NIH Public Access
Author Manuscript
Math Comput Model. Author manuscript; available in PMC 2011 November 1.

Published in final edited form as:
Math Comput Model. 2010 November 1; 52(9-10): 1610–1625. doi:10.1016/j.mcm.2010.06.026.

Standard Error Computations for Uncertainty Quantification in Inverse Problems: Asymptotic Theory vs. Bootstrapping

H. T. Banks, **Kathleen Holm**, and **Danielle Robbins**
Center for Research in Scientific Computation and Center for Quantitative Sciences in Biomedicine North Carolina State University Raleigh, NC 27695-8212

Abstract

We computationally investigate two approaches for uncertainty quantification in inverse problems for nonlinear parameter dependent dynamical systems. We compare the bootstrapping and asymptotic theory approaches for problems involving data with several noise forms and levels. We consider both constant variance absolute error data and relative error which produces non-constant variance data in our parameter estimation formulations. We compare and contrast parameter estimates, standard errors, confidence intervals, and computational times for both bootstrapping and asymptotic theory methods.

Keywords

Uncertainty quantification; parameter estimation; nonlinear dynamic models; bootstrapping; asymptotic theory standard errors; ordinary least squares vs. generalized least squares; computational examples

1 Introduction

One of the more ubiquitous computational problems in all of science and engineering is the inverse problem for estimation of parameters from longitudinal observations of system responses. This is usually formulated in terms of a parameter dependent dynamical mathematical model (ordinary, partial, delay differential or integral equation (see [1,2,3,7,8,10,18,19,20,21,29,30] and the references therein) for which observations of solutions (or certain components of the solutions) are to be used to estimate some unknown parameters q. In particular the general inverse problem for nonlinear parameter dependent ordinary di erential equations is formulated in terms of a system

$$\begin{aligned} \frac{dx}{dt}(t) &= \mathcal{F}(t, x(t), q) \\ x(0) &= x_0, \end{aligned} \tag{1}$$

with observations

NIH-PA Author Manuscript

NIH-PA Author Manuscript

$$y_j = f(t_j, \theta) = Cx(t_j; \theta), \quad j = 1, \ldots, n, \tag{2}$$

where the solutions $x(t; \theta)$ in general depend on unknowns q and the (possibly unknown) initial conditions x_0 so that $\theta = (q, x_0)$. In addition to computing estimates $\hat{\theta}$ for the unknown parameters using observations $\{y_j\}$, it is widely accepted that quantifying the uncertainty in these parameter estimates is equally important. A standard method [3,4,6,8,14,16,25] to do this involves computation of *standard errors* (SE) to be used in *confidence intervals* (CI) for the parameter estimates. Discussions of the fundamental ideas and methods that are accessible to non-statisticians can be found in [3,8].

In this note we investigate computationally and compare the bootstrapping approach and the asymptotic theory approach to computing parameter standard errors corresponding to data with various noise levels (*Nl*). We consider both absolute error (with constant variance) measurement data and relative error (and hence non-constant variance) measurement data sets in obtaining parameter estimates. Both types of errors are found widely in measurement procedures used in science and engineering investigations. We compare and contrast parameter estimates, standard errors, confidence intervals, and computational times for both bootstrapping and asymptotic theory approaches. Our goal is to investigate and illustrate some possible advantages/disadvantages of each approach in treating problems for nonlinear dynamical systems, focusing on computational methods. We discuss these in the context of a simple example for which an analytical solution is available to provide readily obvious distinct qualitative behaviors. We chose this example because its solutions have many features found in those of much more complicated models: regions of rapid change as well as a region near an equilibrium with very little change and di erent regions of relative insensitivity of solutions to different parameters. The model we use to carry out this demonstration analysis is the well known [8,22] logistic model (also called the Verhulst-Pearl growth model) given by

$$\frac{dx(t)}{dt} = rx(t)\left(1 - \frac{x(t)}{K}\right), \quad x(0) = x_0.$$

Here K is the *carrying capacity* as well as the asymptote value (as $t \to \infty$) for solutions and r is the *intrinsic growth rate*.

The solution is given for $\theta = (K, r, x_0)$ by

$$x(t) = f(t, \theta) = \frac{K}{1 + \left(\frac{K}{x_0} - 1\right) e^{-rt}},$$

and is plotted in Figure 1 for $K = 17.5$, $r = .7$ and $x_0 = .1$ and $0 \le t \le 25$.

There is a substantial statistical literature (see [12,15,16,23] and the references therein) on the use of the bootstrap in nonlinear regression problems while Chapter 12 of [25] contains a readable presentation for the corresponding use of asymptotic theory for estimators in nonlinear regression. As will be discussed in illustrative examples below, for quantifying uncertainty in parameter estimates the asymptotic theory is always computationally faster than bootstrapping. This unfortunately is the only definitive general comparative statement that can be made. One of the two methods may be favored over the other dependent on the type of observational data set used and the particular dynamics under investigation. For constant variance data there appears to be little advantage for either method. Bootstrapping may be more reasonable for estimating the standard error in the cases of nonlinear systems so complex as to prohibit computation of the sensitivities needed for asymptotic theory errors. However, if computation time is already a concern and sensitivities can be computed, then asymptotic theory may prove advantageous.

As we shall see, when the relative error statistical model is correct (i.e., one has non-constant variance data), the parameters are most effciently estimated using generalized least squares; in this case the asymptotic theory error estimates and the bootstrapping error estimates may be comparable. In some situations (when local variation in the data is high), the standard errors for bootstrapping are likely to be *larger* (due to a positive additive corrective term) but also more *accurate* than those of the asymptotic theory. As we shall illustrate, in other cases there will be insignificant correction in the bootstrap's estimate of standard error over that of the asymptotic theory. We thus see that the choice between bootstrapping and asymptotic theory approaches depends on the local variation in data in regions of influence in the estimation of a specific parameter. These heuristic conclusions (it is difficult to give rigorous proofs in this area) will be supported by examples in subsequent sections of this paper.

In the next section we give a detailed algorithm for computing the bootstrap estimates for constant variance data. This section is followed by a description for standard error computations using asymptotic theory. Section 4 contains a detailed algorithm for computing bootstrap estimates with non constant variance data, followed by a section with the description of standard error computations using asymptotic theory for non constant variance data. This is followed by a report of the numerical results, and numerical results for data sets where the incorrect variance models are assumed. Section 8 presents remarks on the corrective nature of the bootstrap, followed by a short summary on our findings. Finally in an appendix a brief summary of literature on the asymptotic nature of the bootstrap is given.

