The Prediction Properties of Inverse and Reverse Regression for the Simple Linear Calibration Problem

Peter A. Parker
NASA Langley Research Center
Hampton, Virginia

G. Geoffrey Vining
Sara R. Wilson
John L. Szarka III
Nels G. Johnson
Virginia Tech
Blacksburg, Virginia 24061

Abstract

The calibration of measurement systems is a fundamental but under-studied problem within industrial statistics. The origins of this problem go back to basic chemical analysis based on NIST standards. In today's world these issues extend to mechanical, electrical, and materials engineering. Often, these new scenarios do not provide "gold standards" such as the standard weights provided by NIST.

This paper considers the classic "forward regression followed by inverse regression" approach. In this approach the initial experiment treats the "standards" as the regressor and the observed values as the response to calibrate the instrument. The analyst then must invert the resulting regression model in order to use the instrument to make actual measurements in practice. This paper compares this classical approach to "reverse regression," which treats the standards as the response and the observed measurements as the regressor in the calibration experiment. Such an approach is intuitively appealing because it avoids the need for the inverse regression. However, it also violates some of the basic regression assumptions.

Problem Context

Flight research aircraft at NASA require the calibration and characterization of multiple instruments. For example, the NASA Airborne Subscale Transport Aircraft Research (AirSTAR) program seeks to develop a research testbed for a dynamically scaled vehicle under upset conditions for commercial aircraft. The purpose of this testbed is to allow NASA to evaluate the aircraft response under severe upset conditions without risking full scale aircraft and personal.

The test vehicle contains a comprehensive suite of instrumentation to measure the vehicle's performance under these conditions. For example, each control has an angular measurement that detects its deployment. These instruments transduce angular displacement to an electrical signal, which was calibrated in the laboratory to develop the relationship. In flight, the electrical signals are measured and the actual angle is inferred through the calibration relationship. In the calibration the angles are precisely set, and the electrical signal is measured.

The classical calibration approach develops a forward regression model for signal as a function of angle. The analyst then inverts this model to estimate the actual angle from the electrical signal transmitted in the flight. An alternative method commonly considered directly models the angle as a function of signal in the calibration experiment, thereby reversing the roles of the explanatory variable and the response. Such an approach is attractive because it avoids the need to invert the forward regression model and the necessity of the Delta Method to estimate prediction intervals. However, while this reverse approach is simple and easy to implement in any software, this approach violates certain assumptions, such as the explanatory variables are measured without error while the response is measured with error, and could result in undesirable consequences.

Situations such as the NASA AirSTAR program are not examples of the classical laboratory standards calibration problem. In classical laboratory standards experiments, the precision of the instruments is of the order of 0.01 to 0.1% of the full scale measurement range, which means that the standard deviation of the measurement error of the instrument is 0.01 to 0.1% of the full load measurement. Transducer class instruments, such as the one outlined above, typically have precisions in the range of 0.5 to 5%. However, some applications that should lend themselves to a calibration approach may have much lower precision, on the order of 50% of full-scale instrument range. In these cases, we expect severe problems with the inverse approach required by the classical calibration approach due to having an approximately normally distributed random variable in the denominator.

Background

Consider the problem of calibrating an instrument where we know that we can model the response y as a simple linear function of the factor x. In practice, we collect the data from the instrument at known values of x. We use the resulting data set $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ to model the relationship between x and y. We then use this model to estimate the corresponding unknown value of x for future observations of y from the instrument.

Two commonly used methods for modeling the relationship between x and y are *inverse regression* and *reverse regression*. Inverse regression fits a regression line of y on x. We then estimate future unknown x values by applying the inverse solution to the observed $y's$. Reverse regression treats x as the response and y as the regressor (even though the x's are measured with negligible error) and fits a regression line of x on y. These two methods can lead to very different estimates of the unknown x values, as well as different prediction intervals.

The statistical literature in this area seems somewhat underdeveloped. This calibration problem is common in many engineering and science applications. The structure of the experiment suggests that the inverse regression approach should be most appropriate; however, the statistical properties of this approach are not readily apparent, not very accessible to engineers and scientists, and not generally addressed in much detail in the standard linear regression texts. Our experience, particularly within NASA, indicates that many practitioners naively use the reverse regression approach instead. A natural question then becomes what are the resulting consequences.

Krutchkoff (1967, 1969) compared inverse and reverse regression using Monte Carlo simulation and recommended reverse regression based on the mean squared error. Berkson (1969) Halpern (1970) present significant criticisms to Krutchkoff's work.

From a statistical perspective, the basic problem with inverse regression is that if the random errors follow a normal distribution, then the estimated slope of the regression line is also normally distributed. Inverse regression must use this estimated slope in the denominator for making predictions. Williams (1969) stated that the reciprocal of the slope has infinite variance, which means that the inverse estimator has infinite variance, and hence infinite mean squared error. Williams notes that the reverse estimator has finite variance and mean squared error, and so minimum variance and mean squared error are not suitable criterion on which to base the comparison of the inverse and reverse estimators.

It is important to note that Williams in making this claim is assuming Cauchy-like behavior for the reciprocal of the slope. A Cauchy random variable is the inverse of a standard normal random variable, which has a mean of zero. In practical applications, this issue is not a concern. In every calibration experiment, we can rescale the data

such that the slope has a mean of one. Typically, the standard deviation of the random error is extremely small. We note that the slope of the regression line is not actually normally distributed, but instead follows a truncated normal distribution since it is not possible for extreme values to occur. Assuming a normal distribution with $\sigma = 0.01$, as is reasonable in the calibration problem, zero is one hundred standard deviations away from the mean. Thus, in practice, a slope of zero will not occur, and so the reciprocal of the slope will have finite variance and mean squared error. Berkson (1969) and Halperin (1970) showed that for large samples, the reverse estimator has smaller mean squared error only when x lies within a small interval around $\overline{x}$.

