LINEAR CALIBRATION: A COMPARISON OF AN INVERSE REGRESSION METHOD OF CONFIDENCE INTERVALS WITH THE CLASSICAL REGRESSION METHOD OF CONFIDENCE INTERVALS

by

William T. McClelland, Jr.

Thesis submitted to the Graduate Faculty of the

Virginia Polytechnic Institute

in partial fulfillment for the degree of

MASTER OF SCIENCE

in

Statistics

APPROVED: ______________________________
Dr. R. G. Krutchkoff, Chairman

____
Dr. Boyd Harshbarger

____
Dr. R. H. Myers

November, 1967

Blacksburg, Virginia

TABLE OF CONTENTS

LIST OF TABLES

I. INTRODUCTION

In the discipline of statistics there are many methods, procedures, and general concepts that have become accepted over the years as a result of ingrained habit or because they are at least better than a haphazard guess at the answer to a given problem. Complacency and irresponsibility certainly must be brought to fore when one asks how such conditions could be allowed to occur in the realm of science. However, a certain degree must be attributed to acquiescence on the part of even the most responsible statistician. In many such instances it has been extremely difficult to ascertain the worth of a given procedure because it has been impossible to duplicate exact conditions a large number of times.

With the advent of the modern high-speed computer it is now possible to achieve a desired duplication process or to attain, quite simply, the heretofore insurmountable number of calculations required to determine the intrinsic worth of a given statistical procedure. The Monte Carlo study which we will present in the ensuing pages exhibits one of the many uses of a duplication process that can be achieved by simple computer techniques.

We maintain that in statistical research it is equally as important to ascertain the worth of old methods as it is to pursue new avenues of approach to the ever-expanding wealth of problems which continually confront the statistician. Perhaps the former of these two suggested branches of research may seem more than a little uncomfortable to those who have used unchecked procedures in the past, but to omit the needed investigation is to close one's eyes to reality.

In this spirit we shall consider not an original concept in statistical theory, but rather the worth of an old and widely-used procedure associated with simple linear regression as applied to the process of instrument calibration. We propose to compare the Classical and Inverse least squares methods of linear calibration by Monte Carlo methods using, as the criterion, confidence of the confidence interval $X \varepsilon (\hat{X} \pm L)$.

In a paper, "Classical and Inverse Regression Methods of Calibration", Dr. R. G. Krutchkoff (1967) compared these two methods of calibration by Monte Carlo techniques, using the standard criterion of mean squared error,

$$E(\hat{X} - X)^2 \quad , \qquad (1.1)$$

and showed the Inverse approach to be superior to the Classical approach. In his conclusion, he admitted that "It is evident that this conclusion opens more questions than it settles" and while finding that the Classical method was not optimal for the calibration problem, also pointed out that there was no answer to the question of whether or not the Inverse method could be shown to be the optimal approach.

Since that publication, the first in many years to even consider the point in question, Krutchkoff has found that the Inverse approach is better, or at least no worse, than the Classical approach from the standpoint of expected absolute deviation,

$$E|\hat{X} - X| \ . \tag{1.2}$$

Perhaps the comparisons referenced above have served to convince those who believe in the criteria given by (1.1) and (1.2), but there are many other criteria to consider, one such being confidence intervals. Surely, the employment of a different criterion in this investigation will serve to strengthen the conclusion previously obtained should it show the Inverse approach to be the better of the two.

The exact procedure used to generate the confidence intervals and ascertain the confidences will be given in succeeding chapters. The conclusion obtained is that the Inverse approach is superior to the widely-used Classical approach. In particular, we will show that if one calculates an Inverse confidence interval by substituting the

Inverse estimator, $\hat{X}_{inv}$, for the Classical estimator, $\hat{X}_{cl}$, in the Classical confidence interval $(\hat{X}_{cl} \pm L)$ then this interval will have a higher confidence than the Classical interval for the same length. This is not to say that the Inverse approach is optimal, although that may be the case. There is yet a great deal of research needed to find an optimal procedure with regard to the problem of linear instrument calibration.

Although the calibration problem is quite general and could be illustrated in terms of measuring temperature, pressure, acidity, air speed, and so on, we will use only the pressure gauge as an illustration in this introduction. The calibration of this gauge is to be linear. In simple terms, this implies that the increase or decrease in the gauge marking is proportional to the actual increase or decrease in the pressure to be measured. The procedure used to calibrate such a gauge is to subject it to two or more controlled pressures and note the markings that result. Then, with the use of these markings, the calibration parameters are calculated and the gauge is calibrated. From that point on, the gauge is used to determine unknown pressures simply by reading the numerical values assigned to the calibrated markings.

We will now define the linear relationship between pressure and gauge marking in terms of a statistical model.

For purposes of notation, it is convenient to denote the controlled variable (pressure) by x and the measured response variable (gauge marking) by y. Then, defining the parameters of the linear relationship, intercept and slope, to be α and β respectively, and defining the measurement error to be ε, the model becomes

$$y = \alpha + \beta x + \varepsilon \ . \tag{1.3}$$

In the following two sections, the two methods of calibration will be developed. It is necessary to understand the theoretical aspects of both methods before one is able to make an adequate comparison.

II. THE CLASSICAL APPROACH

Consider n, not necessarily distinct values of x where the measurement errors are assumed to be independent with a mean of zero and a variance of σ^2. The model given by (1.3) may be written as

$$y_i = \alpha + \beta x_i + \varepsilon_i \quad , \quad i = 1, 2, \ldots, n . \tag{2.1}$$

Equivalently, we may rewrite the above model as

$$y_i = \alpha' + \beta(x_i - \bar{x}) + \varepsilon_i \quad , \quad i = 1, 2, \ldots, n \quad , \tag{2.2}$$

where $\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$ and $\alpha' = \alpha + \beta\bar{x}$. Following the procedure of least squares, that is, minimizing the square of the error,

$$L = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} (y_i - \alpha' - \beta(x_i - \bar{x}))^2 \quad , \tag{2.3}$$

by solving the system of equations

$$\left.\begin{aligned} \frac{\partial L}{\partial \alpha'} &= 0 \\ \frac{\partial L}{\partial \beta} &= 0 \end{aligned}\right\} \tag{2.4}$$

we obtain the estimates

$$\hat{\beta} = \sum_{i=1}^{n} y_i(x_i-\bar{x}) / \sum_{i=1}^{n} (x_i-\bar{x})^2 \tag{2.5}$$

and

$$\hat{\alpha}' = \bar{y} . \tag{2.6}$$

The estimator of α in equation (2.1) would be

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} \tag{2.7}$$

where $\bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i$. The resulting least squares line to be used for the purpose of prediction is then

$$y = \hat{\alpha} + \hat{\beta}x \tag{2.8}$$

and the corresponding calibration equation becomes

$$x = \frac{y-\hat{\alpha}}{\hat{\beta}} . \tag{2.9}$$

For purposes of illustration, consider the pressure gauge. Suppose that two attainable and controlled pressures are substituted into equation (2.8) to obtain endpoints on a circular gauge. Then the distance between these two points is divided into certain intervals. However these lines or markings are meaningless in themselves and must have pressures assigned to them to make the gauge useful. By using equation (2.9) and substituting a numerical equivalent of the gauge marking for y, one is able to obtain an estimate of the corresponding unknown pressure, which is

$$\hat{X}_{cl} = \frac{Y-\hat{\alpha}}{\hat{\beta}} . \tag{2.10}$$

III. THE INVERSE APPROACH

Instead of obtaining the estimates for slope and intercept and then inverting the model, we may rewrite the model given by (1.3) to read

$$x = -\alpha/\beta + y/\beta - \varepsilon/\beta \ . \tag{3.1}$$

As before, we have n, not necessarily distinct, observations with measurement errors that we assume to be independent with a mean of zero and a variance of σ^2.

Letting $\gamma = -\alpha/\beta$, $\delta = 1/\beta$, and $\varepsilon' = -\varepsilon/\beta$, the model in (3.1) becomes

$$x_i = \gamma + \delta y_i + \varepsilon_i' \ , \quad i = 1, 2, \ldots, n \ , \tag{3.2}$$

or equivalently, we may rewrite this as

$$x_i = \gamma' + \delta(y_i - \bar{y}) + \varepsilon_i' \ , \quad i = 1, 2, \ldots, n \tag{3.3}$$

where $\gamma' = \gamma + \delta\bar{y}$ and $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$.

By minimizing the square of the error,

$$L = \sum_{i=1}^{n} \varepsilon_i'^2 = \sum_{i=1}^{n} (x_i - \gamma' + \delta(y_i - \bar{y}))^2 \ , \tag{3.4}$$

with respect to the parameters γ' and δ, we obtain, from the system

$$\left.\begin{aligned} \frac{\partial L}{\partial \gamma'} &= 0 \\ \frac{\partial L}{\partial \delta} &= 0 \end{aligned}\right\} , \tag{3.5}$$

the estimators

$$\hat{\delta} = \sum_{i=1}^{n} x_i(y_i-\bar{y}) / \sum_{i=1}^{n} (y_i-\bar{y})^2 \tag{3.6}$$

and

$$\hat{\gamma}' = \bar{x} . \tag{3.7}$$

From this, we may show the estimator of γ in (3.2) to be

$$\hat{\gamma} = \bar{x} - \hat{\delta}\bar{y} . \tag{3.8}$$

The resulting calibration equation becomes

$$x = \hat{\gamma} + \hat{\delta}y , \tag{3.9}$$

and by the same reasoning as was shown in the Classical approach, if one reads Y on the gauge, the estimate of the unknown pressure is

$$\hat{X}_{inv} = \hat{\gamma} + \hat{\delta}Y . \tag{3.10}$$

IV. THE CRITERION: CONFIDENCE INTERVALS

The estimates of the unknown pressure given by equations (2.10) and (3.10) will not, in general, be the same. This brings us to a point where we must choose between the two methods, and in doing so must first decide upon an appropriate criterion to be employed in making this choice.

Shortly after Krutchkoff presented his paper to Technometrics, Dr. John Mandel expressed an opinion that perhaps the length of a confidence interval for X or the probability that $|\hat{X}-X|$ does not exceed a given quantity might be used in lieu of the expected value of $(\hat{X}-X)^2$. In order to answer this question, the criterion used for comparison here will be the confidence of the confidence interval $(\hat{X}\pm L)$.

An important point in the consideration of confidence limits for X is that while there do exist exact limits for the Classical approach, the limits for the Inverse approach may be arrived at only by empirical methods at the present time.

Brownlee [1] derives the exact confidence limits for the Classical approach in a way similar to that given below. Consider a regression equation formed by observing n, not necessarily distinct pairs of observations, (x_i, y_i), with error terms which are independent and normally distributed with a mean of zero and a common variance of σ^2,

$$y = \hat{\alpha}' + \hat{\beta}(x-\bar{x}) \tag{4.1}$$

where $\hat{\alpha}'$, $\hat{\beta}$, and $\bar{x}$ are defined as before and are a direct consequence of the n observations. The problem is to predict the true value of X, by observing a value, Y, as well as to place some form of confidence in this estimate.

In lieu of using equation (2.10) we will invert the model given by (4.1) in order to arrive at the equivalent expression

$$\hat{X}_{cl} = \bar{x} + \frac{Y-\hat{\alpha}'}{\hat{\beta}} \tag{4.2}$$

which will be more useful in the ensuing derivation. If the true value of y is given by the expectation model

$$Ey = \eta = \alpha' + \beta(x-\bar{x}) , \tag{4.3}$$

we also have EY = η and a value of x, say ψ, corresponding to η given by solving equation (4.3),

$$\psi = \bar{x} + \frac{\eta-\alpha'}{\beta} , \tag{4.4}$$

which leads to the equation

$$\eta - \alpha' - \beta(\psi-\bar{x}) = 0 . \tag{4.5}$$

Defining a new variable, say ω, to be such that

$$\omega = Y - \hat{\alpha}' - \hat{\beta}(\psi-\bar{x}) \quad , \qquad (4.6)$$

we may show that

$$\begin{aligned} E\omega &= EY - E\hat{\alpha}' - (\psi-\bar{x})E\hat{\beta} \\ &= \eta - \alpha' - (\psi-\bar{x})\beta \qquad (4.7) \\ &= 0 \end{aligned}$$

and

$$\begin{aligned} V(\omega) &= V(Y) + V(\hat{\alpha}') + (\psi-\bar{x})^2 \, V(\hat{\beta}) \\ &= \sigma^2 + \sigma^2/n + (\psi-\bar{x})^2 \, \sigma^2/\sum_{i=1}^{n}(x_i-\bar{x})^2 \qquad (4.8) \\ &= \sigma^2[1 + 1/n + (\psi-\bar{x})^2/\sum_{i=1}^{n}(x_i-\bar{x})^2] \end{aligned}$$

since the estimators are mutually independent and the error term varies as σ^2. It is also the case that ω is a linear combination of three normally distributed random variables and therefore will be normally distributed itself, i.e.

