GENERALIZED VARIANCE FUNCTIONS FOR A COMPLEX SAMPLE SURVEY

Eugene G. Johnson and Benjamin F. King, Educational Testing Service

ABSTRACT: For a national survey of reading ability among young adults using a multi-stage, stratified probability sample, generalized variance functions (GVF) were estimated. That is, an attempt was made to express the estimated variance of a statistic as a function of that statistic and other characteristics of the variable of interest. With GVFs estimated from a development sample of variables, predictions of sampling variance were made for other variables in a confirmation sample and comparisons made with conventional jackknife estimates. Conclusions were drawn about the feasibility of use of GVFs, with emphasis on the margin of additional estimation error that is introduced.

KEYWORDS: Estimated variance, design effects, jackknife.

INTRODUCTION. In this paper we report the results of an investigation into the feasibility of using generalized variance functions (GVF) for estimation of sampling variances for statistics computed for a large-scale and complex survey. As described in Wolter (1985), the GVF method attempts to model the variance of a survey estimator as a function of the estimate and possibly other variables. If the modelling is successful and the actual computation of the estimated variance by the usual formula is made unnecessary, then considerable cost savings may accrue. An accurate GVF may also be of great value in designing future surveys of the same type. This approach to variance estimation has been adopted by the Bureau of the Census for the Current Population Survey, and also by the National Center for Health Services Research in certain applications. In addition to the references cited in Wolter (1985), the reader should see Cohen (1979), Cohen and Kalsbeek (1981), and Burt and Cohen (1984).

In many of the previous applications the GVF models the relative variance of an estimated subpopulation total. A common specification is

$$\mathrm{relvar}(\hat{X}) = \alpha + \beta/\hat{X}. \qquad (1)$$

This specification is in turn used to derive a model for the relative variance of a ratio or proportion. In the present paper we focus primarily on the direct estimation of variances of proportions -- e.g., for the percentage of subjects who choose a particular distractor in an achievement test -- although we shall also devote some discussion to the modelling of variances of subpopulation totals.

The statistics of interest are from the Young Adult Literacy Survey, conducted by Response Analysis Corporation for the ETS Center for National Assessment of Educational Progress in the summer of 1985. The target population was all persons of age 21 through 25 residing in households in the Continental U.S. The sample design was a fairly conventional stratified, five-stage probability type with pps selection down to the household level. Selection of units in the first three stages was systematic. The design included oversampling of Blacks and Hispanics at an approximately two to one rate. This oversampling was effected at the second stage of selection (roughly Census tracts) -- hence members of each racial group could be selected at more than one rate. Approximately 36,000 housing units were screened for eligible subjects resulting in a final sample of about 3,500 respondents, each of whom provided measures of cognitive and background characteristics. Estimates of means, totals, and proportions obtained for these items involve weights that reflect adjustments for disproportionate sampling, nonresponse, and poststratification to known marginal totals.

The variance of a statistic for this survey will deviate from that under simple random sampling with a fixed sample size for a number of reasons. There are gains in precision over that of srs from stratification by geography and size. These gains, however, are counterbalanced by the effects of nonoptimal disproportionate selection and clustering. The use of weights which are subject to random variability makes the statistics of interest nonlinear. All of these considerations combine to make the estimators of sampling variances more complex and computationally more expensive than the simple srs algorithms.

In this survey the variance estimation procedure is the jackknife applied to forty-nine pairs of ultimate clusters -- see, e.g., Wolter (1985), p. 185. This technique involves the computation of forty-nine pseudo-values of each statistic of interest at nonnegligible expense. The purpose of the present research is to try to develop an alternative estimator of sampling variability that is less computationally intensive, but of adequate precision. The general approach is to fit linear models of functions of the sampling variance to estimates from the survey, using the jackknife variance estimate as the basis for the dependent variable, and various easily-computed statistics as the predictors. For items not used in the development of the model, variances are estimated by prediction from the fitted equation. Presumably in a large-scale survey, a relatively small number of items would be used in GVF development, and the results used for the variance estimation for the remainder.