Graybill (1976) and Seber (1977) developed a $(1-\alpha)100\%$ confidence region for an unknown value of x using the inverse regression approach. In some cases, this method may result in two semi-infinite intervals or the entire real line, rather than a finite interval. Scheffé (1973) proposed a general method for calibration intervals based on inverting simultaneous tolerance intervals. Eberhardt and Mee (1994) developed constant width calibration intervals, which are simple to compute and narrower than the maximum width of the intervals found using Scheffé's (1973) method. Srivastava and Shalabh (1997) examined the asymptotic properties of the inverse and reverse estimators when the errors are not assumed to be normally distributed. Shalabh (2001) compared the inverse and reverse estimators using the balanced loss function.

This paper presents prediction intervals for x using inverse regression based on the Delta Method, as well as prediction intervals using reverse regression. It then derives approximations of the bias for both the inverse and reverse estimators. A simulation study is presented to compare the two estimators. The prediction intervals found using the Delta Method and reverse regression are also compared to the intervals based on Scheffé's (1973) method and the constant width intervals of Eberhardt and Mee (1994). The final section provides some conclusions and additional discussion.

Methods

Assume that a simple linear regression model is appropriate, so the true model is

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \tag{1}$$

where β_0 is the y-intercept, β_1 is the slope, and the ϵ_i's are random errors assumed to be independent and identically distributed as normal with a mean of 0 and a variance of σ^2. Note an important assumption of this model is that the x_i's are measured with negligible error. Let

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

and

$$\overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i.$$

For convenience, we use the centered simple linear model

$$y_i^* = y_i - \overline{y} = \beta_0^* + \beta_1 x_i + \epsilon_i, \tag{2}$$

where $\beta_0^* = \beta_0 - \overline{y}$. The ordinary least squares estimates of β_0^* and β_1 are

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (x_i - \overline{x})(y_i^* - \overline{y^*})}{\sum_{i=1}^n (x_i - \overline{x})^2} = \frac{\sum_{i=1}^n (x_i - \overline{x})(y_i - \overline{y})}{\sum_{i=1}^n (x_i - \overline{x})^2} = \frac{S_{xy}}{S_{xx}},$$

$$\hat{\beta}_0^* = \overline{y^*} - \hat{\beta}_1 \overline{x}$$

However, one can establish that $\overline{y^*} = 0$; thus,

$$\hat{\beta}_0^* = -\hat{\beta}_1 \overline{x}.$$

As a result,

$$\hat{y}_i^* = -\hat{\beta}_1 (x_i - \overline{x}).$$

The corresponding estimate of σ^2 is

$$s_I^2 = \frac{1}{n-2}\sum_{i=1}^{n} (y_i^* - \hat{y}_i^*)^2 = \frac{1}{n-2}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

It is useful to see the relationship between $\hat{\beta}_0$ for the uncentered model and $\hat{\beta}_0^*$ for the centered model:

$$\hat{\beta}_0 = \overline{y} - \hat{\beta}_1 \overline{x} = \overline{y} + \hat{\beta}_0^*.$$

Let y_0 be a future observed value of the response after calibration. The estimated value of x_0 corresponding to y_0 is

$$\hat{x}_{I0} = \frac{y_0 - \overline{y} - \hat{\beta}_0^*}{\hat{\beta}_1} = \frac{y_0 - \hat{\beta}_0}{\hat{\beta}_1}. \tag{3}$$

Next, we want to find a prediction interval for x_0. From equation (2), we know that

$$x_0 = \frac{y_0 - \overline{y} - \beta_0^* - \epsilon_0}{\beta_1} = \frac{y_0 - \beta_0 - \epsilon_0}{\beta_1}.$$

Then, the predicted value is

$$\hat{x}_{I0,pred} = \frac{y_0 - \hat{\beta}_0 - \hat{\epsilon}_0}{\hat{\beta}_1}.$$

Since $\hat{x}_{I0,pred}$ is the ratio of two dependent normal random variables, its properties are difficult to determine. However, we can use the Delta Method to obtain an asymptotic approximation for the variance. Let U and V be two random variables. From Casella and Berger (2002), we note

$$\mathbf{Var}\left(\frac{U}{V}\right) \approx \frac{\mathbf{Var}(U)}{\mathbf{E}(V)^2} + \frac{\mathbf{E}(U)^2}{\mathbf{E}(V)^4}\mathbf{Var}(V) - 2\frac{\mathbf{E}(U)}{\mathbf{E}(V)^3}\mathbf{Cov}(U,V).$$

In this case, $U = y_0 - \hat{\beta}_0 - \hat{\epsilon}_0$ and $V = \hat{\beta}_1$, which gives

$$\mathbf{E}(U) = \beta_1 x_0,$$

$$\mathbf{E}(V) = \beta_1,$$

$$\mathbf{Var}\,(U) = \sigma^2\left(1 + \frac{1}{n} + \frac{\overline{x}^2}{S_{xx}}\right),$$

$$\mathbf{Var}\,(V) = \frac{\sigma^2}{S_{xx}}, \quad \text{and}$$

$$\mathbf{Cov}\,(U,V) = \frac{\sigma^2\overline{x}}{S_{xx}}.$$

Thus,

$$\mathbf{Var}\,(\hat{x}_{I0,pred}) \approx \frac{\sigma^2}{\beta_1^2}\left(1 + \frac{1}{n} + \frac{(x_0 - \overline{x})^2}{S_{xx}}\right),$$

and so a $(1-\alpha)100\%$ prediction interval for x_0 based on the Delta Method is

$$\hat{x}_{I0} \pm t_{1-\frac{\alpha}{2},n-2}\frac{s_I}{\hat{\beta}_1}\sqrt{1 + \frac{1}{n} + \frac{(\hat{x}_{I0} - \overline{x})^2}{S_{xx}}}. \tag{4}$$