2 Bootstrapping Algorithm for Constant Variance Data

Assume we are given experimental data $(y_1, t_1), \dots, (y_n, t_n)$ from an underlying observation process

$$Y_j = f\left(t_j, \theta_0\right) + \mathcal{E}_j, \quad j = 1, \dots, n, \tag{3}$$

where the $\mathcal{E}_j$ are independent identically distributed *(iid)* from a distribution F with mean zero ($E\left(\mathcal{E}_j\right) = 0$) and constant variance σ_0^2, and θ_0 is the "true value" parameter value hypothesized to exist in statistical treatments of data. Associated corresponding realizations $\{y_j\}$ of the random variables $Y = \{Y_j\}$ are given by

$$y_j = f\left(t_j, \theta_0\right) + \epsilon_j.$$

We note that the assumption that the observation errors are uncorrelated across j (i.e., time) may be a reasonable one when the observations are taken with su cient intermittency or when the primary source of error is measurement error. If we define

$$\theta_{\mathrm{OLS}}(Y) = \theta^n_{\mathrm{OLS}}(Y) = \arg\min_{\theta \in \Theta_{ad}} \sum_{j=1}^{n} \left[Y_j - f\left(t_j, \theta\right) \right]^2, \tag{4}$$

NIH-PA Author Manuscript

NIH-PA Author Manuscript

where $\Theta ad \subset \mathbb{R}^p$ is the admissible parameter set, then θ_{OLS} can be viewed as minimizing the distance between the observations and model where all observations are treated as of equal importance. We note that minimizing in (4) corresponds to solving for θ in

$$\sum_{j=1}^{n}\left[Y_j - f\left(t_j,\theta\right)\right]\nabla f\left(t_j,\theta\right)=0. \tag{5}$$

We point out that θ_{OLS} is a *random variable* (because $\mathcal{E}_j=Y_j - f\left(t_j,\theta\right)$ is a random variable); hence if $\left\{y_j\right\}_{j=1}^{n}$ is a realization of the *random process* $\left\{y_j\right\}_{j=1}^{n}$ then solving

$$\widehat{\theta}_{\mathrm{OLS}}=\widehat{\theta}_{\mathrm{OLS}}^{n}=\arg\min_{\theta\in\Theta_{ad}}\sum_{j=1}^{n}\left[Y_j - f\left(t_j,\theta\right)\right]^2 \tag{6}$$

provides a realization $\widehat{\theta}_{\mathrm{OLS}}$ for θ_{OLS}. (A remark on notation: for a random variable or *estimator* θ we will always denote a corresponding *realization* or *estimate* with an over hat, e.g., $\widehat{\theta}$ is an estimate for θ.)

Noting that

$$\sigma_0^2=\frac{1}{n}E\left[\sum_{j=1}^{n}\left[Y_j - f\left(t_j,\theta_0\right)\right]^2\right] \tag{7}$$

suggests that once we have solved for θ_{OLS} in (4), we may obtain an estimate $\widehat{\sigma}_{\mathrm{OLS}}^2$ for σ_0^2. An unbiased estimate is given by

$$\sigma_0^2 \approx \widehat{\sigma}_{OLS}^2=\frac{1}{n-p}\sum_{j=1}^{n}\left(y_j - f\left(t_j,\widehat{\theta}\right)\right)^2.$$

The following algorithm [12,13,15, p. 285–287] can be used to compute the *bootstrapping estimate* $\widehat{\theta}_{\mathrm{boot}}$ of θ_0 and its empirical distribution.

1. First estimate $\widehat{\theta}^0=\left(\widehat{K}^0,\widehat{r}^0,\widehat{x}_0^0\right)$ from the entire sample $\left\{y_j\right\}_{j=1}^{n}$ using OLS.
2. Using this estimate define the standardized residuals

$$\bar{r}_j=\sqrt{\frac{n}{n-p}}\left(y_j - f\left(t_j,\widehat{\theta}^0\right)\right)$$

for $j = 1, \ldots, n$. Then $\left\{\bar{r}_1,\ldots,\bar{r}_n\right\}$ are realizations of *iid* random variables $\bar{R}_j$ with the empirical distribution F_n, and $p = 3$ for the number of parameters. We remark that

$$E\left(\bar{R}_j|F_n\right)=n^{-1}\sum_{j=1}^{n}\bar{r}_j=0,\quad \mathrm{var}\left(\bar{R}_j|F_n\right)=n^{-1}\sum_{j=1}^{n}\bar{r}_j^2=\widehat{\sigma}^2.$$

Set $m = 0$.

3. Create a bootstrap sample of size n using random sampling with replacement from the data (realizations) $\{\bar{r}_1,\dots,\bar{r}_n\}$ to form a bootstrap sample $\{r_1^m,\dots,r_n^m\}$.

4. Create bootstrap sample points

$$y_j^m=f\left(t_j,\widehat{\theta}^0\right)+r_j^m,$$

where $j = 1,\dots,n$.

5. Obtain a new estimate $\widehat{\theta}^{m+1}=\left(\widehat{K}^{m+1},\widehat{r}^{m+1},\widehat{x}_0^{m+1}\right)$ from the bootstrap sample $\{y_j^m\}$ using OLS. Add $\widehat{\theta}^{m+1}$ into the vector Θ, where Θ is a vector of length M which stores the bootstrap estimates.

6. Set $m = m + 1$ and repeat steps 3–5.

7. Carry out the above iterative process M times where M is large (e.g., M=1000), resulting in a vector Θ of length M.

8. We then calculate the mean, standard error, and confidence intervals from the vector Θ using the formulae

$$\begin{aligned}\widehat{\theta}_{\mathrm{boot}}&=\frac{1}{M}\sum_{m=1}^{M}\widehat{\theta}^m,\\ Cov\left(\widehat{\theta}_{\mathrm{boot}}\right)&=\frac{1}{M-1}\sum_{m=1}^{M}\left(\widehat{\theta}^m-\widehat{\theta}_{\mathrm{boot}}\right)^T\left(\widehat{\theta}^m-\widehat{\theta}_{\mathrm{boot}}\right),\\ SE_k\left(\widehat{\theta}_{\mathrm{boot}}\right)&=\sqrt{Cov\left(\widehat{\theta}_{\mathrm{boot}}\right)_{kk}}.\end{aligned}\tag{8}$$

3 Asymptotic Theory for Constant Variance Data

Given the statistical model and realizations described above, we can also compute estimates and standard errors using asymptotic theory. The algorithm [3,8] to obtain these estimates is given below.

