BUREAU OF THE CENSUS
STATISTICAL RESEARCH DIVISION REPORT SERIES
SRD Research Report Number: Census/SRD/RR-85/18

A COMPARISON OF SEVEN IMPUTATION PROCEDURES FOR ISDP

by

Vicki J. Huggins
Statistical Research Division
Bureau of the Census

Recommended by: Paul Biemer

Report completed: November, 1985

Report issued: November, 1985

I. INTRODUCTION

Missing data for longitudinal surveys occur in a variety of patterns which can be sorted and categorized into different classes of missingness depending on the survey unit. For this study, the survey unit is a person. Therefore the missingness that occurs in the data can be person nonresponse, whereby no data is available for a person at any given time period in the survey, record-type nonresponse where an entire module of related data is unavailable, and item nonresponse in which data is missing sporadically throughout the person record. For this study we focused on record-type nonresponse for a single continuous variable. It is important that these types of nonresponse to be addressed as they occur generously throughout a longitudinal survey. Also, simulation of record-type nonresponse provides reasonably sized data files to study and manipulate. It is important to note that the techniques investigated can be employed to compensate for both item and record-type nonresponse.

The objective of this study is to evaluate seven different methods of imputation for continuous data in a longitudinal survey. The methods compared are described below as are the procedures to compare them. In our comparisons, we employed a variety of summary statistics and graphic techniques. The particular findings are detailed in the body of the text and a number of graphs and tables are included in the Appendix to support these findings. No information was observed to support any assumptions of normality in the data studied, and the analysis proceeds using a variety of nonparametric techniques.

In Section II we describe the data used in this study and discuss how it was used. In Section III we discuss each of the alternative imputation strategies that are compared against one another. In Section IV the methods used to compare the different procedures are described and the results of our analysis are presented. Findings are summarized in Sections V and VI, and an Appendix contains the tables, graphs, and summary statistics used in our analysis.

II. SIMULATING MISSING DATA PATTERNS

Twelve-month longitudinal data extracted from the 1979 ISDP (Income Survey Development Program) survey were used in this study. These data were entered into a SIR (Scientific Information Retrieval) database, from which free-format simulation data files were extracted. Subsequent manipulation and evaluation were performed using special purpose FORTRAN programs and the SPSS-X statistical package on a Univac 1100 and IBM-XT at the Bureau of the Census.

For this study, missing data were simulated using records on which the variable of interest was completely reported, and for technical reasons records with zero responses for the variable of interest were excluded. We then had the original values for the misssingness that was simulated in the file to use later in analyzing the properties of imputations obtained by the selected imputation methods. The continuous variable used in the study is wages and salary. The following indicates the simulation procedure used to induce missing data on records.

(1) Define a longitudinal record for wages and salary to be a person record of responses to the question: What were your wages and salaries for month (j), j=1, 12 in 1979?

		J	F	M	A	M	J	J	A	S	O	N	D
Ex:	Rec 1	100	100	150	145	120	200	150	200	100	100	150	175
	Rec 2	10	10	10	10	50	50	50	50	50	50	50	50

(2) Randomly select 500 person records for persons, age $\geq$ 16, with at least one missing response, i.e., month (j) = -1 for some j, and at least one complete response, i.e., month (j) > 0 for some j. (The value "-1" is a place holder for a missing response.)

	J	F	M	A	M	J	J	A	S	O	N	D
Ex:	100	100	-1	-1	-1	100	100	100	150	150	150	150

(3) Select approximately 2,000 person records with complete responses for every month (j), i.e., month (j) >0 for all j=1, 12.

(4) Induce the missing pattern from a record in the set (2) onto a record for a full respondent in set (3) by a nearest match procedure. That is, let $X_{n,j}$=month (j) for some case n from data set (2) and let $Y_{i,j}$ = month (j) for some case i from data set (3), and find the record Y_i in the set (3) to minimize:

$$\sum_{j=1}^{12} (X(n,j) - Y(i,j))^2.$$

We set X(n,j)-Y(i,j) = 0 for X(n,j) missing.

One then induces the n^{th} missing data pattern from (2) onto the i^{th} full respondent in (3) to obtain 500 simulated person records with missing wave responses. In all, 410 unique complete respondents were used in simulating the 500 records with induced missing responses.

III. SEVEN IMPUTATION PROCEDURES

The seven imputation procedures examined in this study are described below. The first three employ regression type techniques which utilize the entire data set to (1) model the missingness that occurs in the entire set of data and (2) derive model-based imputes for the misssing values. The last four procedures implement averaging techniques in which only data for the current case is used in determining an impute for a missing month's value. The regression-based imputation procedures: Iterated Buck, Logarithmic Iterated Buck, and Cube Iterated Buck; and the four averaging techniques: Arithmetic Smoothing (1) and (2) and Multiplicative Smoothing (1) and (2); were tested and evaluated on the simulated data set described above.

(a) Iterated Buck Techniques

The Iterated Buck procedure is a sequential regression technique that estimates regression parameters, derives imputes based on these parameters, and repeats this process until the sequence of estimated parameters converge. For a detailed description and derivation of the Iterated Buck method the reader is referred to papers by S. F. Buck, [2], and Beale and Little, [1], pertaining to missing values in multivariate analysis. The important thing to note here is that Iterated Buck is an EM-Algorithm that gives maximum likelihood estimates of the population parameters when there is the assumption that the data has a multivariate normal distribution.

However, no distributional assumption of normality of the data is justified here, as indicated in Figures 1-4. Histograms of the residuals for Iterated Buck, Logarithmic Iterated Buck and Cube Iterated Buck are presented with a normal overlay represented by the dotted line on the histograms. Comparing the two distributions in each of the histograms suggests that a normality assumption for the data is unjustified. Even in the absence of normality the Iterated Bucks method can be used to derive imputations. Of course, since the data is not normal, our analysis will proceed along nonparametric lines, and considerations especially appropriate to normal data will not be addressed.

We now describe the steps involved in the Iterated Buck procedures. Assume for a set of N observations and n variables that x_{ij} represents the value of the j^{th} variable in the i^{th} observation for $j=1,\ldots,n$ and $i=1,\ldots,N$. Let m_j denote the sample mean value of the j^{th} variable over all complete observations and u_{jk} denote the sample covariance between variables m_j and m_k over all complete observations. The Iterated Buck method uses m_j and u_{jk} to compute:

$$(1) \quad x_{ij} = \begin{cases} x_{ij}, & \text{if } x_{ij} \text{ is observed} \\ \text{a linear combination of the set of variables observed in the } i^{th} & \\ \text{observation,} & \text{otherwise} \end{cases}$$

$$(2) \quad c_{ijk} = \begin{cases} \text{partial covariance of } m_j \text{ and } m_k \text{ if } x_{ij} \text{ and } x_{ik} \text{ are both unknown} \\ 0, \quad \text{otherwise} \end{cases}$$

$$(3) \quad \bar{x}_j = \sum_{i=1}^{N} x_{ij}/N \ ,$$

$$(4) \quad a_{jk} = \sum_{i=1}^{N} (x_{ij} - \bar{x}_j)(x_{ik} - \bar{x}_k) + c_{ijk}.$$

Set $m_j = \bar{x}_j$ and $u_{jk} = a_{jk}/(N-1)$ and repeat (1) thru (4) until there are no further changes in m_j and u_{jk}. The term c_{ijk} is a correction term for the bias that would normally occur in the formation of a_{jk}. The procedure is applied to a longitudinal record for the variable <u>wages and salaries</u> by setting x_{ij} = AMT(i,j) for person record i, i=1,N and month j, j=1,12.

The Logarithmic Iterated Buck is the same algorithm as just described, the only difference is that x_{ij} = log(AMT(i,j)) for the i^{th} person record and j^{th} month. (This is the reason we omitted records containing zero responses.) After the algorithm is satisfied,

x_{ij} is transformed back to original amounts and corresponding imputes. By using the logarithm of amounts of wages and salaries one reduces the impact of skewness in the data and avoids the problem of generating negative imputes. Similarly, Cube Interated Buck operates on $x_{ij} = (\text{AMT }(i,j))^{1/3}$ until closeness criteria are met. The x_{ij} are transformed back to original values and corresponding imputes.

(b) Smoothing Procedures

The two averaging techniques examined here are termed Arithmetic Smoothing and Multiplicative Smoothing because the imputes are based on the arithmetic mean and geometric mean respectively.

Arithmetic Smoothing essentially allocates an equal additive subdivision to each missing value which depends on the length of the interval of missing values in the data record and the reported values on either side of the missing data. For example, suppose March, April, and May values were missing for a particular record, denoted by x_m, then the record looks like the following:

$$\begin{array}{cccccccccccc} J & F & M & A & M & J & J & A & S & O & N & D \\ x_1 & x_2 & \underbrace{x_{m,3} & x_{m,4} & x_{m,5}}_{\text{missing interval}} & x_6 & x_7 & x_8 & x_9 & x_{10} & x_{11} & x_{12} \end{array}$$

We determine the difference in the bounding reported values of the missing interval and divide by the number of subintervals to arrive at

$$d = \frac{x_6 - x_2}{4} \; .$$

We then add d to x_2 consecutively to obtain imputes for $x_{m,3}$, $x_{m,4}$ and $x_{m,5}$. Explicitly,

$$\begin{aligned} x_{m,3} &= x_2 + d \\ x_{m,4} &= x_2 + 2d \\ x_{m,5} &= x_2 + 3d \; . \end{aligned}$$

For the general case, let $\underline{r} = (x_1,\ldots,x_{12})$ be a logitudinal record of amounts. Suppose x_m is a missing response bound below by x_i and above by x_j.

(1) Compute $k = j-i$

(2) Compute $d = (x_j - x_i)/k$

Then (3) $x_m = x_i + (m-i)d$.

Note that $x_j = x_i + k \cdot d$.

One difficulty with this method is that bounds may not exist around missing responses, specifically, when endpoints (month (1) and/or month (12)) of the record are missing. Two solutions to this problem are examined. The first solution is to substitute the arithmetic mean of the record's complete responses, $(\sum_{i=1}^{p} x_i)/p$, where p is the number of reported responses, into the endpoints whenever one or both endpoints of the record is missing. The second solution is to substitute the arithemtic mean of the two nearest values for missing endpoints. Numerical comparisons of both methods are included with all other results at the end of this report.

Multiplicative Smoothing abides basically by the same principles as Arithmetic Smoothing with the difference that the geometric mean of a missing interval's bounding responses is employed, and equal multiplicative subdivisions are allocated to missing values in an interval of missing responses. That is, for Multiplicative Smoothing we determine the quotient of the bounding reported values of the missing interval and base our imputation on that value. For the general case let $\underline{r} = (x_1,\ldots,x_{12})$ be a longitudinal record of amounts and let x_m denote a missing response bound below by x_i and above by x_j.

(1) Compute $k = j-i$

(2) Compute $q = (x_j/x_i)^{1/k}$

Then (3) $x_m = x_i \cdot q^{(m-i)}$.

Note that $x_j = x_i \cdot q^k$.

The two methods used to correct for missing endpoints on a record corresponding to the situation for Arithmetic Smoothing were, (1) use the geometric mean of the record's complete responses, $(\prod_{i=1}^{p} x_i)^{1/p}$, and (2) use the geometric mean of the nearest two values for any missing endpoints.

It should be noted that Multiplicative Smoothing of amounts of wages and salaries and Arithmetic Smoothing of the logarithm of amounts of wages and salaries give identical results. The following shows the relationship between the two procedures.

The basis for Multiplicative Smoothing is that for some missing interval of length k bounded below by x_a and above by x_b, and with x_m missing in that interal, $(a \leq m \leq b)$,

$$(1) \quad x_m = x_a \cdot q^{(m-a)} \text{ where } q = (x_b / x_a)^{1/k} .$$

Taking the logarithm of (1) gives

$$(2) \quad \log x_m = \log x_a + (m-a) \log q \ .$$

and by setting $y_a = \log x_a$ and $y_m = \log x_m$ we get

$$(3) \quad \log q = \frac{\log x_m - \log x_a}{(m-a)}$$

$$= \frac{y_m - y_a}{(m-a)} .$$

Letting log q equal d and substituting into (2) we obtain

$$(4) \quad y_m = y_a + (m-a)d$$

which is the basis for Arithmetic Smoothing as discussed above.

