Model based bootstrapping when correcting for measurement error with application to logistic regression.

John P. Buonaccorsi¹*, Giovanni Romeo² and Magne Thoresen²

¹ Dept. of Mathematics and Statistics, University of Massachusetts, Amherst, MA, USA
² Oslo Centre for Biostatistics and Epidemiology, Department of Biostatistics, University of Oslo, Norway.

*email: johnpb@math.umass.edu

Summary: When fitting regression models, measurement error in any of the predictors typically leads to biased coefficients and incorrect inferences. A plethora of methods have been proposed to correct for this. Obtaining standard errors and confidence intervals using the corrected estimators can be challenging and, in addition, there is concern about remaining bias in the corrected estimators. The bootstrap, which is one option to address these problems, has received limited attention in this context. It has usually been employed by simply resampling observations, which, while suitable in some situations, is not always formally justified. In addition, the simple bootstrap does not allow for estimating bias in non-linear models, including logistic regression. Model based bootstrapping, which can potentially estimate bias in addition to being robust to the original sampling or whether the measurement error variance is constant or not, has received limited attention. However, it faces challenges that are not present in handling regression models with no measurement error. This paper develops new methods for model based bootstrapping when correcting for measurement error in logistic regression with replicate measures. The methodology is illustrated using two examples and a series of simulations are carried out to assess and compare the simple and model based bootstrap methods, as well as other standard methods. While not always perfect, the model based approaches offer some distinct improvements over the other methods.

Key words: Bootstrap, Logistic regression, Measurement error, Replication.

This paper has been submitted for consideration for publication in Biometrics
1. Introduction

In fitting regression models, measurement error in predictors may result in biased estimates of coefficients and other incorrect inferences. This problem has received a tremendous amount of attention (see for example Buonaccorsi (2010), Carroll et al. (2006), Gustafson (2004) and Fuller (1987) for general treatments) and a plethora of correction techniques have been proposed for both linear and non-linear models. Inferences can be challenging for a variety of reasons. While analytical standard errors are available for some methods these are approximate in nature and usually involve some underlying assumptions. Relatedly, Wald type confidence intervals based on these standard errors rely on approximate normality and unbiasedness of the estimator involved. An additional concern is that the corrected estimators are not unbiased; rather, most are either consistent, or approximately consistent, under suitable conditions. A data driven way of assessing potential bias in either the corrected estimators or naive estimators, which ignore the measurement error, is desirable.

One obvious tool for attacking these problems is the bootstrap, which has received limited attention in the measurement error context. The majority of the applications of the bootstrap in measurement error problems have used simple with replacement resampling of observations. This is used in STATA, one of the few statistical software packages that easily handles measurement error in regression. The simple bootstrap is only formally justified with a random sample and “equal sampling effort” in the measurement error process. Equally important is the fact that with non-linear models the simple bootstrap doesn’t necessarily allow for a bootstrap estimate of bias. This is well known without measurement error where an alternative approach is model based bootstrapping.

The goal of this paper is to examine the use of model based bootstrap methods when correcting for additive measurement error in regression with replicate measures of the unobserved true values. Measurement error introduces some substantial challenges that are not
present when using the model based bootstrap in regression without measurement error. Our focus is on logistic regression, an important model which shares with other generalized linear models the feature that implementing the model based bootstrap in the presence of measurement error reduces to two concerns; what to use for true values and how to generate replicates of the error-prone values.

In Section 2 we first describe the regression and measurement error models involved along with laying out the regression calibration method for correcting for measurement error. We concentrate on the most popular correction method, regression calibration, in order to focus the discussion on the bootstrap rather than a comparison of estimators. Section 3 explores the bootstrap methods, first addressing when the simple bootstrap is appropriate and then developing new model based methods. Two examples are presented in Section 4 followed by simulations in Section 5 and a discussion in Section 6.

2. Models and methods

2.1 Models

The logistic model assumes \( P(Y_i = 1|x_i) = E(Y_i|x_i) = m(x_i, \beta) = 1/(1 + e^{-x_i'\beta}) \), where \( Y_i \) is a binary outcome and \( x_i \) denotes the fixed predictors on the \( i \)th observation. When viewed as random, \( X_i \) will denote the random vector and \( x_i \) its realization. We distinguish between the so-called functional case, where the \( x_i \) are treated as fixed, and the “structural” case, taken here to mean the \( X_i \) are independent and identically distributed (i.i.d.) with mean \( \mu_X \) and covariance matrix \( \Sigma_X \).

Instead of observing \( x_i \), we observe \( W_i = x_i + u_i \), with \( E(u_i|x_i) = 0 \) and \( V(u_i|x_i) = \Sigma_{ui} \). The fact that \( E(u_i|x_i) = 0 \) means the measurement error is additive or, equivalently, \( W_i \) is unbiased for \( x_i \). In applications, there are many ways that \( W_i \) may arise and the manner in which it arises will determine both the nature of the measurement error covariance matrix,
\( \Sigma_{ui} \), and how it is estimated. We investigate the frequently studied situation where there are \( k_i \) replicate measures, \( \{ W_{ij}, j = 1 \text{ to } k_i \} \) on the \( i \text{th} \) observation, where \( W_{ij} = x_i + u_{ij} \) and the \( u_{ij} \) are independent with \( E(u_{ij}) = 0 \) and \( \text{Cov}(u_{ij}) = \Sigma_{ui(1)} \) (the measurement error covariance matrix for a single replicate on the \( i \text{th} \) subject). The mean, \( W_i = \sum_{j=1}^{k_i} W_{ij} / k_i \) has expected value \( x_i \), so \( E(u_i) = 0 \) with measurement error covariance matrix \( \Sigma_{ui} = \Sigma_{ui(1)}/k_i \), estimated by \( \hat{\Sigma}_{ui} = S_{Wi}/k_i \), where \( S_{Wi} = \sum_{j=1}^{k_i} (W_{ij} - W_i)(W_{ij} - W_i)'/(k_i - 1) \) is the sample covariance matrix of the replicates on the \( i \text{th} \) observation.

If some predictors are measured perfectly, the elements in the rows and columns of the covariance matrix associated with perfectly measured predictors are all set equal to 0. Although not specified explicitly in the notation, \( \Sigma_{ui(1)} \) is the covariance matrix of a replicate given \( x_i \) and, as such, it could depend on \( x_i \) in some (usually unspecified) manner.