VARIANCE MODELS FOR COGNITIVE ITEMS. In the Young Adult Literacy Survey there were 104 different items, each designed to measure a person's cognitive ability with respect to one of four psychometric scales -- (1) reading proficiency, (2) prose comprehension, (3) document utilization, and (4) practical computation. The principal statistics for any one of these items are the proportions of subjects choosing each of a set of possible response categories. Statistics for each item were produced for each of the 13 domains of the target population. These domains include the total population, each sex, three racial groups, three levels of education, and four geographic regions. Thus we have measures on 104 × 13 or 1,352 variables for which we have arbitrarily chosen the first response category for the present analysis. The items are sufficiently diverse so that observed proportions cover a wide range within each domain. For model development we have further drawn a systematic sample of 897 of the item measures, saving the remainder for validation. The selection was balanced to provide the same 69 items for each of the 13 domains. In the survey, items were administered according to a balanced-incomplete-block spiralling scheme so that not every person was given each item. The average number of responses to an item for the total population was 1,487, and obviously smaller for narrower domains; e.g., for Hispanics the average number of cases was only 167.

The first GVF model for proportions that we consider is derived from that for the estimated total, shown in (1) above. The derivation, based on the assumption that Equation (1) holds for both numerator and denominator of the sample ratio (proportion) and that there is zero correlation between the ratio and its denominator, is discussed in Wolter (1985), p. 204.

Calling the weighted sample proportion $\hat{p}$, and the weighted total $\hat{X}$, we write

$$\mathrm{relvar}(\hat{p}) = \beta(1 - \hat{p})/\hat{p}\hat{X}, \qquad (2)$$

or in terms of the variance,

$$\mathrm{var}(\hat{p}) = \beta\hat{p}(1 - \hat{p})/\hat{X}. \qquad (3)$$

Exhibit 1 shows the results of an ordinary least squares fit of the model in (3) above, where Z is the compact notation for the predictor expression on the right-hand side and the intercept is allowed to be nonzero.

[Exhibit 1 about here].

A noteworthy aspect is the asymmetry about zero in the plot of residuals against predicted values. Furthermore, the large mean squared error indicates poor predictability for the dependent variable.

Exhibit 2 shows the results of a regression of the variance of $\hat{p}$ on the simple random sampling formulation,

$$\mathrm{var}(\hat{p}) = \alpha + \beta\hat{p}(1 - \hat{p})/n. \qquad (4)$$

It can be seen that the linear fit is better than that for (3) above, leading us to doubt the effectiveness of the traditional specification.

[Exhibit 2 about here].

One of the problems with the approach taken thus far is that the variance is modelled according to conventional least squares, the optimality of which depends on underlying normality. Since variances tend to have skewed distributions, one should perhaps not expect symmetry in residuals. There is also discussion in the GVF literature of the necessity in model fitting of correcting for inconstant residual variances. Our approach at this point is to transform to logarithms, with the aim of symmetrizing the errors, making them more homoskedastic, and converting multiplicative relationships to linear.

As an example, we transform the variables in the model in (3) to their natural logarithms, yielding

$$\mathrm{logvar}(\hat{p}) = \alpha + \beta\log[\hat{p}(1 - \hat{p})/\hat{X}]. \qquad (5)$$

With ordinary least squares we obtain the results shown in Exhibit 3. The scatterplot and the normal probability plot indicate that the transformation achieved its purpose, and we shall continue to work in the logarithmic metric.

[Exhibit 3 about here].

Since the model in Equation (4) had a better linear fit than Equation (3) it is natural to examine its logarithmic version:

$$\mathrm{logvar}(\hat{p}) = \alpha + \beta\log[\hat{p}(1 - \hat{p})/n], \qquad (6)$$

the results of which are shown in Exhibit 4.

[Exhibit 4 about here].

A third, and somewhat cruder logarithmic model, is

$$\mathrm{logvar}(\hat{p}) - \log[\hat{p}(1 - \hat{p})/n] = \alpha, \qquad (7)$$

for which least squares yields the mean log design effect as the estimator of α.