IV. COMPARING THE PROCEDURES

There are several questions to be addressed when analzying the effectiveness and efficiency of an imputation procedure, and by focusing on these questions particular imputation procedures can be identified that maximize the desired end results. The final decision as to which imputation strategy is best to use for particular survey items must rest with subject-matter specialists who are familiar with the subject-matter of the survey, the questionnaire form, and the underlying target population. In this report, we present a number of descriptive statistics for each of the procedures described above. These can be compared against one another and serve as a basis for an informed decision as to which procedure is to be preferred. In general, the questions that must be addressed are:

(1) What does a completely reported data record look like? Is it typically reported consistently, erratically, in particular patterns, or does it follow some distribution?

(2) What are the imputations expected to accomplish? Should the derived imputation resemble the reported data, implement a presumed relationship, or smooth over the missingness?

(3) What criteria should be used to evaluate and compare methods?

The data for wages and salary are at times reported consistently across a 12-month period, reported erratically other times, and may or may not follow a particular pattern of responses based on ISDP waves the (3-month interval to which a questionairre refers). Ideally, the optimal imputation procedure would adhere to patterns of consistency or erraticism of the reported data for each individual person record.

As discussed in Section II, we start with completely reported longitudinal records and then blank out responses conforming to missing patterns from a set of longitudinal records having nonresponse. We then impute for the induced nonresponse and compare the imputes with the original values that were blanked out. These comparisons form the basis of our analysis. As noted earlier, normality assumptions are not supported by the data, and accordingly, the analysis is nonparametric.

We let

$$\mathbf{x} = (x_1, x_2, \ldots, x_{12})$$

be a completely reported record, and we assume the value for month j was blanked out, and the imputed value is denoted by $\hat{x}_j$. Thus we have the following:

x_j = The amount of wages and salaries for some month j,

$\hat{x}_j$ = Imputed value of x_j for some imputation procedure,

$r_j = x_j/x_{j+1}$, and

$\hat{r}_j = x_j/x_{j+1}$ where at least one of x_j or x_{j+1} was imputed.

The analytical variables computed and evaluated for each imputation method are

(1) $c_j = x_j - \hat{x}_j$

(2) $c_j = (x_j - \hat{x}_j)/x_j$

(3) $c_j = r_j - \hat{r}_j$

(4) $c_j = (r_j - \hat{r}_j)/r_j$.

Note that:

(a) $x_j - \hat{x}_j$ represents the difference between original value and imputed value,

(b) $(x_j - \hat{x}_j)/x_j$ represents the relative difference,

(c) $r_j - \hat{r}_j$ represents the difference between the ratio of adjacent months when one was imputed, and

(d) $(r_j - \hat{r}_j)/r_j$ measures the relative difference of these ratios.

The statistics we will use to examine these analytic variables are:

(i) $S_1 = \sum_{i=1}^{n} c_i$,

(ii) $S_2 = \sum_{i=1}^{n} c_i^2$,

(iii) $S_3 = (\sum_{i=1}^{n} c_i) / n$,

(iv) $S_4 = \sum_{i=1}^{n} (c_i - \bar{c})^2 / n$,

where m is the total number of cases for which $c_i \neq 0$ and

$$\bar{c} = (\sum_{i=1}^{m} c_i) / m \ .$$

Table 1 contains numerical comparisons for analytical variable $c_j = x_j - \hat{x}_j$. The seven imputation procedures are listed horizontally and the four derived statistics used for evaluation are listed vertically. If one of the smoothing imputation methods has a (1) appended to its name, the method substitutes the mean of all reported months for missing endpoints of a record; if a (2) is appended to the name of the procedure, the mean of the two nearest reported values was substituted for missing endpoints. Table 2 presents the numerical results for the analytical variable $c_j = (x_j - \hat{x}_j)/x_j$ and is set up identical to Table 1. In both Table 1 and Table 2, there are a total of 3183 cases. Tables 3 and 4 contain, respectively the numerical results for the two analytical varaiables $c_j = r_j - \hat{r}_j$ and $c_j = (r_j - \hat{r}_j)/r_j$. A total of 2820 ratios were used in these calculations.

V. OBSERVED RESULTS OF THE COMPARISONS

(a) Tables 1-4

One initial reason for carrying out this study was to determine whether straight Iterated Buck is a better imputation procedure than its counterparts, Logarithmic Iterated Buck and Cube Interated Buck. For each of the analytic variables, the better a procedure simulates an aspect of missing data, the closer the relevant derived statistic (either S_1, S_2, S_3, or S_4) will approach zero.

The most decisive finding in this study is that for every derived statistic, Logarithmic Iterated Buck outperformed Iterated Buck. Using the Logarithm of wages and salary rather than actual amounts provides a two-fold improvement over the Iterated Buck procedure by eliminating negative imputes and increasing the accuracy of the imputes. Moreover, in every statistic except the first and third on Table 1, Cube Iterated Buck outperformed Iterated Buck. From these observations, it is clear that either Logarithmic Iterated Buck or Cube Iterated Buck is superior to the simple Iterated Buck.

Results comparing Logarithmic Iterated Buck with Cube Iterated Buck are mixed. In Tables 3 and 4 Cube Iterated Buck performs better. Most often in Tables 1 and 2, Logarithmic Iterated Buck does better. All in all, the results are close. One interesting observation is for the statistic

$$\sum_{i=1}^{n} ((x_i - \hat{x}_i)/x_i)^2 .$$

Cube Iterated Buck far out performs all other procedures. That is, Cube Iterated Buck seems to do well for scaled residuals. On the other hand, for the statistic

$$\sum_{i=1}^{n} (x_i - \hat{x}_i)^2$$

Logarithmic Iterated Buck does best of all. For the last two analytical statistics presented in Tables 3 and 4 Cube Iterated Buck outperformed all other imputation procedures for each statistic calculated, with Logarithmic Iterated Buck a fairly close second best.

Arithmetic Smoothing (1) and Multiplicative Smoothing (1) using the mean of a record's reported values for missing endpoints virtually tie in comparison to one another and outperform their counterparts Arithmetic Smoothing (2) and Multiplicative Smoothing (2) the majority of the time. Logarithmic Iterated Buck and Cube Iterated Buck do a little better, all in all, then the smoothing techniques. However, the ease in implement either of the two smoothing techniques may strongly argue in their favor.

(b) Figures 4-11

In Figure 4 we present a histogram of the amounts of reported wages and salarys that fall into the range \$0. to \$5,000. Histograms of values produced by each of the seven imputation procedures appear in Figures 5 through 11.

Histograms of the data completed by Logarithmic Iterated Buck in Figure 6, Cube Iterated Buck in Figure 7, Arithmetic Smoothing (1) in Figure 8, and Multiplicative Smoothing (1) in Figure 9 look very much alike and also appear to reasonably resemble Figure 1. Although histograms of Arithmetic and Multiplicative Smoothing (2) in Figures 10 and 11 look somewhat similar to the true data, there appears to be a slight more grouping of the data than in the reported data.

The data for this study was not edited. However one extremely large value for monthly wage and salary amount was deleted as an obvious edit failure as it caused some problems in obtaining informative graphs of the data. Unbounded histograms were produced but offered very little extra information so were not included here.

(c) Figures 12-18

Figures 12 thru 18 present scatterplots of the amounts of wages and salaries versus each imputation procedure in the same order as the histograms are listed. The more linear the relationship the better the imputation procedure is in simulating the reported data. Ideally, we would like the standard error of the estimate

$$\left(\sum_{i=1}^{m} (x_i - \hat{x}_i)^2 / (n-1) \right)^{1/2}$$

to be small, the intercept near zero and the slope close to one. The correlation and R-square values which measure the relationship between the values and the goodness of fit of the linear model respectively, should approach one for the best method. The standard error of the estimate, intercept, and slope of the linear relationship listed at the bottom of each scatterplot all appear best overall for the Logarithmic Iterated Buck procedure, Figure 13. Iterated Buck gives a negative intercept as a result of negative imputes and the standard error of the estimate is the worst of all the methods. Statistics for Logarithmic and Cube Iterated Buck are very close in comparison to each other, with Logarithmic Iterated Buck just slightly better. Scatterplots of the Arthimetic and Multiplicative Smoothing (1) procedures basically have the same statistics and are both better than Iterated Buck except for the slope statistics. Arithmetic and Multiplicative Smoothing (2) have the worst slope and intercept but the best fit based on the R-squared value.

(d) Figures 19-25

Histograms of scaled residuals, that is, $(x_j - \hat{x}_j)/x_j$, are presented in Figures 19 thru 25. The imputation procedure used to get the estimated impute is listed at the top left of each histogram. Iterated Buck and Log Iterated Buck most often overestimate true values and all four of the smoothing techniques most often underestimate true values. However Cube Iterated Buck underestimates more often than any of the other techniques. This is determined by counting the number of negative scaled residuals in each of Figures 19 thru 25 and compare them to the number of positive scaled residuals. The smoothing techniques tend to spike around zero.

(e) Brief Summary of Observations:

Based on the statistics generated as part of this analysis, the four procedures that appear best are: Logarithmic Iterated Buck, Cube Iterated Buck, Arithmetic Smoothing (1) and Multiplicative Smoothing (1). The residual sum of squares presented in Table 1, Row 2 is a traditionally used comparison criterion, and based on this statistic Logarithmic Iterated Buck is the best procedure. When examining histograms of data completed using each of the imputation procedures to the true data, Cube Iterated Buck, Arithmetic and Multiplicative Smoothing (1) appear almost as good as Logarithmic Iterated Buck. Other statistics provided in Tables 1 thru 4 indicate that each of the four methods are favored by different criteria. The issue is to choose comparison criterion that address specific needs of the data problem at hand. Survey-specific needs should be brought to bear in accessing the merit of each of the procedures discussed. The diverse statistics presented in this report may aid in this analysis.

VI. CONCLUDING REMARKS

Of the imputation procedures examined in this report, Logarithmic Iterated Buck and Cube Iterated Buck outperformed straight Iterated Buck. Of the smoothing techniques, Arithmetic Smoothing (1) and Multiplicative Smoothing (1) outperformed Arithmetic Smoothing (2) and Multiplicative Smoothing (2), respectively. All Iterated Buck procedures must consider a sample of cases with missing values to derive parameters for imputing for nonresponse. Both smoothing techniques need only consider one record at a time and bounding values when deriving an imputation for nonresponse. A variety of summary statistics are presented to assist SIPP specialists in the determination of the most appropriate method for SIPP needs.

In this report we did not add variability to the imputes in the form of a residual. To the extent that thisis a comparative study, we felt adding residuals could be omitted at this stage. Of course, in implementing any one of these procedures, one may add some variablity factors. Variability can be computed from the entire data set and added into each impute or computed on a record by record basis where the variability added to the imputes for each record is based on the record under consideration. An alternate form to adding variability on a record by record basis is to split the data file into two or more groups of records. One group might contain cases that report consistently over time and the other group might contain erratic data reporters. The variability added to each record will be determined by the group in which the record lies.

Acknowledgement

I would like to thank Brian Greenberg for suggesting this research and providing a number of helpful recommendations along the way.

REFERENCES

[1] Beale, E.M.L. and Little, R.J.A. (1975). Missing values in multivariate analysis. J.R. Statistics Society, B. 37, 129-146.

[2] Buck, S.F. (1960). A method of estimation of missing values in multivariate data suitable for use with an electronic computer. J.R. Statistics Society. B. 22, 302-306.