2.2 Correction techniques

When correcting for measurement error in non-linear models, a variety of methods have been proposed. As noted in the introduction, we will focus on regression calibration.

The mean \( \bar{W} = \sum_i W_i / n \) is an unbiased estimate of \( \bar{x} = \sum_i x_i / n \) in the functional case and of \( \mu_X \) in the structural case. We can also construct an unbiased estimator of the covariance matrix of the the true predictors,

\[
\hat{\Sigma}_X = S_W - \hat{\Sigma}_u,
\]

where \( \hat{\Sigma}_u = \sum_{i=1}^n \hat{\Sigma}_{ui} / n \) is the average estimated measurement error covariance matrix, and \( S_W \) is the sample covariance matrix of the \( W_i \)'s. Like the mean vector this can be interpreted two ways. In the functional case, \( E(\hat{\Sigma}_X) = S_x = \sum_i (x_i - \bar{x})(x_i - \bar{x})'/(n - 1) \) while in the structural case \( E(\hat{\Sigma}_X) = \Sigma_X \). Note that these results about the expectation of \( \hat{\Sigma}_X \) do not depend on either a constant per-replicate measurement error covariance matrix nor on an equal number of replicates across subjects.

Regression calibration was originally developed in handling measurement error in gen-
eralized linear models, and logistic regression in particular; see Carroll et al. (2006) or Buonaccorsi (2010) for overviews. It is based on the idea that if $X_i$ is random and the measurement error is non-differential (the model for the error in $x$ does not depend on $Y$) then $E(Y_i|W_i) = E(m(\beta, X_i)|W_i)$. In the linear case this means $E(Y_i|W_i) = \beta_0 + \beta'_1 E(X_i|W_i)$.

In non-linear models the approximation $E(Y|W_i) \approx m(\beta, E(X_i|W_i))$ is used. This argues for correcting for measurement error by fitting the regression model but instead of $W_i$ using an imputed value, which is an estimate of $E(X_i|W_i)$.

Formally, regression calibration is based on assuming random $X$'s.

If $X_i$ is distributed normal with mean $\mu_X$ and covariance matrix $\Sigma_X$ and the measurement error is normal with mean 0 and covariance matrix $\Sigma_{ui}$, then $E(X_i|W_i) = \mu_X + \kappa_i(W_i - \mu_X)$, where $\kappa_i = \Sigma_X(\Sigma_X + \Sigma_{ui})^{-1}$. Without the normality assumption $\mu_X + \kappa_i(W_i - \mu_X)$ is the best linear predictor of $X_i$ given $W_k$. The simplest form of regression calibration regresses $Y_i$ on

$$\tilde{x}_i = \bar{W} + \tilde{\kappa}(w_i - \bar{W})$$

where $\tilde{\kappa} = \tilde{\Sigma}_X(\tilde{\Sigma}_X + \tilde{\Sigma}_{ui})^{-1}$ is the estimated reliability matrix for the $i$th subject using the estimated average measurement error covariance matrix $\tilde{\Sigma}_{ui}$ in place of $\Sigma_{ui}$. The estimators using this $\tilde{x}_i$ will be referred to simply as RC estimators.

Note that (2) could be applied with unequal per-replicate covariance matrices and/or unequal numbers of replicates. With unequal numbers of replicates and random $X$'s, an alternative is to use weighted estimates of $\mu_X$ and $\Sigma_X$ (e.g., Carroll et al. (2006, p. 71)) which leads to additional ways to impute. We limit the discussion here to equal numbers of replicates for space considerations.

An alternative to (2) is to replace $\tilde{\kappa}$ with $\tilde{\kappa}_i = \tilde{\Sigma}_X(\tilde{\Sigma}_X + S_{Wi}/k)^{-1}$, assuming $k$ replicates, which utilizes a subject specific per-replicate measurement error covariance matrix. However, we have found through examples and simulations in various settings (see for example
Buonaccorsi (2010) and Buonaccorsi et al. (2016)) that the resulting RC estimators often behave rather poorly.

Analytical approaches are available for obtaining an estimate of the covariance of the RC estimates, and associated standard errors, as outlined in equations (B.9) and (B.10) in Appendix B.3 of Carroll et al (2006). These are implemented in the rcal command in STATA with the default being the “information-type” asymptotic covariance with an option (one we prefer) that uses the robust sandwich type estimated covariance.

3. Bootstrapping

The reasons to consider bootstrapping for inferences are well known (see, for example Efron and Tibshirani (1993) or Davison and Hinkley (1997) among others) including the potential for estimating the bias of a proposed estimator.

There are two bootstrapping strategies that can be employed in regression contexts. The simple refers to simply resampling observations, while the model based bootstrap explicitly incorporates the regression model. Both generate $B$ (large) bootstrap samples with the estimates obtained from the $b$th sample denoted by $\hat{\beta}_b$. The bootstrap estimate of the standard error of the $j$th coefficient, is the standard deviation of the $B$ values $\hat{\beta}_{1j}, \ldots, \hat{\beta}_{Bj}$, sometimes obtained with some amount of trimming to reduce the effect of outliers. For example, STATA’s rcal procedure trims 1% from each end by default.

The simplest bootstrap confidence interval is the percentile method which uses the $(\alpha/2)th$ and $(1 - \alpha/2)th$ percentiles of the $B$ bootstrap values. We will also consider bias corrected (BC) intervals (e.g., Efron and Tibshirani (1994, p. 185)), which are targeted to correcting for median bias. There are also bias corrected accelerated intervals (e.g., Davison and Hinkley (1997, p. 204) or Efron and Tibshirani (1994, p. 184)) but we don’t pursue these further here both because of the computational demands and since determination of the acceleration constant in a complex situation like this is worthy of separate investigation.
3.1 The simple bootstrap.

In each bootstrap sample, one resamples \( n \) times with replacement from the original data. An observation carries with it \( \mathbf{D}_i = (Y_i, W_i, S_{Wi}, k_i) \) and the \( b \)th bootstrap consists of \( (Y_{bi}, W_{bi}, S_{biW}, k_{bi}) \), for \( i = 1 \) to \( n \), chosen with replacement from the original \( \mathbf{D}_1, \ldots, \mathbf{D}_n \).