For the purpose of model comparison, we shall reconsider the measure of goodness of fit of the various models. Our aim in developing a GVF is to predict the variance of a statistic for use in estimation and inference. For many of the purposes for which the results of this survey will be used underestimation of sampling variability is a more serious error than overestimation. Thus we would rather have conservative estimated standard errors than those that are too small. For illustration in this paper, we shall assume that the consequences of an underestimate are three times as severe as those of an overestimate of the same magnitude. We shall also assume the opportunity loss to be linear. A standard result from decision theory -- e.g., Raiffa and Schlaifer (1961) -- shows that the predicted value of the dependent variable that minimizes expected linear opportunity loss of the error of estimation is the fractile of the predictive distribution given by the ratio

$$k_u/(k_u + k_o), \tag{8}$$

where k_u and k_o are the losses of under and overestimation, respectively. Hence, assuming normality, we use as the optimal prediction of the log-variance the expression

$$LV^* = \text{predicted value} + .67 \text{ standard error of prediction}, \tag{9}$$

corresponding to the .75 fractile of the normal probability distribution. We shall evaluate alternative models by comparing the means of absolute residuals from the optimal predicted values, LV*, weighting positive residuals by three. The following table summarizes the results thus far:

Table 1
AVERAGE LOSS OF PREDICTION ERROR

Model	Average Loss
5	0.5248
6	0.4377
7	0.4624

Because Model (6) appears to perform better in prediction, we shall restrict further search to elaborations upon its form.

The next step is to separate the factors in the logarithm of the simple random sampling formula for the variance, as in the following equation:

$$\text{logvar}(\hat{p}) = \alpha + \beta_1\log(\hat{p}) + \beta_2\log(1 - \hat{p}) + \beta_3\log(n). \tag{10}$$

Exhibit 5 shows that the linear fit is slightly improved by this specification, with log(n) playing a lesser role in the determination of the dependent variable than in Model (6). Further elaboration upon the theme is shown in the lower part of Exhibit 5 for the model

$$\text{logvar}(\hat{p}) = \alpha + \beta_1\log(\hat{p}) + \beta_2\log(1 - \hat{p}) + \beta_3\log(n) + \beta_4\log[cv(\hat{X})] \tag{11}$$

where, although the coefficient for the log of the coefficient of variation of the estimated total is significant (LCVWN in the exhibit), its contribution to adjusted R^2 is negligible. The following mean losses may be appended to Table 1 above:

10	0.4351
11	0.4296

[Exhibit 5 about here]

In Exhibit 6 we show plots of the residuals from the ordinary least squares fitting of Model (11) for the thirteen groups of measures corresponding to the domains of the target population, and for the four scale identifiers, where we hoped to see some indication of further association. Although the scatterplot did not promise very much, we experimented with the introduction of group (domain) effects and developed the best fitting model shown in Exhibit 7:

$$\text{logvar}(\hat{p}) = \alpha + \beta_1\log(\hat{p}) + \beta_2\log(1 - \hat{p}) + \beta_3\log(n) + \beta_4 G3 + \beta_5 G4 + \beta_6 G5 + \beta_7 G8 + \beta_8 G9 + \beta_9 G10, \tag{12}$$

where G3, G4, G5, G8, G9, and G10 are one-zero indicator variables for the domains of female, whites, blacks, high school education, greater than high school, and NE region. The exhibit shows that the increase in adjusted R^2 is not great. The average loss using the same predictor and criterion as in the previous models is 0.4224, showing a decrease of less than two percent from that for Model (11).

[Exhibits 6 and 7 about here].

With so little improvement from the introduction of the domain identifier, we decided to stop the data dredging and conclude that there is little hope of developing a GVF model that can perform a great deal better than (7), in which the logvar of the weighted proportion is predicted by the logarithm of the srs formula, plus the average of the logarithm of the design effect. The average logdeff for the set of 897 items is 0.5464 and the standard deviation is .3858. Thus, the confidence interval for the predicted value under that model would have a half-width of about .8 -- not very precise, considering that the average logvar is 1.345.

To understand more fully the implications of the models it is useful to consider goodness of prediction from a different angle. We have stated that we are primarily concerned with underestimation of the standard error, i.e., we would prefer estimates that are too big, rather than too small. For model comparison we have given underestimates three times the weight of overestimates and computed the weighted mean absolute error (average loss). The models are logarithmic, however, and it is reasonable to ask for some measure of performance in terms of the antilogs, i.e., the standard error that we are ultimately interested in knowing. We, therefore, transform the predicted values of the logvariance to predictions of the standard error and compute a relative error of prediction:

$$\text{err} = (\text{jackknife s.e.} - \text{predicted s.e.})/\text{jackknife s.e.} \tag{13}$$

For Model (7) we display in Exhibit 8 the histogram for the relative errors computed as shown in (13). It can be seen that 218 out of the 897 errors (24.3 percent) are positive -- that is, the standard error is underestimated. The maximum relative error of underestimation is 40 percent, but the histogram shows that only 26 out of 897 (2.9 percent) are underestimates greater than 20 percent in relative terms. The median among the errors of underestimation is less than 10 percent. The maximum relative error of overestimation is 176.7 percent, i.e., the predicted standard error was 2.77 times as large as the jackknife estimate. It can be seen, however, that such extreme overpredictions are rare. In only 45 cases (five percent) did the predicted value of the standard error exceed the jackknife estimate by more than 50 percent. Finally, we observe that out of the total of 897 relative errors, 555 or 61.9 percent were less than 20 percent in absolute value.

