random sampling and merely averaged. Moreover, for an estimate of the corrected number of dollars, it was equivalent to 3600 accounts drawn by simple random sampling and raised by the inverse of the probability of selection.

Requests for confirmation were mailed to the total sample of 221 shippers (customers, debtors of record), with a second mailing to nonresponses. The response rate was 75 percent.

## Why Was Direct Confirmation Not Practicable?

Motor carriers have two types of accounts receivable: (1) customer and (2) interline, due from other motor carriers. This study was limited to accounts receivable from customers.

W. EDWARDS DEMING

*is a consulting statistician in Washington whose practice covers wide application of statistical techniques. His work in statistical control of quality in Japan is recognized by the annual Deming Prize and by the award Second Order Medal of the Sacred Treasure from the Emperor in 1960. Dr. Deming is also Professor of Statistics at the Graduate School of Business Administration, New York University.*

T. NELSON GRICE, JR.

*is an auditor-manager in the Winston-Salem office of Ernst & Ernst. He is a graduate of Davidson College and holds an M.B.A. degree from Rutgers University. Mr. Grice is a member of the American Institute of C.P.A.'s and of the National Accounting and Finance Council of the American Trucking Association.*

There has been considerable discussion in recent years about the effectiveness of confirmation of accounts receivable. Most government agencies do not reply to a positive confirmation, and a large number of corporations are today using the voucher system. Increased emphasis has accordingly been placed on negative confirmation and alternative procedures.

The problem of effective confirmation is accentuated in the motor carrier industry and transportation generally. Motor carriers generate a high volume of individual freight bills that are not maintained by customer and which are insignificant in individual amounts. The volume of freight bills at any given date precludes adequate positive confirmation on a judgment basis, and the use of negative confirmation is of doubtful validity.

## What Other Auditing Procedures Were Possible?

Under regulations of the Interstate Commerce Commission, a motor carrier is required to send a bill for freight charges within seven days from the date of delivery, and to collect the charges within seven days from the date of the statement, not counting Saturdays, Sundays, and holidays. Consequently, the rate of turnover of accounts is high.

Review of accounts uncollected at the end of the following month is therefore one of the prime procedures. Seventy percent of accounts receivable will be less than 30 days old.

Use of various tests of activity within the month is a second effective audit procedure. Comparisons of aging and a review of the older categories are the principal remaining procedures.

## Advantages of the Statistical Method

There is greater need to assess the quality and reliability of accounts receivable from customers of motor carriers than in most other accounts because of the volatile nature of the records and the high rate of error in the bills.

The auditing procedures briefly stated in the preceding paragraphs provide evidence of the degree of internal control present in the system and a measure of the effectiveness of collection, but they provide only a limited measure of the accuracy of the accounts receivable from customers.

The degree to which errors are present in the accounts receivable from customers are of vital concern to the auditor. His ability to appraise the effect of these errors on the financial statements by use of judgment samples is severely limited. The selection and testing of a judgment sample of several hundred freight bills would reveal numerous errors both favorable and unfavorable. The results of such a test would not permit the auditor to draw quantitative conclusions covering the number of errors in the field, or the dollar-effect of such errors on the balance. The advantage of the statistical method lies in its ability to produce valid quantitative estimates, with calculable margins of uncertainty from a sample-size that is not too burdensome.

## Brief Description of the Frame

The example presented here came from the accounts receivable from customers of one of the largest domestic motor carriers. The company uses an 8-part prenumbered freight bill of which two copies (delivery receipt and auditor's copy) are punched cards. Each of the 65 terminals in the system has a separate numbering sequence, the first two digits of which denote the terminal of origin. Freight bills are issued to the terminals in blocks, and they are put onto magnetic tape sequentially at the home office as issued, providing a control over future accountability at the terminal.

The delivery receipt, which is signed by the consignee, constitutes proof of delivery. Delivery receipts from all terminals are filed in the main office by serial number.

The delivery receipt and the auditor's copy, plus the office copy, are of primary value for audit. The office copy is the basis for the entry on magnetic tape. The only information recorded on the tape are the serial number, date, weight and charges. The office copies are filed by serial number. This system of billing and filing is characteristic of the industry.

The auditor's copy clears the customer receivable when collection is made. The date of deposit is punched in this copy, and a reference file is maintained by serial number. Files of remittance memoranda by bank account and date of deposit are also maintained.

The company does not make from the tape-record periodic prints of accounts receivable. Because of the geographically scattered operation of the company, the close of the year is not completed until the fourth week of the following month. Hence the file of auditor's copies of the freight bills at the end of the month prior to the end of the fiscal year was selected as the frame for circulation.