In reverse regression, we treat x as the response and y as the regressor. In this case, we express our estimated model by

$$\hat{x}_i = \hat{\gamma}_0 + \hat{\gamma}_1(y_i - \overline{y}), \tag{5}$$

where

$$\hat{\gamma}_0 = \overline{x},$$

$$\hat{\gamma}_1 = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sum_{i=1}^{n}(y_i - \overline{y})^2} = \frac{S_{xy}}{S_{yy}}, \quad \text{and}$$

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

Note this method violates the simple linear regression assumption that the regressor is measured with negligible error. For a future observed value y_0, the estimate of x_0 using reverse regression is

$$\hat{x}_{R0} = \hat{\gamma}_0 + \hat{\gamma}_1(y_0 - \overline{y}). \tag{6}$$

A $(1-\alpha)100\%$ prediction interval for x_0 based on reverse regression is

$$\hat{x}_{R0} \pm t_{1-\frac{\alpha}{2},n-2} s_R \sqrt{1 + \frac{1}{n} + \frac{(y_0 - \overline{y})^2}{S_{yy}}}. \tag{7}$$

It is important to note that y_0, $\overline{y}$, and S_{yy} are all random, unlike the usual regression case where they are fixed effects. Consequently, there is more variability to the prediction interval for reverse regression than the naive user realizes. Typically, one runs a single calibration experiment and then uses the results for several, if not many, predictions. The user has no basis for evaluating the full impact of the additional randomness upon the quality of the resulting prediction interval.

Bias in Prediction

Both the inverse and reverse estimators involve the ratio of two dependent normal random variables, which makes the derivation of the expected values difficult. However, the Delta Method allows us to obtain an asymptotic approximation. From Pham-Gia, Turkkan, and Marchand (2006), we note that for two random variables U and V,

$$\mathbf{E}\left(\frac{U}{V}\right) \approx \frac{\mathbf{E}(U)}{\mathbf{E}(V)} + \frac{\mathbf{E}(U)}{\mathbf{E}(V)^3}\mathbf{var}(V) - \frac{\mathbf{cov}(U,V)}{\mathbf{E}(V)^2}.$$

In the case of the inverse estimator in equation (3), $U = y - \hat{\beta}_0$ and $V = \hat{\beta}_1$, which gives

$$\mathbf{E}(U) = \beta_1 x,$$

$$\mathbf{E}(V) = \beta_1,$$

$$\mathbf{Var}\,(V) = \frac{\sigma^2}{S_{xx}}, \quad \text{and}$$

$$\mathbf{Cov}\,(U, V) = \frac{\sigma^2 \overline{x}}{S_{xx}}.$$

Thus, an approximation of the expected value of the inverse estimator using the Delta Method is

$$\mathbf{E}\,(\hat{x}_I) \approx x + \frac{(x - \overline{x})\sigma^2}{\beta_1^2 S_{xx}}.$$

The bias is $\mathbf{E}\,(\hat{x}_I) - x$; thus,

$$\mathbf{bias}(\hat{x_I}) \approx x + \frac{(x - \overline{x})\sigma^2}{\beta_1^2 S_{xx}} - x = \frac{(x - \overline{x})\sigma^2}{\beta_1^2 S_{xx}} \tag{8}$$

This relationship suggests that there should be no bias at $\overline{x}$. It also suggests that the bias is negative for $x < \overline{x}$ and is positive for $x > \overline{x}$. Finally, the rate of change in the bias as x increases is

$$\frac{(x - \overline{x})\sigma^2}{\beta_1^2 S_{xx}}.$$

For the reverse estimator in equation (6), it is trivial to show that

$$\mathbf{E}(\hat{\gamma}_0) = \overline{x}.$$

We observe that $\hat{\gamma}_1$ and $\overline{y}$ are independent. We also note that $\hat{\gamma}_1$ and y are independent since y was not used to estimate γ_1. As a result, one can show that

$$\mathbf{E}(\hat{x}_R) = \overline{x} + (x - \overline{x})\beta_1 \mathbf{E}(\hat{\gamma}_1). \tag{9}$$

Since $\hat{\gamma}_1$ is the ratio of two random variables, the Delta Method allows us to approximate the expected value. In this case, $U = S_{xy}$, and $V = S_{yy}$. One can show that

$$\mathbf{E}(U) = \beta_1 S_{xx},$$

$$\mathbf{E}(V) = (n-1)\sigma^2 + \beta_1^2 S_{xx},$$

$$\mathbf{var}(V) = 2(n-1)\sigma^4 + 4\sigma^2\beta_1^2 S_{xx} \quad \text{and}$$

$$\mathbf{cov}(U, V) = 2\beta_1\sigma^2 S_{xx}.$$

Thus,

$$\mathrm{E}(\hat{\gamma}_1) \approx \frac{\beta_1 S_{xx}}{(n-1)\sigma^2 + \beta_1^2 S_{xx}} - \frac{2\beta_1\sigma^2 S_{xx}}{[(n-1)\sigma^2 + \beta_1^2 S_{xx}]^2} + \frac{[\beta_1 S_{xx}][2(n-1)\sigma^4 + 4\sigma^2\beta_1^2 S_{xx}]}{[(n-1)\sigma^2 + \beta_1^2 S_{xx}]^3}$$