$$\frac{\omega-E\omega}{\sqrt{V(\omega)}} \cap N(0,1) \qquad (4.9)$$

where $V(\omega)$ and $E\omega$ are given above and $E\omega = 0$. Thus, using the usual estimate of σ^2,

$$s^2 = \frac{1}{n-2} \sum_{i=1}^{n} (y_i - \hat{\alpha} - \hat{\beta} x_i)^2 , \tag{4.11}$$

the random variable in (4.9) will be distributed as a Student's -t with n-2 degrees of freedom, or, more specifically,

$$\frac{Y - \hat{\alpha}' - \hat{\beta}(\psi - \bar{x})}{s\sqrt{\left(1 + \frac{1}{n}\right) + \frac{(\psi - \bar{x})^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2}}} \cap t(n-2) . \tag{4.12}$$

Observe that we also have the probability statement

$$\Pr\left(-t_{\alpha/2} \leq \frac{Y - \hat{\alpha}' - \hat{\beta}(\psi - \bar{x})}{s\sqrt{\left(1 + \frac{1}{n}\right) + \frac{(\psi - \bar{x})^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2}}} \leq t_{\alpha/2} \right) = \alpha . \tag{4.13}$$

Let ψ_{1o} be the lower confidence limit of ψ, and it follows that

$$t_{\alpha/2} = \frac{Y - \hat{\alpha}' - \hat{\beta}(\psi_{1o} - \bar{x})}{s\sqrt{\left(1 + \frac{1}{n}\right) + \frac{(\psi_{1o} - \bar{x})^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2}}} \tag{4.14}$$

where $t_{\alpha/2}$ is used in order to simplify calculations because of the symmetry involved. Solving (4.14) for ψ_{1o}, by first obtaining the quadratic equation

$$\psi_{lo}^2\left[\hat{\beta}^2 - \frac{t^2s^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}\right] - 2\psi_{lo}\left[\frac{t^2s^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}x - \hat{\beta}^2\bar{x} - \hat{\beta}(Y-\hat{\alpha}')\right]$$

$$+ (Y - \hat{\alpha}' + \hat{\beta}\bar{x})^2 - t^2s^2\left[1 + \frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}\right] = 0 \qquad (4.15)$$

and finally,

$$\psi_{lo} = \bar{x} + \frac{\hat{\beta}(Y-\hat{\alpha}')}{\hat{\beta}^2 - \frac{t^2s^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}} - \frac{t\,s}{\hat{\beta}^2 - \frac{t^2s^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}}$$

$$\cdot\left\{\left[\hat{\beta}^2 - \frac{t^2s^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}\right]\left(1 + \frac{1}{n}\right) + \frac{(Y-\hat{\alpha}')^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}\right\}^{\frac{1}{2}}, \qquad (4.16)$$

where $t_{\alpha/2} = t$. Because of the symmetry involved, the upper confidence limit is derived in exactly the same way with the exception that ψ_{up} and $-t_{\alpha/2}$ are used. Using (4.16) and replacing t with -t for ψ_{up}, we have

$$Pr(\psi_{lo} \le \psi \le \psi_{up}) = \alpha \qquad (4.17)$$

and the derivation is complete.

Brownlee points out that a very interesting and useful result is that when

$$Q = \left(\frac{t_{\alpha/2}\hat{V}(\hat{\beta})}{\hat{\beta}} \right)^2 = \frac{t^2_{\alpha/2} s^2}{\hat{\beta}^2 \sum_{i=1}^{n} (x_i - \bar{x})^2} << 1.0 \qquad (4.18)$$

the following approximations to the exact limits are quite acceptable:

$$\psi_{up} \approx \bar{x} + \frac{Y - \hat{\alpha}'}{\hat{\beta}} + \frac{t_{\alpha/2} s}{|\hat{\beta}|} \left[\left(1 + \frac{1}{n}\right) + \frac{(Y - \hat{\alpha}')^2}{\hat{\beta}^2 \sum_{i=1}^{n} (x_i - \bar{x})^2} \right]^{\frac{1}{2}} \qquad (4.19)$$

$$\psi_{lo} \approx \bar{x} + \frac{Y - \hat{\alpha}'}{\hat{\beta}} - \frac{t_{\alpha/2} s}{|\hat{\beta}|} \left[\left(1 + \frac{1}{n}\right) \quad \frac{(Y - \hat{\alpha}')^2}{\hat{\beta}^2 \sum_{i=1}^{n} (x_i - \bar{x})^2} \right]^{\frac{1}{2}} .$$

These approximations follow directly from (4.16) as Q becomes small. Assuming that Q will be such that (4.19) may be used, we have

$$\Pr\left(\left| \psi - \hat{X}_{cl} \right| \leq \frac{t_{\alpha/2} s}{|\hat{\beta}|} \left[\left(1 + \frac{1}{n}\right) + \frac{(Y - \hat{\alpha}')}{\hat{\beta}^2 \sum_{i=1}^{n} (x_i - \bar{x})^2} \right]^{\frac{1}{2}} \right) = \alpha \qquad (4.20)$$

where $\hat{X}_{cl} = \bar{x} + \frac{Y-\hat{\alpha}'}{\hat{\beta}}$

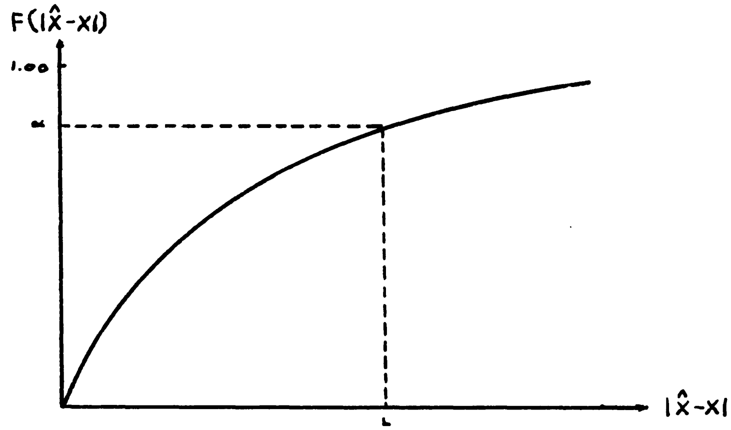
Suppose that a symmetric interval for the Classical approach was of the form

$$\Pr\left(\hat{X}_{cl} - L \leq X \leq \hat{X}_{cl} + L\right) = \alpha \ . \qquad (4.21)$$

This could be written as

$$\Pr\left(|\hat{X}_{cl} - X| \leq L\right) = \alpha \ , \qquad (4.22)$$

a form which may assist one in understanding the criterion to be used in the Monte Carlo investigation. The empirical cumulative distribution function for $|\hat{X}_{cl}-X|$ may be derived by Monte Carlo methods and would be similar to the graph given below:



(4.23)

The cumulative distribution function, $F\left(\left|\hat{X}_{inv}-X\right|\right)$, may also be derived empirically and comparing the two curves is identical to comparing the two confidences for some fixed value of L. This fact will prove useful when we discuss the results in the section dealing with comparisons. Note that L in (4.23) compares with the upper limit of $|\psi-\hat{X}_{cl}|$ in (4.20) where α is the confidence.

The majority of our comparisons will deal with the approximate confidence interval where we have $Q \ll 1.0$. The cases where Q becomes too large will be rare and will be dealt with when they occur. More specifically, what we will show is that by replacing the Classical estimate, $\hat{X}_{cl}$, by the Inverse estimate, $\hat{X}_{inv}$, in the approximate confidence interval so that we have

$$\Pr\left(|X-\hat{X}_{cl}| \leq \frac{t_{\alpha/2}s}{|\hat{\beta}|}\left[\left(1+\frac{1}{n}\right) + \frac{(Y-\hat{\alpha}')}{\hat{\beta}^2 \sum_{i=1}^{n}(x_i-\bar{x})^2}\right]^{\frac{1}{2}}\right) = p_1 \qquad (4.24)$$

and

$$\Pr\left(|X-\hat{X}_{inv}| \leq \frac{t_{\alpha/2}s}{|\hat{\beta}|}\left[\left(1+\frac{1}{n}\right) + \frac{(Y-\hat{\alpha}')}{\hat{\beta}^2 \sum_{i=1}^{n}(x_i-\bar{x})^2}\right]^{\frac{1}{2}}\right) = p_2 \ , \qquad (4.25)$$

for fixed and identical values of L, p_2 will be uniformly larger than p_1 for all values of X, where

$$L = \frac{t_{\alpha/2}s}{|\hat{\beta}|}\left[\left(1+\frac{1}{n}\right) + \frac{(Y-\hat{\alpha}')}{\hat{\beta} \sum_{i=1}^{n}(x_i-\bar{x})^2}\right]^{\frac{1}{2}} . \qquad (4.26)$$

Note that (4.24) and (4.25) may be written as

$$\Pr\left(\hat{X}_{cl} - L \leq X \leq \hat{X}_{cl} + L\right) = p_1 \tag{4.27}$$

and

$$\Pr\left(\hat{X}_{inv} - L \leq X \leq \hat{X}_{inv} + L\right) = p_2 \tag{4.28}$$

and our comparisons will actually involve confidences such that (4.27) and (4.28) cover the true value of X .

V. MONTE CARLO TECHNIQUES

Reference has been made to a so-called Monte Carlo method to be used in the comparison of the two approaches to the problem of linear instrument calibration. We will now give the details of exactly what was involved in the Monte Carlo technique.

As a basis of our investigation we used the model given by (1.3), i.e. $y = \alpha + \beta x + \varepsilon$, with parameters $\alpha = 0.0$ and $\beta = 0.5$ (a line through the origin with a slope of 30°). The range, only a scale factor in this problem, was set equal to [0,1] and the standard deviation was assumed to be 10% of the range (i.e. $\sigma = 0.1$). Ten observations were taken at each of the endpoints, $X = 0$ and $X = 1$, and the twenty values of y that resulted were used to calculate $\hat{\alpha}$, $\hat{\beta}$, $\hat{\delta}$, and $\hat{\gamma}$ from equations (2.5), (2.7), (3.6), and (3.8) respectively.

Following this, six known values of x ($X = 0.0, 0.2, 0.4, 0.6, 0.8, 1.0$) were used to obtain six values of Y from (1.3). For each of these Y values, two estimates were calculated according to (2.10) and (3.10) so that we had the Inverse and Classical estimates associated with each of the six true values of x within our range.

At this point we had access to six true values of x, as well as two estimates for each, either of which might be used if x were unknown. The value of Q , given by (4.18), was then computed using the percentage point of the t-distribution corresponding to the intended theoretical confidence. The value of Q did not determine whether or not the approximate confidence limits would be used, but was calculated in order to give some insight as to the accuracy of the approximations. We then calculated L , according to equation (4.26), and determined whether or not we had $X \varepsilon \left(\hat{X}_{cl} \pm L\right)$ and/or $X \varepsilon \left(\hat{X}_{inv} \pm L\right)$.

This entire procedure was repeated 10,000 times, each time computing new estimates of the parameters, new values of Q and L , and determining whether or not the intervals covered X for each of the six true values of x . The percentage of time that X was covered by $(\hat{X} \pm L)$ was then calculated for each approach using the six x-values. The standard errors for each of these percentages were also found. This was accomplished by using 100 sets, of 100 percentages each, and the following formula:

$$s = \sqrt{\frac{\Sigma x_i^2 - \frac{(\Sigma x_i)^2}{100}}{9900}} \qquad (5.1)$$

where x_i = the percentage of times $(\hat{X} \pm L)$ covered the true X for each set of 100 repetitions.

The above procedure was repeated for each of five theoretical confidences, .95, .90, .80, .70, and .60 .

Since the value of L , which is one-half of the length of the symmetric confidence interval, is different for each set of estimates, only its average was reported in the tables.

This, in all essence, is a basic outline of the Monte Carlo procedure used to obtain the results that follow. The endpoint design was used to calculate the estimates of α , β , δ and γ , ten observations being taken at each endpoint of the range [0,1], and representative values of x within the range of the endpoints were simulated in order to compare the Inverse and Classical estimators. The investigation was undertaken at the Virginia Polytechnic Institute using an I.B.M. 7040 computer and a program generating pseudo-normal random numbers. Monte Carlo experiments of 10,000 sets each were conducted for the theoretical confidences given above. As an illustration of the results Table I, which will be used as a basis for most of the subsequent comparisons, shows the different confidences that the interval $(\hat{X} \pm L)$ will cover the true value of x , using both the Classical and Inverse methods.