## Content of the Inquiry

There were three main questions directed to the customer for every freight bill in the sample, drawn from the frame:

1. Is the bill paid?
2. Was there an error in this bill (whether paid or not)?
3. If yes, what is the magnitude of the error?

An answer to Questions 1 and 2 is yes or no, coded 0 and 1 respectively. An answer to Question 3 is a number of dollars, which we designate as $x_{jk}$ for Bill k in Stratum j. Of course, if there was no error, $x_{jk} = 0$.

The delivery receipt was obtained (the office copy was used where the delivery receipt could not be found) for every bill in the sample, and a copy made. The address of the debtor was verified and a letter requesting confirmation was prepared and mailed with a copy of the delivery receipt enclosed.

An additional question, directed to the carrier, is whether the delivery receipt for a selected shipment is in place in the file.

## Design of the Sample

Allocation of the sample was an attempt to maximize the accuracy obtainable for the allowable effort to be expended. Allocation was based on the distribution of accounts receivable at the end of the preceding fiscal year. This was not just a matter of convenience to gain time in the preparation of the procedure of selection. The distribution of accounts by dollar-bracket does not vary much from year to year. Moreover, as theory shows, small or moderate departures from optimum allocation cause only negligible loss in efficiency.[1]

Of the 221 confirmations mailed, replies came from 113 prior to the date set for mailing out second requests. There were 53 additional replies to second requests. The total number of replies received was 166 or approximately 75 percent of the original number sent out.

The percentage of replies was a pleasant surprise (based on past experience). Several factors appeared to be responsible:

1. Most important was the confirmation of a single freight bill in contrast with a statement, and attachment of a photocopy of the signed delivery receipt.

[1] W. Edwards Deming, *Sample Design in Business Research*, John Wiley, New York, 1960, p. 295.

Table 1

CALCULATED AND ADOPTED VALUES OF THE ZONING INTERVALS

| Class or stratum j | Class intervals | Estimate of the dollars receivable a year ago $N_j x_j$ | $n_j$ Calculated by Eq. 5 | $N_j/n_j$ Calculated by Eq. 6 | $Z_j$ Zoning interval adopted | $n_j$ Expected this year per subsample* |
|---|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 1 | Under $10 | 460,000 | 1.4 | 65,000 | 32,000 | 2.3 |
| 2 | $ 10 to 24.99 | 575,000 | 1.8 | 22,000 | 16,000 | 2.1 |
| 3 | 25 to 49.99 | 660,000 | 2.0 | 8,750 | 8,000 | 1.8 |
| 4 | 50 to 99.99 | 760,000 | 2.4 | 4,360 | 4,000 | 2.2 |
| 5 | 100 to 199.99 | 830,000 | 2.5 | 2,180 | 2,000 | 2.3 |
| 6 | 200 to 399.99 | 1,330,000 | 4.0 | 1,100 | 1,000 | 3.8 |
| 7 | 400 to 799.99 | 1,450,000 | 4.5 | 550 | 500 | 4.0 |
| 8 | 800 to 1,499.99 | 480,000 | 1.5 | 275 | 200 | 1.1 |
| 9 | 1,500 or over | | 10 | | | 24 |
| Total | | 6,545,000 | | | | |

* Calculated as $n_j = N_j/Z_j$, by use of the $N_j$ in Table 2.

2. Careful determination of the correct address for every account in the sample produced a minimum of undelivered returns and necessity to re-address the enquiry.
3. The explanation in the accompanying letter that the request was vital to a statistical survey appears to have generated interest and compliance.

Alternative procedures on non-replies consisted of examination of signed delivery receipts and, in most cases, examination of memoranda of remittance. In addition, all freight bills in the sample were verified by company personnel. These people were not informed of errors reported by customers. Four such errors were not detected during this verification. Two more were alleged but differed from the customer's report, and the customer was correct—illustrative of the complexity of the rate and commodity structure.

The rule adopted here was that the sample in any stratum should be roughly proportionate to the dollars receivable in that stratum, with two exceptions: (a) large accounts would be in the sample with certainty, and (b) the sample of smallest accounts would be about double the size calculated by proportionality.[2]

A frequency distribution constructed a year ago gave counts of the number of bills receivable in the size-classes shown in Table 1. What appears as Strata 7 and 8 in Table 1 appeared as one stratum in the count made a year ago. No recommendation is implied here that 9 is the best number of strata. However, we should point out that there would be little economy in using fewer strata, when one works with a tape. There is certainly no point in using more than 9 strata.