Formally, the simple bootstrap is only justified if the \( \mathbf{D}_1, \ldots, \mathbf{D}_n \) are i.i.d. This means that we have a random sample of units from some population, with the same method and “equal sampling effort” for each observations to obtain the mismeasured value \( W_i \) and its estimated measurement error covariance matrix. With replication, equal sampling effort means either an equal number of replicates for all observations or, if the number of replicates are unequal, they are random in a way that doesn’t depend on the position in the sample, \( i \).

3.2 Model based bootstrapping.

Here the bootstrap is implemented by explicitly mimicking the various steps leading to the data; generating responses from the regression model as well as generating replicate values of the error-prone predictors. To implement this we need a set of estimated true values, \( \mathbf{x}_T_1, \ldots, \mathbf{x}_T_n \) which we discuss further after the overall strategy is outlined.

For the \( b \)th bootstrap sample and for \( i = 1 \) to \( n \) first generate

\[
Y_{bi} \sim g(\hat{x}_{Ti}, \hat{\beta}),
\]

where \( g \) is a specified distribution corresponding to a generalized linear model with mean \( m(\hat{x}_{Ti}, \hat{\beta}) \). The distribution is just the Bernoulli distribution for logistic regression with \( P(Y_{bi} = 1) = m(\hat{x}_{Ti}, \hat{\beta}) \). If there were no measurement error then \( \hat{x}_{Ti} = x_i \) and the model based bootstrap would stop here. Other generalized linear models can be handled in similar fashion. However models of the form \( Y_i = m(x_i, \beta) + \epsilon_i \) where the distribution of the \( \epsilon_i \) is not fully specified by knowing \( x_i \) and \( \beta \) (as happens for GLIMs) require explicitly generating an error term; see Buonaccors (2010, p. 217) for further discussion.

With replication, the next step generates
\[ W_{b_{ij}} = \tilde{x}_{T_{ij}} + U_{bij} \] (4)

for \( j = 1 \) to \( k_i \), where the \( U_{bij} \) are i.i.d. with mean 0 and covariance matrix \( \tilde{\Sigma}_{ui(1)} \), or the true \( \Sigma_{ui(1)} \) itself (see the empirical option below). Recall that \( k_i \) is the number of replicates for the \( i \)th observation and \( \tilde{\Sigma}_{ui(1)} \) is the estimated per-replicate covariance matrix for the \( i \)th observation. From this obtain \( W_{bi} = \sum_j W_{b_{ij}}/k_i \), \( S_{Wbi} = \sum_{j=1}^{k_i} (W_{b_{ij}} - W_{bi})(W_{b_{ij}} - W_{bi})'/(k_i - 1) \) and \( \tilde{\Sigma}_{bui} = S_{Wbi}/k_i \). There are two issues that need to be addressed.

The first issue is what to use for the \( \tilde{x}_{T_{ij}} \)’s, the estimated true values. Treating the \( x \)’s as fixed, Buonaccorsi (2010) discussed two options, one to simply use \( W_i \) (as done in a number of his examples), the other to use the \( \tilde{x}_i \)’s employed in regression calibration. However neither of these turn out to be a very good choice. Ideally we’d like each \( \tilde{x}_{T_{ij}} \) to equal the corresponding true \( x_i \), which is of course impossible with measurement error. In lieu of that, one strategy is to choose the \( \tilde{x}_{T_{ij}} \)’s so they at least have a mean and covariance equal to the estimated mean and covariance of the true \( x_i \)’s (or \( X_i \)’s in the structural case). If we use \( \tilde{x}_{T_{ij}} = W_i \), these values have a covariance matrix \( S_W \) that overestimates \( S_x \) or \( \Sigma_X \), while use of the \( \tilde{x}_i \)’s from (2) leads to underestimation. Our proposal is to use

\[ \tilde{x}_{Pi} = \tilde{W} + (\tilde{\kappa})^{1/2}(W_i - \tilde{W}) \] (5)

where \( \tilde{\kappa}^{1/2} = (S_W - \tilde{\Sigma}_u)^{1/2}S_W^{-1/2} = \tilde{\Sigma}_X^{1/2}S_W^{-1/2} \). These values have mean \( \tilde{W} \) and covariance matrix \( \tilde{\Sigma}_X = S_W - \tilde{\Sigma}_u \). The above is easily modified to accommodate unequal numbers of replicates.

For the univariate setting, the use of the estimated true values in (5) for model based bootstrapping was proposed by Buonaccorsi et al. (2016). It was also used by Zheng and Frey (2005) for making inferences in a one-way random effects model and by Thomas et al. (2011) when using moment matching to impute in measurement error problems. Hutchison et al. (2003) use a related model based approach to try and estimate the bias in the naive estimator and then correct for it in a linear mixed model. Linder and Babu (1994) also
considered a version of a model based bootstrap using modified residuals but in the limited situation where there is no error in the equation and the ratio of the measurement error variances is known.

The second major issue is how to generate the error $\mathbf{U}_{bij}$ in (4). We explore two approaches.

- **Parametric model based (PMB).** Assume the measurement errors are normal and generate replicates $\mathbf{U}_{bij}$ distributed normal with mean $\mathbf{0}$ and covariance matrix $\Sigma_{ui(1)} = S_{Wi}$ (an estimate of the per-replicate measurement covariance for the $ith$ observation).

- **Empirical nonparametric (EMP).** This approach borrows the idea behind the empirical SIMEX approach proposed by Devanarayan and Stefanski (2002). Consider $\mathbf{U}_{bij} = \sum_{m=1}^{k_i} c_{bim} \mathbf{W}_{im}$, where $\mathbf{C}_{bij} = (c_{bij1}, \ldots, c_{bijk_i})'$ is a random normalized contrast vector with $\mathbf{C}'_{sj1} \mathbf{1} = 0$ and $\mathbf{C}'_{bijn} \mathbf{C}_{bij} = 1$. This results in $\mathbf{U}_{bij}$ having mean $\mathbf{0}$ and covariance matrix $\Sigma_{ui(1)}$ (exactly, without knowing what $\Sigma_{ui(1)}$ is!). There is no parametric assumption on the measurement errors here but if the original measurement errors are normally distributed then $\mathbf{U}_{bij}$ will also be normally distributed.