1. First obtain the estimate, $\widehat{\theta}=\left(\widehat{K},\widehat{r},\widehat{x}_0\right)$ using OLS.

2. Compute the sensitivity matrix $\chi=\frac{\partial f}{\partial\theta}$, and variance estimate $\widehat{\sigma}^2$ as follows. The logistic model can be described in term of the differential equation:

$$\frac{dx(t)}{dt}=\mathcal{F}\left(x(t,\theta),\theta\right)=rx(t)\left(1-\frac{x(t)}{K}\right),$$

with corresponding sensitivity equations

$$\frac{d}{dt}\frac{\partial x}{\partial \theta}=\frac{\partial \mathcal{F}}{\partial x}\frac{\partial x}{\partial \theta}+\frac{\partial \mathcal{F}}{\partial \theta}.$$

The solution to this ordinary di erential equation provides the sensitivity matrix

$$\chi_{j,k}=\frac{\partial x\left(t_j\right)}{\partial \theta_k}=\frac{\partial f\left(t_j,\theta\right)}{\partial \theta_k}, \quad \text{for } j=1,\ldots,n,\ k=1,\ldots,p.$$

Note that $\chi = \chi^n$ is an $n \times p$ matrix. Again an unbiased estimate of the constant variance can be obtained by

$$\sigma_0^2 \approx \widehat{\sigma}_{OLS}^2=\frac{1}{n-p}\sum_{j=1}^{n}\left(y_j - f\left(t_j,\widehat{\theta}\right)\right)^2$$

3. Estimate the covariance matrix.

   The approximate true covariance matrix,

$$\Sigma_0^n=\sigma_0^2\left[\chi^T\left(\theta_0\right)\chi\left(\theta_0\right)\right]^{-1},$$

   is unknown since the true parameters θ_0 and variance σ_0^2 are unknown. It can be shown [25, p. 570] that under certain conditions, the estimate of θ is asymptotically normal:

$$\widehat{\theta}_n \sim \mathcal{N}_p\left(\theta_0,\sigma_0^2\left[\chi^T\left(\theta_0\right)\chi\left(\theta_0\right)\right]^{-1}\right).$$

   We estimate the covariance matrix using $\widehat{\theta}$ and $\widehat{\sigma}_{OLS}^2$ by

$$\Sigma_0^n \approx \widehat{\Sigma}^n\left(\widehat{\theta}\right)=\widehat{\sigma}_{OLS}^2\left[\chi^T\left(\widehat{\theta}\right)\chi\left(\widehat{\theta}\right)\right]^{-1}$$

4. Compute the standard error using $\widehat{\Sigma}^n\left(\widehat{\theta}\right)$ as

$$SE_k\left(\widehat{\theta}\right)=\sqrt{\widehat{\Sigma}_{kk}^n\left(\widehat{\theta}\right)}.$$

4 Bootstrapping Algorithm for Non-constant Variance Data

We suppose now that we are given experimental data $(y_1, t_1), \ldots, (y_n, t_n)$ from the following underlying observation process

$$Y_j=f\left(t_j,\theta_0\right)\left(1+\mathcal{E}_j\right) \tag{9}$$

where $j = 1, \ldots, n$ and the $\mathcal{E}_j$ are *iid* from a distribution F with mean zero and non-constant variance. Note that $E(Y_j) = f(t_j, \theta_0)$ and $\mathrm{var}\left(Y_j\right)=\sigma_0^2 f^2\left(t_j,\theta_0\right)$, with associated corresponding realizations of Y_j given by

$$y_j=f\left(t_j,\theta_0\right)\left(1+\epsilon_j\right).$$

We see that the variance generated in this fashion is model dependent and hence generally is longitudinally non-constant variance. The appropriate method to use to estimate θ_0 and σ_0^2 is a particular form of the Generalized Least Squares (GLS) method [3,16]. To define the *random variable* θ_{GLS} the following equation must be solved for the estimator θ_{GLS}:

$$\sum_{j=1}^{n} w_j\left[Y_j - f\left(t_j,\theta_{\mathrm{GLS}}\right)\right]\nabla f\left(t_j,\theta_{\mathrm{GLS}}\right)=0, \tag{10}$$

where Y_j obeys (9) and $w_j = f^{-2}$(tj , θGLS). We note these are the normal equations (obtained by equating to zero the gradient of the weighted least squares criterion in the case where the weights w_j are not dependent on θ). The quantity θ_{GLS} is a random variable, hence if $\left\{y_j\right\}_{j=1}^{n}$ is a *realization* of the random process Y_j then solving

$$\sum_{j=1}^{n} f^{-2}\left(t_j,\hat{\theta}\right)\left[y_j - f\left(t_j,\hat{\theta}\right)\right]\nabla f\left(t_j,\hat{\theta}\right)=0, \tag{11}$$

for $\hat{\theta}$ we obtain an estimate $\hat{\theta}_{\mathrm{GLS}}$ for θ_{GLS}.

An estimate $\hat{\theta}_{\mathrm{GLS}}$ can be solved for iteratively. The iterative procedure as described in [16] is summarized as follows:

1. Estimate $\hat{\theta}_{\mathrm{GLS}}$ by $\hat{\theta}^{(0)}$ using the OLS equation (4). Set $k = 0$.
2. Form the weights $\hat{w}_j=f^{-2}\left(t_j,\hat{\theta}^{(k)}\right)$.
3. Re-estimate $\hat{\theta}$ by solving

   $$\hat{\theta}^{(k+1)}=\arg\min_{\theta\in\Theta_{ad}}\sum_{j=1}^{n}\hat{w}_j\left(y_j - f\left(t_j,\theta\right)\right)^2$$

   to obtain the $k + 1$ estimate $\hat{\theta}^{(k+1)}$ for $\hat{\theta}_{\mathrm{GLS}}$.
4. Set $k = k + 1$ and return to 2. Terminate the process when two of the successive estimates for $\hat{\theta}_{\mathrm{GLS}}$ are suffciently close.