FIGURE 1

HISTOGRAM OF RESIDUALS

```
ITERATED BUCK

  COUNT   MIDPOINT    ONE SYMBOL EQUALS APPROXIMATELY 16.00 OCCURRENCES

      1   -1450.00
      2   -1350.00
      4   -1250.00                                            * = AMOUNT-IMPUTE
     11   -1150.00   :
      5   -1050.00   .
      5    -950.00    .
     11    -850.00   *  .
      8    -750.00   *   .
     90    -650.00   ******.
    191    -550.00   *********:**
    179    -450.00   ***********.
    190    -350.00   ************  .
    241    -250.00   *************** .
    510    -150.00   ******************:*************
    672     -50.00   ******************:***********************
    434      50.00   ****************:**********
    196     150.00   ************   .
    119     250.00   *******      .
     82     350.00   *****      .
     62     450.00   ****    .
     39     550.00   **    .
     28     650.00   **  .
     20     750.00   * .
      7     850.00    .
     16     950.00   :
     12    1050.00   :
      8    1150.00   *
      2    1250.00
      4    1350.00
      4    1450.00
                    I....+....I....+....I....+....I....+....I....+....I
                    0        160       320       480       640       800
                                 HISTOGRAM FREQUENCY

VALID CASES    3182      MISSING CASES     0
```

FIGURE 2

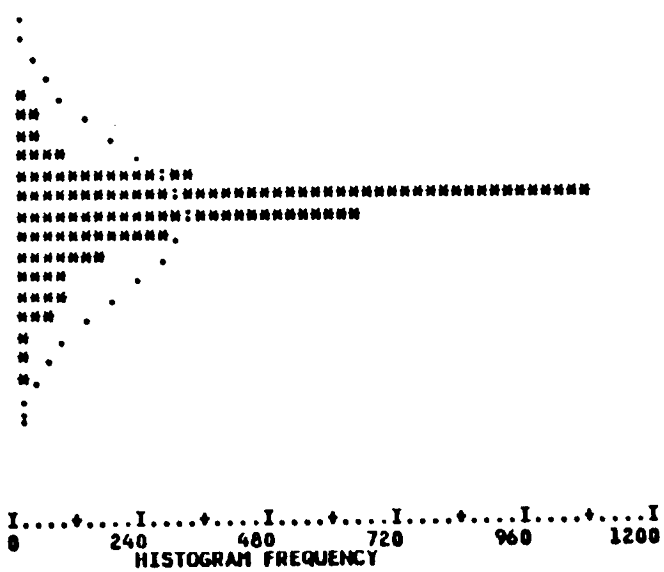
HISTOGRAM OF RESIDUALS

LOG ITERATED BUCK

* = AMOUNT-IMPUTE

ONE SYMBOL EQUALS APPROXIMATELY 24.00 OCCURRENCES

COUNT	MIDPOINT
2	-1450.00
1	-1350.00
1	-1250.00
10	-1150.00
5	-1050.00
2	-950.00
10	-850.00
9	-750.00
10	-650.00
19	-550.00
37	-450.00
55	-350.00
105	-250.00
332	-150.00
1090	-50.00
648	50.00
276	150.00
178	250.00
105	350.00
93	450.00
63	550.00
32	650.00
16	750.00
17	850.00
7	950.00
12	1050.00
7	1150.00
4	1250.00
2	1350.00
5	1450.00



VALID CASES 3182 MISSING CASES 0

FIGURE 3

HISTOGRAM OF RESIDUALS

```
FILE:
CUBE  ITERATED BUCK
     COUNT   MIDPOINT    ONE SYMBOL EQUALS APPROXIMATELY 24.00 OCCURRENCES

         0   -1450.00
         4   -1350.00                                     * = AMOUNT-IMPUTE
         1   -1250.00
        10   -1150.00
         4   -1050.00
         5    -950.00   .
         5    -850.00   .
         6    -750.00    .
        11    -650.00     .
        28    -550.00   *  .
        25    -450.00   *    .
        58    -350.00   **     .
       100    -250.00   ****     .
       246    -150.00   **********.
      1030     -50.00   ************:****************************
       689      50.00   ************:****************
       309     150.00   ************:
       204     250.00   *********  .
       136     350.00   ******   .
        88     450.00   ****   .
        62     550.00   ***  .
        38     650.00   **  .
        23     750.00   * .
        18     850.00   *.
        15     950.00   :
        16    1050.00   :
        11    1150.00
         4    1250.00
         4    1350.00
         1    1450.00
                        I....+....I....+....I....+....I....+....I....+....I
                        0        240       480       720       960      1200
                                      HISTOGRAM FREQUENCY

   VALID CASES     3182       MISSING CASES      0
```

FIGURE 4

HISTOGRAM OF REPORTED AMOUNTS

```
FILE:

AMOUNT
   COUNT   MIDPOINT   ONE SYMBOL EQUALS APPROXIMATELY 12.00 OCCURRENCES

     436      50.00  ************************************
     516     150.00  *******************************************        * = REPORTED AMOUNTS
     501     250.00  ******************************************
     489     350.00  *****************************************
     470     450.00  ***************************************
     415     550.00  ***********************************
     503     650.00  ******************************************
     320     750.00  ***************************
     351     850.00  *****************************
     244     950.00  ********************
     245    1050.00  ********************
     169    1150.00  **************
     263    1250.00  **********************
      93    1350.00  ********
     101    1450.00  ********
     104    1550.00  *********
      78    1650.00  *******
      72    1750.00  ******
      42    1850.00  ****
      91    1950.00  ********
      57    2050.00  *****
      31    2150.00  ***
      33    2250.00  ***
      44    2350.00  ****
      32    2450.00  ***
      35    2550.00  ***
      14    2650.00  *
      20    2750.00  **
      28    2850.00  **
      12    2950.00  *
      26    3050.00  **
      21    3150.00  **
       2    3250.00
       9    3350.00  *
       1    3450.00
       7    3550.00  *
       7    3650.00  *
       7    3750.00  *
       3    3650.00
       6    3950.00  *
      15    4050.00  *
       0    4150.00
      17    4250.00  *
       1    4350.00
       7    4450.00  *
       5    4550.00
       0    4650.00
       5    4750.00
       8    4850.00  *
       8    4950.00  *
                     I....+....I....+....I....+....I....+....I....+....I
                     0        120       240       360       480       600
                              HISTOGRAM FREQUENCY

VALID CASES     5999      MISSING CASES     0
```

FIGURE 5

HISTOGRAM OF DATA COMPLETED BY IMPUTATION

FILE:

ITERATED BUCK

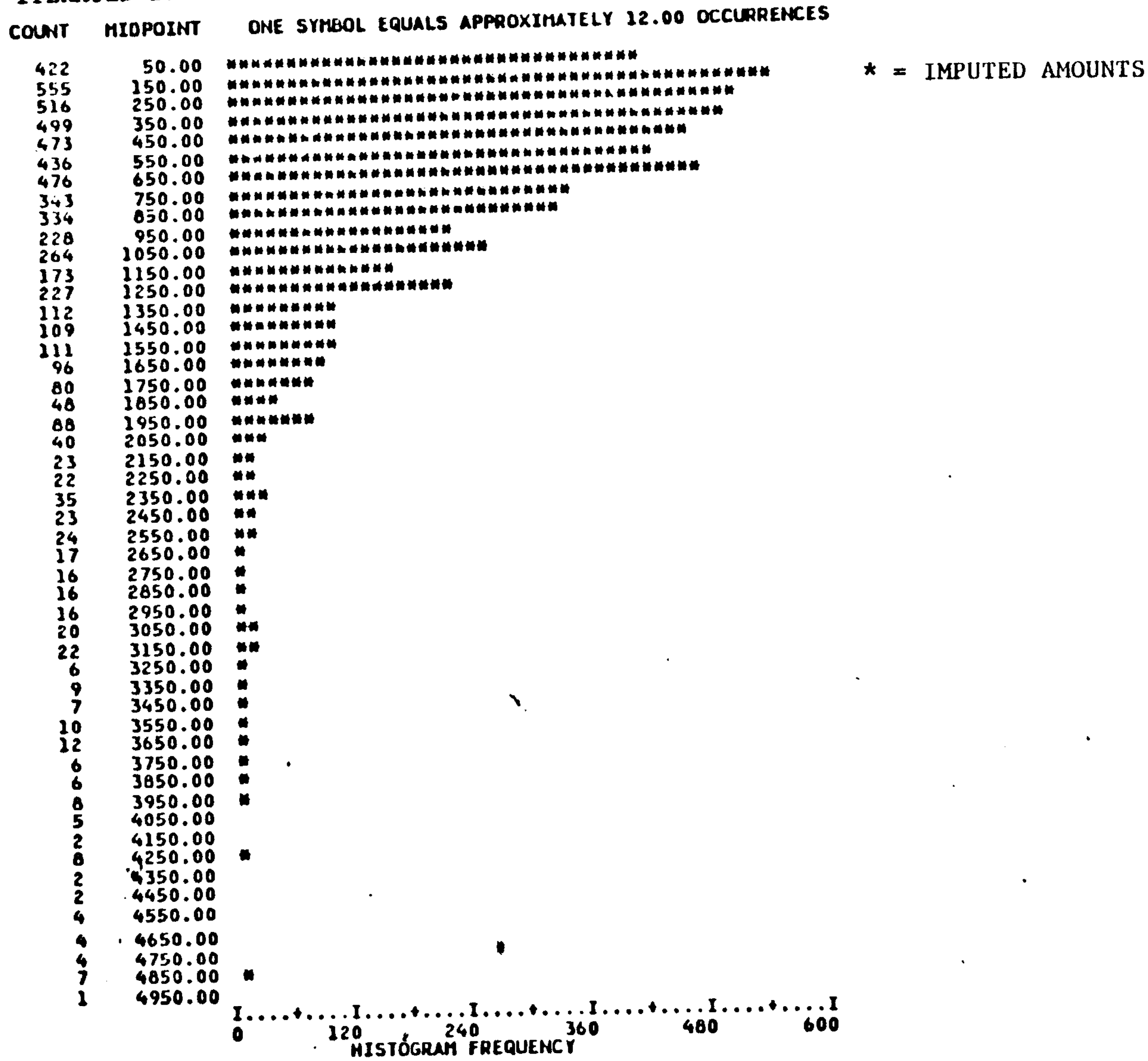
```
COUNT  MIDPOINT   ONE SYMBOL EQUALS APPROXIMATELY 12.00 OCCURRENCES

  198     50.00   *****************
  392    150.00   *********************************           * = IMPUTED AMOUNTS
  444    250.00   *************************************
  480    350.00   ****************************************
  390    450.00   *********************************
  415    550.00   ***********************************
  563    650.00   ***********************************************
  493    750.00   *****************************************
  457    850.00   **************************************
  311    950.00   **************************
  303   1050.00   *************************
  239   1150.00   ********************
  232   1250.00   *******************
  128   1350.00   ***********
  137   1450.00   ***********
  125   1550.00   **********
   83   1650.00   *******
   76   1750.00   ******
   45   1850.00   ****
   62   1950.00   *****
   31   2050.00   ***
   28   2150.00   **
   26   2250.00   **
   40   2350.00   ***
   21   2450.00   **
   20   2550.00   **
   19   2650.00   **
   22   2750.00   **
   16   2850.00   *
    7   2950.00   *
   18   3050.00   **
   20   3150.00   **
    9   3250.00   *
    5   3350.00
    2   3450.00
    9   3550.00   *
    8   3650.00   *
    8   3750.00   *
    3   3850.00
    6   3950.00   *
    7   4050.00   *
    2   4150.00
    9   4250.00   *
    1   4350.00
    6   4450.00   *
    6   4550.00   *
    0   4650.00
    7   4750.00   *
    7   4850.00   *
    1   4950.00
                 I....+....I....+....I....+....I....+....I....+....I
                 0        120       240       360       480       600
                           HISTOGRAM FREQUENCY

VALID CASES    5999      MISSING CASES     0
```

FIGURE 6

HISTOGRAM OF DATA COMPLETED BY IMPUTATION

FILE:

LOG ITERATED BUCK



VALID CASES 5999 MISSING CASES 0

FIGURE 7

HISTOGRAM OF DATA COMPLETED BY IMPUTATION

```
FILE:
CUBE  ITERATED BUCK
    COUNT   MIDPOINT    ONE SYMBOL EQUALS APPROXIMATELY 12.00 OCCURRENCES
      397      50.00  *********************************
      558     150.00  ***********************************************
      546     250.00  **********************************************
      537     350.00  *********************************************
      500     450.00  ******************************************
      424     550.00  ***********************************
      455     650.00  **************************************
      331     750.00  ****************************
      354     850.00  ******************************
      246     950.00  *********************
      263    1050.00  **********************
      168    1150.00  **************
      197    1250.00  ****************
      105    1350.00  *********
      113    1450.00  *********
      115    1550.00  **********
       98    1650.00  ********
       72    1750.00  ******
       37    1850.00  ***
       79    1950.00  *******
       36    2050.00  ***
       24    2150.00  **
       22    2250.00  **
       35    2350.00  ***
       25    2450.00  **
       22    2550.00  **
       14    2650.00  *
       18    2750.00  **
       16    2850.00  *
       15    2950.00  *
       23    3050.00  **
       21    3150.00  **
        5    3250.00
       14    3350.00  *
        4    3450.00
        9    3550.00  *
        8    3650.00  *
        6    3750.00  *
        7    3850.00  *
        3    3950.00
        9    4050.00  *
        0    4150.00
        9    4250.00  *
        3    4350.00
        2    4450.00
        5    4550.00
        3    4650.00
        4    4750.00
        8    4850.00  *
        2    4950.00
                      I....+....I....+....I....+....I....+....I....+....I
                      0        120       240       360       480       600
                                HISTOGRAM FREQUENCY

 VALID CASES     5999       MISSING CASES      0
```