We note then

$$\mathrm{E}(\hat{\gamma}_1) \approx \frac{\beta_1 S_{xx}}{(n-1)\sigma^2 + \beta_1^2 S_{xx}} + o\left(\frac{1}{n}\right).$$

As a result, a reasonable approximation for moderate sample size is

$$\mathrm{E}(\hat{\gamma}_1) \approx \frac{\beta_1 S_{xx}}{(n-1)\sigma^2 + \beta_1^2 S_{xx}},$$

which we may rewrite as

$$\mathrm{E}(\hat{\gamma}_1) \approx \frac{1}{\frac{(n-1)\sigma^2}{\beta_1 S_{xx}} + \beta_1}. \tag{10}$$

Substituting (10) into (9), we obtain

$$\mathbf{E}(\hat{x}_R) \approx \overline{x} + \frac{(x-\overline{x})\beta_1}{\frac{(n-1)\sigma^2}{\beta_1 S_{xx}} + \beta_1}. \tag{11}$$

With some algebra, one can show that the resulting bias is

$$\mathbf{bias}(\hat{x}_R) \approx \overline{x} + \frac{(x-\overline{x})\beta_1}{\frac{(n-1)\sigma^2}{\beta_1 S_{xx}} + \beta_1} - x = \frac{-(x-\overline{x})}{1 + \frac{\beta_1^2 S_{xx}}{(n-1)\sigma^2}}. \tag{12}$$

This relationship suggests that the bias is zero for $x = \overline{x}$. It also suggests that the bias is positive for $x < \overline{x}$ and that the bias is negative for $x > \overline{x}$. Finally, this relationship suggests that the rate of change for the bias as x increases is

$$\frac{-1}{1 + \frac{\beta_1^2 S_{xx}}{(n-1)\sigma^2}}$$

If we scale the x's used in the calibration experiment to be in the range $[-1, 1]$, then the maximum size for S_{xx} is n. However, this maximum occurs when half the points are at -1, and half are at 1. One immediate consequence is that $\overline{x} = 0$. Putting runs

into the calibration experiment to test for lack-of-fit reduces S_{xx}, which increases the bias in $\hat{x}_I$ and $\hat{x}_R$. Finally, assume that $\beta_1 = 1$ and that n is moderately large. If we choose the x's in the calibration experiment to maximize S_{xx}, then

$$\textbf{bias}(\hat{x}_I) \approx \frac{x \cdot \sigma^2}{n},$$

and

$$\textbf{bias}(\hat{x}_R) \approx -\frac{x}{1 + \frac{1}{\sigma^2}},$$

which gives the basis for the lower bounds on the bias. These relationships help explain the bias behavior observed in the simulation study. The bias for inverse regression is lower than the bias for reverse regression whenever $n > \sigma^2 + 1$, which is true for all realistic calibration experiments. Replicates improve the bias for inverse regression, but surprisingly, they do not help reverse regression. Of course, replicates help both inverse and reverse regression in terms of the width of their intervals.

It is no surprise that the inverse regression approach produces slightly biased estimates. However, many analysts do not realize that the forward regression approach also produces biased estimates. This bias is the direct result of reversing the roles of the regressor and the response. Most analysts who uses the reverse regression approach naively assume that one can arbitrarily reverse the roles, since the mathematics appear to allow it. As our results point out, the reality is much more subtle.

These bias results suggest that one can create bias correction factors. However, it is important to consider the size of this bias for most calibration experiments. An adequate instrument typically has a σ on the order of 0.01 (a 1% instrument). An extremely marginal instrument has a σ on the order of .1. For an adequate instrument, the bias for inverse regression is approximately

$$\frac{.0001 \cdot x}{n} < .0001,$$

and the bias for reverse regression is

$$\frac{1}{1 + 1/.0001} = \frac{1}{10001}.$$

Both biases are practically zero. For a marginal instrument, the bias for inverse regression is

$$\frac{.01 \cdot x}{n} < .01,$$

and the bias for reverse regression is

$$\frac{1}{1+1/.01} = \frac{1}{101}.$$

In light of the reliability of the measurement from a marginal instrument, these biases are minor. One could attempt to correct for these biases; however, the variability introduced from estimating σ^2, which bias corrections would require, would certainly render any improvement negligible.

Simulation Study

The analytic results presented in this paper are all asymptotic. As a result, we conducted a simulation study to compare the prediction intervals based on the Delta Method and the reverse regression approach for small sample sizes. The calibration experiment used n known values of x and obtained the corresponding values of y using equation (1) under the assumption that $\beta_0 = 0$ and $\beta_1 = 1$. Essentially, this study assumed an appropriate centering and scaling, which is always an option in these kinds of experiments. The study used these n pairs of observations to compute $\hat{\beta}_0$, $\hat{\beta}_1$, s_I^2, $\hat{\gamma}_0$, $\hat{\gamma}_1$, and s_R^2. Next, the study observed a value of y for each of x = -1, -0.5, 0, 0.5, and 1, again using equation (1). The values of x corresponding to these five values of y were then estimated using both the inverse estimator in equation (3) and the reverse estimator in (5), and the biases were computed. Finally, the study found 95% prediction intervals using the Delta Method in equation (4) and reverse regression (7). The study also computed intervals based on the methods of Scheffé (1973) and Eberhardt and Mee (1994). For a given set of coefficients, these estimates and intervals were computed M times, and the mean bias, mean interval widths, and capture probabilities were examined. This entire process was then repeated N times.