A standard algorithm [12,13,15, p. 287–290] can be used to compute the *bootstrapping estimate* $\hat{\theta}_{\mathrm{boot}}$ of θ_0 and its empirical distribution. We treat the general case for nonlinear dependence of the model output on the parameters θ. (For the special case of linear dependence, one can consult the references given above.)

1. First obtain the estimate $\hat{\theta}^0 = \left(\hat{K}^0, \hat{r}^0, \hat{x}_0^0\right)$ from the entire sample $\{y_j\}$ using GLS.
2. Define the non-constant variance standardized residuals

$$\bar{s}_j = \frac{y_j - f\left(t_j, \hat{\theta}^0\right)}{f\left(t_j, \hat{\theta}^0\right)}.$$

   Then $\left\{\bar{s}_1, \ldots, \bar{s}_n\right\}$ are realizations of *iid* random variables $\bar{S}_j$ with empirical distribution F_n. In this nonlinear weighted case the desired mean and variance conditions only hold approximately

$$E\left(\bar{S}_j | F_n\right) = n^{-1} \sum_{j=1}^{n} \bar{s}_j \approx 0, \quad \operatorname{var}\left(\bar{S}_j | F_n\right) = n^{-1} \sum_{j=1}^{n} \bar{s}_j^2 \approx \hat{\sigma}^2.$$

   Set $m = 0$.
3. Create a bootstrap sample of size n using random sampling with replacement from the data (realizations) $\left\{\bar{s}_1, \ldots, \bar{s}_n\right\}$ to form a bootstrap sample $\left\{\bar{s}_1^m, \ldots, \bar{s}_n^m\right\}$.
4. Create bootstrap sample points

$$y_j^m = f\left(t_j, \hat{\theta}^0\right) + f\left(t_j, \hat{\theta}^0\right) s_j^m,$$

   where $j = 1, \ldots, n$.
5. Obtain a new estimate $\hat{\theta}^{m+1} = \left(\hat{K}^{m+1}, \hat{r}^{m+1}, \hat{x}_0^{m+1}\right)$ from the bootstrap sample $\left\{y_j^m\right\}$ using GLS. Add $\hat{\theta}^{n+1}$ into the vector Θ, where Θ is a vector of length M which stores the bootstrap estimates.
6. Set $m = m + 1$ and repeat steps 3–5.
7. Carry out the above iterative process M times where M is large (e.g., M=1000), resulting in a vector Θ of length M.
8. We then calculate the mean, standard error, and confidence intervals from the vector Θ using the same formulae (8) as before.

If bootstrapping samples $\left\{y_1^m, \ldots, y_n^m\right\}$ resemble the data $\{y_1, \ldots, y_n\}$ in terms of the empirical error distribution, F_n, then the parameter estimates are expected to be consistent. The modification of the standardized residuals allows each of the bootstrapping samples to have an empirical distribution with the same mean and variance as the original F_n. We used both definitions for the standardized residuals (for linear and non-linear models) when doing our analysis. Although the definition of the standardized residual for the non-linear model gives an approximation to the conditions needed to hold, it performs comparably to the standardized residual as defined for the linear model. As a result, we chose the nonlinear standardized residuals for our simulation comparison studies.

5 Asymptotic Theory for Non-constant Variance Data

Assume we are given experimental data $(y_1, t_1), \dots, (y_n, t_n)$ from the following underlying observation process

$$Y_j = f\left(t_j, \theta_0\right)\left(1+\mathcal{E}_j\right),$$

where $j = 1, \dots, n$ and the E_j are *iid* with non-constant variance. Note that $E(Y_j) = f(t_j, \theta_o)$ and $\mathrm{var}\left(Y_j\right) = \sigma_0^2 f^2\left(t_j, \theta_0\right)$, with associated corresponding realizations of Y_j given by

$$y_j = f\left(t_j, \theta_0\right)\left(1+\epsilon_j\right).$$

When using asymptotic theory [3,8], we obtain the $\widehat{\theta}$ using the GLS algorithm. Then σ_0^2 is approximated by

$$\sigma_0^2 \approx \widehat{\sigma}_{GLS}^2 = \frac{1}{n-p} \sum_{j=1}^{n} \frac{1}{f^2\left(t_j, \widehat{\theta}\right)} \left(f\left(t_j, \widehat{\theta}\right) - y_j\right)^2.$$

We estimate the covariance matrix using $\widehat{\theta}$ and $\widehat{\sigma}_{GLS}^2$ by

$$\Sigma_0^n \approx \widehat{\Sigma}^n\left(\widehat{\theta}\right) = \widehat{\sigma}_{GLS}^2 \left[\chi^T\left(\widehat{\theta}\right) W\left(\widehat{\theta}\right) \chi\left(\widehat{\theta}\right)\right]^{-1},$$

where $W^{-1}(\theta) = diag(f^2(t_1, \theta), \dots, f^2(t_n, \theta))$. We compute the standard error using $\widehat{\Sigma}^n\left(\widehat{\theta}\right)$ and

$$SE_k\left(\widehat{\theta}\right) = \sqrt{\widehat{\Sigma}_{kk}^n\left(\widehat{\theta}\right)}.$$

6 Results of Numerical Simulations

6.1 Simulated Noisy Data Sets

We created noisy data sets using model simulations and a time vector of length $n = 50$, $t = (0, \dots, 25)$. The underlying model with the true parameter values $\theta_0 = (17.5, 0.7, 0.1)$ was solved for $f(t_j, \theta_0)$ using MATLAB's function *ode*45. A noise vector ε of length n with noise level Nl, was taken from a random number generator for $N(0, Nl^2)$. The constant variance data sets were obtained from the equation

$$y_j = f\left(t_j, \theta_0\right) + \epsilon_j.$$

Similarly, for non-constant variance data sets we used

$$y_j = f\left(t_j, \theta_0\right)\left(1+\epsilon_j\right).$$

Constant variance and non-constant variance data sets were created for 1%, 5%, and 10% noise, i.e., Nl = .01, .05, and .1.