* = IMPUTED AMOUNTS

FIGURE 8

HISTOGRAM OF DATA COMPLETED BY IMPUTATION

```
FILE:
ARITHMETIC SMOOTHING (1)
    COUNT   MIDPOINT   ONE SYMBOL EQUALS APPROXIMATELY 12.00 OCCURRENCES

     508      50.00  ******************************************
     565     150.00  ***********************************************
     496     250.00  *****************************************
     529     350.00  ********************************************
     503     450.00  ******************************************          * = IMPUTED AMOUNTS
     394     550.00  *********************************
     446     650.00  *************************************
     294     750.00  *************************
     366     850.00  *******************************
     249     950.00  *********************
     238    1050.00  ********************
     174    1150.00  allocated***************
     198    1250.00  *****************
     113    1350.00  *********
      89    1450.00  *******
      99    1550.00  ********
      82    1650.00  *******
      89    1750.00  *******
      38    1850.00  ***
     109    1950.00  *********
      51    2050.00  ****
      19    2150.00  **
      16    2250.00  *
      35    2350.00  ***
      25    2450.00  **
      29    2550.00  **
       7    2650.00  *
      26    2750.00  **
      21    2850.00  **
      18    2950.00  **
      20    3050.00  **
      17    3150.00  *
       4    3250.00
       6    3350.00  *
       3    3450.00
       8    3550.00  *
       4    3650.00
       9    3750.00  *
       7    3850.00  *
       3    3950.00
      11    4050.00  *
      11    4150.00  *
      11    4250.00  *
       3    4350.00
       5    4450.00
       4    4550.00
       0    4650.00
       7    4750.00  *
       6    4850.00  *
       2    4950.00
                     I....+....I....+....I....+....I....+....I....+....I
                     0        120       240       360       480       600
                               HISTOGRAM FREQUENCY

VALID CASES    5999      MISSING CASES     0
```

FIGURE 9

HISTOGRAM OF DATA COMPLETED BY IMPUTATION

FILE:
MULTIPLICATIVE SMOOTHING (1)

```
 COUNT   MIDPOINT   ONE SYMBOL EQUALS APPROXIMATELY 12.00 OCCURRENCES

   504     50.00  ******************************************
   562    150.00  ***********************************************
   495    250.00  *****************************************
   534    350.00  *********************************************
   498    450.00  ******************************************
   410    550.00  **********************************
   436    650.00  ************************************
   297    750.00  *************************
   351    850.00  *****************************
   251    950.00  *********************
   230   1050.00  *******************
   172   1150.00  **************
   188   1250.00  ****************
   116   1350.00  **********
    89   1450.00  *******
    98   1550.00  ********
    82   1650.00  *******
    80   1750.00  *******
    45   1850.00  ****
   103   1950.00  *********
    53   2050.00  ****
    20   2150.00  **
    21   2250.00  **
    41   2350.00  ***
    25   2450.00  **
    31   2550.00  ***
     9   2650.00  *
    23   2750.00  **
    30   2850.00  ***
    17   2950.00  *
    20   3050.00  **
    11   3150.00  *
     4   3250.00
     7   3350.00  *
     2   3450.00
    20   3550.00  **
     9   3650.00  *
     8   3750.00  *
     5   3850.00
     8   3950.00  *
    12   4050.00  *
     5   4150.00
    12   4250.00  *
     3   4350.00
     7   4450.00  *
     3   4550.00
     1   4650.00
     8   4750.00  *
     5   4850.00
     2   4950.00
                 I....+....I....+....I....+....I....+....I....+....I
                 0        120       240       360       480       600
                           HISTOGRAM FREQUENCY

VALID CASES     5999      MISSING CASES      0
```

* = IMPUTED AMOUNTS

FIGURE 10

HISTOGRAM OF DATA COMPLETED BY IMPUTATION

```
FILE:
ARITHMETIC SMOOTHING (2)
    COUNT   MIDPOINT   ONE SYMBOL EQUALS APPROXIMATELY 12.00 OCCURRENCES

     514      50.00  *******************************************
     544     150.00  *********************************************
     507     250.00  ******************************************
     521     350.00  *******************************************
     523     450.00  ********************************************
     375     550.00  *******************************
     456     650.00  **************************************
     270     750.00  ***********************
     374     850.00  *******************************
     269     950.00  **********************
     237    1050.00  ********************
     164    1150.00  **************
     227    1250.00  *******************
      97    1350.00  ********
      94    1450.00  ********
      90    1550.00  ********
      91    1650.00  ********
      79    1750.00  *******
      45    1850.00  ****
     104    1950.00  *********
      48    2050.00  ****
      16    2150.00  *
      20    2250.00  **
      39    2350.00  ***
      17    2450.00  *
      29    2550.00  **
      18    2650.00  **
      17    2750.00  *
      19    2850.00  **
      19    2950.00  **
      20    3050.00  **
      19    3150.00  **
       3    3250.00
       6    3350.00  *
       2    3450.00
       9    3550.00  *
      17    3650.00  *
       5    3750.00
       3    3850.00
       6    3950.00  *
      11    4050.00  *
       7    4150.00  *
      10    4250.00  *
       3    4350.00
       5    4450.00
       3    4550.00
       1    4650.00
       7    4750.00  *
       5    4850.00
       4    4950.00
                    I....+....I....+....I....+....I....+....I....+....I
                    0        120       240       360       480       600
                              HISTOGRAM FREQUENCY

VALID CASES     6000      MISSING CASES      0
```

* = IMPUTED AMOUNTS

FIGURE 11

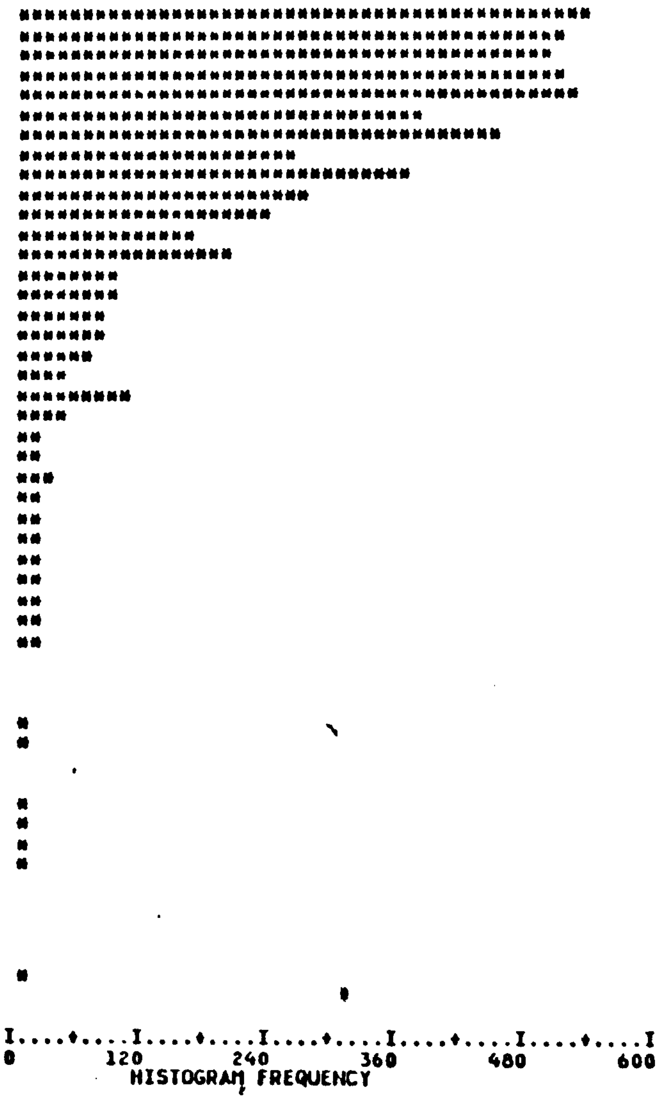
HISTOGRAM OF DATA COMPLETED BY IMPUTATION

FILE:

MULTIPLICATIVE SMOOTHING (2)

ONE SYMBOL EQUALS APPROXIMATELY 12.00 OCCURRENCES

COUNT	MIDPOINT
542	50.00
510	150.00
501	250.00
514	350.00
528	450.00
383	550.00
457	650.00
269	750.00
371	850.00
270	950.00
239	1050.00
168	1150.00
208	1250.00
95	1350.00
95	1450.00
89	1550.00
89	1650.00
77	1750.00
46	1850.00
110	1950.00
48	2050.00
20	2150.00
22	2250.00
37	2350.00
29	2450.00
29	2550.00
19	2650.00
18	2750.00
19	2850.00
20	2950.00
18	3050.00
19	3150.00
4	3250.00
5	3350.00
2	3450.00
7	3550.00
17	3650.00
5	3750.00
3	3850.00
7	3950.00
12	4050.00
8	4150.00
11	4250.00
3	4350.00
5	4450.00
3	4550.00
1	4650.00
7	4750.00
5	4850.00
5	4950.00