[Exhibit 8 about here].

Exhibit 9 shows the relative errors for Model (10) above. The maximum relative error of underestimation is now only 35 percent and only 20 (2.2 percent) exceed 20 percent. Out of the total of 897, 571 (63.7 percent) are less than 20 percent in absolute value.

[Exhibit 9 about here].

VALIDATION. The prediction equations estimated for Models (7) and (10) were applied to the 455 cases that had been held in reserve for validation. (See pages 3-4, above). For Model (7) the average loss in the validation run was 0.5017, as opposed to 0.4627 in the original fitting. Curiously, only 16.7 percent of the relative errors computed according to Expression (13) were underestimates. The maximum relative underestimate was 34.58 percent. For Model (10), the average loss for the validation sample was 0.4805, as opposed to 0.4351 originally. The errors of underestimation constituted 21.1 percent of the total, with a maximum value of 39.2 percent. The validation run confirms our earlier conclusion that there is no great advantage in prediction accuracy of the more complicated Model (10) over the simple approach of adding the average logdeff to the logarithm of pq/n.

ESTIMATION OF DOMAIN TOTALS. In addition to the cognitive items, for which we have been discussing GVFs for variances of proportions, the Young Adult Literacy Survey provides information on 214 background items, covering the 13 different subpopulations of interest. Thus, there is a large number of weighted estimates of domain totals with their corresponding jackknife estimates of variance. A systematic sample of 947 estimated totals (72 or 73 values for each of the 13 domains) was selected for analysis, with many other values held in reserve for subsequent exploration and validation.

[To comply with restrictions on the length of papers for this publication we have omitted the remainder of this section, including Exhibits 10 through 13. A full copy of the paper is available upon request.]

USING PRIOR KNOWLEDGE OF THE DESIGN EFFECT. As a final exercise with the cognitive items we shall introduce the design effect into the right-hand side of the model. To do so exactly would lead to a perfect fit by tautology, but to be a bit more realistic, we assume that the analyst has a rough prior idea of the magnitude of the effect. In the discussion of theoretical motivations for Models (3) and (4), Wolter (1985) suggests that the specification is consistent with a constant deff for groups of items. The logarithm of the deff ranges from -1.23 to 1.814 in these data. Assume that it is possible a priori to place a proportion in one of the four categories of log-deff: (1) less than -.9, (2) greater than or equal to -.9 and less than zero, (3) greater than or equal to zero and less than +.9, (4) greater than or equal to +.9. In Exhibit 14 the variables LD1, LD2, and LD3 are indicators that the item falls into the first three of the log-deff categories above. It can be seen that with this important information the mean squared error is reduced by more than two thirds from that of previous models. The mean loss with the 3.1 penalty for underestimation that we have been using falls to 0.2618, only 62 percent of the previous minimum. This superiority is further borne out in the examination of relative errors of estimation of the standard error. The range is -41.7 to 24.4 percent, with 777 out of 897 (86.6 percent) of the relative errors less than 20 percent in absolute value.

[Exhibit 14 about here].

To be even more realistic (but not realistic enough for the real world) in Exhibit 15 we assume that the best that the analyst can do _a priori_ is to place the proportion in the two categories: negative vs. nonnegative logdeff -- in other words, deff less than, or greater than or equal to one. The categorical variable is LDSIGN. Exhibit 15 shows that the fit is still better than the models that do not involve log-deff. The mean loss of estimation error is calculated to be 0.3807. The range of relative error of estimation of the standard error is -74.8 to 36.2 percent with about two-thirds falling within plus or minus 20 percent.

[Exhibit 15 about here].

In summary, it would be very nice if we could identify some surrogate for the design effect of a variable that would enable us to classify it in at least a rough manner. In further research we shall experiment with the imposition of various amounts of measurement error on the classification variable to gauge the effects.