There are several significant features of Table I that should be discussed at this time. The true values of the parameters involved in the compilation of this representative table were $\alpha = 0.0$, $\beta = 0.5$, and $\sigma = 0.1$, which are given at the top of the table. The confidence, or more

TABLE I: COMPARISON OF AVERAGE CONFIDENCE FOR CLASSICAL AND INVERSE CALIBRATION CONFIDENCE INTERVALS

$\alpha=0.0$, $\beta=0.5$, $\sigma=0.1$, $\beta/\sigma=5.0$
Design: 10(X=0), 10(X=1)

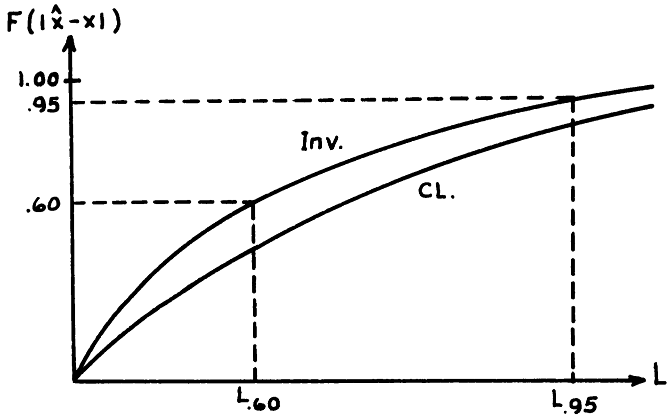
	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.95}$	0.951	0.955	0.948	0.952	0.954	0.952
$(\hat{X}_{inv}\pm L)_{.95}$	0.966	0.970	0.968	0.970	0.969	0.963
Average L	0.441	0.434	0.430	0.430	0.434	0.440
	All Standard Errors are 0.002					
$(\hat{X}_{cl}\pm L)_{.90}$	0.902	0.905	0.902	0.906	0.906	0.902
$(\hat{X}_{inv}\pm L)_{.90}$	0.921	0.932	0.935	0.939	0.932	0.922
Average L	0.363	0.358	0.355	0.355	0.358	0.363
	All Standard Errors are between 0.002 and 0.003					
$(\hat{X}_{cl}\pm L)_{.80}$	0.807	0.805	0.800	0.797	0.798	0.800
$(\hat{X}_{inv}\pm L)_{.80}$	0.833	0.852	0.854	0.850	0.842	0.828
Average L	0.278	0.274	0.272	0.272	0.274	0.278
	All Standard Errors are 0.004					
$(\hat{X}_{cl}\pm L)_{.70}$	0.702	0.699	0.709	0.701	0.704	0.695
$(\hat{X}_{inv}\pm L)_{.70}$	0.736	0.751	0.766	0.763	0.752	0.730
Average L	0.224	0.220	0.218	0.218	0.220	0.224
	All Standard Errors are between 0.004 and 0.005					
$(\hat{X}_{cl}\pm L)_{.60}$	0.602	0.606	0.597	0.598	0.593	0.593
$(\hat{X}_{inv}\pm L)_{.60}$	0.636	0.656	0.661	0.660	0.648	0.627
Average L	0.180	0.177	0.176	0.176	0.177	0.180
	All Standard Errors are between 0.004 and 0.005					

specifically, the value of the percentage point of the Student's t-distribution was varied so as to detect a prevalent pattern in a range from 60% to 95% confidence.

Of primary importance is the value of Q, as given by (4.18). Recall that we required that $Q << 1.0$ in order that the approximate interval was acceptable. The data reported in Table I involved five Monte Carlo experiments of 10,000 repetitions each, and of the 50,000 Q values calculated, only 18 were above 0.1, with a maximum of 0.163. This shows that the use of the approximation was reasonable in this case.

The first segment of Table I considers a theoretical confidence of 95% developed through the use of a Student's t-distribution with 18 degrees of freedom. The Classical method shows confidences that are consistently within one or two standard errors of the actual 0.950 value while, for the identical length, the Inverse method has a uniformly larger confidence. Notice that the Inverse confidence is between one and one and a half percent above the Classical confidence at the end points of the range and about two percent higher in the middle of the range where the average length of the interval had decreased. At the 90% level, the improvement associated with the Inverse approach is much more noticable. The difference between the two methods is about two percent at the end points of the range and increases to about three and a half percent in the middle of the range where the difference between the true X and its estimator would be more easily detected. The confidences for the Classical approach are again within two standard errors of the theoretical value.

As the intended confidence decreases toward 60%, the Inverse approach gives evidence of being substantially more reliable than the Classical approach. Notice that at X = 0.4 where the theoretical confidence is 60%, the Inverse confidence is six percent higher than the Classical confidence. The average length of the interval decreases in the middle of the range for all confidences without exception. If we were to plot the cumulative distribution function for each method, the graph would be similar to that given below.



(5.2)

where the greatest difference would be toward the middle of the confidence range.

The tables which will be presented in the ensuing pages will follow the same general format as Table I and will serve to substantiate the supposition that the proposed

Inverse approach is the better method to employ in the problem of linear instrument calibration from the standpoint of the confidence interval criterion.

Using the Monte Carlo techniques to ascertain the effects of certain parameters on our conclusion requires only minor changes of the model parameters. To investigate the effect of slope on the conclusion, for instance, we will hold all other parameters constant, say $\alpha = 0.0$ and $\sigma = 0.1$, and design at 10(X=0), 10(X=1). Then the slope will be changes by representative and meaningful increments. From the results we will show whether or not the slope does effect the conclusion as to which of the two approaches is the better.

In the following sections such investigations will be made and the results tabulated quite extensively. The parameters α, β, and σ will be varied, as well as the design, so as to show how different variations effect the conclusion. Robustness to certain assumptions will also be dealt with, but any discussion of this aspect will be left until the appropriate time.

VI. COMPARISONS

Throughout this section we intend to show that the probability of $(\hat{X}_{inv} \pm L)$ covering the true value of X will exceed the probability that $(\hat{X}_{cl} \pm L)$ covers the true X, regardless of the parameters and design involved. Using Monte Carlo methods similar to those discussed in the previous section, we will vary one parameter at a time, keeping all others constant, so as to determine what effect each has on the conclusion derived from Table I.

(A) The Effect of Intercept

Table I gave confidences in each of the two approaches to the calibration problem using the parameters $\alpha = 0.0$, $\beta = 0.5$ and $\sigma = 0.1$ with a design of ten observations per endpoint of the range [0,1]. We will now hold everything constant except the intercept in order to ascertain its effect on our conclusion.

It would seem reasonable that changing α should not effect the conclusion of Table I since the location parameter will cancel out in the differences $(\hat{X}_{cl}-X)$ and $(\hat{X}_{inv}-X)$. The proof of this is the following:

Recalling equation (2.10), we have

$$\hat{X}_{cl} - X = \frac{Y-\hat{\alpha}}{\hat{\beta}} - X \qquad (6.1)$$

which, from (1.3) and (2.7), becomes

$$\hat{X}_{cl} - X = \frac{\alpha' + \beta(X-\bar{x}) + \varepsilon' - \bar{y} + \hat{\beta}\bar{x}}{\hat{\beta}} - X \qquad (6.2)$$

$$= \left(\frac{\beta}{\hat{\beta}} - 1\right)(X-\bar{x}) + \frac{\varepsilon'}{\hat{\beta}} + \frac{\alpha'-\bar{y}}{\hat{\beta}} \quad ,$$

where ε' is the error made in observing Y. But since

$$\bar{y} = \alpha + \beta\bar{x} + \bar{\varepsilon} \quad , \qquad (6.3)$$

it follows that

$$\frac{\alpha'-\bar{y}}{\hat{\beta}} = \frac{(\alpha + \beta\bar{x}) - (\alpha + \beta\bar{x} + \bar{\varepsilon})}{\hat{\beta}} = -\bar{\varepsilon}/\hat{\beta} \ . \qquad (6.4)$$

The identity in (6.1) is then

$$\hat{X}_{cl} - X = \left(\frac{\beta}{\hat{\beta}} - 1\right)(X-\bar{x}) + \frac{1}{\hat{\beta}}(\varepsilon'-\bar{\varepsilon}) \quad , \qquad (6.5)$$

where $\bar{\varepsilon}$ is the average of the errors made in the calibration of the instrument.

For the Inverse approach, using (3.10), we have

$$\hat{X}_{inv} - X = \hat{\gamma} + \hat{\delta}Y - X \ . \qquad (6.6)$$

Using (2.5) and (3.6), it may easily be verified that

$$\hat{\delta} = \frac{r^2}{\hat{\beta}} \qquad (6.7)$$

where r^2 , the sample correlation coefficient between x and

y, is defined to be

$$r^2 = \frac{\left[\sum_{i=1}^{n} (x_i-\overline{x})(y_i-\overline{y})\right]^2}{\sum_{i=1}^{n} (x_i-\overline{x})^2 \sum_{i=1}^{n} (y_i-\overline{y})^2} \quad . \tag{6.8}$$

Then, from (3.8), (6.6), and (6.7) we have

$$\hat{X}_{inv} - X = \frac{\hat{\beta}\overline{x} - r^2\overline{y} + r^2 Y}{\hat{\beta}} - X \quad , \tag{6.9}$$

which reduces to

$$\hat{X}_{inv} - X = \left(\frac{\beta r^2}{\hat{\beta}} - 1\right)\left(X-\overline{x}\right) + \frac{r^2\varepsilon'}{\hat{\beta}} + \frac{r^2(\alpha'-\overline{y})}{\hat{\beta}} \tag{6.10}$$

and finally to

$$\hat{X}_{inv} - X = \left(\frac{\beta r^2}{\hat{\beta}} - 1\right)\left(X - \overline{x}\right) + \frac{r^2}{\hat{\beta}}(\varepsilon'-\overline{\varepsilon}) \quad , \tag{6.11}$$

where ε', $\overline{\varepsilon}$, and r^2 are defined as before.

Therefore, by (6.5) and (6.11), we have shown that $(\hat{X}_{cl}-X)$ and $(\hat{X}_{inv}-X)$ are independent of the intercept. This says that $F(|\hat{X}_{inv}-X|)$ and $F(|\hat{X}_{cl}-X|)$ will be independent of α, which is identical with saying that the resulting confidences will be independent of α.

Equation (4.26) may appear to say that the length of the interval does depend on the choice of α. This is easily cleared up by the fact that we have

$$\begin{aligned}(Y-\hat{\alpha}') &= Y - \overline{y} \\ &= \alpha + \beta X + \varepsilon' - \alpha - \beta X - \overline{\varepsilon} \qquad (6.12) \\ &= \beta(X-\overline{x})+(\varepsilon'-\overline{\varepsilon}) \ ,\end{aligned}$$

and we have shown that the length of the symmetric interval, $(\hat{X}\pm L)$, as well as the confidences, is completely independent of α .

Because of the above theoretical results, an extensive investigation of the effect of intercept is not needed. However, we have chosen two extreme values, $\alpha = 10.0$ and $\alpha = -5.0$, to report at this time. Table II was obtained by using Monte Carlo experiments of 10,000 repetitions for each of five theoretical confidences associated with the value of α in question. The other parameters and design were unchanged from the values assigned in Table I. It is interesting to note that in a total of 100,000 values of Q which were calculated for Table II, only 24 were above 0.1, the maximum being 0.183.

Since all of the entries in Table II are within three standard errors of their counterparts in Table I, and since we know by our theoretical results that α will not effect either method, there is no discussion of Table II necessary. The conclusion that we obtained in Table I has remained unchanged by the choice of intercept.

TABLE II: COMPARISON OF AVERAGE CONFIDENCE FOR CLASSICAL AND INVERSE CALIBRATION CONFIDENCE INTERVALS; EFFECT OF INTERCEPT

α=10.0, β=0.5, σ=0.1, β/σ=5.0
Design: 10(X=0), 10(X=1)

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.95}$	0.956	0.957	0.958	0.953	0.955	0.957
$(\hat{X}_{inv}\pm L)_{.95}$	0.965	0.970	0.972	0.972	0.970	0.967
Average L	0.441	0.434	0.431	0.431	0.434	0.441
All Standard Errors are 0.002						
$(\hat{X}_{cl}\pm L)_{.90}$	0.905	0.911	0.911	0.899	0.903	0.904
$(\hat{X}_{inv}\pm L)_{.90}$	0.922	0.935	0.942	0.933	0.930	0.920
Average L	0.363	0.357	0.355	0.355	0.357	0.363
All Standard Errors are between 0.002 and 0.003						
$(\hat{X}_{cl}\pm L)_{.80}$	0.797	0.804	0.803	0.807	0.807	0.800
$(\hat{X}_{inv}\pm L)_{.80}$	0.828	0.842	0.854	0.856	0.851	0.831
Average L	0.278	0.274	0.272	0.272	0.274	0.278
All Standard Errors are 0.004						
$(\hat{X}_{cl}\pm L)_{.70}$	0.698	0.707	0.701	0.701	0.696	0.706
$(\hat{X}_{inv}\pm L)_{.70}$	0.731	0.754	0.762	0.759	0.746	0.737
Average L	0.223	0.220	0.218	0.218	0.220	0.223
All Standard Errors are between 0.004 and 0.005						
$(\hat{X}_{cl}\pm L)_{.60}$	0.603	0.606	0.602	0.605	0.604	0.601
$(\hat{X}_{inv}\pm L)_{.60}$	0.636	0.656	0.664	0.671	0.655	0.641
Average L	0.181	0.178	0.177	0.177	0.178	0.181
All Standard Errors are between 0.005 and 0.006						

$\alpha=-5.0$, $\beta=0.5$, $\sigma=0.1$, $\beta/\sigma=5.0$
Design: 10(X=0), 10(X=1)

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.95}$	0.951	0.952	0.955	0.955	0.949	0.952
$(\hat{X}_{inv}\pm L)_{.95}$	0.964	0.968	0.973	0.973	0.966	0.961
Average L	0.441	0.434	0.431	0.431	0.434	0.441

All Standard Errors are 0.002

$(\hat{X}_{cl}\pm L)_{.90}$	0.904	0.900	0.903	0.906	0.906	0.906
$(\hat{X}_{inv}\pm L)_{.90}$	0.923	0.930	0.933	0.937	0.933	0.928
Average L	0.364	0.358	0.356	0.356	0.358	0.364

All Standard Errors are 0.003

$(\hat{X}_{cl}\pm L)_{.80}$	0.795	0.809	0.800	0.802	0.804	0.805
$(\hat{X}_{inv}\pm L)_{.80}$	0.825	0.852	0.852	0.854	0.847	0.833
Average L	0.279	0.275	0.272	0.273	0.275	0.279

All Standard Errors are 0.004

$(\hat{X}_{cl}\pm L)_{.70}$	0.698	0.701	0.703	0.696	0.701	0.709
$(\hat{X}_{inv}\pm L)_{.70}$	0.734	0.753	0.766	0.760	0.751	0.738
Average L	0.223	0.219	0.218	0.218	0.219	0.223

All Standard Errors are between 0.004 and 0.005

$(\hat{X}_{cl}\pm L)_{.60}$	0.599	0.603	0.599	0.607	0.611	0.596
$(\hat{X}_{inv}\pm L)_{.60}$	0.639	0.655	0.663	0.668	0.659	0.636
Average L	0.180	0.178	0.176	0.176	0.178	0.180

All Standard Errors are 0.005

(B) The Effect of Slope

As the second step in our investigation of the different model parameters with regard to the existing criterion, we varied the slope of the line. Values of $\beta = 0.25$, 0.30, 0.75, 1.00, 1.50, 5.00, and 20.00 were investigated at five confidence levels with Monte Carlo experiments of 10,000 repetitions each. All of the other parameters and the design remained as in Table I, which allows us to use that table for comparison.