### 3.3 Estimating bias

The easiest (and standard) way to estimate bias in the $jth$ estimated coefficient is with $\hat{\text{Bias}}_j = (\sum_{b=1}^{B} \hat{\beta}_{bj}/B) - \hat{\beta}_j$, where $\hat{\beta}_j$ is the original estimate of interest. The mean bias can be sensitive to outliers so an alternative is to estimate median bias via $\hat{\text{Bias}}_{\text{med},j} = \text{MED}_j - \hat{\beta}_j$, where $\text{MED}_j$ is the median of the $B$ estimates of the $jth$ coefficient, $\hat{\beta}_{1j}, \ldots, \hat{\beta}_{Bj}$.

The corresponding bias corrected estimator (e.g., Efron and Tibshirani (1993, p. 138)) is $\hat{\beta}_j - \hat{\text{Bias}}_j = 2\hat{\beta}_j - (\sum_{b=1}^{B} \hat{\beta}_{bj}/B)$ (or similarly $\hat{\beta}_j - \hat{\text{Bias}}_{\text{med},j}$).

Even without measurement error the simple bootstrap does not provide a valid estimate of bias in non-linear models, including logistic regression. The reason is that the estimators are not ‘plug-in” estimators; see, for example, Efron and Tibshirani (1993). The model based
Model Based Bootstrapping

4. Examples

4.1 Cholesterol/Heart disease

The first example, presented in Section 7.2.2 of Buonaccorsi (2010), uses a subset of 1475 individuals from the Framingham Heart Study in which there were two measures of cholesterol on each individuals, treated as replicates. A logistic model is used for the occurrence of heart disease as a function of true mean cholesterol. The estimated mean cholesterol ranged between 175 and 325 mg/dl and the per-replicate variability varied greatly among individuals with the estimated measurement error standard deviation (per-replicate sd/$\sqrt{2}$) associated with the mean ranging from 4.5 to 86.

We present results only for $\beta_1$. Table 1 shows results first using the naive method with standard analytical standard errors (AN) and associated Wald inferences followed by simple and model based bootstrap analyses. This is followed by similar results for the RC estimate. The correction for measurement error leads to a substantial 30% change in $\beta_1$, from the naive estimate of .0102 to the corrected estimate of .0133. (This corresponds, for example, to a change in the odds ratio of 2.705 to 3.785 for a 100 mg/dl change in cholesterol.)
The CIs associated with the analytical robust (ROB) and information based (INF) are Wald intervals. The model based methods, B-PMB and B-EMP, both use the $\hat{x}_{T_i}$'s (5) as true values, with replicate measurement errors generated as described at the end of Section 3.2. Recall that PMB uses normal replicates while the EMP method is nonparametric. The B-MW method is like B-PMB but with the observed $W_i$ as true values, as used by Buonaccorsi (2010, p.238). As noted in Section 3.2 this underestimates the standard error.

Looking at the results for the RC estimate, there is a remarkable amount of agreement among the results for the simple and two model based bootstraps using the $\hat{x}_{T_i}$'s; see the supplemental material for a plot of fitted bootstrap distributions. In addition, the analytical standard errors are very close to those from the different bootstrap methods. The amount of agreement is somewhat surprising given the differences in the motivation for the different methods. Notice that all of the methods estimate no bias in the RC estimator using either the mean, trimmed mean or median of the bootstrap samples.

Using either the mean or median, the simple bootstrap estimates the bias of the naive estimator to be .0102-.0102 = .0000. However, this is not a valid way to estimate bias, as noted earlier. Using the B-PMB and B-EMP methods to estimate the bias of the naive estimate (see the end of Section 3.3) leads to an estimated bias of .0102 - .0133 = - .0031, corresponding exactly to the change observed in moving from the naive estimate to the RC estimate.

4.2 Gypsy moths and defoliation

Ecological examples abound where logistic regression is used to model a binary outcome (absence of a species, success or failure of a nest, etc.) as a function of habitat variables associated with the unit of interest. Frequently however those habitat variables are estimated with an average over sampled subareas within the spatial region of interest. To illustrate the
behavior of the bootstrap with a small sample size (typical in many ecological studies) with substantial among replicate variability, we examine the relationship between gypsy moth egg mass density ($x$) and defoliation over $n = 18$ plots, each 60 ha in size where 10, 15 or 20 replicates (subsamples of size .01ha) are sampled in a plot. These data are courtesy of Sandy Liebhold and discussed in more detail in Buonaccorsi (2010, p. 75). For our purposes we use only the first 10 replicates ($k_i = 10$) since we are concentrating on the case with an equal number of replicates. $W_i$ is the mean egg density over the ten subplots/replicates on the $ith$ plot. For illustration we dichotomize the outcome defoliation by setting it equal to 1 if the average defoliation is over 50 on the plot (we’ll call this “severe” defoliation) and 0 otherwise, and ignore any potential misclassification. For $Y = 0$, the $W_i$ (mean egg mass density) ranged from 8.3 to 72.0 with a mean of 33.4 and a median of 29.4, with corresponding values of 22.1 to 88.3, 46.1 and 29.2 respectively, for $Y = 1$. The measurement error standard deviation, $S_{W_i}/\sqrt{10}$, ranged from 2.31 to 28.32.

Results for estimating $\beta_1$ appear in Table 2, showing the correction for measurement error leading to a dramatic change in the fitted model. For the RC estimates, the estimated standard errors differ considerably between analytical and bootstrap values. There are also some substantial differences among bootstrap methods and within each method trimming makes a big difference due to outliers. The trimmed standard errors are all still substantially larger than the analytical standard errors. The percentile intervals from the two model based bootstrap methods are somewhat close, but quite different than the percentile intervals from the simple bootstrap or the Wald intervals using analytical standard errors.

[Table 2 about here.]

The bootstrap estimates of mean bias in the RC estimates differ wildly among the simple and model based methods due to outliers. The estimated median biases of the model based
methods (which are more valid than those from the simple bootstrap) are .008 (PMB) and .007 (EMP), pointing towards moderate positive bias.

As with the first example, estimating the bias of the naive estimator based on the simple bootstrap is misleading. The model based estimates based on the median leads to bias estimates of .027 - .0327 = -.006 (PMB) and .0263 - .0327 = -.007 (EMP) which are more in line with the correction of .010 in going from the naive (.023) to RC (.033) estimate.

The story here is more complicated than our first example, showing the challenges in dealing with a small sample size with substantial measurement error. We recommend the use of the model based bootstrap results for these data.