We required a total sample of about 200, to be replicated in 10 sub-samples. Each subsample gave a valid estimate for the entire frame. Replication facilitates calculation of estimates of standard errors.[3]

The dollar amounts $B_j$ in column 3 were not counted a year ago, so for allocation of the sample approximations were made by setting $B_j = N_j \bar{x}_j$ where $N_j$ is the number of bills in Stratum j a year ago and $\bar{x}_j$ is the mid-point of Stratum j (or, for large strata, something less). We introduce a factor g of proportionality and write

$$n_j = g B_j = g N_j \bar{x}_j \qquad (1)$$

Summation gives

$$n = g \Sigma N_i \bar{x}_i = 6{,}545{,}000\, g \qquad (2)$$

As each subsample would contain about 20 accounts,

$$g = 20/6{,}545{,}000 = 30.6 \times 10^{-7} \qquad (3)$$

$$1/g = 327{,}000 \qquad (4)$$

Hence, the number $n_j$ of accounts expected in each subsample is

$$n_j = g N_j \bar{x}_j = N_j \bar{x}_j / 327{,}000 \qquad (5)$$

The zoning interval in Stratum j may then be calculated as

$$Z_j = N_j/n_j = 327{,}000/\bar{x}_j \qquad (6)$$

[2] Hansen, Hurwitz, and Madow, *Sample Survey Methods and Theory*, John Wiley, New York, 1953, Chapter 5.

[3] W. Edwards Deming, *Op. cit.*, Chapter 11.

Table 2

ACCOUNTS RECEIVABLE ACCORDING TO THE TAPE

| Class or stratum | | Recorded in the tape: Amount receivable | Recorded in the tape: $N_j$, number of accounts |
|---|---|---|---|
| 1 | 2 | 3 | 4 |
| 1 | Under $10 | 501,454.43 | 73,520 |
| 2 | $ 10 to 24.99 | 532,813.18 | 34,273 |
| 3 | 25 to 49.99 | 518,888.06 | 14,766 |
| 4 | 50 to 99.99 | 615,465.89 | 8,822 |
| 5 | 100 to 199.99 | 650,088.20 | 4,676 |
| 6 | 200 to 399.99 | 1,110,683.38 | 3,838 |
| 7 | 400 to 799.99 | 1,021,106.80 | 1,982 |
| 8 | 800 to 1,499.99 | 214,117.62 | 224 |
| 9 | 1,500 or over | 48,643.75 | 24 |
| | Total | 5,213,261.31 | 142,125 |

Columns 4 and 5 in Table 1 show $n_j$ and the zoning interval $Z_j$ calculated by the above equations. Column 6 shows the zoning interval adopted, and Column 7 shows the number $n_j$ of accounts expected in each subsample, calculated later with the zoning interval adopted and with the actual count of the numbers $N_j$ in the tape for the present year (Table 2). Table 2 was produced after the close of the present year.

The next step was to provide 10 random starts for each zone (except for Stratum 9, which was 100% and a member of every subsample). The bill whose serial number contains the random start for Subsample 1 in Stratum j and every bill thereafter that contains a dollar formed by repeated addition of the adopted zoning interval $Z_j$ to the random start also belonged to Subsample 1 in that stratum. Similarly for Subsamples 2, 3, and onward to 10. It was possible to instruct the machine to serialize the accounts in each stratum and to print those designated for the sample.

## Estimates Provided by the Audit

The results of the sample were used in the formulas that follow to obtain estimates of what the results would have been had the sample of accounts receivable been 100 percent of the frame (tape). For example,

$$X = 73520\, x_1/n_1 + 34273\, x_2/n_2 + \ldots + 24\, x_9/24 \qquad (7)$$

Where the numbers 73520, 34723, etc., are the actual counts $N_j$ in the frame (Table 2) and $x_j$ is the number of accounts reported defective among the $n_j$ accounts in the sample in Stratum j. Table 2 was produced at the close of the year. The procedure of selection (Columns 1-6 of Table 1) was frozen before the end of the year.

The above formula gave an estimate of the number of accounts that would be reported defective in either direction (against the carrier, in favor of the carrier, or either way) were the whole tape studied with the same care as the accounts in the sample. The proportion of accounts defective in the whole frame would then be estimated as

$$p = X/N \qquad (8)$$

where N = 142, 125, the number of accounts in the whole frame (Table 2 or Table 3).