REFERENCES

Burt, V.L., and Cohen, S.B. (1984), "A Comparison of Methods to Approximate Standard Errors for Complex Survey Data." _Review of Public Data Use_, 12:159-168.

Cohen, S.B. (1979), "An Assessment of Curve Smoothing Strategies Which Yield Variance Estimates from Complex Survey Data," _Proceedings of the Survey Research Methods Section of the American Statistical Association_.

______, and Kalsbeek, W.D. (1981), NMCES Estimation and Sampling Variances in the Household Survey. National Center for Health Services Research.

Raiffa, H., and Schlaifer, R. (1961), Applied Statistical Decision Theory. Boston: Harvard University.

Wolter, K.M. (1985), Introduction to Variance Estimation. New York: Springer-Verlag.

EXHIBIT 1

UNWEIGHTED LEAST SQUARES LINEAR REGRESSION FOR VAR

PREDICTOR VARIABLES	COEFFICIENT	STD ERROR	STUDENT'S T	P
CONSTANT	1.857	1.841E-01	10.08	0.0000
Z	6.449E+03	161.2	40.02	0.0000

```
CASES INCLUDED          897        MISSING CASES      0
DEGREES OF FREEDOM      895
OVERALL F             1.601E+03    P VALUE        0.0000
ADJUSTED R SQUARED    0.6411
R SQUARED             0.6415
MEAN SQUARED ERROR     17.79
```

```
                         RESIDU VS PRED
RESIDU

  40.0 +
       :
       :          +     +
       :
       :
  20.0 +        +          +
       :       + + + + 2+                    +
       :       +  2    +23 +
       :      247*7+3++342+                   2
       :    2***6975223334 + + ++    +        +++++
   0.0 +  *****5**773343 3 +   2
       :  +8**736774642563+ +++ +        +  3 +  +
       :           ++  2++36++  +++   +  + +     3
       :                ++3+ +    ++ +        +
       :                                 +
 -20.0 +
       -+---------+---------+---------+---------+-
       0.0       9.0      18.0      27.0      36.0
              PRED   897 CASES PLOTTED
```

EXHIBIT 2

UNWEIGHTED LEAST SQUARES LINEAR REGRESSION FOR VAR

PREDICTOR VARIABLES	COEFFICIENT	STD ERROR	STUDENT'S T	P
CONSTANT	-3.662E-01	1.741E-01	-2.10	0.0336
PQN	2.094	3.925E-02	53.35	0.0000

```
CASES INCLUDED          897        MISSING CASES      0
DEGREES OF FREEDOM      895
OVERALL F             2.847E+03    P VALUE        0.0000
ADJUSTED R SQUARED    0.7605
R SQUARED             0.7608
MEAN SQUARED ERROR     11.87
```

```
                         RESIDU VS PRED
RESIDU

  40.0 +
       :
       :
       :               +   +
       :
  20.0 +           +
       :                   +   +      +
       :          + +2+ 2  +3 +
       :        +222 2222   2++2         +2
       :    +43444575557+532+223 + +    ++ ++
   0.0 +  *************675424 622            +
       :        23**84689985 6+23++   + 2 2 +2
       :                3++52++3 ++ + 2       2
       :                     2+       +  + + +
       :
 -20.0 +
       -+---------+---------+---------+---------+-
      -2.0       7.0      16.0      25.0      34.0
              PRED   897 CASES PLOTTED
```

EXHIBIT 3

UNWEIGHTED LEAST SQUARES LINEAR REGRESSION FOR LV

PREDICTOR VARIABLES	COEFFICIENT	STD ERROR	STUDENT'S T	P
CONSTANT	8.454	1.066E-01	79.33	0.0000
LZ	9.108E-01	1.351E-02	67.41	0.0000

```
CASES INCLUDED          897        MISSING CASES      0
DEGREES OF FREEDOM      895
OVERALL F             4.544E+03    P VALUE        0.0000
ADJUSTED R SQUARED    0.8353
R SQUARED             0.8354
MEAN SQUARED ERROR    2.102E-01
```

```
                         RESIDU VS PRED
RESIDU

   3.0 +
       :
       :
       :
       :                         +++
   1.0 +                    ++++23+++ +
       :                2553*********44*
       :            347**************8*4252
       :        +++3 678*********46***63356
       :       +  325243 325357*435878*435+
  -1.0 +        ++  +  2 ++++2+2++2+ 37+++
       :             +        +++  2 ++4
       :                  +
       :                 +
       :
  -3.0 +
       -+---------+---------+---------+---------+-
      -3.0      -1.0       1.0       3.0       5.0
              PRED   897 CASES PLOTTED
```

```
                       RANKITS VS RESIDU
RANKITS
   5.0 +
       :
       :
       :                                        ++
       :                                   2+3++
   2.0 +                                 **3
       :                               2**8
       :                             ***
       :                           ***
       :                       ****2
  -1.0 +                   4****
       :               +8***
       :          3+3+75
       :      +  +3
       :  +
  -4.0 +
       -+---------+---------+---------+---------+-
      -2.1      -1.2      -0.3       0.6       1.5
                          RESIDU
      APPROX. WILK-SHAPIRO 0.9769   897 CASES PLOTTED
```