5. Simulations

Here we provide some initial assessment of the performance of the various bootstrap methods and of the analytical standard errors and associated Wald confidence intervals. There are obviously an endless number of scenarios that could be considered. We’ll discuss the situation with random $X$ and $k = 2$ replicates in full here and summarize some results from additional settings in Section 5.5.

The $X_i$ are random with mean $\mu_X = 250$ and standard deviation $\sigma_X = 25$ with $X$ either normally distributed or $X = 250 + 25*Z$ where $Z$ follows a normalized chi-square distribution with 3 degrees of freedom, leading to a skewed distribution. Based on the cholesterol example, we take $\beta_0 = -5$ while $\beta_1 = .0133$ (the corrected estimate from the example) or .03. The sample size is $n = 100$ or $n = 500$ with a per-replicate standard deviation of measurement error of $\sigma_u(1) = 17.677$ (overall reliability = .8) or $\sigma_u(1) = 34.355$ (reliability = .5). Lastly, the measurement error is either normal with mean 0 and standard error $\sigma_u(1)$ or comes from a standardized chi-square distribution with 3 degrees of freedom, rescaled to have standard deviation $\sigma_u(1)$. Overall there are 32 combinations, numbered as indicated in the results in Table 3. For each combination, 1000 data sets are generated. However, for obtaining
the “true” standard errors and biases used in Sections 5.2 and 5.3 and the performance of
the estimators discussed in Section 5.1, 5000 simulations were used (but the computational
demands were excessive to use 5000 simulations for everything). All of the bootstrap methods
use \( B = 1000 \) bootstrap samples.

5.1 Performance of estimators

A comparison of the estimators is not our main objective. Boxplots in the supplemental
material show the performance of the naive and RC estimators of \( \beta_1 \) along with that of the
estimator based on true values, as a benchmark and the first portion of Table 3 shows the
median and the relative median bias of the RC estimator. The mean and median bias were
often similar but outliers influenced the mean in a couple of scenarios (see the supplemental
material) so the median is given here. These results show the well known fact that the naive
estimator can be significantly biased downward and that the RC estimator has less bias but
more variability. While in many cases the RC estimator has a bias similar to that based on
using true \( x \)’s, if they could be observed, there are scenarios (chosen intentionally in order
to assess how well we can estimate bias) where there is a substantial amount of bias in the
RC estimator; see in particular simulations 11-16 and 29-32. The four most severe cases
(15,16,31 and 32) have low reliability (large measurement error), a big effect (\( \beta_1 = .03 \)) and
the distribution of the measurement error is skewed.

5.2 Estimating the standard error of the RC estimator

With random \( X \) and constant measurement error variance, the simple bootstrap is applicable.
As noted earlier there are some problems with outliers so we use the trimmed standard
deviation from 5000 simulated data sets as the “true” standard error. The top of Figure
1 shows the median of the estimated standard error using the robust analytical and three
bootstrap methods (also using 1% trimming) versus the “true” standard error over all 32
scenarios. Figure 2 shows further details for simulations 9 - 16 with additional figures in the
supplementary material. The robust estimator of the standard error always does rather well and the estimated SE from the simple bootstrap also does fairly well. Here model based estimates tend to overestimate the standard error when it is large, with some reasons for this touched on in the discussion. We note that this problem largely disappears with more replicates; see the bottom of Figure 1.

5.3 Estimating bias

The mean bias is sometimes quite erratic due to outliers, so Figure 3 shows the median of the estimated median bias of the RC estimator versus the “true” median bias for the RC estimator (from 5000 simulations). Further details on the bias estimation is shown in the boxplots in the supplementary material. The model based bootstrap methods are preferred over the simple bootstrap for estimating bias since the latter may estimate that there is negligible bias, even when there is some. This problem is even more severe when using the simple bootstrap to estimate the bias of a naive estimator; see the supplementary material for details. In those places where the simple bootstrap is at its worse, when there is substantial negative bias in the RC estimator, the model based bootstrap methods also encounter some difficulty but they still do better than the simple bootstrap. The corresponding bias corrected estimators based on the model based bootstrap do result in less biased, but somewhat more variable, estimators, with some extreme values, which are directly related to a corresponding extreme value of the RC estimator. Supporting figures appear in the supplementary material.

5.4 Performance of confidence intervals for $\beta_1$

Table 3 shows the estimated coverage rates for a variety of confidence intervals for $\beta_1$, with a target rate of .95. The true and naive intervals use standard Wald intervals based on the true and error prone values, respectively, while the others are based on the RC estimator.

[Table 3 about here.]
The confidence intervals from using the true values always do quite well, as expected and, also as expected, the naive interval can do quite badly, sometimes disastrously so. The Wald interval using the robust standard error does fairly well in most cases but does encounter some serious problems in simulations 13-16 and 29-32. The Wald-STATA intervals are sometimes better and sometimes worse than the Wald interval. In the eight aforementioned cases it is better in some cases, but in others, offers little, or no, improvement.

Focusing on the percentile intervals, the simple bootstrap does not do quite as well as the model based methods. There are cases where the coverage rates from the simple bootstrap are too small (simulations 5,15, 29 and 32) but both the model based methods do fairly well. All three bootstrap methods fall substantially short of the desired .95 in simulations 14 and 16. The bias corrected intervals did offer some improved coverage rates when the standard percentile methods fell short but did not always improve matters; e.g., simulations 14 and 30 among others.

The supplementary material provides summary statistics on the length of the various intervals. For $n = 500$ the lengths are fairly similar for the various methods, although the model based bootstrap intervals are a bit larger when the measurement error distribution is skewed and $\kappa = .5$. When $n = 100$ the Wald and Wald-S intervals have shorter mean and median length than the other procedures, but the Wald intervals sometimes lead to wildly wide intervals in some scenarios. Still with $n = 100$, the model based intervals are generally longer throughout than those based on the simple bootstrap, which is what leads to better coverage rates. The bias corrected percentile intervals, while similar in median length to their non-bias corrected counterparts in some cases, they are longer in some others and more susceptible to outliers.
5.5 Additional simulations

A number of additional simulations were run with details on the set-up as well as extensive results appearing in the supplementary material. The first of these were identical to the preceding simulations except with $k = 5$ replicates and the per-replicate variance rescaled so the reliability stayed at .5 or .8. As noted earlier (see Figure 1), the model based methods do a better job of estimating the SE than with $k = 2$. The model based methods still have some trouble estimating bias, although they do somewhat better than when $k = 2$ and they still do better than the simple bootstrap. When the error associated with a replicate is skewed, the error for the mean becomes more normal with $k = 5$ than with $k = 2$ and so the performance of the confidence intervals there are similar to results for corresponding cases with $U$ distributed normal in Table 3. The lengths of the confidence intervals behave similarly as for $k = 2$ with the important exception that the model based bootstrap intervals are now either similar in length or generally shorter than those from the simple bootstrap. The Wald intervals are not as susceptible to outliers.