To estimate the average error per dollar in the tape (e.g., in both directions) we form the ratio

$$r = \frac{73520\, x_1/n_1 + 34273\, x_2/n_2 + \ldots + 24\, x_9/24}{73520\, y_1/n_1 + 34273\, y_2/n_2 + \ldots + 24\, x_9/24} \qquad (9)$$

The symbol $y_j$ in the denominator is the number of dollars receivable as recorded in the tape for the $n_j$ bills in the sample drawn from Stratum j. Thus, the fraction $y_1/n_1$ is the average number of dollars per account receivable, as recorded in the tape, in the $n_1$ accounts drawn from Stratum 1.

The numerator of Eq. 9 is the estimate given by Eq. 7 for the total error in dollars in both directions in the whole tape. The denominator is an estimate of the number of dollars receivable as recorded in the whole tape. (This figure is of course known from Table 2, but we use an estimate thereof in Eq. 9 to capture the efficiency of the ratio-estimator shown as Eq. 10 for the total error in dollars in the whole tape.)

The total error in dollars in the whole tape would then be estimated as

$$X = 5{,}213{,}261\ r \qquad (10)$$

the figure 5,213,261 being the total number of dollars claimed in the whole tape (Table 2).

The same formulas (Eqs. 9 and 10 in combination) will provide an estimate of the corrected dollars in the frame if $x_j$ be re-defined as the corrected dollars in the $n_j$ bills in Stratum j (see Table 5).

The estimates in Tables 3, 4 and 5 were calculated for the whole sample and for each subsample. The accepted result comes from the whole sample. The only reason to calculate results by subsample is to facilitate estimation of standard errors.

## Replication for Estimation of Standard Errors

An estimate of the variance of any estimate x may be calculated as

$$\hat{\mathrm{Var}}\, x = \frac{1}{10 \times 9} \sum_{1}^{10} (x_i - x)^2 \qquad (11)$$

where x is the estimate derived from the full sample, and $x_i$ is the estimate derived from Subsample i.

A simpler and satisfactory estimate of the standard error comes from the range of the 10 results by use of the formula[4]

$$\hat{\sigma}_x = \frac{1}{10}\,[\max - \min] \qquad (12)$$

For example, Eq. 11 gives

$$\mathrm{Var}\, p = \frac{1}{90}\ .2989 = .003321 \qquad (11a)$$

hence,

$$\hat{\sigma}_p = \sqrt{.003321} = .058 \qquad (11b)$$

For comparison, Eq. 12 gives

$$\hat{\sigma}_p = \frac{1}{10}\,(.556 - .010) = .055 \qquad (12a)$$

in good agreement.

[4]Nathan Mantel, "On a rapid estimation of standard errors for the means of small samples," *American Statistician*, Vol. 5, October 1951, pp. 26-27; M. H. Quenouille, *Rapid Statistical Calculations*, Hafner, 1959, pp. 5-7.

Table 3

ESTIMATES BY SUBSAMPLE OF THE ACCOUNTS THAT WOULD SHOW VARIANCE WITH REPLIES FROM SHIPPERS

| Subsample i | $X_i$, estimated number of accounts in error in either direction (Eq. 7) | $N_i$, number of accounts in frame (tape) | $P_i = X_i/N_i$ proportion of accounts in error (Eq. 8) |
|---|---|---|---|
| 1 | 2 | 3 | 4 |
| 1 | 3,354 | 142,125 | .024 |
| 2 | 59,281 | 142,125 | .417 |
| 3 | 43,089 | 142,125 | .303 |
| 4 | 78,982 | 142,125 | .556 |
| 5 | 24,497 | 142,125 | .173 |
| 6 | 19,190 | 142,125 | .135 |
| 7 | 1,469 | 142,125 | .010 |
| 8 | 14,990 | 142,125 | .106 |
| 9 | 55,561 | 142,125 | .391 |
| 10 | 44,162 | 142,125 | .311 |
| Whole sample | 34,437 | 142,125 | .242 |
| Standard error | 7,751 | 0 | .058 |

## Gains from Stratification

It may be interesting to compare this result with the over-worked formula $\sqrt{pq/n}$ for the standard error of p, applicable if and only if the sample of accounts had been drawn by random numbers without stratification, and the straight average used as an estimate. Under such a plan, the estimate of the standard error would be

$$\sqrt{pq/n} = \sqrt{.242\ (1 - .242)/10} = \sqrt{.01834} = .135 \qquad (13)$$

Table 4

ESTIMATES BY SUBSAMPLE OF THE AVERAGE ERROR AND OF THE TOTAL ERROR

| Subsample | Numerator of Eq. 9 (000) | Denominator of Eq. 9 (000) | Ratio Proportionate error r | Error in the whole tape, in either direction, by Eq. 10, in thousands of dollars |
|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 |
| 1 | 70 | 4865 | .0144 | 75 |
| 2 | 268 | 5093 | .0527 | 275 |
| 3 | 375 | 5366 | .0699 | 364 |
| 4 | 195 | 4980 | .0391 | 204 |
| 5 | 581 | 5175 | .1122 | 585 |
| 6 | 196 | 5005 | .0391 | 204 |
| 7 | 86 | 5536 | .0155 | 81 |
| 8 | 185 | 5211 | .0356 | 185 |
| 9 | 446 | 5468 | .0815 | 425 |
| 10 | 469 | 4637 | .1010 | 527 |
| Whole sample | 260 | 5102 | .0510 | 266 |
| Standard error | | | .0067 | 57 |