EXHIBIT 4

UNWEIGHTED LEAST SQUARES LINEAR REGRESSION FOR LV

PREDICTOR VARIABLES	COEFFICIENT	STD ERROR	STUDENT'S T	P
CONSTANT	4.675E-01	1.618E-02	28.90	0.0000
LPQN	1.099	1.288E-02	85.33	0.0000

```
CASES INCLUDED          897        MISSING CASES      0
DEGREES OF FREEDOM      895
OVERALL F             7.281E+03    P VALUE        0.0000
ADJUSTED R SQUARED    0.8904
R SQUARED             0.8905
MEAN SQUARED ERROR    1.398E-01
```

EXHIBIT 5

UNWEIGHTED LEAST SQUARES LINEAR REGRESSION FOR LV

PREDICTOR VARIABLES	COEFFICIENT	STD ERROR	STUDENT'S T	P
CONSTANT	-1.316	1.837E-01	-7.16	0.0000
LP	1.193	1.619E-02	73.70	0.0000
LQ	1.222	1.909E-02	64.02	0.0000
LN	-9.322E-01	2.121E-02	-43.94	0.0000

```
CASES INCLUDED          897        MISSING CASES      0
DEGREES OF FREEDOM      893
OVERALL F             2.711E+03    P VALUE        0.0000
ADJUSTED R SQUARED    0.9007
R SQUARED             0.9011
MEAN SQUARED ERROR    1.266E-01
```

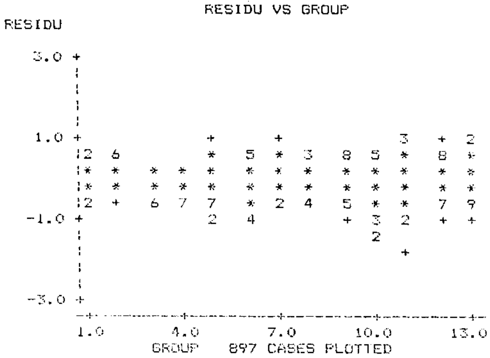
UNWEIGHTED LEAST SQUARES LINEAR REGRESSION FOR LV

PREDICTOR VARIABLES	COEFFICIENT	STD ERROR	STUDENT'S T	P
CONSTANT	-1.404	1.863E-01	-7.53	0.0000
LP	1.195	1.616E-02	73.97	0.0000
LQ	1.227	1.913E-02	64.15	0.0000
LN	-8.697E-01	3.221E-02	-27.00	0.0000
LCVWN	1.371E-01	5.325E-02	2.57	0.0098

CASES INCLUDED 897 MISSING CASES 0
DEGREES OF FREEDOM 892
OVERALL F 2.048E+03 P VALUE 0.0000
ADJUSTED R SQUARED 0.9014
R SQUARED 0.9018
MEAN SQUARED ERROR 1.258E-01