The simulation study considered five different designs for the calibration experiment: $x = \{-1, 1\}$ with 3, 5, and 10 replicates at each point, $x = \{-1, -1, 0, 0, 1, 1\}$, and $x = \{-1, 0, 0, 0, 0, 1\}$. The true standard deviations σ used were 0.01, which represents a "1%" or reasonably precise device, 0.1, which represents a "10%" or at best a borderline precision device, and 0.5, where one should expect the inverse regression to break down. Simulations were run with $N = 100,000$ and $M = 1$, so that the least squares estimates of the coefficients were recalculated each time the intervals were computed. However, in practical applications, the same estimated model would be used many times, so simulations were also run with $N = 1,000$ and $M = 100$. Results for some of the simulations are shown in Tables 1 - 6.

For $\sigma = 0.01$ or 0.1, the intervals based on the Delta Method and reverse regression had approximately the same mean widths and capture probabilities of 95%. Both Scheffé's (1973) and Eberhardt and Mee's (1994) methods are very conservative and resulted in intervals with capture probabilities of more than 99%, and mean widths

much wider than those of the intervals found using the Delta Method and reverse regression. In fact, intervals computed using Scheffé's (1973) were usually more than twice as wide. In the case of high variability, i.e. $\sigma = 0.5$, Scheffé's (1973) method did not work for most designs, and prediction intervals found using the Delta Method were much wider than those using reverse regression. The simulation results also show that both the inverse and reverse estimators are biased and that these biases increase as σ increases.

Table 1 A. Mean interval widths for $x = \{-1, -1, 0, 0, 1, 1\}$, $N = 1,000$ and $M = 100$.

Sigma	Method	-1	-0.5	0	0.5	1
0.01	Scheffé	0.200	0.182	0.174	0.182	0.200
	EM	0.121	0.121	0.121	0.121	0.121
	Delta	0.061	0.057	0.055	0.057	0.061
	Reverse	0.061	0.057	0.055	0.057	0.061
0.1	Scheffé	2.349	2.214	2.172	2.213	2.347
	EM	1.227	1.227	1.227	1.227	1.227
	Delta	0.620	0.577	0.562	0.577	0.619
	Reverse	0.615	0.573	0.558	0.573	0.615
0.5	Scheffé	.	.	.	.	.
	EM	6.923	6.923	6.923	6.923	6.923
	Delta	6.495	5.027	4.370	4.802	6.465
	Reverse	2.868	2.692	2.632	2.694	2.871

Table 1 B. Mean bias for $x = \{-1, -1, 0, 0, 1, 1\}$, $N = 1,000$ and $M = 100$.

Sigma	Method	-1	-0.5	0	0.5	1	Intercept	Slope
0.01	Inverse	1.197e-4	1.422e-4	1.248e-4	8.738e-5	5.415e-5	1.057e-4	-3.718e-5
	Reverse	2.334e-5	9.401e-5	1.248e-4	1.355e-4	1.505e-4	1.056e-4	5.916e-5
0.1	Inverse	0.005	0.004	0.003	0.002	0.001	0.003	-0.002
	Reverse	-0.005	-0.001	0.003	0.007	0.011	0.003	0.008
0.5	Inverse	0.125	0.062	-0.004	-0.065	-0.131	-0.0026	-0.1278
	Reverse	-0.179	-0.093	-0.004	0.087	0.172	-0.0034	0.1764

Table 1 C. Mean variance of the interval widths for $x = \{-1, -1, 0, 0, 1, 1\}$, $N = 1,000$ and $M = 100$.

Sigma	Method	-1	-0.5	0	0.5	1
0.01	Inverse	5.164e-4	4.480e-4	4.252e-4	4.480e-4	5.165e-4
	Reverse	5.162e-4	4.479e-4	4.251e-4	4.479e-4	5.163e-4
0.1	Inverse	0.051	0.044	0.041	0.044	0.051
	Reverse	0.049	0.042	0.040	0.042	0.049
0.5	Inverse	10023	4335	1995	3186	9593
	Reverse	1.375	1.063	0.959	1.078	1.378

Table 1 D. Capture probabilities for $x = \{-1, -1, 0, 0, 1, 1\}$, $N = 1,000$ and $M = 100$.

Sigma	Method	-1	-0.5	0	0.5	1
0.01	Scheffé	0.999	0.999	0.999	0.998	0.998
	EM	0.994	0.995	0.996	0.995	0.993
	Delta	0.945	0.946	0.946	0.945	0.946
	Reverse	0.945	0.946	0.946	0.945	0.946
0.1	Scheffé	0.999	0.999	0.999	0.998	0.999
	EM	0.995	0.997	0.997	0.997	0.996
	Delta	0.950	0.951	0.950	0.951	0.950
	Reverse	0.950	0.951	0.951	0.951	0.950
0.5	Scheffé	.	.	.	.	.
	EM	0.993	0.995	0.996	0.995	0.994
	Delta	0.952	0.961	0.966	0.963	0.956
	Reverse	0.954	0.959	0.962	0.962	0.958

Table 2 A. Mean interval widths for $x = \{-1, 0, 0, 0, 0, 1\}$, $N = 1,000$ and $M = 100$.

Sigma	Method	-1	-0.5	0	0.5	1
0.01	Scheffé	0.226	0.195	0.180	0.195	0.226
	EM	0.123	0.123	0.123	0.123	0.123
	Delta	0.067	0.059	0.056	0.059	0.067
	Reverse	0.067	0.059	0.056	0.059	0.067
0.1	Scheffé	3.284	3.053	2.994	3.053	3.283
	EM	1.256	1.256	1.256	1.256	1.256
	Delta	0.690	0.606	0.575	0.606	0.690
	Reverse	0.679	0.597	0.568	0.597	0.678
0.5	Scheffé	.	.	.	.	.
	EM	7.658	7.658	7.658	7.658	7.658
	Delta	16.926	11.057	9.572	13.165	20.347
	Reverse	2.782	2.526	2.434	2.522	2.772

Table 2 B. Mean bias for $x = \{-1, 0, 0, 0, 0, 1\}$, $N = 1,000$ and $M = 100$.