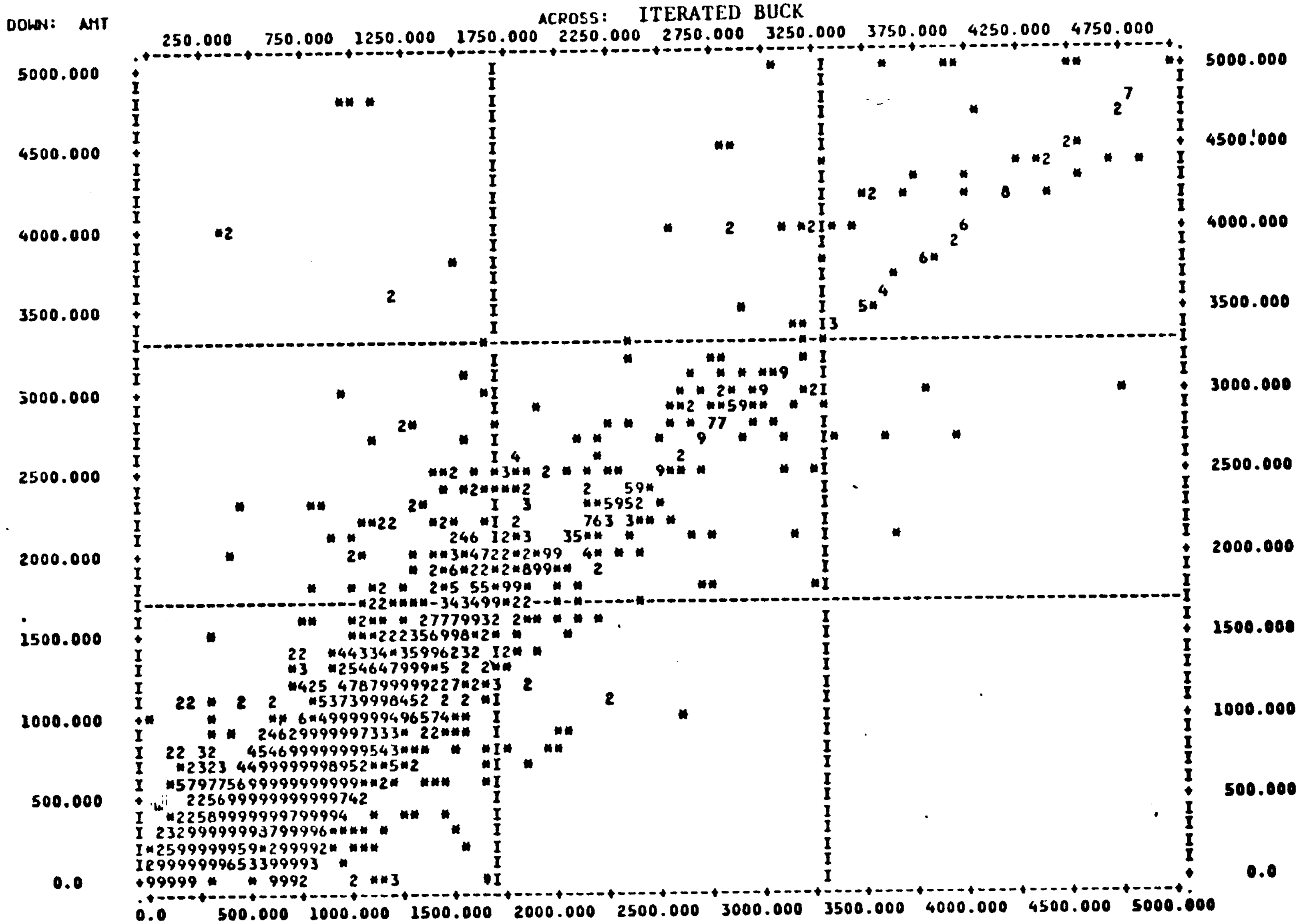


* = IMPUTED AMOUNTS

VALID CASES 6000 MISSING CASES 0

FIGURE 12

REPORTED AMOUNTS BY IMPUTED AMOUNTS

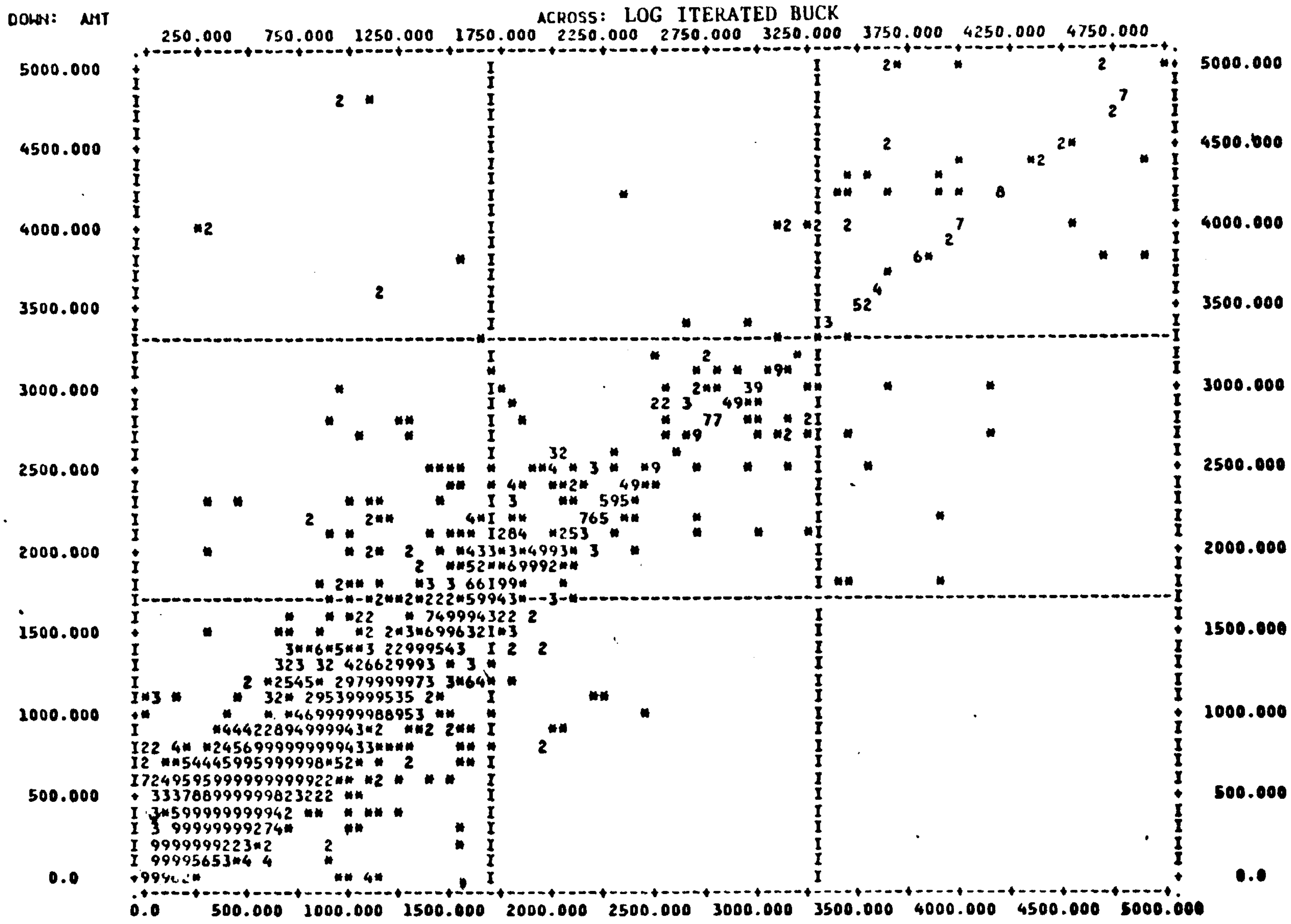


STATISTICS..					
CORRELATION (R)-	.92884	R SQUARED -	.86274	SIGNIFICANCE -	.00000
STD ERR OF EST -	284.54653	INTERCEPT (A) -	-58.87632	SLOPE (B) -	1.02211
PLOTTED VALUES -	5949	EXCLUDED VALUES-	50	MISSING VALUES -	0

COEFFICIENT CANNOT BE COMPUTED.

FIGURE 13

REPORTED AMOUNTS BY IMPUTED AMOUNTS



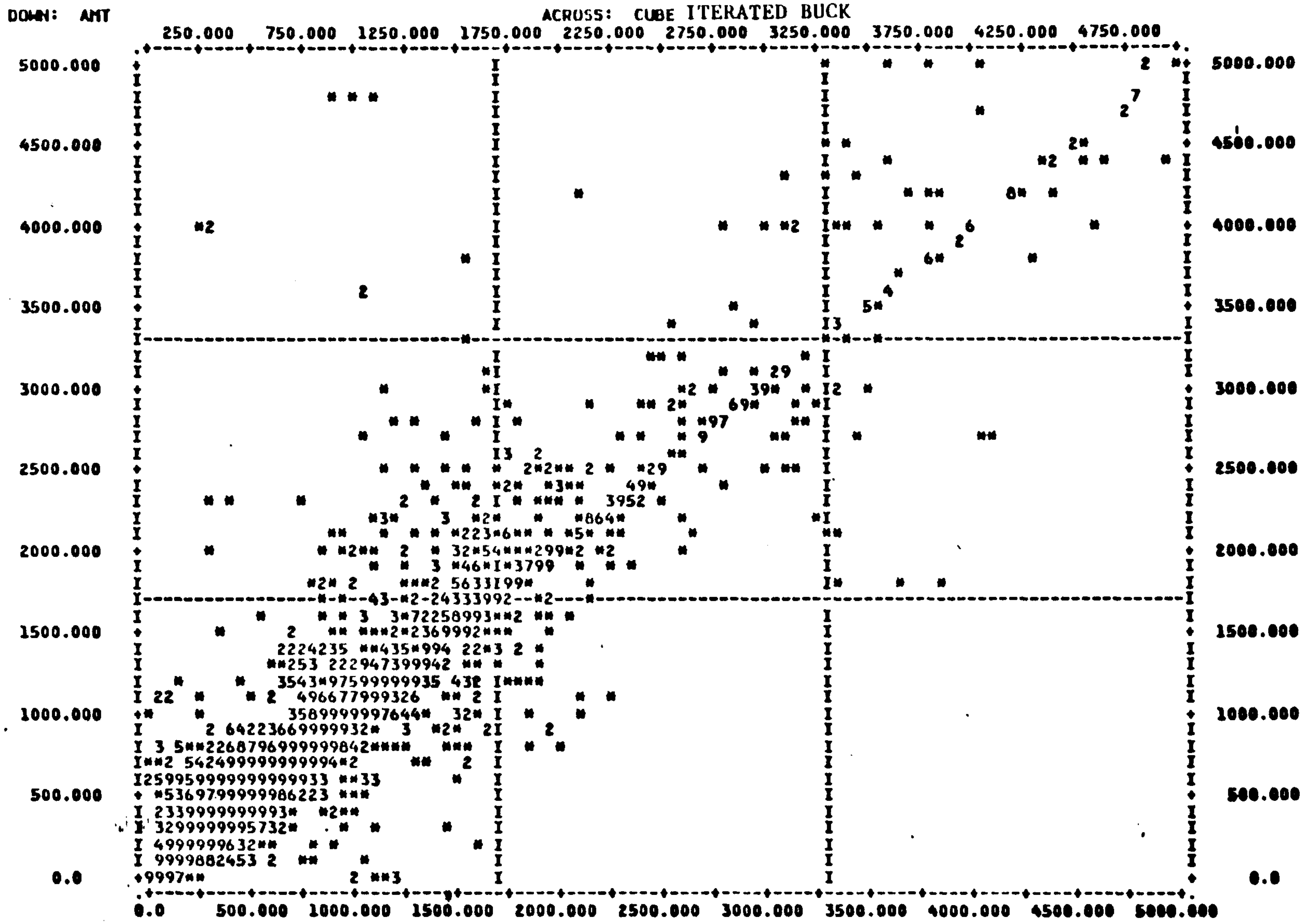
STATISTICS..

CORRELATION (R)-	.94601	R SQUARED -	.89493	SIGNIFICANCE -	.00000
STD ERR OF EST -	248.21565	INTERCEPT (A) -	29.57185	SLOPE (B) -	.99223
PLOTTED VALUES -	5958	EXCLUDED VALUES-	41	MISSING VALUES -	0

'********' IS PRINTED IF A COEFFICIENT CANNOT BE COMPUTED.

FIGURE 14

REPORTED AMOUNTS BY IMPUTED AMOUNTS



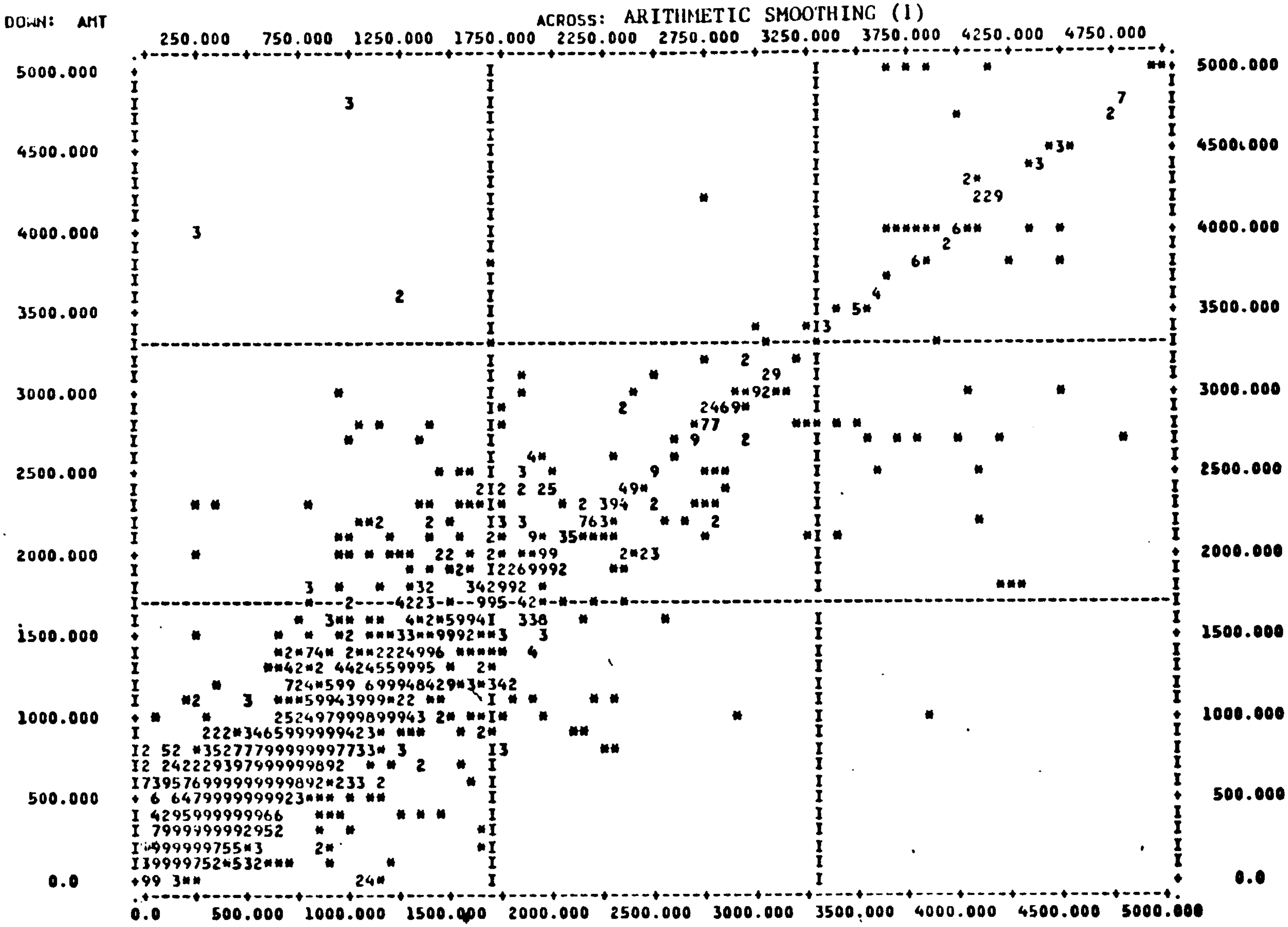
STATISTICS..