EXHIBIT 6

RESIDU VS GROUP



RESIDU VS SCALE

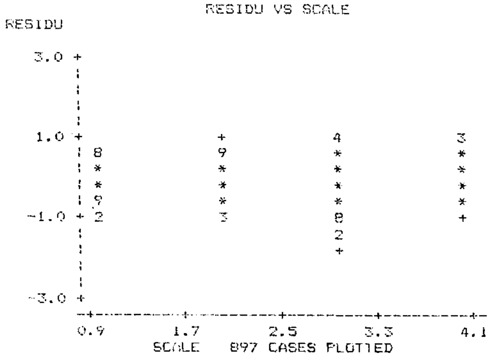


EXHIBIT 7

UNWEIGHTED LEAST SQUARES LINEAR REGRESSION FOR LV

PREDICTOR VARIABLES	COEFFICIENT	STD ERROR	STUDENT'S T	P
CONSTANT	-1.307	1.995E-01	-6.55	0.0000
LP	1.185	1.636E-02	72.43	0.0000
LQ	1.210	1.967E-02	61.49	0.0000
LN	-9.119E-01	2.450E-02	-37.23	0.0000
G3	-1.881E-01	4.755E-02	-3.96	0.0001
G4	-1.564E-01	4.795E-02	-3.26	0.0013
G5	-8.411E-02	4.515E-02	-1.86	0.0594
G8	-9.531E-02	4.667E-02	-2.04	0.0390
G9	-6.647E-02	4.637E-02	-1.43	0.1479
G10	-2.462E-01	4.598E-02	-5.35	0.0000

CASES INCLUDED 897 MISSING CASES 0
DEGREES OF FREEDOM 887
OVERALL F 951.8 P VALUE 0.0000
ADJUSTED R SQUARED 0.9052
R SQUARED 0.9062
MEAN SQUARED ERROR 1.209E-01

EXHIBIT 8

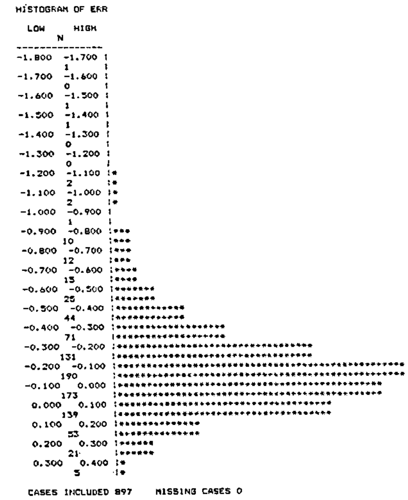


EXHIBIT 9

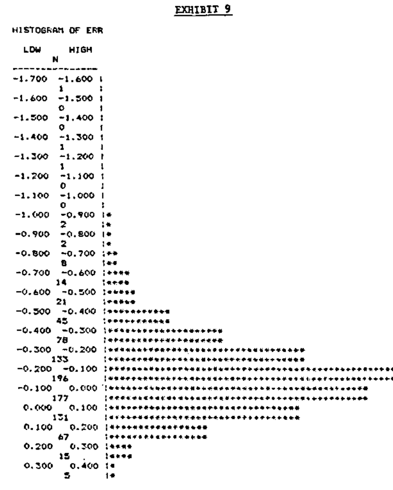


EXHIBIT 14

UNWEIGHTED LEAST SQUARES LINEAR REGRESSION FOR LV

PREDICTOR VARIABLES	COEFFICIENT	STD ERROR	STUDENT'S T	P
CONSTANT	4.252E-01	1.151E-01	3.69	0.0003
LP	1.056	1.003E-02	105.24	0.0000
LQ	1.073	1.168E-02	91.86	0.0000
LN	-0.969	1.257E-02	-77.07	0.0000
LD1	-2.107	1.212E-01	-17.38	0.0000
LD2	-1.240	3.225E-02	-38.45	0.0000
LD3	-5.547E-01	1.906E-02	-29.10	0.0000

CASES INCLUDED 897 MISSING CASES 0
DEGREES OF FREEDOM 890
OVERALL F 4.304E+03 P VALUE 0.0000
ADJUSTED R SQUARED 0.9665
R SQUARED 0.9667
MEAN SQUARED ERROR 4.279E-02

EXHIBIT 15

UNWEIGHTED LEAST SQUARES LINEAR REGRESSION FOR LV

PREDICTOR VARIABLES	COEFFICIENT	STD ERROR	STUDENT'S T	P
CONSTANT	-1.054	1.518E-01	-6.95	0.0000
LP	1.110	1.393E-02	79.73	0.0000
LQ	1.135	1.627E-02	69.74	0.0000
LN	-0.996	1.774E-02	-56.16	0.0000
LDSIGN	7.899E-01	3.829E-02	20.63	0.0000

CASES INCLUDED 897 MISSING CASES 0
DEGREES OF FREEDOM 892
OVERALL F 3.106E+03 P VALUE 0.0000
ADJUSTED R SQUARED 0.9327
R SQUARED 0.9330
MEAN SQUARED ERROR 8.583E-02