6.2 Constant Variance Data with OLS

We used the constant variance (CV) data with OLS to carry out the parameter estimation calculations. The bootstrap estimates were computed with M = 1000 and of course, θ_0 = (17.5, 0.7, 0.1). We choose M = 1000 as recommend by Davidian in [15]. We also use M = 1000 because we are computing confidence intervals and not only estimates and standard errors, Efron and Tibirshani recommend that M = 1000 when confidence intervals are to be computed [17]. The standard errors SE_k and corresponding confidence intervals $\left[\widehat{\theta}_k - 1.96SE_k, \widehat{\theta}_k+1.96SE_k\right]$ are listed in Tables 1, 2 and 3. In Figure 2 we plot the empirical distributions for the case Nl = .05 corresponding to the results in Table 2; plots in the other two cases are quite similar.

The parameter estimates and standard errors are comparable between the asymptotic theory and the bootstrapping theory for this case of constant variance. However, the computational times (given in Table 4) are two to three orders of magnitude greater for the bootstrapping method as compared to those for the asymptotic theory. For this reason, the asymptotic approach would appear (as expected from the theory) to be the more advantageous method.

6.3 Non-constant Variance Data with GLS

We carried out a similar set of computations for the case of non-constant variance (NCV) data using a GLS formulation (in these calculations we used 1 GLS iteration). The bootstrap estimates were again computed with M = 1000. Standard errors and corresponding confidence intervals are listed in Tables 5, 6 and 7. In Figure 3 we plot the empirical distributions for the case Nl = .05 corresponding to the results in Table 6; again, plots in the other two cases are quite similar.

We observe that the standard errors computed from the bootstrapping methodare very similar to the standard errors computed using asymptotic theory. In each of these cases the standard errors for the parameter K are one to two orders of magnitudes greater than the standard errors for r and x_0. The computational time is also slower for the bootstrapping method, thus asymptotic theory may again be the method of choice.

7 Using Incorrect Assumptions on Errors

In practice, one rarely knows the form of the statistical error with any degree of certainty, so that the assumed models (3) and (9) may well be incorrect for a given data set. To obtain some information on the effect, if any, of incorrect error model assumptions on comparisons between bootstrapping and use of asymptotic theory in computing standard errors we carried out further computations.

In this section, we repeat the comparisons above for asymptotic theory and bootstrapping generated standard errors, but with incorrect assumptions about the error (constant or non-constant variance). Using the same data sets as created previously, we computed parameter estimates and standard errors for the constant variance data for asymptotic theory and bootstrapping using a GLS formulation, which usually is employed if non-constant variance data is suspected. Similarly, we estimated parameters and standard errors for both approaches in an OLS formulation with the non-constant variance data. In summary, we demonstrate below that if incorrect assumptions are made about the statistical model for error contained within the data, one cannot ascertain whether the correct assumption about the error has been made simply from examining the estimated standard errors. The residual plots must be examined to determine if the error model for constant or non-constant variance is reasonable. This is discussed more fully in [3] and [8, Chapter 3].

7.1 Constant Variance Data, Using GLS

The bootstrap estimates were computed with $M = 1000$ and 1 GLS iteration, with the findings reported in the tables below in a format similar to those in the previous sections. No empirical distribution plots are given here because we found they added little in the way of notable new information.

As can be summarized from the results obtained, when the data has constant variance, but the incorrect assumption is made about the error, the asymptotic theory error estimates are again comparable to the bootstrapping error estimates. Also, K has large standard error as Nl increases, specifically it goes from approximately .043 to .5 for noise levels 1% to 10%, respectively. The r and x_0 standard errors also increase as the noise levels increase. These incorrectly obtained standard error estimates are larger in comparison to those obtained using an OLS formulation with the correct assumption about the error, comparing Tables 9-11 to Tables 1-3 respectively.

7.2 Non-constant Variance Data, Using OLS

The bootstrap estimates were again computed with $M = 1000$. When the error estimates for data sets with non-constant variance are estimated using OLS, the estimates are comparable for the asymptotic theory and bootstrapping methods. In comparing these estimates to the error estimates obtained under accurate assumptions (comparing Tables 13-15 to Tables 5-7 respectively), we observe that under the incorrect error model assumption, the standard error for K is always smaller (though comparable) regardless of the noise level, while the corresponding r and x_0 standard errors are always larger. For x_0, the standard error under OLS is an order of magnitude larger as compared to the GLS case (non-constant variance).

8 The Corrective Nature of Bootstrapping Covariance Estimates

When estimating θ using GLS, there may be more error in the estimated covariance matrix, $\widehat{\Sigma}^n\left(\widehat{\theta}\right)$ due to the estimation of weights. Recalling Theorem 4 from Carroll, Wu, and Ruppert [13, p. 1048], originally given in [24, p. 815], we can write the bootstrap estimate of the covariance matrix as

$$Cov\left(\widehat{\theta}_{\text{boot}}\right)=\widehat{\Sigma}^n\left(\widehat{\theta}\right)+n^{-1}\sigma_0^2\Lambda\left(F\right)+O_p\left(n^{-3/2}\right),$$

where (F) is an unknown positive-definite matrix depending on the distribution F for the *iid* errors E_j. As a result the bootstrapping covariance matrix is generally thought to be more accurate than the asymptotic theory estimate due to the corrective term $n^{-1}\sigma_0^2\Lambda(F)$. The corrective term is discussed for the linear model in [13, p. 1050], and for the nonlinear model in [12, p. 28] where the standardized residuals are defined by

$$r_i=\frac{y_i-f\left(t_i,\theta\right)}{\sigma_g\left(f\left(t_i,\theta\right),t_i,\theta\right)}.$$

Although in our results there are minimal differences between the two estimates obtained via bootstrapping and asymptotic theory, Carroll, Wu, and Ruppert report bootstrapping to be more favorable when estimating weights that have a large effect on the parameter estimates [13]. This corrective term does not arise with the OLS method where no weights are being estimated.