The second set of simulations return to $k = 2$, but use a fixed set of true values ($n = 100$ or $500$), based on selected cases from the original Framingham data. The per-replicate measurement error variance is either constant or changing over observations. The conclusions for estimating bias or the standard error and the performance of confidence intervals are very similar to those from the random $X$ setting. The performance of the simple bootstrap in estimating the standard error is somewhat surprising given the $x$ values are fixed, but this is mostly likely due to the fact that there are many (100 or 500) different $x$ values; see Chernick (2008, Ch. 4) also for some discussion on the robustness of simple bootstrapping in regression problems. However, despite the fact that the model based method does not always do well estimating the standard errors the model based percentile confidence intervals routinely outperform the intervals from the simple bootstrap and the standard Wald interval.
6. Discussion

We have developed some new methodology for implementing model based bootstrap methods when correcting for measurement error in logistic regression (more generally, generalized linear models) with replicates. These are an alternative to the simple bootstrap, which is not always justified and is also not designed to estimate bias in non-linear models. To implement the model based methods we propose the use of an estimated set of true values which match the estimated mean and variance of the underlying true values (see equation 5) and suggested two ways to generate replicates, one using a normal model with estimated measurement error variances, the other an empirical method which exploits a clever technique originally introduced for use with SIMEX.

In a series of simulations we found i) The model based bootstrap improved on the simple bootstrap in estimating bias in either the naive or RC estimator. As expected, the simple bootstrap tended to estimate there was little bias, even when bias was present. There were a few settings, however where the model based methods, while better than the simple bootstrap, still had some difficulty in estimating bias. The corresponding model based bias corrected estimators reduced some of the bias in the RC estimator, when present, although they were a bit more variable. ii) While often adequate, the model based methods did not do as well as the simple bootstrap and analytical methods in estimating a large standard error. However, this did not always translate into poor performance of the associated percentile or bias corrected confidence intervals. iii) The model based bootstrap percentile intervals generally had the best overall coverage rates. This improved performance was achieved without giving up much, if anything, in terms of having wider intervals. There were a number of settings where the Wald intervals based on either the analytical standard error or a trimmed standard error from the simple bootstrap, or the percentile interval from the simple bootstrap, fell considerably short in their coverage rates. Still there were some settings
where the model based percentile intervals, as well at their bias corrected counterpart, also fell somewhat short, essentially when there was substantial bias in the RC estimator.

Overall the model based bootstrap methods offer some definite advantages over the other, more standard, methods of inference and we recommend their use. However, there are still some shortcomings and there are a few potential reasons for that. One is uncertainty in the estimated coefficients that enter into (3). This is a potential problem even in regression problems without measurement error where the bias may be a function of the true coefficients. A second reason is the uncertainty in the estimated true values, which enter into generating both the $Y$’s in (3) and the replicates in (4). Based on some simulations (results not shown) where we used the true $x$’s in place of the $\hat{x}_{Ti}$’s but the estimation of bias did not improve much, this doesn’t appear to be a major factor. A third reason is uncertainty in the replicate measures that go into the EMP method and into estimating the measurement error variances used in the normal based PMB method. There is an additional issue with the EMP method, which we expected to outperform the PMB method, especially when the errors were non-normal. While it does generate replicates that unconditionally have measurement error covariance matrix $\Sigma_{ui(1)}$ (even without knowing what it is) and the replicates are conditionally independent, they are unconditionally dependent since they each use the common estimated true value $\hat{x}_{Ti}$. (It is easy, using double expectations, to show that unconditionally $\text{Cov}(W_{bi1}, W_{bi2}) = \text{Cov}(\hat{x}_{Ti}))$. This has the potential to compromise the performance of the EMP method somewhat.

Our goal was to introduce some relatively basic model based methods, but there are a number of additional strategies beyond the scope of this paper that could be explored to try and improve the methods. For improving estimation of bias, one alternative is the use of the double bootstrap (Davison and Hinkley (1997), p. 104). Another is to try and model the bias over plausible values of the underlying parameters and then use the resulting model
for bias correction (see, for example, Pfefferman and Correa (2012)). If one was comfortable with the assumption of constant per-replicate variance then a pooled estimate of the per-replicate standard deviation rather than the subject specific values could be used in the PMB method. There are also other nonparametric model based bootstraps that could be considered including the use of an estimate of the distribution of the true values with random \( X \)’s to generate true values. This was used by Carroll et al. (2011) in estimating a shape-constrained nonparametric density and regression.

**Supplementary Materials**

Tables and Figures referenced in Sections 4 and 5 are available with this paper at the Biometrics website on Wiley Online Library.”

**Acknowledgments**

We are grateful to Sandy Liebhold and Ray Carroll for use of data going into the defoliation and cholesterol examples.

**7. References**


[Figure 1 about here.]
Figure 1. Random X. Median of estimated standard error (bootstrap trimmed) versus “true” SE (trimmed). r = analytical robust standard error; s = simple bootstrap; m = model based with normal measurement errors; e = empirical bootstrap.
Figure 2. Random $X$, $k = 2$. Boxplots showing performance of estimated standard errors (trimmed) for simulation 9-16; gray dot = mean, black line = median. Dashed line = simulated SE from 5000 simulations, untrimmed; solid line = simulated SE, trimmed.
Figure 3. Random $X, k = 2$. Median of estimated median bias versus “true” median bias for RC estimator. $s =$ simple bootstrap; $m =$ model based with normal measurement errors; $e =$ empirical bootstrap.
Inferences for $\beta_1$ for cholesterol example. SE = standard error; SE(T) = trimmed standard error for bootstrap methods; AN = analytical based SE; ROB = analytical robust SE; INF = analytical information based SE; B = bootstrap; SIM = simple; PMB = parametric normal model based; EMP = empirical; MW = normal model based with W’s as true values; CI = confidence interval; CI(BC) = bias corrected CI for bootstrap methods.