Sigma	Method	-1	-0.5	0	0.5	1	Intercept	Slope
0.01	Inverse	3.697e-4	1.244e-4	-6.728e-5	-3.166e-4	-5.785e-4	-9.366e-5	-4.675e-4
	Reverse	1.689e-4	2.401e-5	-6.730e-5	-2.163e-4	-3.777e-4	-9.368e-5	-2.667e-4
0.1	Inverse	0.009	0.003	0.001	-0.003	-0.008	4.000e-4	-0.008
	Reverse	-0.012	-0.007	0.001	0.007	0.013	4.000e-4	0.0128
0.5	Inverse	0.223	0.112	-0.008	-0.116	-0.228	-0.0034	-0.226
	Reverse	-0.316	-0.156	0.004	0.166	0.327	0.005	0.3216

Table 2 C. Mean variance of the interval widths for $x = \{-1, 0, 0, 0, 0, 1\}$, $N = 1,000$ and $M = 100$.

Sigma	Method	-1	-0.5	0	0.5	1
0.01	Inverse	6.205e-4	4.807e-4	4.339e-4	4.805e-4	6.202e-4
	Reverse	6.200e-4	4.803e-4	4.337e-4	4.801e-4	6.197e-4
0.1	Inverse	0.069	0.052	0.046	0.052	0.069
	Reverse	0.063	0.048	0.043	0.048	0.063
0.5	Inverse	1.703e5	6.515e4	4.697e4	1.100e5	2.590e5
	Reverse	1.314	0.898	0.752	0.904	1.325

Table 3 A. Mean interval widths for $x = \{-1, -1, -1, 1, 1, 1\}$, $N = 1,000$, and $M = 100$.

Sigma	Method	-1	-0.5	0	0.5	1
0.01	Scheffé	0.193	0.180	0.175	0.180	0.193
	EM	0.122	0.122	0.122	0.122	0.122
	Delta	0.060	0.057	0.056	0.057	0.060
	Reverse	0.060	0.057	0.056	0.057	0.060
0.1	Scheffé	2.171	2.071	2.039	2.071	2.171
	EM	1.241	1.241	1.241	1.241	1.241
	Delta	0.607	0.578	0.568	0.578	0.608
	Reverse	0.605	0.576	0.566	0.576	0.605
0.5	Scheffé	.	.	.	.	.
	EM	6.385	6.385	6.385	6.385	6.385
	Delta	3.240	3.060	2.994	3.053	3.228
	Reverse	2.841	2.715	2.670	2.710	2.831

Table 3 B. Mean bias for $x = \{-1, -1, -1, 1, 1, 1\}$, $N = 1,000$ and $M = 100$.

Sigma	Method	-1	-0.5	0	0.5	1	Intercept	Slope
0.01	Inverse	2.988e-4	1.640e-4	-1.020e-6	-1.070e-4	-2.281e-4	2.534e-5	-2.650e-4
	Reverse	2.330e-4	1.311e-4	-1.018e-6	-7.414e-5	-1.623e-4	2.533e-5	-1.992e-4
0.1	Inverse	6.408e-4	-4.455e-4	-8.823e-4	-0.003	-0.004	-1.537e-3	-2.367e-3
	Reverse	-0.006	-0.004	-8.814e-4	7.777e-4	0.003	-1.421e-3	4.556e-3
0.5	Inverse	0.048	0.029	-0.012	-0.006	-0.024	0.007	-0.036
	Reverse	-0.118	-0.054	-0.010	0.074	0.138	0.006	0.128

Table 3 C. Mean variance of the interval widths for $x = \{-1, -1, -1, 1, 1, 1\}$, $N = 1,000$ and $M = 100$.

Sigma	Method	-1	-0.5	0	0.5	1
0.01	Inverse	4.778e-4	4.329e-4	4.180e-4	4.329e-4	4.777e-4
	Reverse	4.776e-4	4.328e-4	4.179e-4	4.328e-4	4.776e-4
0.1	Inverse	0.053	0.047	0.046	0.047	0.053
	Reverse	0.051	0.046	0.045	0.046	0.051
0.5	Inverse	2.784	2.175	1.960	2.163	2.781
	Reverse	1.224	1.047	0.981	1.033	1.196

Table 4 A. Mean interval widths for $x = \{-1, 1\}$ with 5 replicates at each point, $N = 1,000$, and $M = 100$.

Sigma	Method	-1	-0.5	0	0.5	1
0.01	Scheffé	0.119	0.111	0.108	0.111	0.119
	EM	0.086	0.086	0.086	0.086	0.086
	Delta	0.049	0.048	0.047	0.048	0.049
	Reverse	0.049	0.048	0.047	0.048	0.049
0.1	Scheffé	1.211	1.144	1.119	1.144	1.211
	EM	0.857	0.857	0.857	0.857	0.857
	Delta	0.492	0.477	0.471	0.477	0.492
	Reverse	0.490	0.474	0.469	0.474	0.490
0.5	Scheffé	.	.	.	.	.
	EM	4.269	4.269	4.269	4.269	4.269
	Delta	2.489	2.407	2.379	2.409	2.494
	Reverse	2.220	2.160	2.140	2.161	2.223

Table 4 B. Mean bias for $x = \{-1, 1\}$ with 5 replicates at each point, $N = 1,000$ and $M = 100$.