CORRELATION (R)-	.94384	R SQUARED -	.89083	SIGNIFICANCE -	.00000
STD ERR OF EST -	254.02342	INTERCEPT (A) -	34.55968	SLOPE (B) -	.99963
PLOTTED VALUES -	5959	EXCLUDED VALUES-	40	MISSING VALUES -	0

'********' IS PRINTED IF A COEFFICIENT CANNOT BE COMPUTED.

FIGURE 15

REPORTED AMOUNTS BY IMPUTED AMOUNTS



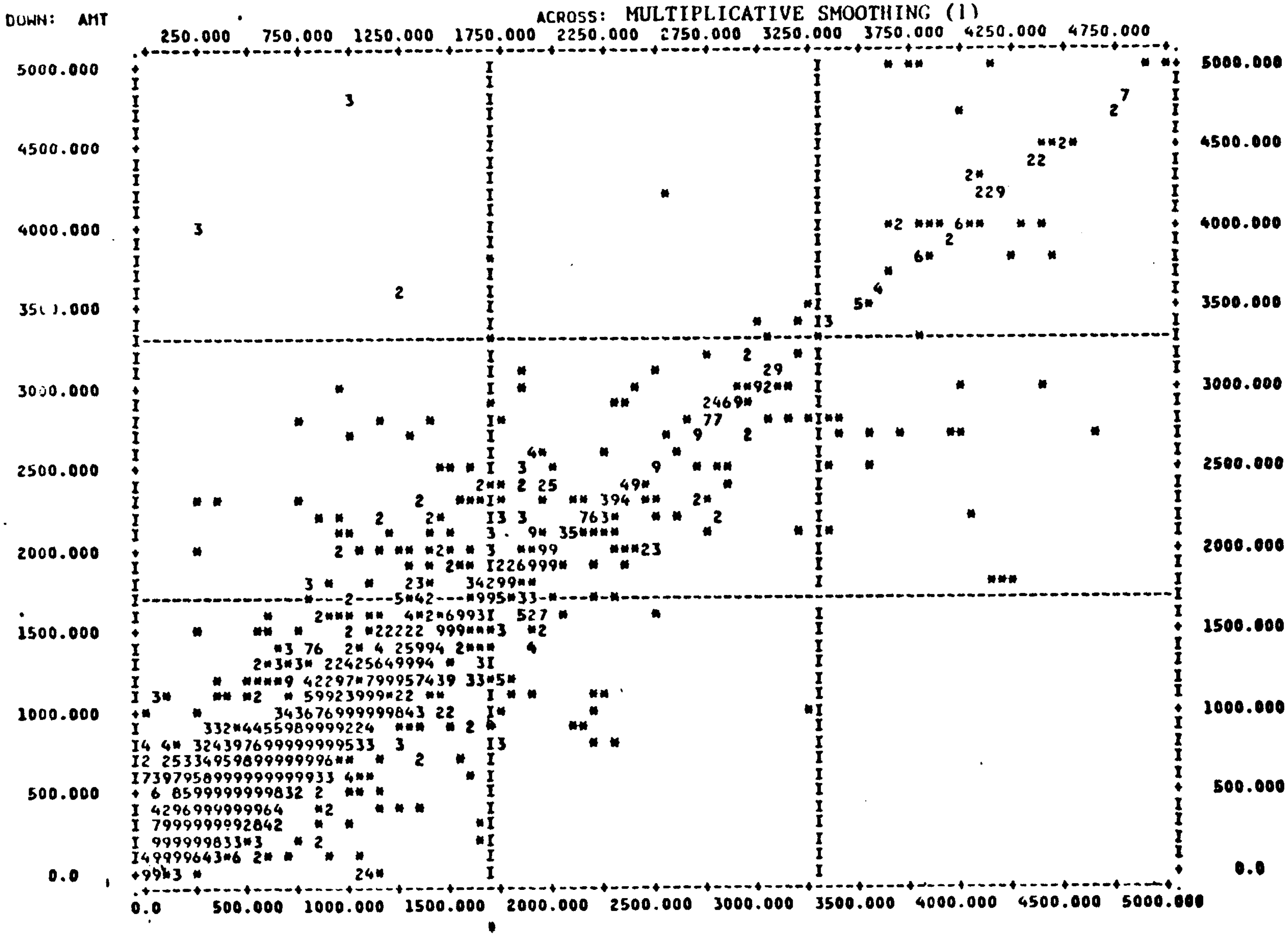
STATISTICS..

CORRELATION (R)-	.94223	R SQUARED -	.88780	SIGNIFICANCE -	.00000
STD ERR OF EST -	255.96089	INTERCEPT (A) -	61.27008	SLOPE (B) -	.95075
PLOTTED VALUES -	5957	EXCLUDED VALUES-	42	MISSING VALUES -	0

'********' IS PRINTED IF A COEFFICIENT CANNOT BE COMPUTED.

FIGURE 16

REPORTED AMOUNTS BY IMPUTED AMOUNTS



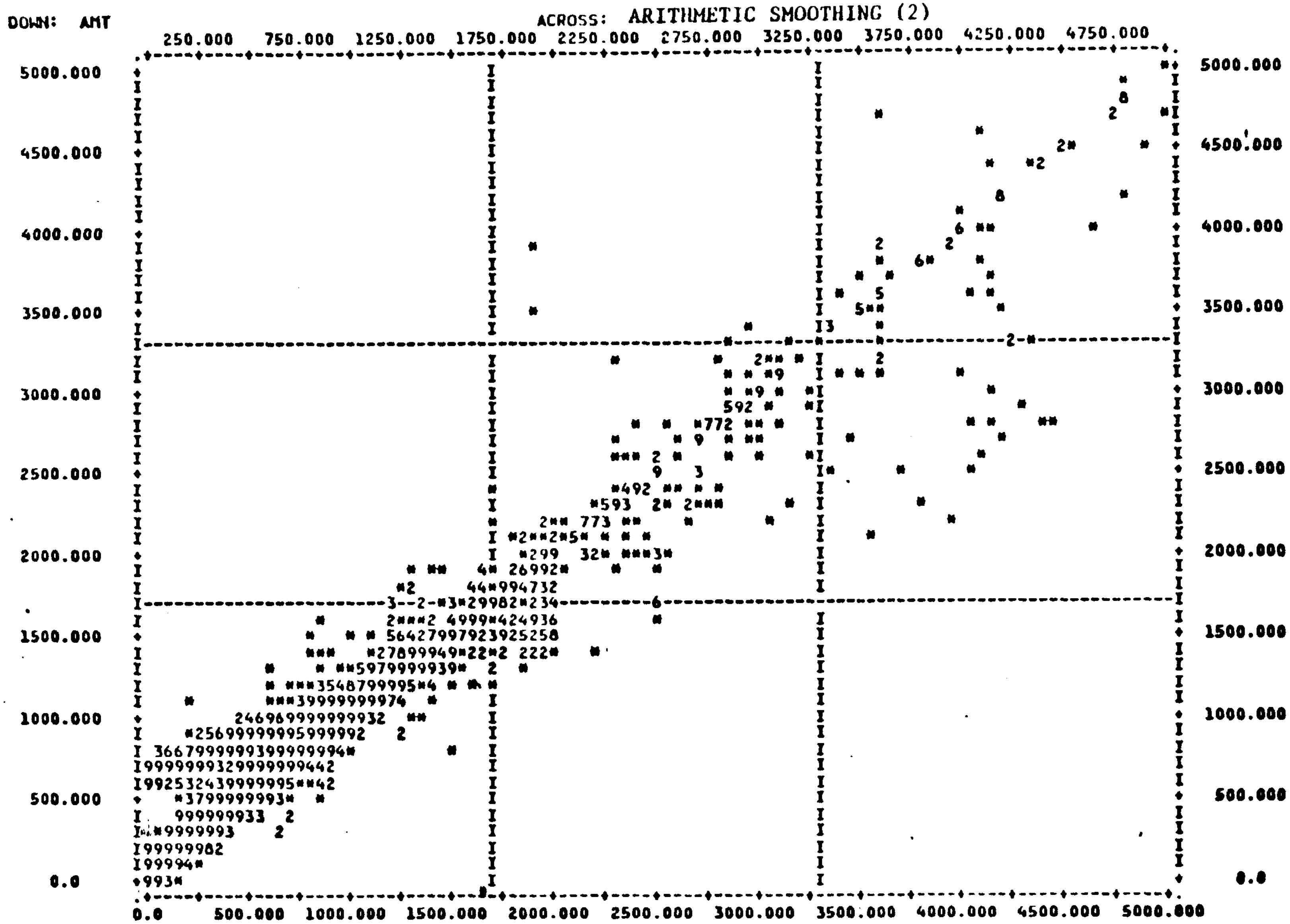
STATISTICS..

CORRELATION (R)-	.94273	R SQUARED -	.88874	SIGNIFICANCE -	.00000
STD ERR OF EST -	254.88986	INTERCEPT (A) -	63.88467	SLOPE (B) -	.95477
PLOTTED VALUES -	5957	EXCLUDED VALUES-	42	MISSING VALUES -	0

'********' IS PRINTED IF A COEFFICIENT CANNOT BE COMPUTED.

FIGURE 17

REPORTED AMOUNTS BY IMPUTED AMOUNTS



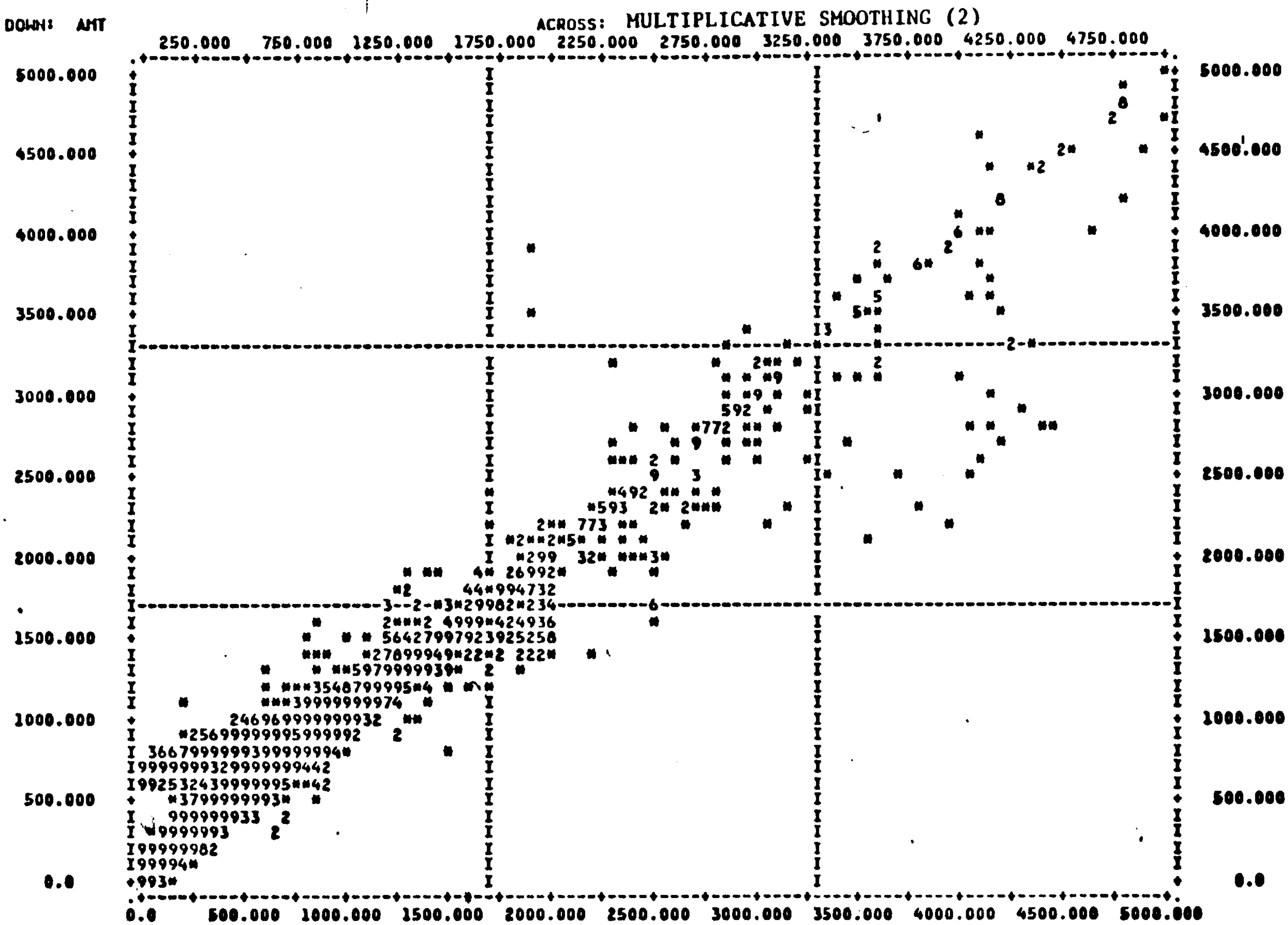
STATISTICS..