From Table III it is quite apparent that the effect of slope is significant. Just as apparent is the fact that it effects each of the two approaches differently. Notice, for example, that in Table I, where $\beta = 0.5$ and $X = 0.4$ for a theoretical confidence of 60%, the Classical confidence is 0.597 and the Inverse confidence is 0.661, while the average value of L is 0.176. When the slope is 0.25 under the same conditions in Table III, the Classical confidence is 0.604 and the Inverse confidence is 0.805 for an average L of 0.367.

For this table, the number of times Q exceeded 0.2 is given, along with the maximum value it obtained. This is done to indicate where we should interpret the results with caution. Observe that as the value of slope tends toward zero (i.e. the line becomes more horizontal) Q, and hence the average value of L, tends to be large. This is exactly what one should expect since the estimate of slope in equation (4.26) is in the denominator. Thus, as the slope decreases to 0.25 the approximation to the interval is not generally valid since one-fourth of the Q-values are above 0.2 with a maximum of 5.92 for the 95% interval. One should therefore interpret the results in this case with caution.

Although the average L is inflated for the case where $\beta = 0.25$, there is still much to be learned by noting the results. While the confidence in the Classical approach increases slightly, there is a substantial increase in the confidence associated with the Inverse method. For lower theoretical confidences the number of times Q exceeded 0.2 decreases and the results become more reliable. Observe that for confidences of 80% and below, there is no question as to the reliability of the approximate interval.

It can hardly go without notice that for these lower confidences the gap between the two methods becomes significantly larger. At the 80% confidence level, the Inverse confidence is about 9% above the Classical confidence at the endpoints and about 13% higher in the middle of the range. At the 60% level, the Inverse confidence is higher by 11-12% at the endpoints and as much as 20% in the middle of the range.

As the slope is increased to 0.30, the number of times Q exceeds 0.2 decreases significantly and the use of the approximate interval becomes more reasonable. The average L decreases and the gap between the two methods becomes smaller, but there is still a substantial increase in the Inverse confidence as opposed to little or no increase in the Classical confidence. When the theoretical confidence is set at 95%, the Inverse confidence is about 2% above the Classical at the endpoints and 3% higher at mid-range. At the 60% level, the Inverse confidence is 9% higher at the endpoints and 16% higher in the middle of the range.

Increasing the slope past 0.5 yields a still smaller length as one would expect from (4.26). The confidences for both methods begin to converge on the theoretical confidence although the Inverse approach retains a uniformly higher confidence for $\beta \leq 1.0$. Consider the 60% level for $\beta = 1.0$ where we would expect the greatest difference between the two methods. The Inverse confidence is only 1% higher than the Classical confidence at the endpoints and 2% higher at mid-range. At the 95% level, the differences are negligible for the entire range. When the true value of β is 20.0, rather an extreme case, the interval length reduces substantially and the confidences are almost identical.

Although the effect of slope significantly effects both methods, the confidence for the Inverse approach remains uniformly larger, or at least no smaller, than the confidence for the Classical approach for the same length.

Thus, since one will never know the true value of slope, the Inverse method and its confidence interval would be the better of the two because it protects against the effect of a small slope and is certainly no worse for a large slope.

Since we have investigated the effect of slope and intercept, the next consideration should be error variance. However, investigating the effect of the error term independently of slope is superfluous. Krutchkoff (1967) noted that "the effect of slope and variance depend only on the ratio β/σ." We will now demonstrate that this is the case here too.

In Table IV, the ratio was varied and Monte Carlo experiments of 10,000 repetitions for each of five theoretical confidences were conducted. From that table; $\beta/\sigma = 5.0$ is to be compared to $\beta/\sigma = 5.0$ of Table I, $\beta/\sigma = 7.5$ is to be compared to $\beta/\sigma = 7.5$ of Table III, and $\beta/\sigma = 10.0$ is to be compared to $\beta/\sigma = 10.0$ of Table III.

Observe that the entries in Table I are within one or two standard errors of the corresponding entries in Table IV where $\beta/\sigma = 5.0$. The same correspondance exists between Table III and IV where $\beta/\sigma = 7.5$ and $\beta/\sigma = 10.0$. Because of the fact that the effects of slope and error variance depend only on β/σ , the discussion of slope is adequate for the effect of error variance and the latter will not

TABLE III: COMPARISON OF AVERAGE CONFIDENCE FOR CLASSICAL AND INVERSE CALIBRATION CONFIDENCE INTERVALS; EFFECT OF SLOPE

α=0.0, β=0.25, σ=0.1, β/σ=2.5
Design: 10(X=0), 10(X=1)

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.95}$	0.961	0.961	0.965	0.964	0.959	0.961
$(\hat{X}_{inv}\pm L)_{.95}$	0.983	0.988	0.993	0.994	0.989	0.984
Average L	0.922	0.907	0.899	0.899	0.908	0.921

All Standard Errors are between 0.001 and 0.002
2237 Q values were above 0.2, max Q = 5.920

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.90}$	0.914	0.913	0.914	0.917	0.916	0.917
$(\hat{X}_{inv}\pm L)_{.90}$	0.958	0.975	0.979	0.983	0.977	0.961
Average L	0.760	0.749	0.742	0.742	0.749	0.760

All Standard Errors are between 0.002 and 0.003
700 Q values were above 0.2, max Q = 0.744

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.80}$	0.808	0.819	0.810	0.812	0.815	0.816
$(\hat{X}_{inv}\pm L)_{.80}$	0.890	0.929	0.943	0.942	0.926	0.891
Average L	0.583	0.573	0.568	0.568	0.573	0.582

All Standard Errors are between 0.003 and 0.004
89 Q values were above 0.2, max Q = 0.523

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.70}$	0.707	0.709	0.710	0.705	0.707	0.716
$(\hat{X}_{inv}\pm L)_{.70}$	0.813	0.859	0.886	0.882	0.855	0.813
Average L	0.465	0.458	0.454	0.454	0.458	0.465

All Standard Errors are between 0.004 and 0.005
15 Q values were above 0.2, max Q = 0.453

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.60}$	0.605	0.607	0.604	0.612	0.616	0.603
$(\hat{X}_{inv}\pm L)_{.60}$	0.718	0.777	0.805	0.813	0.781	0.724
Average L	0.376	0.370	0.367	0.367	0.370	0.377

All Standard Errors are between 0.004 and 0.005
2 Q values were above 0.2, max Q = 0.258

α=0.0, β=0.3, σ=0.1, β/σ=3.0
Design: 10(X=0), 10(X=1)

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.95}$	0.960	0.961	0.962	0.959	0.960	0.961
$(\hat{X}_{inv}\pm L)_{.95}$	0.979	0.987	0.989	0.989	0.985	0.981
Average L	0.754	0.742	0.736	0.736	0.742	0.754

All Standard Errors are between 0.001 and 0.002
542 Q values were above 0.2, max Q = 0.697

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.90}$	0.910	0.918	0.919	0.906	0.910	0.911
$(\hat{X}_{inv}\pm L)_{.90}$	0.946	0.962	0.972	0.968	0.962	0.945
Average L	0.621	0.610	0.606	0.606	0.611	0.620

All Standard Errors are between 0.002 and 0.003
91 Q values were above 0.2, max Q = 0.416

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.80}$	0.804	0.809	0.809	0.813	0.813	0.806
$(\hat{X}_{inv}\pm L)_{.80}$	0.870	0.901	0.919	0.921	0.906	0.869
Average L	0.475	0.468	0.464	0.464	0.467	0.475

All Standard Errors are between 0.002 and 0.003
6 Q values were above 0.2, max Q = 0.244

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.70}$	0.702	0.710	0.706	0.705	0.699	0.709
$(\hat{X}_{inv}\pm L)_{.70}$	0.783	0.827	0.845	0.844	0.821	0.790
Average L	0.381	0.375	0.372	0.372	0.375	0.381

All Standard Errors are between 0.004 and 0.005

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.60}$	0.607	0.608	0.605	0.608	0.608	0.606
$(\hat{X}_{inv}\pm L)_{.60}$	0.692	0.738	0.763	0.767	0.737	0.695
Average L	0.308	0.303	0.301	0.301	0.303	0.308

All Standard Errors are between 0.004 and 0.005

$\alpha=0.0$, $\beta=.75$, $\sigma=0.1$, $\beta/\sigma=7.5$
Design: 10(X=0), 10(X=1)

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.95}$	0.954	0.955	0.956	0.950	0.953	0.953
$(\hat{X}_{inv}\pm L)_{.95}$	0.959	0.963	0.963	0.961	0.961	0.960
Average L	0.292	0.287	0.285	0.285	0.287	0.292

All Standard Errors are 0.002

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.90}$	0.903	0.910	0.910	0.896	0.900	0.901
$(\hat{X}_{inv}\pm L)_{.90}$	0.911	0.921	0.925	0.913	0.914	0.910
Average L	0.240	0.237	0.235	0.235	0.237	0.240

All Standard Errors are 0.003

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.80}$	0.794	0.802	0.801	0.806	0.806	0.797
$(\hat{X}_{inv}\pm L)_{.80}$	0.811	0.821	0.828	0.830	0.827	0.814
Average L	0.184	0.181	0.180	0.180	0.181	0.184

All Standard Errors are 0.004

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.70}$	0.697	0.705	0.699	0.700	0.695	0.705
$(\hat{X}_{inv}\pm L)_{.70}$	0.713	0.726	0.731	0.728	0.717	0.719
Average L	0.148	0.145	0.144	0.144	0.146	0.148

All Standard Errors are 0.005

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.60}$	0.602	0.605	0.602	0.604	0.603	0.600
$(\hat{X}_{inv}\pm L)_{.60}$	0.617	0.625	0.630	0.634	0.625	0.620
Average L	0.120	0.119	0.117	0.117	0.118	0.120

All Standard Errors are between 0.005 and 0.006

$\alpha=0.0$, $\beta=1.0$, $\sigma=0.1$, $\beta/\sigma=10.0$
Design: 10(X=0), 10(X=1)

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.95}$	0.948	0.950	0.953	0.953	0.946	0.950
$(\hat{X}_{inv}\pm L)_{.95}$	0.953	0.954	0.959	0.959	0.952	0.952
Average L	0.218	0.215	0.213	0.213	0.215	0.218

All Standard Errors are 0.002

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.90}$	0.902	0.899	0.899	0.903	0.903	0.903
$(\hat{X}_{inv}\pm L)_{.90}$	0.906	0.906	0.909	0.913	0.911	0.909
Average L	0.180	0.178	0.176	0.176	0.178	0.180

All Standard Errors are 0.003

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.80}$	0.793	0.807	0.798	0.800	0.803	0.802
$(\hat{X}_{inv}\pm L)_{.80}$	0.801	0.818	0.812	0.814	0.814	0.810
Average L	0.138	0.136	0.135	0.135	0.136	0.138

All Standard Errors are 0.004

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.70}$	0.695	0.699	0.701	0.694	0.699	0.707
$(\hat{X}_{inv}\pm L)_{.70}$	0.706	0.714	0.719	0.711	0.713	0.715
Average L	0.110	0.109	0.108	0.108	0.109	0.110

All Standard Errors are 0.005

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.60}$	0.599	0.602	0.598	0.606	0.611	0.596
$(\hat{X}_{inv}\pm L)_{.60}$	0.608	0.615	0.616	0.621	0.618	0.609
Average L	0.089	0.088	0.087	0.087	0.088	0.089