Therefore, as might be expected, there seems to be no advantages in implementing the bootstrapping method over asymptotic theory for estimation of parameters and standard errors from constant variance data [15, p. 287]. To examine the effects of this second order correction, we re-ran our analysis for $n = 10$ data points for non-constant variance (10% noise) using GLS. The results were similar comparing asymptotic theory to bootstrapping, however the standard errors from bootstrapping were slightly smaller (less conservative) than the standard errors from asymptotic theory for all parameters. This is unexpected due to the correction term described above. We note that in repeated simulations different parameter estimates resulted in smaller (less conservative) SE some of the times, and other times larger (more conservative) SE for bootstrapping than for the corresponding asymptotically computed SE. However, the bootstrapping standard errors were never all larger (more conservative) than asymptotic theory standard errors. (We note that in general one desires small standard errors in assessing uncertainty in estimates but *only if they are accurate*; in some cases–essentially when they are more accurate– *more conservative* (larger) SE are desirable.) For fewer than $n = 10$ data points, the inverse problem using bootstrapping did not converge in a reasonable time period, though it did for asymptotic theory. This may be due to the resampling with replacement of the residuals causing the bootstrap samples to sometimes have a large estimate of K, which in turn makes the estimated GLS weights very small.

In simulations reported in previous sections, we found that our results did not necessarily meet expectations consistent with the theory presented by Carroll, Wu and Ruppert [13]. After further examination of the theory, we find that the theory is presented and discussed in detail for a linear model. To further explore this, we linearized our original model ($f(t, \theta)$) about $\theta = \theta^*$, and re-ran our simulations. As a result of the linearization we have the following new model

$$y(t) = \chi_0(t) + \chi_1(t)(K - K^*) + \chi_2(t)(r - r^*) + \chi_3(t)\left(x_0 - x_0^*\right), \tag{12}$$

where $\chi_0 = f(t, \theta^*)$, $\chi_1(t) = \frac{\partial f(t,\theta^*)}{\partial K}$, $\chi_2(t) = \frac{\partial f(t,\theta^*)}{\partial r}$, and $\chi_3(t) = \frac{\partial f(t,\theta^*)}{\partial x_0}$. Note that $\theta^* = (K^*, r^*, x_0^*) = (17.5, 0.7, 0.1)$. We considered this new model and performed the same computational analysis as described earlier. At each simulation while the values for the bootstrapping standard errors are similar, their specific values would vary in comparison to the asymptotic theory standard errors, which are the same for repeated simulations using the same set of simulated data each time. For example, during the first run all bootstrapping computed standard errors for K, r, x_0 would be larger in comparison to the asymptotic theory estimates, while on the next run only K, r would have corresponding larger estimates. We performed a Monte Carlo analysis, for 1000 trials, to determine if the corrective nature of the bootstrap was present on average. For each of the Monte Carlo trials we computed the estimates and standard errors using asymptotic theory and bootstrapping with M=250. We performed these Monte Carlo simulations on the same time interval T ε [0, 20], at 10% noise with relative error, generating new simulated data for each Monte Carlo simulation. Each Monte Carlo analysis used a fixed n for the 1000 trails, and the analysis was repeated for n = 20, 25, 30, 35, 40, 45, and 50 time points. The results for these Monte Carlo simulation are given in Tables 17-23.

When comparing this average bootstrapping standard error with the average asymptotic standard error for K, r, x_0 at n = 20, 25, 30, 35, 40, 45, and 50, we observe that the bootstrapping estimates are larger (as expected with the correction term) for parameters K, r, but slightly smaller (not expected) for the parameter x_0 (see Tables 17-23). The relative error model has very little noise around $t = 0$, due to the model have the smallest function value in that region. Therefore, we do not expect for the weights from GLS to have a large impact on the standard errors for x_0. This may explain why we do not observe a correction from the bootstrapping method for x_0.

We would expect that the theory presented in [13] would be observed at lower values of *n* due to the correction term $n^{-1}\sigma_0^2\Lambda(F)$; however, the inverse problem is very unstable at this sample size because it appeared that at least 9 data points are needed for each parameter that is being estimated in this problem. (We remark that in general the correlation between the number of data points required per parameter estimated is not easy to compute and is very much problem dependent–this is only one of a number of diffculties arising in such problems–see [9, p. 87] for discussions. Indeed this problem is one of active research interest in the areas of sensitivity functions [4,5,6,28] and design of experiments.) We expected to observe theory-consistent SE at the higher values of *n* ($n = 20$), because we have three parameters estimates. For the parameters that did show a correction from Bootstrapping, *K*, *r* , Table 24 shows the average correction term for each *n*, given by $SE_{avg}\left(\widehat{\theta}_{\text{boot}}\right) - SE_{avg}\left(\widehat{\theta}_{asy}\right)$.

Table 24 confirms that there is a larger correction for smaller *n*. This is expected given the correction term $n^{-1}\sigma_0^2\Lambda(F)$. We also observe that the correction for *K* is consistently larger than the correction for *r*. This implies that the estimation of GLS weights has a greater effect on the standard errors of *K* than for *r*.

To better understand the fluctuations in bootstrapping standard errors among the three parameters, we ran the Monte Carlo simulations for the estimation of just one parameter. When only the parameter *r* was estimated, the bootstrapping standard errors were larger than those of asymptotic theory for every case, using $n = 10$, 15, 20 and 25 time points. These simulations supported the theory of the corrective nature of bootstrapping. We also examined the standard errors when only estimating x_0, again for $n = 10$, 15, 20 and 25. For x_0, the standard errors were consistently smaller by an order of magnitude for the bootstrapping method, as compared to asymptotic theory, thus not displaying the corrective nature of the bootstrap. This example provides some evidence that the corrective nature of the bootstrap may only be realized in certain situations.

Carroll and Ruppert [12, p. 15–16] describe the following result for the behavior of GLS with a particular non-constant variance model. If the data points have higher variance in a neighborhood of importance for estimating a specific parameter (for our example the region of rapidly increasing slope is important for estimating *r*), then the weights in GLS heavily influence the estimate and the standard error. Based on our computational findings, we are inclined to think that this local variance also influences whether or not the bootstrapping estimate will be "corrective". If the GLS weights are not as important, then the bootstrap error estimate will not exhibit the corrective term of the covariance properly.