CORRELATION (R)-	.96460	R SQUARED -	.93046	SIGNIFICANCE -	.00000
STD ERR OF EST -	184.66221	INTERCEPT (A) -	155.05095	SLOPE (B) -	.89032
PLOTTED VALUES -	5955	EXCLUDED VALUES-	45	MISSING VALUES -	0

'********' IS PRINTED IF A COEFFICIENT CANNOT BE COMPUTED.

FIGURE 18

REPORTED AMOUNTS BY IMPUTED AMOUNTS



STATISTICS..					
CORRELATION (R)-	.96460	R SQUARED -	.93046	SIGNIFICANCE -	.00000
STD ERR OF EST -	184.66222	INTERCEPT (A) -	155.05096	SLOPE (B) -	.89032
PLOTTED VALUES -	5955	EXCLUDED VALUES-	45	MISSING VALUES -	0

'********' IS PRINTED IF A COEFFICIENT CANNOT BE COMPUTED.

FIGURE 19

HISTOGRAM OF SCALED DIFFERENCES

ITERATED BUCK

Count	Midpoint
10	-1.725
12	-1.675
13	-1.625
10	-1.575
21	-1.525
12	-1.475
10	-1.425
19	-1.375
18	-1.325
12	-1.275
14	-1.225
14	-1.175
21	-1.125
14	-1.075
29	-1.025
33	-.975
15	-.925
32	-.875
28	-.825
26	-.775
43	-.725
45	-.675
46	-.625
49	-.575
66	-.525
55	-.475
61	-.425
62	-.375
84	-.325
81	-.275
103	-.225
125	-.175
170	-.125
192	-.075
182	-.025
194	.025
174	.075
141	.125
112	.175
103	.225
55	.275
59	.325
48	.375
26	.425
19	.475
22	.525
23	.575
13	.625
15	.675
11	.725
6	.775
10	.825
7	.875
5	.925
1	.975

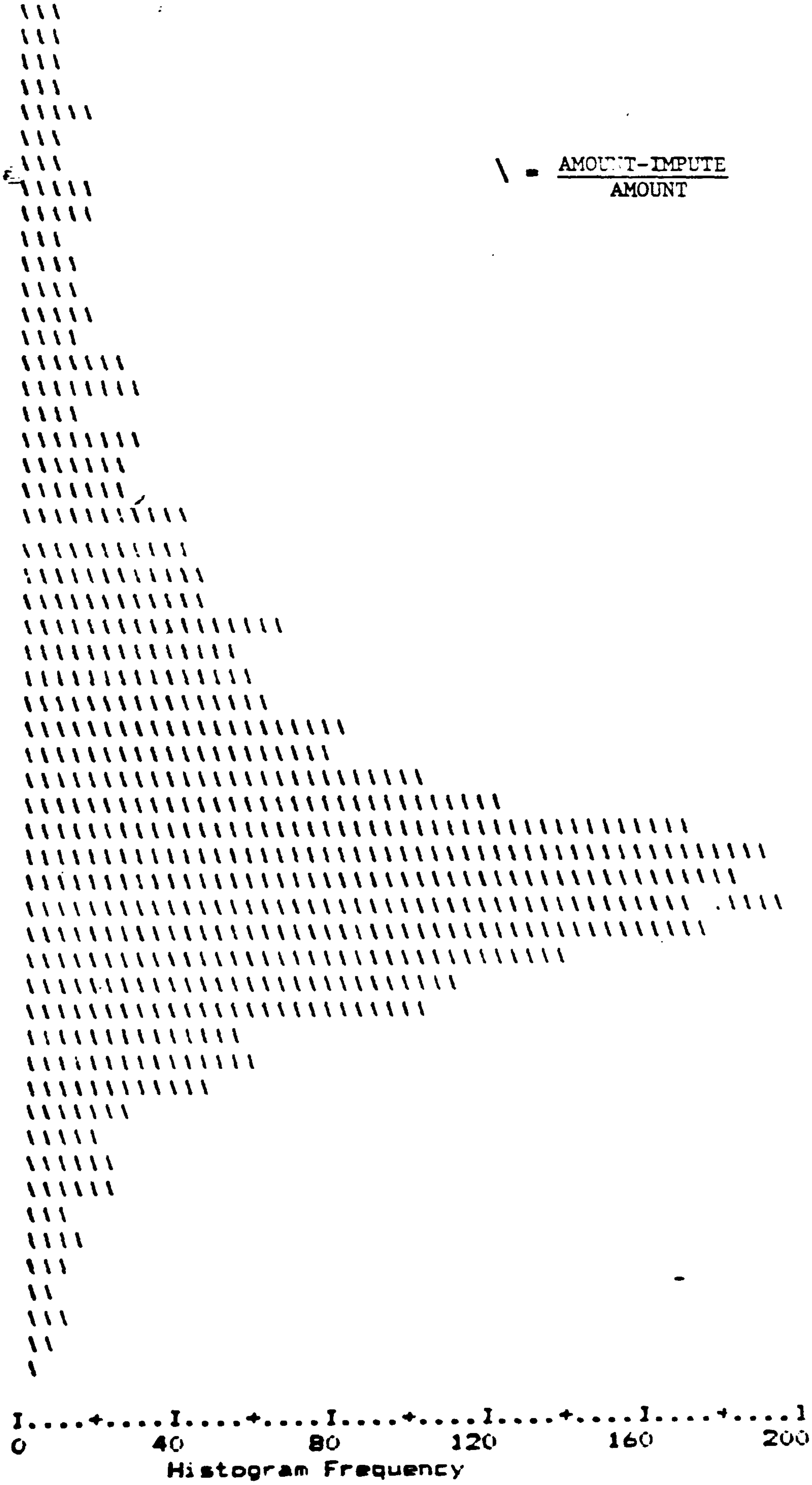


FIGURE 20

HISTOGRAM OF SCALED DIFFERENCES

$\backslash = \dfrac{\text{AMOUNT-IMPUTE}}{\text{AMOUNT}}$

LOG ITERATED	BUCK
3	-1.725
3	-1.675
4	-1.625
2	-1.575
2	-1.525
4	-1.475
1	-1.425
3	-1.375
6	-1.325
4	-1.275
3	-1.225
10	-1.175
8	-1.125
8	-1.075
12	-1.025
6	-.975
13	-.925
13	-.875
14	-.825
20	-.775
19	-.725
26	-.675
29	-.625
28	-.575
34	-.525
32	-.475
58	-.425
76	-.375
89	-.325
98	-.275
125	-.225
166	-.175
194	-.125
241	-.075
272	-.025
258	.025
227	.075
203	.125
141	.175
115	.225
102	.275
62	.325
57	.375
64	.425
41	.475
51	.525
35	.575
29	.625
16	.675
5	.725
19	.775
20	.825
8	.875
6	.925
18	.975

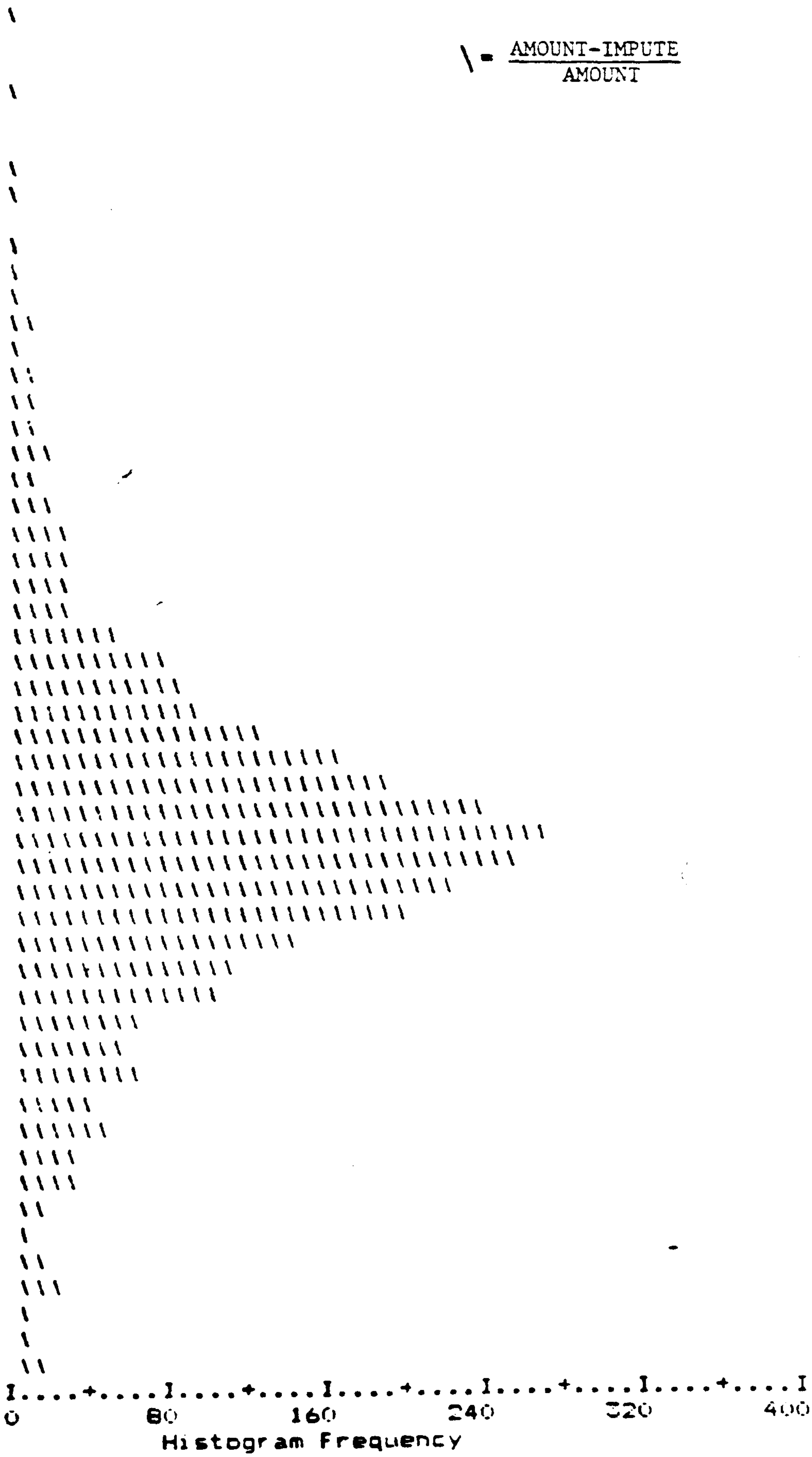


FIGURE 21

HISTOGRAM OF SCALED DIFFERENCES

CUBE ITERATED BUCK

$$\backslash = \frac{\text{AMOUNT-IMPUTE}}{\text{AMOUNT}}$$

Count	Midpoint
0	-1.7750
5	-1.7250
0	-1.6750
6	-1.6250
1	-1.5750
0	-1.5250
5	-1.4750
8	-1.4250
2	-1.3750
7	-1.3250
7	-1.2750
9	-1.2250
6	-1.1750
9	-1.1250
15	-1.0750
13	-1.0250
9	-.9750
11	-.9250
15	-.8750
20	-.8250
12	-.7750
11	-.7250
22	-.6750
21	-.6250
32	-.5750
35	-.5250
48	-.4750
51	-.4250
49	-.3750
73	-.3250
91	-.2750
116	-.2250
174	-.1750
168	-.1250
187	-.0750
234	-.0250
249	.0250
250	.0750
219	.1250
183	.1750
126	.2250
121	.2750
95	.3250
60	.3750
60	.4250
66	.4750
48	.5250
24	.5750
25	.6250
12	.6750
16	.7250
25	.7750
20	.8250
4	.8750
14	.9250
8	.9750

```
I....+....I....+....I....+....I....+....I....+....I
0        50        100       150       200       250
              Histogram Frequency
```

lid Cases 3183 Missing Cases 0

FIGURE 22

ARITHMETIC SMOOTHING (1)