All Standard Errors are 0.005

α=0.0, β=1.5, σ=0.1, β/σ=15.0
Design: 10(X=0), 10(X=1)

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.95}$	0.948	0.949	0.953	0.952	0.956	0.949
$(\hat{X}_{inv}\pm L)_{.95}$	0.949	0.951	0.955	0.955	0.949	0.951
Average L	0.145	0.143	0.142	0.142	0.143	0.145

All Standard Errors are 0.002

$(\hat{X}_{cl}\pm L)_{.90}$	0.901	0.897	0.899	0.902	0.903	0.903
$(\hat{X}_{inv}\pm L)_{.90}$	0.902	0.902	0.904	0.908	0.905	0.905
Average L	0.120	0.118	0.117	0.117	0.118	0.120

All Standard Errors are 0.003

$(\hat{X}_{cl}\pm L)_{.80}$	0.794	0.806	0.797	0.799	0.802	0.802
$(\hat{X}_{inv}\pm L)_{.80}$	0.797	0.810	0.804	0.806	0.807	0.807
Average L	0.092	0.091	0.090	0.090	0.091	0.092

All Standard Errors are 0.004

$(\hat{X}_{cl}\pm L)_{.70}$	0.695	0.699	0.701	0.694	0.698	0.707
$(\hat{X}_{inv}\pm L)_{.70}$	0.701	0.704	0.710	0.703	0.704	0.710
Average L	0.073	0.072	0.072	0.072	0.072	0.073

All Standard Errors are 0.005

$(\hat{X}_{cl}\pm L)_{.60}$	0.599	0.602	0.597	0.606	0.611	0.596
$(\hat{X}_{inv}\pm L)_{.60}$	0.602	0.606	0.606	0.612	0.612	0.600
Average L	0.059	0.059	0.058	0.058	0.059	0.059

All Standard Errors are 0.005

α=0.0, β=5.0, σ=0.1, β/σ=50.0
Design: 10(X=0), 10(X=1)

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl} \pm L)_{.95}$	0.950	0.952	0.955	0.949	0.947	0.951
$(\hat{X}_{inv} \pm L)_{.95}$	0.950	0.952	0.955	0.949	0.948	0.951
Average L	0.043	0.043	0.043	0.043	0.043	0.043

All Standard Errors are 0.002

$(\hat{X}_{cl} \pm L)_{.90}$	0.899	0.896	0.905	0.901	0.901	0.896
$(\hat{X}_{inv} \pm L)_{.90}$	0.898	0.896	0.905	0.901	0.901	0.897
Average L	0.036	0.035	0.035	0.035	0.035	0.036

All Standard Errors are 0.003

$(\hat{X}_{cl} \pm L)_{.80}$	0.793	0.802	0.801	0.802	0.794	0.806
$(\hat{X}_{inv} \pm L)_{.80}$	0.795	0.802	0.802	0.802	0.794	0.805
Average L	0.028	0.027	0.027	0.027	0.027	0.028

All Standard Errors are 0.004

$(\hat{X}_{cl} \pm L)_{.70}$	0.703	0.693	0.696	0.706	0.695	0.702
$(\hat{X}_{inv} \pm L)_{.70}$	0.703	0.693	0.698	0.707	0.695	0.703
Average L	0.022	0.022	0.022	0.022	0.022	0.022

All Standard Errors are between 0.004 and 0.005

$(\hat{X}_{cl} \pm L)_{.60}$	0.600	0.591	0.599	0.597	0.608	0.602
$(\hat{X}_{inv} \pm L)_{.60}$	0.598	0.590	0.599	0.598	0.607	0.602
Average L	0.018	0.018	0.017	0.017	0.018	0.018

All Standard Errors are 0.005

α=0.0, β=20.0, σ=0.1, β/σ=200.0
Design: 10(X=0), 10(X=1)

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl} \pm L)_{.95}$	0.952	0.954	0.954	0.950	0.951	0.952
$(\hat{X}_{inv} \pm L)_{.95}$	0.952	0.954	0.954	0.950	0.951	0.952
Average L	0.011	0.011	0.011	0.011	0.011	0.011

All Standard Errors are 0.002

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl} \pm L)_{.90}$	0.900	0.908	0.908	0.895	0.899	0.898
$(\hat{X}_{inv} \pm L)_{.90}$	0.900	0.908	0.909	0.895	0.899	0.898
Average L	0.009	0.009	0.009	0.009	0.009	0.009

All Standard Errors are 0.003

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl} \pm L)_{.80}$	0.792	0.801	0.799	0.805	0.805	0.797
$(\hat{X}_{inv} \pm L)_{.80}$	0.792	0.800	0.799	0.805	0.804	0.797
Average L	0.007	0.007	0.007	0.007	0.007	0.007

All Standard Errors are 0.004

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl} \pm L)_{.70}$	0.797	0.703	0.698	0.699	0.694	0.705
$(\hat{X}_{inv} \pm L)_{.70}$	0.696	0.704	0.698	0.699	0.694	0.704
Average L	0.006	0.005	0.005	0.005	0.005	0.006

All Standard Errors are 0.005

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl} \pm L)_{.60}$	0.602	0.605	0.601	0.603	0.603	0.599
$(\hat{X}_{inv} \pm L)_{.60}$	0.601	0.605	0.601	0.603	0.603	0.600
Average L	0.005	0.004	0.004	0.004	0.004	0.005

All Standard Errors are 0.005

be given here. Note that Table IV can be converted to a table showing the effect of error variance, by using a constant slope of $\beta = 0.5$ and solving for σ in the ratio β/σ .

(C) The Effect of Design

Throughout the preceding section, the design was held constant at ten observations per endpoint on the range [0,1]. In order to ascertain the effect of design, we investigated two separate aspects of the problem. The first of these was the effect of the number of observations to be taken at each endpoint and the second was the effect of the points at which the calibration was to be performed.

For the investigation of the number of observations per endpoint, the parameters of the linear model were held constant at the values used to generate Table I. The results are presented in Table V for the designs 2(X=0), 2(X=1); 7(X=0), 7(X=1); 9(X=0), 9(X=1); and 20(X=0), 20(X=1). Monte Carlo experiments of 10,000 repetitions for each of five theoretical confidence levels were used.

Tables I and V illustrate that there is a significant effect associated with the number of observations taken in the design of the calibration experiment. The average value of L increases considerably as the number of observations decreases, which is to be expected since the number

TABLE IV: COMPARISON OF AVERAGE CONFIDENCE FOR CLASSICAL AND INVERSE CALIBRATION CONFIDENCE INTERVALS; EFFECT OF ERROR VARIANCE

$\alpha=0.0,\ \beta=1.0,\ \sigma=0.2,\ \beta/\sigma=5.0$
Design: 10(X=0), 10(X=1)

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.95}$	0.952	0.954	0.956	0.952	0.951	0.954
$(\hat{X}_{inv}\pm L)_{.95}$	0.961	0.968	0.974	0.970	0.969	0.964
Average L	0.440	0.434	0.430	0.430	0.434	0.440
	All Standard Errors are 0.002					
$(\hat{X}_{cl}\pm L)_{.90}$	0.902	0.900	0.907	0.903	0.904	0.900
$(\hat{X}_{inv}\pm L)_{.90}$	0.924	0.929	0.939	0.937	0.931	0.922
Average L	0.363	0.358	0.355	0.355	0.358	0.363
	All Standard Errors are 0.003					
$(\hat{X}_{cl}\pm L)_{.80}$	0.798	0.804	0.803	0.805	0.796	0.808
$(\hat{X}_{inv}\pm L)_{.80}$	0.834	0.848	0.851	0.852	0.840	0.832
Average L	0.279	0.275	0.273	0.273	0.275	0.279
	All Standard Errors are 0.004					
$(\hat{X}_{cl}\pm L)_{.70}$	0.705	0.696	0.699	0.708	0.696	0.705
$(\hat{X}_{inv}\pm L)_{.70}$	0.742	0.747	0.765	0.768	0.744	0.736
Average L	0.223	0.220	0.218	0.218	0.220	0.223
	All Standard Errors are between 0.004 and 0.005					
$(\hat{X}_{cl}\pm L)_{.60}$	0.601	0.592	0.601	0.599	0.608	0.602
$(\hat{X}_{inv}\pm L)_{.60}$	0.639	0.647	0.663	0.663	0.659	0.641
Average L	0.180	0.178	0.176	0.176	0.178	0.181
	All Standard Errors are 0.005					

α=0.0, β=0.5, σ=2/30, β/σ=7.5
Design: 10(X=0), 10(X=1)

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.95}$	0.950	0.954	0.949	0.953	0.949	0.952
$(\hat{X}_{inv}\pm L)_{.95}$	0.955	0.960	0.960	0.962	0.956	0.957
Average L	0.291	0.287	0.285	0.285	0.287	0.291
	All Standard Errors are 0.002					
$(\hat{X}_{cl}\pm L)_{.90}$	0.898	0.900	0.895	0.896	0.898	0.900
$(\hat{X}_{inv}\pm L)_{.90}$	0.908	0.913	0.912	0.913	0.912	0.910
Average L	0.239	0.235	0.234	0.234	0.235	0.239
	All Standard Errors are 0.003					
$(\hat{X}_{cl}\pm L)_{.80}$	0.802	0.804	0.803	0.801	0.801	0.802
$(\hat{X}_{inv}\pm L)_{.80}$	0.816	0.825	0.830	0.826	0.824	0.816
Average L	0.184	0.182	0.180	0.180	0.182	0.185
	All Standard Errors are 0.004					
$(\hat{X}_{cl}\pm L)_{.70}$	0.696	0.699	0.697	0.701	0.703	0.700
$(\hat{X}_{inv}\pm L)_{.70}$	0.715	0.725	0.726	0.726	0.729	0.717
Average L	0.148	0.146	0.144	0.144	0.145	0.148
	All Standard Errors are 0.005					
$(\hat{X}_{cl}\pm L)_{.60}$	0.601	0.598	0.598	0.604	0.591	0.603
$(\hat{X}_{inv}\pm L)_{.60}$	0.614	0.622	0.629	0.633	0.610	0.615
Average L	0.119	0.117	0.117	0.117	0.117	0.119
	All Standard Errors are between 0.004 and 0.005					

$\alpha=0.0$, $\beta=0.5$, $\sigma=0.05$, $\beta/\sigma=10.0$
Design: 10(X=0), 10(X=1)

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.95}$	0.950	0.954	0.951	0.951	0.953	0.955
$(\hat{X}_{inv}\pm L)_{.95}$	0.954	0.958	0.958	0.956	0.957	0.956
Average L	0.218	0.215	0.213	0.213	0.215	0.218

All Standard Errors are 0.002

$(\hat{X}_{cl}\pm L)_{.90}$	0.900	0.897	0.898	0.896	0.894	0.902
$(\hat{X}_{inv}\pm L)_{.90}$	0.907	0.905	0.909	0.907	0.901	0.908
Average L	0.180	0.177	0.176	0.176	0.177	0.180

All Standard Errors are 0.003

$(\hat{X}_{cl}\pm L)_{.80}$	0.799	0.801	0.801	0.794	0.806	0.802
$(\hat{X}_{inv}\pm L)_{.80}$	0.806	0.810	0.816	0.808	0.820	0.809
Average L	0.138	0.136	0.135	0.135	0.136	0.138

All Standard Errors are 0.004

$(\hat{X}_{cl}\pm L)_{.70}$	0.698	0.707	0.711	0.694	0.696	0.692
$(\hat{X}_{inv}\pm L)_{.70}$	0.706	0.725	0.726	0.712	0.714	0.704
Average L	0.111	0.109	0.108	0.108	0.109	0.111

All Standard Errors are between 0.004 and 0.005

$(\hat{X}_{cl}\pm L)_{.60}$	0.595	0.600	0.604	0.599	0.598	0.596
$(\hat{X}_{inv}\pm L)_{.60}$	0.607	0.612	0.620	0.616	0.615	0.608
Average L	0.089	0.088	0.087	0.087	0.088	0.089

All Standard Errors are 0.005

of observations directly effects the degrees of freedom in the t-distribution and the associated variance. As the number of observations per endpoint increases the average value of L decreases and confidence for the Inverse method increases. Throughout this investigation, the confidence for the Classical approach remained unchanged.

Observe that for the design 2(X=0), 2(X=1), the confidences are not significantly different for theoretical confidences of 95% or 80%, while the interval length is unreasonably large. Even when the theoretical confidence is 60% the Inverse confidence is only 3% above the Classical confidence in the middle of the calibrated range. For the design 20(X=0), 20(X=1), there is a significant increase in the Inverse confidence accompanied by a decrease in the average value of L. The Inverse confidence is 1½% above the Classical confidence at the endpoints and 2% higher in the middle of the range for a theoretical confidence of 95%. At the 60% level, the Inverse confidence is 4% higher at the endpoints and 6% higher at mid-range.