<table>
<thead>
<tr>
<th>Method</th>
<th>Estimate</th>
<th>SE</th>
<th>SE(T)</th>
<th>mean</th>
<th>median</th>
<th>CI</th>
<th>CI(BC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Naive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AN</td>
<td>.0102</td>
<td>.0029</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>[.0045,.0159]</td>
<td>.</td>
</tr>
<tr>
<td>B-SIM</td>
<td>.0102</td>
<td>.0028</td>
<td>.0026</td>
<td>.0102</td>
<td>.0102</td>
<td>[.0047,.0155]</td>
<td>[.0047,.0156]</td>
</tr>
<tr>
<td>B-PMB</td>
<td>.0102</td>
<td>.0029</td>
<td>.0028</td>
<td>.0102</td>
<td>.0102</td>
<td>[.0045,.0159]</td>
<td>[.0109,.0202]</td>
</tr>
<tr>
<td>B-EMP</td>
<td>.0102</td>
<td>.0029</td>
<td>.0027</td>
<td>.0102</td>
<td>.0102</td>
<td>[.0046,.0158]</td>
<td>[.0109,.0202]</td>
</tr>
<tr>
<td>ROB</td>
<td>.0133</td>
<td>.0036</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>[.0063,.0204]</td>
<td>.</td>
</tr>
<tr>
<td>INF</td>
<td>.0133</td>
<td>.0038</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>[.0058,.0209]</td>
<td>.</td>
</tr>
<tr>
<td>B-SIM</td>
<td>.0133</td>
<td>.0036</td>
<td>.0034</td>
<td>.0133</td>
<td>.0133</td>
<td>[.0062,.0203]</td>
<td>[.0062,.0204]</td>
</tr>
<tr>
<td>B-PMB</td>
<td>.0133</td>
<td>.0039</td>
<td>.0036</td>
<td>.0133</td>
<td>.0133</td>
<td>[.0059,.0208]</td>
<td>[.0060,.0208]</td>
</tr>
<tr>
<td>B-EMP</td>
<td>.0133</td>
<td>.0038</td>
<td>.0035</td>
<td>.0133</td>
<td>.0133</td>
<td>[.0060,.0209]</td>
<td>[.0063,.0211]</td>
</tr>
<tr>
<td>B-MW</td>
<td>.0133</td>
<td>.0032</td>
<td>.0030</td>
<td>.0133</td>
<td>.0133</td>
<td>[.0070,.0195]</td>
<td>[.0071,.0196]</td>
</tr>
</tbody>
</table>
Table 2

Inferences for $\beta_1$ for egg mass/defoliation example. Labeling as in Table 1.

<table>
<thead>
<tr>
<th>Method</th>
<th>Est</th>
<th>SE</th>
<th>SE(T)</th>
<th>Bmean</th>
<th>Bmed</th>
<th>CI</th>
<th>CI(BC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Naive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AN</td>
<td>.0229</td>
<td>.0210</td>
<td></td>
<td></td>
<td></td>
<td>[.0183, .0641]</td>
<td></td>
</tr>
<tr>
<td>B-SIM</td>
<td>.0229</td>
<td>.0547</td>
<td>.0228</td>
<td>.0274</td>
<td>.0249</td>
<td>[.0291, .0783]</td>
<td>[.0370, .0708]</td>
</tr>
<tr>
<td>B-PMB</td>
<td>.0229</td>
<td>.3964</td>
<td>.0313</td>
<td>.0461</td>
<td>.0269</td>
<td>[.0219, .1276]</td>
<td>[.0101, .2048]</td>
</tr>
<tr>
<td>B-EMP</td>
<td>.0229</td>
<td>.0365</td>
<td>.0298</td>
<td>.0323</td>
<td>.0263</td>
<td>[.0179, .1169]</td>
<td>[.0076, .1572]</td>
</tr>
<tr>
<td>RC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ROB</td>
<td>.0327</td>
<td>.0238</td>
<td></td>
<td></td>
<td></td>
<td>[.0139, .0794]</td>
<td></td>
</tr>
<tr>
<td>INF</td>
<td>.0327</td>
<td>.0281</td>
<td></td>
<td></td>
<td></td>
<td>[.0224, .0877]</td>
<td></td>
</tr>
<tr>
<td>B-SIM</td>
<td>.0327</td>
<td>.0887</td>
<td>.0356</td>
<td>.0355</td>
<td>.037</td>
<td>[.0604, .1146]</td>
<td>[.1004, .0934]</td>
</tr>
<tr>
<td>B-PMB</td>
<td>.0327</td>
<td>.4849</td>
<td>.0558</td>
<td>.0703</td>
<td>.0407</td>
<td>[.0350, .2416]</td>
<td>[.048, .1632]</td>
</tr>
<tr>
<td>B-EMP</td>
<td>.0327</td>
<td>.2124</td>
<td>.0519</td>
<td>.0580</td>
<td>.0400</td>
<td>[.029, .2053]</td>
<td>[.0473, .1619]</td>
</tr>
</tbody>
</table>
Table 3

Random X, k = 2. 'Med = 'True" median and RBias = relative median bias of RC estimator based on 5000 simulations. Followed by estimated coverage rates of confidence intervals (1000 simulations). T = Wald interval/true values; N = Wald interval/naive values; W = \( \hat{\beta}_1 + 1.96 SE \), \( \hat{\beta}_1 = RC \) estimate, SE = robust standard error; W-S (Wald STATA) = \( \hat{\beta}_1 + 1.96 SE \), SE = trimmed bootstrap standard error from simple bootstrap; SIM, PMB and EMP = percentile intervals from simple, normal model based and empirical model based respectively; BC = bias corrected. N indicates normal distribution, C indicates use of a standardized chi-square distribution.