Sigma	Method	-1	-0.5	0	0.5	1	Intercept	Slope
0.01	Inverse	-6.943e-5	-1.050e-4	-1.211e-4	-3.010e-5	1.145e-5	-6.284e-5	4.733e-5
	Reverse	-1.509e-4	-1.457e-4	-1.210e-4	1.063e-5	9.292e-5	-6.281e-5	1.288e-4
0.1	Inverse	0.001	2.540e-4	4.547e-4	-7.550e-4	-8.652e-5	1.734	-6.364e-4
	Reverse	-0.007	-0.004	4.410e-4	0.003	0.007	-1.118e-4	0.007
0.5	Inverse	0.012	0.004	-0.007	-0.019	-0.030	-0.008	0.021
	Reverse	-0.157	-0.080	-0.005	0.069	0.144	-0.006	0.150

Table 4 C. Mean variance of the interval widths for $x = \{-1, 1\}$ with 5 replicates at each point, $N = 1,000$ and $M = 100$.

Sigma	Method	-1	-0.5	0	0.5	1
0.01	Inverse	0.313	0.282	0.272	0.284	0.318
	Reverse	0.205	0.189	0.184	0.190	0.208
0.1	Inverse	0.015	0.014	0.013	0.014	0.014
	Reverse	0.014	0.013	0.013	0.013	0.014
0.5	Inverse	0.594	0.524	0.503	0.530	0.608
	Reverse	0.303	0.278	0.270	0.280	0.308

Table 5 A. Mean interval widths for $x = \{-1, 1\}$ with 10 replicates at each point, $N = 1,000$, and $M = 100$.

Sigma	Method	-1	-0.5	0	0.5	1
0.01	Scheffé	0.081	0.077	0.075	0.077	0.081
	EM	0.065	0.065	0.065	0.065	0.065
	Delta	0.044	0.043	0.043	0.043	0.044
	Reverse	0.044	0.043	0.043	0.043	0.044
0.1	Scheffé	0.815	0.774	0.758	0.774	0.815
	EM	0.647	0.647	0.647	0.647	0.647
	Delta	0.438	0.430	0.428	0.430	0.438
	Reverse	0.435	0.428	0.426	0.428	0.435
0.5	Scheffé	5.468	5.385	5.360	5.385	5.468
	EM	3.281	3.281	3.281	3.281	3.281
	Delta	2.235	2.195	2.181	2.194	2.235
	Reverse	1.983	1.955	1.945	1.955	1.983

Table 5 B. Mean bias for $x = \{-1, 1\}$ with 10 replicates at each point, $N = 1,000$ and $M = 100$.

Sigma	Method	-1	-0.5	0	0.5	1	Intercept	Slope
0.01	Inverse	1.365e-4	9.139e-5	1.051e-4	4.927e-5	4.111e-5	8.467e-5	-4.658e-5
	Reverse	4.480e-5	4.555e-5	1.051e-4	9.511e-5	1.328e-4	8.467e-5	1.080e-5
0.1	Inverse	0.001	9.699e-4	8.257e-4	3.806e-4	3.063e-5	6.414e-4	-5.056e-4
	Reverse	-0.008	-0.004	8.200e-4	0.005	0.009	5.640e-4	0.009
0.5	Inverse	0.015	0.011	-3.283e-4	-0.004	-0.013	0.002	-0.014
	Reverse	-0.180	-0.087	-4.835e-4	0.092	0.181	0.001	0.180

Table 5 C. Mean variance of the interval widths for $x = \{-1, 1\}$ with 10 replicates at each point, $N = 1,000$ and $M = 100$.

Sigma	Method	-1	-0.5	0	0.5	1
0.01	Inverse	5.581e-5	5.391e-5	5.327e-5	5.391e-5	5.581e-5
	Reverse	5.579e-5	5.389e-5	5.326e-5	5.389e-5	5.580e-5
0.1	Inverse	5.365e-3	5.176e-3	5.114e-3	5.179e-3	5.369e-3
	Reverse	5.210e-3	5.032e-3	4.973e-3	5.034e-3	5.214e-3
0.5	Inverse	0.246	0.229	0.222	0.228	0.246
	Reverse	0.118	0.112	0.110	0.112	0.117

Table 6 A. Mean interval widths for $x = \{-1, -1, 0, 0, 1, 1\}$, $N = 100,000$ and $M = 1$.

Sigma	Method	-1	-0.5	0	0.5	1
0.01	Scheffé	0.204	0.186	0.178	0.186	0.204
	EM	0.124	0.124	0.124	0.124	0.124
	Delta	0.062	0.058	0.056	0.058	0.062
	Reverse	0.062	0.058	0.056	0.058	0.062
0.1	Scheffé	2.376	2.241	2.199	2.241	2.376
	EM	1.236	1.236	1.236	1.236	1.236
	Delta	0.624	0.581	0.566	0.581	0.624
	Reverse	0.619	0.577	0.562	0.577	0.619
0.5	Scheffé	.	.	.	.	.
	EM	6.741	6.741	6.741	6.741	6.741
	Delta	9.029	12.018	10.734	16.613	22.927
	Reverse	2.881	2.706	2.644	2.705	2.880

Table 6 B. Mean bias for $x = \{-1, -1, 0, 0, 1, 1\}$, $N = 100,000$ and $M = 1$.