Table 5

ESTIMATES BY SUBSAMPLE OF THE CORRECTED AMOUNTS

| Subsample | Numerator of Eq. 9 (000) | Denominator of Eq. 9 (000) | Ratio r by Eq. 9 | Estimate of dollars actually receivable, Eq. 10 (000) |
|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 |
| 1 | 4839 | 4865 | .9947 | 5185 |
| 2 | 5183 | 5093 | 1.0177 | 5306 |
| 3 | 5393 | 5366 | 1.0049 | 5237 |
| 4 | 5015 | 4980 | 1.0070 | 5250 |
| 5 | 4607 | 5175 | .8910 | 4641 |
| 6 | 5010 | 5005 | 1.0011 | 5219 |
| 7 | 5582 | 5536 | 1.0083 | 5257 |
| 8 | 5028 | 5211 | .9650 | 5031 |
| 9 | 5481 | 5469 | 1.0030 | 5225 |
| 10 | 4205 | 4637 | .9070 | 4727 |
| Whole sample | 5018 | 5099 | .9840 | 5130 |
| Standard error | 132 | xxx | .013 | 75 |

Table 6

ESTIMATES FOR THE FRAME OF ACCOUNTS RECEIVABLE

| Item | Estimate | Standard error |
|---|---|---|
| Proportion of accounts in error in either direction | .24 | .06 |
| Defective dollars, either direction | 266,000 | 57,000 |
| Corrected dollars receivable | 5,130,000 | 75,000 |
| Dollars receivable in tape | 5,213,261 | 0 |

Thus, the stratified plan with Eq. 8 as the formula for estimation of the proportion of accounts in error has an efficiency estimated as .01834/.003321 = 5.52 compared with unstratified sampling, with the straight average used as an estimate. That is, the sample of 221 accounts described here is the equivalent of a sample of 5.52 x 221 or 1220 accounts selected and processed as a simple random sample.

## Gains from the Ratio Estimate

We pause to observe also the gain in efficiency from use of the ratio-estimator in Eq. 10 for the corrected dollars in the frame. The result that it gave, viz., 5130 (in thousands of dollars), and its standard error, 75, are shown in Col. 5 of Table 5. Eq. 7, an unbiased estimator of the same numerical characteristic of the frame, gave the figure 5018 with standard error 132, as shown in Col. 2 of Table 5. The efficiency of the ratio-estimator in Eq. 10 over the unbiased estimator in Eq. 7 is thus $(132/75)^2$, or about 3. This gain is over and above the gain from stratification already possessed by Eq. 7.

In contrast, a judgment sample does not furnish a basis for any objective statement about the margin of uncertainty in a result, nor any basis to diminish or enlarge the sample next time.

## Results

The results are in Tables 3 to 6. The estimate of the proportion of accounts with error in either direction is 24 percent with a standard error of 6 percent (from Table 3 or Table 6). The estimate of the number of defective dollars in the frame is 266,000 (from Table 4 or Table 6). The corrected total amount receivable in the tape is $5,130,068 (from Table 5), against $5,213,261 claimed by summation of the tape (see Table 2). The difference is 1.6 percent, not significant, as the relative standard error is 1.3 percent (Col. 4 of Table 5).

The auditor may assume rationally that if he had increased the size of sample to 100 percent, maintaining the same standard of workmanship, were such a thing possible, the result of the sample would not differ in either direction from the estimates provided by the sample by more than 2.3 standard errors. The statistician's risk in this interpretation is about 1 experience in 20.

The standard errors were estimated by

*Continued on page 34*

formulas that are appropriate to the sampling procedure. The standard error includes the effect of accidental errors of judgment in rating freight-bills, random variations in the interpretation of illegible entries and other small independent accidental variations, including the use of sampling. All of these uncertainties, except for the use of sampling, would be present were the size of the sample increased from the one described here to 100 percent.

One must always remember that some large error in accounts receivable may escape detection by any sample less than 100 percent.