<table>
<thead>
<tr>
<th>Sim</th>
<th>n</th>
<th>RC</th>
<th>W</th>
<th>T</th>
<th>N</th>
<th>W-S</th>
<th>SIM</th>
<th>PMB</th>
<th>EMP</th>
<th>Dist</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>100</td>
<td>.8</td>
<td>.0133</td>
<td>N</td>
<td>N</td>
<td>0.0139</td>
<td>4.5</td>
<td>.956</td>
<td>.938</td>
<td>.930</td>
</tr>
<tr>
<td>2</td>
<td>500</td>
<td>.8</td>
<td>.0133</td>
<td>N</td>
<td>N</td>
<td>0.0133</td>
<td>-0.05</td>
<td>.947</td>
<td>.909</td>
<td>.949</td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td>.5</td>
<td>.0133</td>
<td>N</td>
<td>N</td>
<td>0.0133</td>
<td>0.12</td>
<td>.951</td>
<td>.878</td>
<td>.959</td>
</tr>
<tr>
<td>4</td>
<td>500</td>
<td>.5</td>
<td>.0133</td>
<td>N</td>
<td>N</td>
<td>0.0133</td>
<td>-0.21</td>
<td>.957</td>
<td>.556</td>
<td>.962</td>
</tr>
<tr>
<td>5</td>
<td>100</td>
<td>.8</td>
<td>.0300</td>
<td>N</td>
<td>N</td>
<td>0.0305</td>
<td>1.52</td>
<td>.939</td>
<td>.893</td>
<td>.933</td>
</tr>
<tr>
<td>6</td>
<td>500</td>
<td>.8</td>
<td>.0300</td>
<td>N</td>
<td>N</td>
<td>0.0296</td>
<td>-1.40</td>
<td>.955</td>
<td>.729</td>
<td>.941</td>
</tr>
<tr>
<td>7</td>
<td>100</td>
<td>.5</td>
<td>.0300</td>
<td>N</td>
<td>N</td>
<td>0.0305</td>
<td>1.59</td>
<td>.960</td>
<td>.544</td>
<td>.967</td>
</tr>
<tr>
<td>8</td>
<td>500</td>
<td>.5</td>
<td>.0300</td>
<td>N</td>
<td>N</td>
<td>0.0291</td>
<td>-2.93</td>
<td>.968</td>
<td>.025</td>
<td>.952</td>
</tr>
<tr>
<td>9</td>
<td>100</td>
<td>.8</td>
<td>.0133</td>
<td>N</td>
<td>C</td>
<td>0.0138</td>
<td>3.35</td>
<td>.946</td>
<td>.936</td>
<td>.915</td>
</tr>
<tr>
<td>10</td>
<td>500</td>
<td>.8</td>
<td>.0133</td>
<td>N</td>
<td>C</td>
<td>0.0133</td>
<td>0.26</td>
<td>.960</td>
<td>.915</td>
<td>.953</td>
</tr>
<tr>
<td>11</td>
<td>100</td>
<td>.5</td>
<td>.0133</td>
<td>N</td>
<td>C</td>
<td>0.0125</td>
<td>-5.70</td>
<td>.958</td>
<td>.882</td>
<td>.966</td>
</tr>
<tr>
<td>12</td>
<td>500</td>
<td>.5</td>
<td>.0133</td>
<td>N</td>
<td>C</td>
<td>0.0128</td>
<td>-4.74</td>
<td>.951</td>
<td>.541</td>
<td>.957</td>
</tr>
<tr>
<td>13</td>
<td>100</td>
<td>.8</td>
<td>.0300</td>
<td>N</td>
<td>C</td>
<td>0.0314</td>
<td>4.68</td>
<td>.958</td>
<td>.931</td>
<td>.953</td>
</tr>
<tr>
<td>14</td>
<td>500</td>
<td>.8</td>
<td>.0300</td>
<td>N</td>
<td>C</td>
<td>0.0304</td>
<td>1.48</td>
<td>.939</td>
<td>.762</td>
<td>.940</td>
</tr>
<tr>
<td>15</td>
<td>100</td>
<td>.5</td>
<td>.0300</td>
<td>N</td>
<td>C</td>
<td>0.0326</td>
<td>8.68</td>
<td>.953</td>
<td>.626</td>
<td>.966</td>
</tr>
<tr>
<td>16</td>
<td>500</td>
<td>.5</td>
<td>.0300</td>
<td>N</td>
<td>C</td>
<td>0.0318</td>
<td>4.55</td>
<td>.953</td>
<td>.087</td>
<td>.957</td>
</tr>
<tr>
<td>17</td>
<td>100</td>
<td>.8</td>
<td>.0133</td>
<td>C</td>
<td>N</td>
<td>0.0138</td>
<td>3.45</td>
<td>.974</td>
<td>.966</td>
<td>.944</td>
</tr>
<tr>
<td>18</td>
<td>500</td>
<td>.8</td>
<td>.0133</td>
<td>C</td>
<td>N</td>
<td>0.0139</td>
<td>4.16</td>
<td>.956</td>
<td>.922</td>
<td>.944</td>
</tr>
<tr>
<td>19</td>
<td>100</td>
<td>.5</td>
<td>.0133</td>
<td>C</td>
<td>N</td>
<td>0.0142</td>
<td>6.93</td>
<td>.969</td>
<td>.905</td>
<td>.970</td>
</tr>
<tr>
<td>20</td>
<td>500</td>
<td>.5</td>
<td>.0133</td>
<td>C</td>
<td>N</td>
<td>0.0145</td>
<td>8.85</td>
<td>.953</td>
<td>.617</td>
<td>.946</td>
</tr>
<tr>
<td>21</td>
<td>100</td>
<td>.8</td>
<td>.0300</td>
<td>C</td>
<td>N</td>
<td>0.0270</td>
<td>-10.10</td>
<td>.950</td>
<td>.870</td>
<td>.901</td>
</tr>
<tr>
<td>22</td>
<td>500</td>
<td>.8</td>
<td>.0300</td>
<td>C</td>
<td>N</td>
<td>0.0260</td>
<td>-13.27</td>
<td>.951</td>
<td>.665</td>
<td>.893</td>
</tr>
<tr>
<td>23</td>
<td>100</td>
<td>.5</td>
<td>.0300</td>
<td>C</td>
<td>N</td>
<td>0.0248</td>
<td>-17.46</td>
<td>.955</td>
<td>.537</td>
<td>.923</td>
</tr>
<tr>
<td>24</td>
<td>500</td>
<td>.5</td>
<td>.0300</td>
<td>C</td>
<td>N</td>
<td>0.0234</td>
<td>-22.07</td>
<td>.937</td>
<td>.025</td>
<td>.845</td>
</tr>
</tbody>
</table>