Sigma	Method	-1	-0.5	0	0.5	1	Intercept	Slope
0.01	Inverse	6.767e-5	6.382e-5	6.398e-6	3.003e-5	-4.129e-5	2.533e-5	-5.034e-5
	Reverse	-3.270e-5	1.363e-5	6.392e-6	8.020e-5	5.906e-5	2.446e-5	5.174e-5
0.1	Inverse	0.002	5.279e-4	-2.070e-4	-0.001	-0.003	-3.358e-4	-2.306e-4
	Reverse	-0.008	-0.004	-2.007e-4	0.004	0.007	-2.401e-4	0.008
0.5	Inverse	0.088	0.038	-1.064e-4	-0.051	-0.095	-0.004	-0.091
	Reverse	-0.173	-0.085	0.004	0.087	0.176	0.002	0.174

Table 6 C. Mean variance of the interval widths for $x = \{-1, -1, 0, 0, 1, 1\}$, $N = 100,000$ and $M = 1$.

Sigma	Method	-1	-0.5	0	0.5	1
0.01	Inverse	5.136e-5	4.456e-5	4.229e-5	4.455e-5	5.135e-5
	Reverse	5.134e-5	4.454e-5	4.227e-5	4.454e-5	5.133e-5
0.1	Inverse	0.053	0.046	0.043	0.046	0.053
	Reverse	0.051	0.044	0.042	0.044	0.051
0.5	Inverse	2.870e6	7.551e6	5.665e6	1.764e7	3.704e7
	Reverse	1.280	1.012	9.223	1.016	1.280

Conclusions

In general, the prediction intervals for both delta and reverse regression approaches, on the average, have essentially correct coverage probabilities. We show the results for only one simulation case, but all of the other simulation cases are similar. The width of the prediction interval for reverse regression is slightly less variable than for inverse regression, which suggests some preference for using reverse regression. As expected, the Scheffé and the Eberhardt and Mee methods are too conservative.

For a reasonably precise instrument, $\sigma = 0.01$, the delta and reverse regression prediction interval widths are virtually the same, while the Scheffé and Eberhardt and Mee prediction intervals are two to three times larger. The bias is practically zero for both the inverse and reverse regression methods. As a result, the mean square error is dominated by the variance.

For a borderline precision instrument $\sigma = 0.1$, the reverse regression prediction interval widths are slightly smaller than delta method. Once again, the Scheffé and the Eberhardt and Mee prediction intervals are two to three times larger. The bias depends on location, but the inverse regression results appear to be generally slightly better. Once again, the variance dominates the bias. As a result, the mean squared error is effectively the variance.

The $\sigma = 0.5$ case illustrates a situation where anything based on inverse regression breaks down. The Scheffé intervals are not computable in most cases. The delta method prediction intervals are significantly wider than reverse regression; however, they still are smaller than the Eberhardt and Mee intervals.

The simulation study confirms that both the inverse and the reverse regression approaches lead to biased predictions. As expected, inverse regression has less bias than reverse regression as we predict further away from 0. In both cases the bias depends on location and follows a linear trend. Both approaches give unbiased predictions at 0. The bias for inverse regression has a positive slope while the bias for reverse regression has a negative slope. Increasing the number of design points at the interval

boundaries decreases the bias for inverse regression; however, it has no effect on the bias for reverse regression. In general, inverse regression tends to be less biased than reverse regression. The one case where forward regression appears to have less bias is the rather unrealistic situation with only two points on the interval boundaries and many points at the interval center. Such a situation does provide good ability to test for lack-of-fit, but it fails to estimate the basic relationship well.

The bottom-line conclusions are somewhat mixed. Reverse regression appears to have a slight edge in terms of the width of the prediction intervals for reasonable and borderline instruments. On the other hand, inverse regression appears to have a definite edge in terms of bias for calibration experiments that replicate the interval boundaries.

It is important to note that this study only considers a simple linear relationship. Future research intends to study more complicated calibration relationships including more than one calibration factor and higher-order models.

References

Berkson, J. (1969). "Estimation of a Linear Function for a Calibration Line; Consideration of a Recent Proposal," Technometrics, 11, 649-660.

Casella, G. and Berger, R.L. (2002). *Statistical Inference*, 2nd ed., Duxbury, CA.

Eberhardt, K.R., and Mee, R.W. (1994). "Constant-Width Calibration Intervals for Linear Regression," Journal of Quality Technology, 26, 21-29.

Graybill, F. (1976). *Theory and Application of the Linear Model*, Duxbury, North Scituate, MA.

Halperin, M. (1970). "On Inverse Estimation in Linear Regression," Technometrics, 12, 727-736.

Krutchkoff, R.G. (1967). "Classical and Inverse Regression Methods of Calibration," Technometrics, 9, 425-439.

Krutchkoff, R.G. (1969). "Classical and Inverse Regression Methods of Calibration in Extrapolation," Technometrics, 11, 605-608.

Pham-Gia, T., Turkkan, N., and Marchand, E. (2006). "Density of the Ratio of Two Normal Random Variables and Applications," Communications in Statistics - Theory and Methods, 35, 1569-1591.

Scheffe, H. (1973). "A Statistical Theory of Calibration," The Annals of Statistics, 1, 1-37.

Seber, G.A.F. (1977). *Linear Regression Analysis*, Wiley, NY.

Shalabh. (2001). "Least Squares Estimators in Measurement Error Models under the Balanced Loss Function," Sociedad de Estadistica e Investigacion Operativa, 10, 301-308.

Srivastava, A.K., and Shalabh. (1997). "Asymptotic Efficiency Properties of Least squares in an Ultrastructural Model," Sociedad de Estadistica e Investigacion Operativa, 6, 419-431.

Williams, E.J. (1969). "A Note on Regression Methods in Calibration," Technometrics, 11, 189-192.