Comparing 2(X=0), 2(X=1) with 20(X=0), 20(X=1) we see that the Inverse confidence increases from 2-3% almost everywhere when 18 observations are added to each endpoint while the Classical confidence remains unchanged. Thus, since the Classical confidence does not change and the Inverse confidence can only be improved by additional

TABLE V: COMPARISON OF AVERAGE CONFIDENCES FOR CLASSICAL AND INVERSE CALIBRATION CONFIDENCE INTERVALS; EFFECT OF THE NUMBER OF OBSERVATIONS AT EACH DESIGN POINT

$\alpha=0.0$, $\beta=0.5$, $\sigma=0.1$, $\beta/\sigma=5.0$
Design: 2(X=0), 2(X=1)

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.95}$	0.952	0.953	0.951	0.950	0.951	0.953
$(\hat{X}_{inv}\pm L)_{.95}$	0.952	0.954	0.952	0.951	0.952	0.954
Average L	1.002	0.938	0.903	0.904	0.936	1.002
All Standard Errors are 0.002						
$(\hat{X}_{cl}\pm L)_{.90}$	0.908	0.904	0.907	0.909	0.911	0.909
$(\hat{X}_{inv}\pm L)_{.90}$	0.909	0.906	0.910	0.912	0.914	0.910
Average L	0.688	0.645	0.622	0.623	0.645	0.687
All Standard Errors are 0.003						
$(\hat{X}_{cl}\pm L)_{.80}$	0.810	0.813	0.815	0.810	0.807	0.813
$(\hat{X}_{inv}\pm L)_{.80}$	0.816	0.823	0.824	0.820	0.817	0.816
Average L	0.443	0.415	0.400	0.400	0.414	0.442
All Standard Errors are 0.004						
$(\hat{X}_{cl}\pm L)_{.70}$	0.709	0.706	0.711	0.711	0.710	0.714
$(\hat{X}_{inv}\pm L)_{.70}$	0.723	0.723	0.727	0.731	0.726	0.725
Average L	0.324	0.304	0.293	0.293	0.304	0.325
All Standard Errors are 0.005						
$(\hat{X}_{cl}\pm L)_{.60}$	0.613	0.616	0.606	0.617	0.610	0.613
$(\hat{X}_{inv}\pm L)_{.60}$	0.633	0.643	0.634	0.642	0.636	0.632
Average L	0.250	0.235	0.226	0.226	0.234	0.250
All Standard Errors are 0.005						

$\alpha=0.0$, $\beta=0.5$, $\sigma=0.1$, $\beta/\sigma=5.0$
Design: 7(X=0), 7(X=1)

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.95}$	0.952	0.955	0.952	0.952	0.953	0.953
$(\hat{X}_{inv}\pm L)_{.95}$	0.960	0.966	0.966	0.968	0.965	0.962
Average L	0.463	0.453	0.448	0.448	0.453	0.463

All Standard Errors are 0.002
2 Q values were above 0.2, max Q = 0.223

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.90}$	0.906	0.901	0.908	0.904	0.907	0.904
$(\hat{X}_{inv}\pm L)_{.90}$	0.920	0.925	0.934	0.931	0.931	0.921
Average L	0.380	0.372	0.368	0.368	0.372	0.380

All Standard Errors are between 0.002 and 0.003

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.80}$	0.801	0.801	0.797	0.804	0.805	0.806
$(\hat{X}_{inv}\pm L)_{.80}$	0.826	0.842	0.846	0.850	0.844	0.832
Average L	0.288	0.282	0.279	0.779	0.282	0.288

All Standard Errors are between 0.003 and 0.004

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.70}$	0.708	0.702	0.705	0.701	0.697	0.701
$(\hat{X}_{inv}\pm L)_{.70}$	0.742	0.753	0.763	0.755	0.750	0.735
Average L	0.231	0.226	0.224	0.224	0.226	0.231

All Standard Errors are between 0.004 and 0.005

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.60}$	0.603	0.592	0.607	0.596	0.598	0.598
$(\hat{X}_{inv}\pm L)_{.60}$	0.633	0.643	0.666	0.658	0.648	0.640
Average L	0.186	0.182	0.180	0.180	0.182	0.186

All Standard Errors are 0.005

$\alpha=0.0$, $\beta=0.5$, $\sigma=0.1$, $\beta/\sigma=5.0$
Design: 9(X=0), 9(X=1)

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.95}$	0.953	0.955	0.954	0.953	0.954	0.951
$(\hat{X}_{inv}\pm L)_{.95}$	0.963	0.968	0.972	0.970	0.968	0.964
Average L	0.448	0.440	0.437	0.437	0.440	0.448

All Standard Errors are 0.002

$(\hat{X}_{cl}\pm L)_{.90}$	0.899	0.903	0.909	0.910	0.904	0.904
$(\hat{X}_{inv}\pm L)_{.90}$	0.917	0.927	0.940	0.940	0.932	0.921
Average L	0.367	0.361	0.358	0.358	0.361	0.367

All Standard Errors are between 0.002 and 0.003

$(\hat{X}_{cl}\pm L)_{.80}$	0.806	0.804	0.810	0.808	0.802	0.800
$(\hat{X}_{inv}\pm L)_{.80}$	0.835	0.848	0.858	0.855	0.845	0.827
Average L	0.281	0.276	0.274	0.274	0.276	0.281

All Standard Errors are between 0.003 and 0.004

$(\hat{X}_{cl}\pm L)_{.70}$	0.709	0.704	0.708	0.703	0.697	0.709
$(\hat{X}_{inv}\pm L)_{.70}$	0.743	0.751	0.766	0.758	0.750	0.742
Average L	0.225	0.221	0.220	0.220	0.221	0.225

All Standard Errors are 0.004

$(\hat{X}_{cl}\pm L)_{.60}$	0.596	0.602	0.607	0.596	0.600	0.598
$(\hat{X}_{inv}\pm L)_{.60}$	0.636	0.655	0.668	0.659	0.654	0.634
Average L	0.182	0.179	0.178	0.178	0.179	0.182

All Standard Errors are 0.005

$\alpha=0.0$, $\beta=0.5$, $\sigma=0.1$, $\beta/\sigma=5.0$
Design: 20(X=0), 20(X=1)

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.95}$	0.949	0.952	0.953	0.952	0.950	0.951
$(\hat{X}_{inv}\pm L)_{.95}$	0.963	0.971	0.976	0.974	0.969	0.965
Average L	0.415	0.412	0.410	0.410	0.412	0.415

All Standard Errors are 0.002

$(\hat{X}_{cl}\pm L)_{.90}$	0.901	0.898	0.895	0.903	0.894	0.901
$(\hat{X}_{inv}\pm L)_{.90}$	0.923	0.933	0.933	0.941	0.928	0.921
Average L	0.345	0.343	0.341	0.341	0.343	0.345

All Standard Errors are 0.003

$(\hat{X}_{cl}\pm L)_{.80}$	0.809	0.799	0.798	0.801	0.798	0.811
$(\hat{X}_{inv}\pm L)_{.80}$	0.839	0.848	0.856	0.856	0.848	0.838
Average L	0.267	0.265	0.264	0.264	0.265	0.267

All Standard Errors are 0.004

$(\hat{X}_{cl}\pm L)_{.70}$	0.701	0.703	0.704	0.695	0.696	0.701
$(\hat{X}_{inv}\pm L)_{.70}$	0.733	0.758	0.766	0.764	0.751	0.735
Average L	0.215	0.213	0.213	0.213	0.213	0.215

All Standard Errors are between 0.004 and 0.005

$(\hat{X}_{cl}\pm L)_{.60}$	0.594	0.595	0.599	0.593	0.591	0.594
$(\hat{X}_{inv}\pm L)_{.60}$	0.632	0.648	0.663	0.661	0.648	0.631
Average L	0.174	0.173	0.172	0.172	0.173	0.174

All Standard Errors are 0.005

observations, our conclusion that the Inverse approach is the better of the two does not change.

The effect of different design points was investigated by using the designs 3(X=0), 3(X=.2), 3(X=.4), 3(X=.6), 3(X=.8), 3(X=1); 6(X=0), 6(X=.5), 6(X=1); and 9(X=.15), 9(X=.85). It is necessary to compare these designs with 9(X=0), 9(X=1) of Table V so as to keep the total number of observations constant. Table VI was generated using Monte Carlo experiments of 10,000 repetitions for each of five theoretical confidences, where the parameters remained constant at the values assigned in Table V.

Comparing the uniform design, with observations at each of six points, to the endpoint design, 9(X=0), 9(X=1) of Table V, we see that although there is a slight increase in the average value of L, it is accompanied by an increase in the confidence associated with the Inverse approach while the Classical confidence remains unchanged. In the middle of the range at the 95% level, the Inverse confidence is about 3% higher than the Classical confidence, an increase of 1% from the design 9(X=0), 9(X=1). At the 60% level, the Inverse confidence is 12% higher in the middle of the range, while it was only 6% above the Classical confidence for the design 9(X=0), 9(X=1).

For the midpoint design, 6(X=0), 6(X=.5), 6(X=1), of Table VI, the increase in L is negligible, but the increase in the Inverse confidence is not as great as before. The Inverse confidence is 8-9% above the Classical confidence in the middle of the range for the 60-70% confidence levels, which is only a 2% increase over the difference shown using the design 9(X=0), 9(X=1). At the 95% level, the difference in the two designs is only one standard error which indicates that there is, for all practical purposes, no change at all.

Designing in from the endpoints, 9(X=.15), 9(X=.85), yields a very substantial increase in the middle of the range, accompanied by a slight decrease at the endpoints, for the Inverse confidence. The average L increased slightly and the Classical confidence, as in all three cases, stayed within standard error of the results for 9(X=0), 9(X=1). In the middle of the range at the 95% level, the Inverse confidence increased by 1% over the result for 9(X=0), 9(X=1), and was a total of 3% above the Classical confidence. At the 60% level, the Inverse confidence for 9(X=.15), 9(X=.85) was about 5% above its counterpart for 9(X=0), 9(X=1), which was 12% above the Classical confidence for the middle of the range. At the endpoints for the 90-95% levels, there was no significant change in the Inverse confidence, but at the 60-70% levels the Inverse confidence for 9(X=0), 9(X=1) was 2-3% higher than the Inverse confidence for 9(X=.15), 9(X=.85).

Although the investigation of design is by no means complete, we have shown how each method reacts to minor changes in design. The question of finding the optimal design for the Inverse approach is beyond the scope of this investigation. We have, however, shown that our conclusion remains unchanged by any variation in design and that there do exist better designs for the Inverse method than the endpoint design.

TABLE VI: COMPARISON OF AVERAGE CONFIDENCE FOR CLASSICAL AND INVERSE CALIBRATION CONFIDENCE INTERVALS; EFFECT OF DESIGN POINTS

$\alpha=0.0$, $\beta=0.5$, $\sigma=0.1$, $\beta/\sigma=5.0$
Design: 3(X=0),3(X=0.2),3(X=0.4),3(X=0.6),3(X=0.8),3(X=1.0)

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.95}$	0.955	0.956	0.958	0.958	0.953	0.958
$(\hat{X}_{inv}\pm L)_{.95}$	0.960	0.974	0.982	0.984	0.971	0.963
Average L	0.468	0.452	0.443	0.443	0.452	0.468

All Standard Errors are 0.002
266 Q values were above 0.2, max Q = 0.525

$(\hat{X}_{cl}\pm L)_{.90}$	0.909	0.910	0.908	0.916	0.905	0.910
$(\hat{X}_{inv}\pm L)_{.90}$	0.916	0.944	0.957	0.962	0.942	0.917
Average L	0.384	0.370	0.364	0.364	0.371	0.384

All Standard Errors are between 0.002 and 0.003
31 Q values were above 0.2, max Q = 0.383

$(\hat{X}_{cl}\pm L)_{.80}$	0.810	0.798	0.807	0.813	0.808	0.810
$(\hat{X}_{inv}\pm L)_{.80}$	0.821	0.867	0.895	0.898	0.868	0.823
Average L	0.294	0.283	0.278	0.278	0.283	0.294

All Standard Errors are 0.004

$(\hat{X}_{cl}\pm L)_{.70}$	0.704	0.703	0.697	0.699	0.705	0.701
$(\hat{X}_{inv}\pm L)_{.70}$	0.731	0.785	0.808	0.810	0.776	0.717
Average L	0.236	0.228	0.223	0.223	0.227	0.236

All Standard Errors are 0.005

$(\hat{X}_{cl}\pm L)_{.60}$	0.602	0.597	0.607	0.602	0.601	0.600
$(\hat{X}_{inv}\pm L)_{.60}$	0.627	0.681	0.729	0.724	0.689	0.617
Average L	0.192	0.185	0.181	0.181	0.185	0.192

All Standard Errors are 0.002

$\alpha=0.0$, $\beta=0.5$, $\sigma=0.1$, $\beta/\sigma=5.0$
Design: 6(X=0), 6(X=0.5), 6(X=1.0)

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.95}$	0.954	0.951	0.953	0.956	0.956	0.954
$(\hat{X}_{inv}\pm L)_{.95}$	0.963	0.969	0.975	0.978	0.972	0.964
Average L	0.455	0.443	0.438	0.438	0.443	0.455

All Standard Errors are 0.002
8 Q values were above 0.2, max Q = 0.288

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.90}$	0.901	0.910	0.903	0.903	0.906	0.897
$(\hat{X}_{inv}\pm L)_{.90}$	0.916	0.940	0.945	0.943	0.938	0.916
Average L	0.374	0.365	0.360	0.360	0.365	0.374