HISTOGRAM OF SCALED DIFFERENCES

$\backslash$ = $\frac{\text{AMOUNT-IMPUTE}}{\text{AMOUNT}}$

Count	Midpoint
0	-1.725
0	-1.675
0	-1.625
1	-1.575
4	-1.525
4	-1.475
2	-1.425
7	-1.375
3	-1.325
4	-1.275
1	-1.225
2	-1.175
5	-1.125
8	-1.075
11	-1.025
15	-.975
10	-.925
10	-.875
9	-.825
12	-.775
9	-.725
14	-.675
13	-.625
25	-.575
25	-.525
35	-.475
32	-.425
47	-.375
61	-.325
111	-.275
102	-.225
115	-.175
154	-.125
181	-.075
427	-.025
234	.025
223	.075
159	.125
227	.175
150	.225
130	.275
56	.325
83	.375
70	.425
56	.475
46	.525
32	.575
25	.625
26	.675
19	.725
16	.775
23	.825
16	.875
14	.925
14	.975

FIGURE 23

HISTOGRAM OF SCALED DIFFERENCES

MULTIPLICATIVE SMOOTHING (1)

\ = AMOUNT-IMPUTE / AMOUNT

Count	Midpoint	
2	-1.725	
2	-1.675	
1	-1.625	
1	-1.575	
3	-1.525	
4	-1.475	
3	-1.425	
6	-1.375	\
4	-1.325	
2	-1.275	
1	-1.225	
4	-1.175	
5	-1.125	\
5	-1.075	\
11	-1.025	\
9	-.975	\
6	-.925	\
13	-.875	\
6	-.825	\
14	-.775	\
[illegible]	-.725	\
9	-.675	\
14	-.625	\
20	-.575	\\
26	-.525	\\\
33	-.475	\\\
24	-.425	\\
44	-.375	\\\\
61	-.325	\\\\\\
99	-.275	\\\\\\\\\\
109	-.225	\\\\\\\\\\\
114	-.175	\\\\\\\\\\\
140	-.125	\\\\\\\\\\\\\\
191	-.075	\\\\\\\\\\\\\\\\\\\
421	-.025	\\
233	.025	\\\\\\\\\\\\\\\\\\\\\\\
228	.075	\\\\\\\\\\\\\\\\\\\\\\\
164	.125	\\\\\\\\\\\\\\\\
223	.175	\\\\\\\\\\\\\\\\\\\\\\
144	.225	\\\\\\\\\\\\\\
128	.275	\\\\\\\\\\\\\
95	.325	\\\\\\\\\\
80	.375	\\\\\\\\
76	.425	\\\\\\\\
49	.475	\\\\\
58	.525	\\\\\\
42	.575	\\\\
24	.625	\\
36	.675	\\\\
28	.725	\\\
16	.775	\\
23	.825	\\
16	.875	\\
13	.925	\
20	.975	\\

I....+....I....+....I....+....I....+....I....+....I

0 100 200 300 400 500

FIGURE 24

HISTOGRAM OF SCALED DIFFERENCES

ARITHMETIC SMOOTHING (2)

\ = $\frac{\text{AMOUNT-IMPUTE}}{\text{AMOUNT}}$

Count	Midpoint
2	-1.725
3	-1.675
0	-1.625
2	-1.575
4	-1.525
6	-1.475
1	-1.425
6	-1.375
6	-1.325
2	-1.275
6	-1.225
3	-1.175
6	-1.125
9	-1.075
14	-1.025
14	-.975
6	-.925
8	-.875
8	-.825
11	-.775
14	-.725
10	-.675
15	-.625
23	-.575
32	-.525
28	-.475
20	-.425
42	-.375
47	-.325
91	-.275
115	-.225
109	-.175
174	-.125
186	-.075
471	-.025
194	.025
209	.075
164	.125
239	.175
116	.225
126	.275
92	.325
79	.375
77	.425
66	.475
34	.525
45	.575
24	.625
30	.675
42	.725
23	.775
22	.825
16	.875
14	.925
20	.975

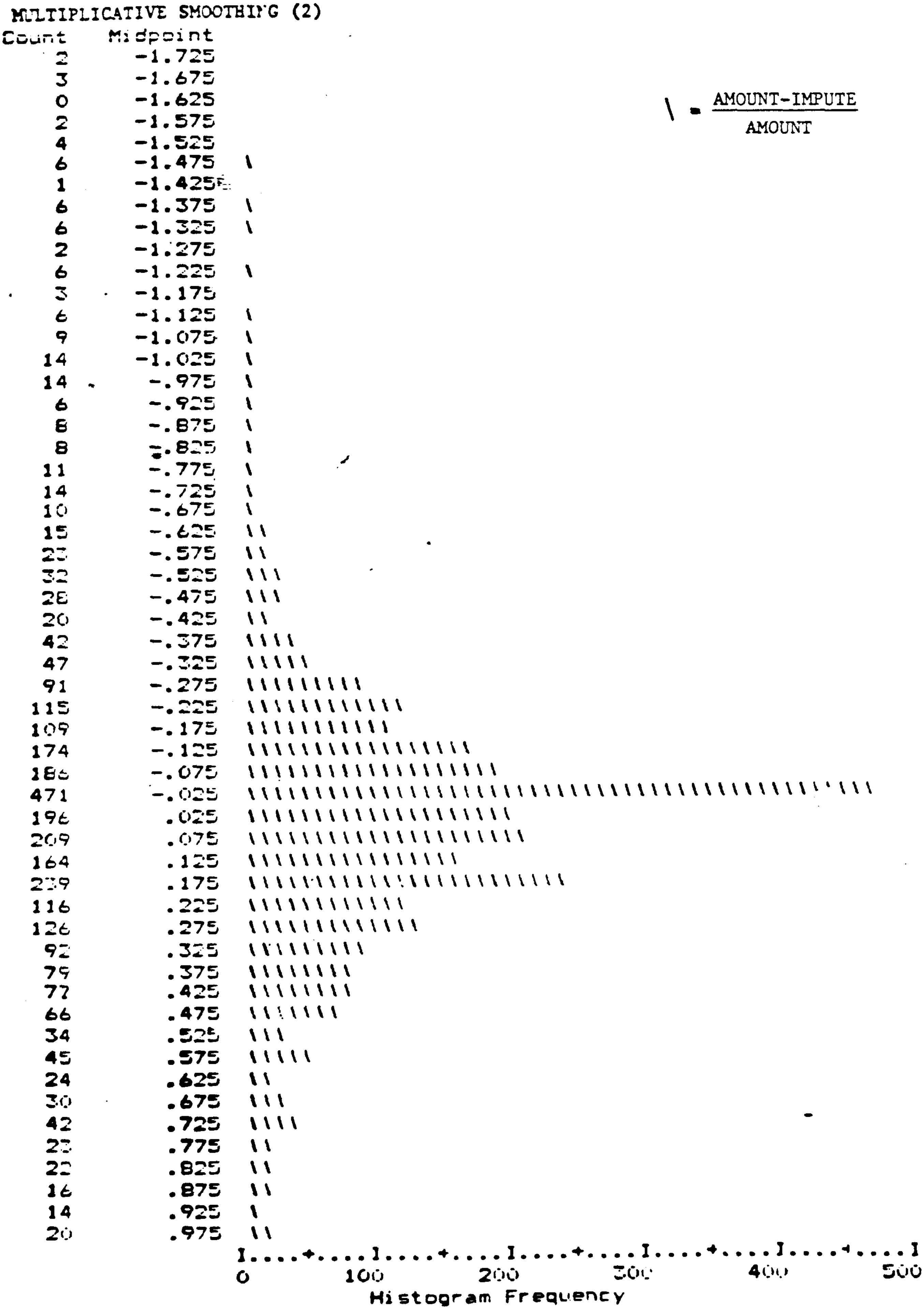
I....+....I....+....I....+....I....+....I....+....I

0 100 200 300 400 500

Histogram Frequency

FIGURE 25

HISTOGRAM OF SCALED DIFFERENCES



$c_i = x_i - \hat{x}_i$	Iterated Buck	Logarithmic Iterated Buck	Cube Iterated Buck	Arithmetic Smoothing (1)	Multiplicative Smoothing (1)	Arithmetic Smoothing (2)	Multiplicative Smoothing (2)
$\sum_i c_i$	-198,198.8	182,713.8	244,565.5	150,246.2	191,657.9	177,842.6	198,243.3
$\sum_i c_i^2$	881,607,100	791,591,000	835,393,408	836,990,100	823,631,700	880,732,300	874,453,700
$\frac{\sum_i c_i}{N}$	-62.268	57.403	76.835	47.203	60.213	55.873	62.282
$\frac{\sum_i (c_i - \bar{c})^2}{N}$	273,096.5	245,398.1	256,550.8	260,729.8	255,135	273,577.6	270,848.4

3183 cases

i = only imputed values

TABLE 1

$c_i = \frac{x_i - \hat{x}_i}{x_i}$	Iterated Buck	Logarithmic Iterated Buck	Cube Iterated Buck	Arithmetic Smoothing (1)	Multiplicative Smoothing (1)	Arithmetic Smoothing (2)	Multiplicative Smoothing (2)
$\sum_i c_i$	-5792.08	-1001.723	-1120.027	-734.073	-661.089	-699.351	-661.261
$\sum_i c_i^2$	3,594,429	952,807.2	58973.7	1,024,158	1,007,097	1,022,134	1,013,160
$\frac{\sum_i c_i}{N}$	-1.820	-.315	-.352	-.231	-.208	-.220	-.208
$\frac{\sum_i (c_i - \bar{c})^2}{N}$	148.434	16.454	18.404	16.904	16.827	17.016	16.985

3163 cases
i = only imputed values

TABLE 2

$c_i = r_i - \hat{r}_i$	Iterated Buck	Logarithmic Iterated Buck	Cube Iterated Buck	Arithmetic Smoothing (1)	Multiplicative Smoothing (1)	Arithmetic Smoothing (2)	Multiplicative Smoothing (2)
$\sum_i c_i$	-2,252.849	-1040.282	-957.756	-1482.371	-1514.899	-1512.631	-1522.117
$\sum_i c_i^2$	4,071,898	122,167.8	119,478.695	202,261.6	221,854.2	213,856.469	226,826.703
$\frac{\sum_i c_i}{N}$	-.799	-.369	-.340	-.526	-.537	-.536	-.540
$\frac{\sum_i (c_i - \bar{c})^2}{N}$	1,443.303	43.186	42.253	71.448	78.383	75.547	80.144

i = only imputed values

TABLE 3

$e_i = \frac{r_i - \hat{r}_i}{r_i}$	Iterated Buck	Logarithmic Iterated Buck	Cube Iterated Buck	Arithmetic Smoothing (1)	Multiplicative Smoothing (1)	Arithmetic Smoothing (2)	Multiplicative Smoothing (2)
$\sum_i e_i$	-2960.78	-1756.161	-1679.103	-2258.004	-2,305.211	-2295.991	-2320.885
$\sum_i e_i^2$	4,074,304	123,128.9	121,931.687	215,218.4	232,074.2	227,100.922	237,474.719
$\frac{\sum_i e_i}{N}$	-1.050	-.623	-.595	-.801	-.817	-.814	-.823
$\frac{\sum_i (e_i - \bar{e})^2}{N}$	1443.693	43.275	42.884	75.677	81.628	79.868	83.533

i = only imputed values

TABLE 4