We ran some simulations for both the linearized and nonlinear models to test this hypothesis. For our example, in the solution to the logistic equation there are three regions each of importance for estimating one of the parameters. Region I is important for estimation of x_0 and is the region where $x(t)$ is near x_0, located near the initial time points before the slope begins to increase significantly. Region II data influences heavily the estimation of *r* and is located in the area of the increasing slope. Region III is located where the solution approaches the steady state $x(\infty) = K$. Due to the manner in which our simulated non-constant variance data was modeled, Region I had little variation in comparison to Regions II and III. This low variation in Region I led us to believe that increasing the variance in this region would demonstrate the corrective nature of the bootstrapping estimate. We subsequently created data sets (linear and nonlinear) with larger variance in Region I and the variance remained the same in Regions II and III. For these new data sets all of the bootstrapping standard errors were larger than the asymptotic standard errors; however, the estimates for the intrinsic growth rate, *r*, become unreasonable. Although it appears the theoretically predicted corrective nature of the bootstrap is exhibited, there is exists a break down in the inverse problem due to the unreasonable estimate for *r*. This break down occurs as a result of the added variance in the Region I, thus changing the assumptions on the variance model for the inverse problem, which are now incorrect. While there are issues observing the effect of the addition of variance into regions with small local variance, the regions that naturally have a larger amount of local variance exhibit the corrective nature of the bootstrap. This leads us to believe that *local variance* strongly influences the presence or not of the corrective nature of the bootstrapping estimate.

9 Summary

Based on our computational experience with a known example we make the following summary remarks/conclusions. Asymptotic theory and bootstrapping are distinct methods for quantifying uncertainty in parameter estimates. Asymptotic theory is always faster computationally than bootstrapping; however, this is the only definitive comparison that can be made. Depending on the type of data set, one method may be more favorable than the other. For constant variance data using OLS there is no clear advantage in using bootstrapping over asymptotic theory. However, for a complex system it may be too complicated to compute the sensitivities needed for asymptotic theory; then bootstrapping may be more reasonable for estimating the standard error. If computation time is already a concern, but sensitivities can be computed, then asymptotic theory may be advantageous.

Given that the statistical model correctly assumes non-constant variance, the parameters are estimated using GLS, and the asymptotic theory error estimates and the bootstrapping error estimates are comparable. If local variation in the data points is large in a region of importance for estimation of a parameter, then the bootstrapping covariance estimate may contain a corrective term of significance. In that situation, the standard errors for bootstrapping will be *larger* than those of asymptotic theory, and hence *more conservative* and also more *accurate*. If the local variation in the data is low in a region of importance for the estimation of a parameter, then there will be insignificant correction in the bootstrap's estimate of standard error over that of the asymptotic theory. Thus for non-constant variance data, the choice between bootstrapping and asymptotic theory depends on *local variation* in the data in regions of importance for the estimation of the parameters.

Acknowledgments

This research was supported by the National Institutes of Health in part under grant number R01AI071915-07 from the National Institute of Allergy and Infectious Diseases and in part by the Air Force Office of Scientific Research under grant number FA955009-1-0226. The research of D.R. was also supported in part by a Graduate Fellowship under the Initiative for Maximizing Student Diversity (IMSD) program grant PAR-06-553 funded by the NIH.

A Appendix: Summary of Literature on the Asymptotic Nature of the Bootstrapping Estimate

For the sake of completeness, we include here a brief review of some of the theoretical underpinnings for bootstrapping similar to the convergence properties in the asymptotic theory for standard errors as given in [25, Chapter 12]. Bickel and Freedman [11] discuss an early result for the asymptotic theory of the specific case where the parameters are estimated directly from the sample (ex: mean, and variance). In this situation, there is no model for the data therefore no inverse problem is needed to estimate the parameters. The results that follow are of interest for the general bootstrapping method, but are not directly applicable for parameter estimation for a nonlinear model using bootstrapping. Bickel and Freedman [11] give insight to asymptotic theory for the general bootstrapping method in a population $Y_1, \ldots, Y_n$, where $Y_i = t_i + \varepsilon i$ and *iid* with empirical distribution F_n. It is assumed that F_n has finite mean, μ_n, and finite positive variance σ_n^2 where the standard deviation is

$$s_n^2 = \frac{1}{n} \sum_{i=1}^{n} (Y_i - \mu_n)^2.$$

Then a conditionally independent bootstrapping sample of size n_b sampled from the population $\{Y_1, \ldots, Y_n$ is defined as $\left\{Y_1^*, \ldots, Y_{n_b}^*\right\}$, Where $s_{n_b}^* = \frac{1}{n_b} \sum_{i=1}^{n_b} \left(Y_1^* - \mu_{n_b}^*\right)^2$, and $\mu_{n_b}^* = \frac{1}{n_b} \sum_{i=1}^{n_b} Y_i^*$ is an estimate for μ_n. From here a pivotal quantity $Q_n = \sqrt{n}\,(\mu_n - \mu)\,/\,s_n$ can be estimated from $Q_{n_b}^* = \sqrt{n_b}\left(\mu_{n_b}^* - \mu_n\right) / s_{n_b}^*$ Note that n is the number of data points in the population while n_b is the bootstrapping sample size which does not have to equal n, and Q_n converges weakly to $\mathcal{N}(0, 1)$, given μ is known, by the Central Limit Theorem. Bickel and Freedman report the asymptotic properties of $Q_{n_b}^*$ and find its asymptotic convergence to be in accordance with the convergence of Q_n. The following results are due to the previously discussed findings. Given $Y_1, Y_2, \ldots$ are *iid* and have finite positive variance σ^2. Along almost all sample sequences Y_1, $Y_2, \ldots$, given $(Y_1, .., Y_n)$, as n and n_b tend to ∞ we have the following result:

- **a.** The conditional distribution of $\sqrt{n_b}\left(\mu_{n_b}^* - \mu_n\right)$ converges weakly to $\mathcal{N}\left(0, \sigma^2\right)$ and
- **b.** $s_{n_b}^*$ converges to σ in conditional probability [11, p. 1197] .