All Standard Errors are between 0.002 and 0.003

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.80}$	0.803	0.803	0.803	0.808	0.802	0.804
$(\hat{X}_{inv}\pm L)_{.80}$	0.833	0.861	0.876	0.874	0.858	0.828
Average L	0.287	0.280	0.276	0.276	0.280	0.287

All Standard Errors are 0.004

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.70}$	0.700	0.699	0.700	0.705	0.707	0.699
$(\hat{X}_{inv}\pm L)_{.70}$	0.738	0.762	0.787	0.789	0.769	0.732
Average L	0.230	0.224	0.221	0.221	0.224	0.230

All Standard Errors are between 0.004 and 0.005

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.60}$	0.598	0.605	0.602	0.600	0.606	0.600
$(\hat{X}_{inv}\pm L)_{.60}$	0.630	0.671	0.695	0.686	0.672	0.673
Average L	0.185	0.181	0.178	0.178	0.181	0.185

All Standard Errors are 0.005

α=0.0, β=0.5, σ=0.1, β/σ=5.0
Design: 9(X=.15), 9(X=.85)

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl} \pm L)_{.95}$	0.954	0.954	0.958	0.956	0.958	0.960
$(\hat{X}_{inv} \pm L)_{.95}$	0.959	0.974	0.983	0.982	0.975	0.962
Average L	0.465	0.450	0.442	0.442	0.450	0.465

All Standard Errors are 0.002
171 Q values were above 0.2, max Q = 0.700

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl} \pm L)_{.90}$	0.907	0.906	0.910	0.911	0.911	0.911
$(\hat{X}_{inv} \pm L)_{.90}$	0.917	0.943	0.961	0.958	0.942	0.918
Average L	0.383	0.370	0.364	0.363	0.370	0.383

All Standard Errors are between 0.002 and 0.003
19 Q values were above 0.2, max Q = 0.287

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl} \pm L)_{.80}$	0.803	0.814	0.810	0.802	0.809	0.805
$(\hat{X}_{inv} \pm L)_{.80}$	0.819	0.868	0.894	0.889	0.866	0.826
Average L	0.294	0.284	0.279	0.279	0.284	0.294

All Standard Errors are 0.004
1 Q value was above 0.2, max Q = 0.212

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl} \pm L)_{.70}$	0.705	0.714	0.711	0.698	0.704	0.699
$(\hat{X}_{inv} \pm L)_{.70}$	0.722	0.786	0.818	0.813	0.777	0.724
Average L	0.236	0.228	0.224	0.224	0.228	0.236

All Standard Errors are between 0.004 and 0.005

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl} \pm L)_{.60}$	0.605	0.606	0.597	0.612	0.602	0.606
$(\hat{X}_{inv} \pm L)_{.60}$	0.621	0.684	0.715	0.725	0.686	0.627
Average L	0.190	0.183	0.180	0.180	0.184	0.190

All Standard Errors are 0.005

VII. ROBUSTNESS

(A) The Effect of an Ignored Quadratic Term:

We have previously discussed the effect of the various parameters involved in the linear model as well as the effect of various designs. An investigation of robustness to the linear assumption entails a few minor changes in the Monte Carlo procedure. We will discuss this rather simple variation before the actual effect is investigated.

Consider the model given by (1.3), $y = \alpha + \beta x + \varepsilon$. Suppose that this was assumed to be the correct model and used to calibrate the instrument in question, when the true relationship between x and y is given by

$$y = \alpha + \beta x + \theta x^2 + \varepsilon \ , \qquad (7.1)$$

a quadratic model. By using (1.3) to calibrate the experiment and calculate the estimates of α , β , γ and δ , and then by using (7.1) to obtain the Y's which lead to the estimates, $\hat{X}_{inv}$ and $\hat{X}_{cl}$, one may easily ascertain the effect of the ignored quadratic term by leaving the remaining portion of the Monte Carlo procedure intact.

Krutchkoff (1967), in his investigation of this effect, designed in from the endpoints to gain some degree of protection against the ignored quadratic term. Following his technique, we will use the design 10(X=.15), 10(X=.85). Monte Carlo experiments of 10,000 repetitions for each of five theoretical confidences were conducted for $\theta = \pm 0.1$, ± 0.2, 0.5, keeping all other parameters constant at the values used to generate Table I. Both the number of times Q exceeded 0.2 and the maximum value of Q were reported. We did not investigate larger values of θ because we assume that when θ becomes large one will most surely be aware of its presence.

Recalling equations (6.5) and (6.11) from the investigation of the effect of intercept, we may see, theoretically, what effect we should expect from varying the magnitude of θ . It may easily be verified that when Y is given by (7.1) instead of (1.3), (6.5) and (6.11) become

$$\hat{X}_{cl} - X = \left(\frac{\beta}{\hat{\beta}} - 1\right)\left(X-\overline{x}\right) + \frac{1}{\hat{\beta}}(\varepsilon'-\overline{\varepsilon}) + \frac{\theta x^2}{\hat{\beta}} \tag{7.2}$$

and

$$\hat{X}_{inv} - X = \left(\frac{\beta r^2}{\hat{\beta}} - 1\right)\left(X-\overline{x}\right) + \frac{r^2}{\hat{\beta}}(\varepsilon'-\overline{\varepsilon}) + \frac{r^2\theta x^2}{\hat{\beta}} \; .$$

For all practical purposes, we will have $r^2 < 1$. Thus, for $\theta > 0$, $(\hat{X}_{cl}-X)$ will become larger than $(\hat{X}_{inv}-X)$, while for $\theta < 0$, $(\hat{X}_{cl}-X)$ will become smaller than $(\hat{X}_{inv}-X)$. Clearly, $(\hat{X}_{cl}-X)$ is more sensitive to the quadratic term than $(\hat{X}_{inv}-X)$ because of the magnitude of r^2 . The results in Table VII show that this is the case.

Observe that along with a minor increase in the average value of L for $\theta = 0.1$, the Inverse approach improves at the upper end of the range while the Classical confidence decreases. This is most apparent for a theoretical confidence of 60% where the Inverse confidence is 2% higher at $X = 0$, and 29% higher at $X = 1$. One should expect a more significant effect for large X since the quadratic term involves the square of that value.

At $\theta = 0.2$, the interval length becomes larger and the Inverse confidence remains uniformly higher than the Classical confidence although both decrease significantly at the upper end of the range. Observe that at the 95% level the Classical confidence is 0.685 at $X = 1$ while the Inverse confidence is 0.938. At the 60% level, the Classical confidence drops to 0.165 at $X = 1$, while the Inverse confidence only drops to 0.510, 34% above the Classical confidence.

The same pattern is present for $\theta = 0.5$ as was seen for $\theta = 0.2$, although to a much greater extent. At the 95% level for $X = 1$, both confidences are extremely low. The Classical confidence is 0.019 as compared to 0.378 for the Inverse method. As the theoretical confidence decreases, the percentage of time that $(\hat{X} \pm L)$ covers X is almost negligible for $X \geq 0.8$.

In the case of the negative quadratic coefficient, the average value of L decreases at the upper end of the range and the Classical confidence becomes higher than the Inverse confidence, although both show a significant decrease at the 95% level of $\theta = -0.1$, the Classical confidence is 3% higher than the Inverse confidence for $X = 1$, while at the 60% level, it is 12% higher. For the case of $\theta = -0.2$ and a 95% theoretical confidence, the Classical confidence is 2% higher than the Inverse confidence. At the 60% level the Classical confidence is 0.151 at $X = 1$, as compared to 0.073 for the Inverse approach.

Our conclusion is therefore slightly altered by these results. The Inverse approach is uniformly better except where θ is negative and the true value of X lies at the upper end of the calibrated range. In spite of this exception, we maintain that, for all practical purposes, the Inverse approach is the superior of the two.

(B) The Effect of a Uniform Error Term:

The only investigation of the robustness to normality that we considered was the effect of a uniform error distribution with a mean of zero and $\sigma = 0.1$. In order to ascertain the effect of this distribution, only the pseudo-random number generator used in the Monte Carlo procedure was changed. A generator which produces random numbers from a uniform distribution was substituted for the generator which produced random numbers from the normal distribution,

TABLE VII: COMPARISON OF AVERAGE CONFIDENCE FOR CLASSICAL AND INVERSE CALIBRATION CONFIDENCE INTERVALS; EFFECT OF IGNORED QUADRATIC TERM

$\alpha=0.0$, $\beta=0.5$, $\sigma=0.1$, $\beta/\sigma=5.0$, $\theta=0.1$
Design: 10(X=.15), 10(X=.85)

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.95}$	0.957	0.960	0.956	0.946	0.931	0.898
$(\hat{X}_{inv}\pm L)_{.95}$	0.962	0.977	0.981	0.982	0.984	0.985
Average L	0.458	0.444	0.437	0.439	0.453	0.479

All Standard Errors are 0.002
65 Q values were above 0.2, max Q = 0.376

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.90}$	0.906	0.912	0.910	0.890	0.865	0.820
$(\hat{X}_{inv}\pm L)_{.90}$	0.916	0.941	0.955	0.956	0.959	0.962
Average L	0.377	0.365	0.360	0.361	0.373	0.393

All Standard Errors are between 0.002 and 0.003
7 Q values were above 0.2, max Q = 0.251

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.80}$	0.801	0.806	0.802	0.786	0.737	0.661
$(\hat{X}_{inv}\pm L)_{.80}$	0.819	0.862	0.884	0.899	0.897	0.898
Average L	0.289	0.280	0.275	0.277	0.285	0.302

All Standard Errors are 0.003

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.70}$	0.703	0.707	0.700	0.679	0.603	0.549
$(\hat{X}_{inv}\pm L)_{.70}$	0.723	0.776	0.801	0.812	0.812	0.826
Average L	0.232	0.225	0.221	0.222	0.229	0.242

All Standard Errors are between 0.004 and 0.005

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.60}$	0.604	0.605	0.599	0.580	0.520	0.438
$(\hat{X}_{inv}\pm L)_{.60}$	0.623	0.672	0.709	0.726	0.727	0.731
Average L	0.188	0.182	0.179	0.180	0.185	0.196

All Standard Errors are 0.005

$\alpha=0.0$, $\beta=0.5$, $\sigma=0.1$, $\beta/\sigma=5.0$, $\theta=0.2$
Design: 10(X=.15), 10(X=.85)

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.95}$	0.955	0.953	0.949	0.917	0.831	0.685
$(\hat{X}_{inv}\pm L)_{.95}$	0.960	0.970	0.978	0.974	0.963	0.938
Average L	0.459	0.444	0.437	0.442	0.464	0.503

All Standard Errors are between 0.002 and 0.003
76 Q values were above 0.2, max Q = 0.716

$(\hat{X}_{cl}\pm L)_{.90}$	0.906	0.902	0.884	0.840	0.717	0.518
$(\hat{X}_{inv}\pm L)_{.90}$	0.918	0.940	0.941	0.937	0.917	0.880
Average L	0.378	0.366	0.360	0.364	0.383	0.416

All Standard Errors are between 0.003 and 0.004
5 Q values were above 0.2, max Q = 0.243

$(\hat{X}_{cl}\pm L)_{.80}$	0.800	0.810	0.784	0.712	0.549	0.326
$(\hat{X}_{inv}\pm L)_{.80}$	0.811	0.862	0.866	0.858	0.825	0.756
Average L	0.290	0.280	0.276	0.279	0.293	0.318

All Standard Errors are between 0.004 and 0.005

$(\hat{X}_{cl}\pm L)_{.70}$	0.700	0.697	0.682	0.599	0.427	0.227
$(\hat{X}_{inv}\pm L)_{.70}$	0.719	0.762	0.780	0.756	0.714	0.636
Average L	0.231	0.224	0.220	0.223	0.234	0.254

All Standard Errors are between 0.004 and 0.005

$(\hat{X}_{cl}\pm L)_{.60}$	0.601	0.603	0.578	0.502	0.348	0.165
$(\hat{X}_{inv}\pm L)_{.60}$	0.615	0.671	0.674	0.674	0.667	0.510
Average L	0.187	0.181	0.179	0.181	0.189	0.206

All Standard Errors are between 0.005 and 0.006

$\alpha=0.0$, $\beta=0.5$, $\sigma=0.1$, $\beta/\sigma=5.0$, $\theta=0.5$
Design: 10(X=.15), 10(X=.85)

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.95}$	0.955	0.948	0.894	0.662	0.233	0.019**
$(\hat{X}_{inv}\pm L)_{.95}$	0.960	0.963	0.946	0.862	0.655	0.378
Average L	0.458	0.443	0.437	0.455	0.510	0.605

All Standard Errors are between 0.003 and 0.004
76 Q values were above 0.2, max Q = 0.716

$(\hat{X}_{cl}\pm L)_{.90}$	0.906	0.898	0.808	0.515	0.118	0.004**
$(\hat{X}_{inv}\pm L)_{.90}$	0.918	0.928	0.882	0.751	0.480	0.216
Average L	0.378	0.365	0.361	0.375	0.421	0.499

All Standard Errors are 0.003
5 Q values were above 0.2, max Q = 0.243

$(\hat{X}_{cl}\pm L)_{.80}$	0.800	0.802	0.681	0.344	0.049	0.001**
$(\hat{X}_{inv}\pm L)_{.80}$	0.811	0.842	0.775	0.574	0.276	0.077
Average L	0.290	0.280	0.276	0.288	0.322	0.383

All Standard Errors are between 0.004 and 0.005

$(\hat{X}_{cl}\pm L)_{.70}$	0.700	0.692	0.572	0.248	0.026	0.0003*
$(\hat{X}_{inv}\pm L)_{.70}$	0.719	0.737	0.658	0.439	0.159	0.029
Average L	0.231	0.223	0.220	0.230	0.257	0.305

All Standard Errors are 0.005

$(\hat{X}_{cl}\pm L)_{.60}$	0.601	0.594	0.476	0.175	0.015	0.0001*
$(\hat{X}_{inv}\pm L)_{.60}$	0.615	0.645	0.549	0.328	0.096	0.013
Average L	0.187	0.181	0.179	0.186	0.208	0.247

All Standard Errors are 0.005

* Standard Error is 0.0001
** Standard Error is 0.001

α=0.0, β=0.5, σ=0.1, β/σ=5.0, θ=-0.1
Design: 19(X=.15), 10(X=.85)

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.95}$	0.955	0.958	0.959	0.941	0.908	0.840
$(\hat{X}_{inv}\pm L)_{.95}$	0.957	0.976	0.985	0.973	0.928	0.817
Average L	0.458	0.444	0.438	0.436	0.439	0.444

All Standard Errors are between 0.003 and 0.004
82 Q values were above 0.2, max Q = 0.363

$(\hat{X}_{cl}\pm L)_{.90}$	0.907	0.904	0.909	0.884	0.833	0.752
$(\hat{X}_{inv}\pm L)_{.90}$	0.918	0.944	0.963	0.939	0.856	0.695
Average L	0.378	0.367	0.361	0.360	0.362	0.366

All Standard Errors are 0.003
9 Q values were above 0.2, max Q = 0.330

$(\hat{X}_{cl}\pm L)_{.80}$	0.801	0.808	0.804	0.771	0.712	0.616
$(\hat{X}_{inv}\pm L)_{.80}$	0.823	0.876	0.898	0.856	0.729	0.517
Average L	0.291	0.282	0.277	0.276	0.278	0.281

All Standard Errors are between 0.004 and 0.005

$(\hat{X}_{cl}\pm L)_{.70}$	0.708	0.696	0.697	0.686	0.609	0.504
$(\hat{X}_{inv}\pm L)_{.70}$	0.730	0.783	0.818	0.776	0.618	0.388
Average L	0.232	0.225	0.222	0.221	0.222	0.225

All Standard Errors are between 0.994 and 0.005

$(\hat{X}_{cl}\pm L)_{.60}$	0.604	0.595	0.596	0.570	0.516	0.428
$(\hat{X}_{inv}\pm L)_{.60}$	0.618	0.685	0.728	0.670	0.503	0.305
Average L	0.188	0.182	0.179	0.179	0.180	0.182

All Standard Errors are 0.005

α=0.0, β=0.5, σ=0.1, β/σ=5.0, θ=-0.2
Design: 10(X=.15), 10(X=.85)

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.95}$	0.955	0.959	0.954	0.903	0.779	0.546
$(\hat{X}_{inv}\pm L)_{.95}$	0.957	0.977	0.984	0.954	0.819	0.525
Average L	0.458	0.445	0.438	0.436	0.436	0.437

All Standard Errors are between 0.002 and 0.003
82 Q values were above 0.2, max Q = 0.363

$(\hat{X}_{cl}\pm L)_{.90}$	0.907	0.904	0.901	0.829	0.666	0.414
$(\hat{X}_{inv}\pm L)_{.90}$	0.918	0.947	0.961	0.899	0.695	0.360
Average L	0.378	0.367	0.362	0.360	0.360	0.360

All Standard Errors are between 0.003 and 0.004
9 Q values were above 0.2, max Q = 0.330

$(\hat{X}_{cl}\pm L)_{.80}$	0.801	0.807	0.791	0.694	0.518	0.280
$(\hat{X}_{inv}\pm L)_{.80}$	0.823	0.879	0.896	0.782	0.518	0.194
Average L	0.291	0.282	0.278	0.277	0.277	0.277

All Standard Errors are between 0.003 and 0.004

$(\hat{X}_{cl}\pm L)_{.70}$	0.708	0.696	0.679	0.605	0.418	0.198
$(\hat{X}_{inv}\pm L)_{.70}$	0.730	0.788	0.813	0.693	0.397	0.116
Average L	0.232	0.225	0.222	0.221	0.221	0.221

All Standard Errors are between 0.003 and 0.004

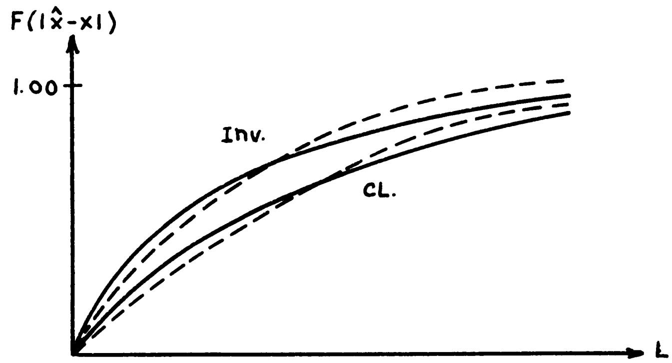
$(\hat{X}_{cl}\pm L)_{.60}$	0.604	0.596	0.581	0.491	0.323	0.151
$(\hat{X}_{inv}\pm L)_{.60}$	0.618	0.692	0.726	0.574	0.284	0.073
Average L	0.188	0.182	0.180	0.179	0.179	0.179

All Standard Errors are 0.005

while other portions of the Monte Carlo procedure remained unchanged.

Table VIII was generated using the parameters and design from Table I, while the distribution of the error term was uniform. Monte Carlo experiments of 10,000 repetitions were conducted for each of the five theoretical confidence levels.

The results indicated by Table VIII show that the average value of L decreases substantially, while both methods are significantly effected by the uniform error distribution. Both methods improve for theoretical confidences above 80% while both confidences decrease below 80% when compared to Table I. This is quite reasonable when one considers that the cumulative distribution function for the uniform case, similar to (45), will be less curved.



(7.3)

The dotted lines represent the distribution functions involving uniform errors and the solid lines represent the distribution functions involving normal errors. The graph shows the results of Table VIII pictorially.

Although both methods are effected by the non-normal error term, the Inverse method retains a uniformly higher confidence for the identical interval length than does the Classical method.

TABLE VIII: COMPARISON OF AVERAGE CONFIDENCE FOR CLASSICAL AND INVERSE CALIBRATION CONFIDENCE INTERVALS; EFFECT OF UNIFORM ERROR TERM

$\alpha=0.0$, $\beta=0.5$, $\sigma=0.1$, $\beta/\sigma=5.0$
Design: 10(X=0), 10(X=1)

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.95}$	0.981	0.984	0.987	0.986	0.982	0.981
$(\hat{X}_{inv}\pm L)_{.95}$	0.982	0.986	0.989	0.987	0.984	0.981
Average L	0.152	0.150	0.148	0.148	0.150	0.152

All Standard Errors are 0.001

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.90}$	0.934	0.939	0.941	0.939	0.933	0.932
$(\hat{X}_{inv}\pm L)_{.90}$	0.933	0.944	0.948	0.944	0.936	0.934
Average L	0.125	0.124	0.123	0.123	0.124	0.125

All Standard Errors are between 0.002 and 0.003

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.80}$	0.791	0.786	0.783	0.784	0.781	0.790
$(\hat{X}_{inv}\pm L)_{.80}$	0.798	0.794	0.793	0.796	0.791	0.794
Average L	0.096	0.095	0.094	0.094	0.095	0.096

All Standard Errors are 0.004

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.70}$	0.639	0.630	0.633	0.627	0.622	0.649
$(\hat{X}_{inv}\pm L)_{.70}$	0.649	0.644	0.645	0.639	0.634	0.666
Average L	0.077	0.076	0.076	0.076	0.076	0.077

All Standard Errors are 0.005

	X=0.0	X=0.2	X=0.4	X=0.6	X=0.8	X=1.0
$(\hat{X}_{cl}\pm L)_{.60}$	0.520	0.511	0.508	0.509	0.507	0.518
$(\hat{X}_{inv}\pm L)_{.60}$	0.527	0.519	0.517	0.520	0.516	0.530
Average L	0.063	0.062	0.061	0.061	0.062	0.063

All Standard Errors are 0.005

VIII. CONCLUSION

In retrospect, the Monte Carlo investigation of the problem of linear instrument calibration has substantiated the supposition which was suspected at the outset, that is, that the Inverse method exhibits a uniformly larger confidence for the identical interval length than does the Classical method. There does exist a slight exception connected with high X values for an ignored negative quadratic term. Notwithstanding this exception, the investigation has demonstrated the superiority of the Inverse approach to the widely-used Classical approach with regard to the problem of linear instrument calibration.

IX. ACKNOWLEDGEMENTS

The author wishes to express his appreciation to Dr. Richard G. Krutchkoff for his suggestion of this problem and his encouragement throughout the investigation. The author would also like to thank Dr. Raymond H. Myers for reading the manuscript and contributing to it through his suggestions.

The author would like to express his sincere gratitude to the National Aeronautics and Space Administration for the financial support which enabled him to complete his graduate program.

The author would like to sincerely thank Dr. Boyd Harshbarger for continued encouragement throughout his undergraduate and graduate programs, and for making possible the financial assistance from NASA.

X. BIBLIOGRAPHY

1. Brownlee, K. A., "Statistical Theory and Methodology in Science and Engineering," John Wiley and Sons, New York, 1960, pp.346-349.

2. Davis, O. L., "Statistical Methods in Research and Production," Hafner Publishing Company, New York, 1961, pp.169-175.

3. Kendall, M. C. and Stuart, A., "The Advanced Theory of Statistics," Hafner Publishing Company, New York, 1946, Vol. 2, pp.162-165.

4. Krutchkoff, R. G., "Classical and Inverse Regression Methods of Calibration," Technometrics, (1967), Vol. 9, No. 3, pp.425-439.

5. Ott, L., "Optimal Designs for Inverse Estimation in Regression Analysis," Virginia Polytechnic Institute, Ph.D. Dissertation, 1966.

6. Waugh, F. V., "Choice of the Dependent Variable in Regression Analysis," Journal of the American Statistical Association, (1943), Vol. 38, pp. 210-216.

The vita has been removed from the scanned document

LINEAR CALIBRATION: A COMPARISON OF AN INVERSE REGRESSION METHOD OF CONFIDENCE INTERVALS WITH THE CLASSICAL REGRESSION METHOD OF CONFIDENCE INTERVALS

by

William T. McClelland, Jr.

ABSTRACT

This thesis contains an investigation of the Classical and Inverse regression approaches to the problem of linear instrument calibration. A comparison was achieved through Monte Carlo techniques using confidence intervals as the criterion. Confidences that $(\hat{X}_{inv} \pm L)$ and $(\hat{X}_{cl} \pm L)$ covered the true X were calculated where

$$y = \alpha + \beta x + \varepsilon$$

and the estimates are

$$\hat{X}_{cl} = \frac{Y-\hat{\alpha}}{\hat{\beta}}$$

and

$$\hat{X}_{inv} = \hat{\gamma} + \hat{\delta} Y \ .$$

The interval length was 2L where L is given by

$$L = \frac{ts}{|\hat{\beta}|} \left[\left(1 + \frac{1}{n}\right) + \frac{(Y-\hat{\alpha}')}{\hat{\beta} \sum_{i=1}^{n} (x_i - \bar{x})^2} \right]^{\frac{1}{2}} .$$

It was shown that the confidence that $(\hat{X}_{cl} \pm L)$ covered X was uniformly lower than the confidence that $(\hat{X}_{inv} \pm L)$ covered X, regardless of the values of the model parameters or the calibration design. It was verified that the effect of β and σ depended on β/σ and, moreover, that the Inverse approach improved significantly for small β/σ while the Classical approach remained unchanged, and that both methods became identical for large β/σ. The number of observations taken at each endpoint of the design was also investigated, as well as the use of the endpoint design itself. The Inverse approach improved with an increase in observations and for various different designs. At no time was the Classical confidence higher than the Inverse confidence.

Robustness to the assumption of linearity was investigated with regard to an ignored quadratic term, as well as robustness to the assumption of normality with regard to a uniform error term. Both confidences were effected and, in general, the Inverse confidence was the higher of the two. There was a slight exception for large X and a negative quadratic coefficient, where the true model is

$$y = \alpha + \beta x + \theta x^2 + \varepsilon \ .$$

In this case, the Classical confidence was slightly higher than the Inverse confidence, but this did not significantly detract from the superiority of the Inverse approach.

The overall conclusion was that the Inverse approach was the superior of the two, although neither approach was assumed to be optimal for the calibration problem.