Next if $Y_1, Y_2, \ldots$, are independent with a common distribution in R^d, and . $E\{\|y_1\|^2\} < \infty$. Let F_n be the empirical distridution of $Y_1, \ldots, Y_n$, and let $Y_1^*, \ldots, Y_{n_b}^*$ be conditionally independent, with common distribution F_n. Then along almost all sample sequences, as n_b and n tend to infinity we have the following result:

- **a.** The conditional distribution of $\sqrt{n_b}\left(\frac{1}{n_b}\sum_{i=1}^{n_b} Y_i^* - \frac{1}{n}\sum_{j=1}^{n} Y_j\right)$ converges weakly to the d-dimensional normal distribution Pwith mean 0, and variance-covariance matrix equal to the theoretical variance covariance matrix Y_1
- **b.** The empirical variance-covariance matrix of $Y_1^*, \ldots, Y_{n_b}^*$ converges in conditional probability to the theoretical variance-covariance matrix of Y_1. [11, p. 1197]

Later asymptotic theory results are for the same bootstrapping on the residuals method that we described above, but are restricted to a model that is linear in the parameters. Shao and Tu [26] discuss the asymptotic theory of the bootstrapping on the residuals method for a linear model: $y_i = (t^{(i)})'\theta + \epsilon_i$ for $i = 1, \ldots, n$, where $t^{(i)}$ is a p-vector of explanatory variables associated with the data (for example $t^{(i)} = \left(1, t_i, t_i^2, \ldots t_i^{p-1}\right)'$, and θ is a p-vector of unknown parameters. Let $T' = (t^{(1)}, \ldots, t^{(n)})$, $h_i = t_i'\,(T'T)^{-1}t^{(i)}$ and $h_{max} = max_{i \le n} h_i$. Let v_{boot} be the variance estimator of $\widehat{\theta}_{\text{boot}}$, the bootstrapping estimate of θ. The OLS estimate of θ is denoted by $\widehat{\theta}_{OLS}$.

In addition, we define the Mallows' distance (as shown in [26, p. 73]) on $\mathcal{F}_{r,s} = \left\{G \in \mathcal{F}_{\mathbb{R}^s} : \int \|t\|^r dG(t) \langle \infty\right\}$. The two distribution H and G in $\mathcal{F}_{r,s}$ have Mallows' distance,

$$d_{\text{mal}}(H, G) = \inf_{\tau_{X,Y}} (E\|X - Y\|^r)^{1/r},$$

where $T_{X,Y}$ is the collection of all possible joint distributions of (X, Y), if X and Y have marginal distributions H and G, respectively.

Shao and Tu give us the following two results that prove the bootstrapping estimator of variance of the parameter is consistent, as is the bootstrapping distribution estimator [26]. To observe consistency of v_{boot} (the variance of the parameter, the follow must hold $T'T \to \infty$, $h_{max} \to 0$, and we have constant variance ($\sigma_i^2 = \sigma^2$ for all i), then $v_{\text{boot}} / \text{var}\left(c'\widehat{\theta}_{OLS}\right) \to_p 1$ 1, where c is a vector of arbitrary constants. Additionally, we have that $\text{bias}\,(v_{\text{boot}})\,/\text{var}\left(c'\widehat{\theta}_{OLS}\right) = O\left(n^{-1}\right)$ and $\text{var}\,(v_{\text{boot}})\,/\left[\text{var}\left(c'\widehat{\theta}_{OLS}\right)\right]^2 = O\left(n^{-1}\right)$, if $E\left(\epsilon_i^4 \langle \infty\right)$[26, p. 317].

Next to observe consistency of the distribution estimator based on bootstrapping the residuals, assume $T'T \to \infty$, $h_{max} \to 0$, and the $\in_i$ are *i.i.d.*. Then

$$d_{\text{mal}}\,(H_{\text{boot}}, H_n) \to_{a.s.} 0,$$

where H_n is the distribution of $\left(T'T\right)^{1/2}\left(\widehat{\theta}_{OLS} - \theta\right)$ and H_{boot} is the bootstrap distribution of $\left(T'T\right)^{1/2}\left(\widehat{\theta}_{\text{boot}} - \widehat{\theta}_{OLS}\right)$[26, p. 320]: .

Shao and Tu also state that these consistency results can be generalized for nonlinear models in [26, p. 331, 335–337], though no rigorous proof is given.

A.1 Iterative Bootstrap Asymptotics

Earlier we discussed asymptotics for bootstrapping methods in reference to sample size, n, now we will highlight some discussion on the asymptotics in reference to the number M of bootstrapping samples. Efron and Tibshirani state the bootstrap method is asymptotically e cient in [17, p. 395] , meaning that as $M \to \infty$, $\widehat{SE}^{(M)}\left(\widehat{\theta}_{\text{boot}}\right) \to \widehat{SE}^{\infty}\left(\widehat{\theta}_{\text{boot}}\right)$ the ideal bootstrap estimate of standard error in [17, p. 50]. In addition, for a suffciently large number M of bootstrap samples, bootstrapping will give accurate parameter estimates [17, p. 395]. For a general bootstrap method, $M = 50$ is suffcient to estimate θ, $M = 200$ is su cient to estimate the standard error, and $M = 1000$ is su cient to estimate confidence intervals for θ as reported in [17, p. 52, 164].

A.2 Determining Iterative Size for the Bootstrapping Estimate

Let $H_{boot}(x)$ represent the distribution of $\widehat{\theta}_{\text{boot}}$ conditional to the empirical distribution F . Let $H_{\text{boot}}^M(x)$ be the approximation of $H_{boot}(x)$ at the M^{th} iterate. The following result holds

$$\sup_x |H_{\text{boot}}^M(x) - H_{\text{boot}}(x)| = \epsilon_n + \sqrt{M^{-1}\log\log M}$$

and

$$\sup_{0\langle t \langle 1} |\left(H_{\text{boot}}^M\right)^{-1}(t) - (H_{\text{boot}})^{-1}(t)| = O\left(\epsilon_n + \sqrt{M^{-1}\log\log M}\right).$$

where $\epsilon_n = \sup_x |H_{\text{boot}}^M(x) - H_{\text{boot}}(x)|$, established by Shi, Wu and Chen [27] as shown in Shao and Tu [26, p. 210]. In addition M is determined using the following relationship: $M^{-1}\log\log M = o\left(\epsilon_n^2\right)$.

References

[1]. Anger, G. Inverse Problems in Differential Equations. Plenum Press; New York: 1990.

[2]. Banks, HT.; Buksas, MW.; Lin, T. Electromagnetic Material Interrogation Using Conductive Interfaces and Acoustic Wavefronts. SIAM FR21; Philadelphia: 2002.

[3]. Banks, HT.; Davidian, M.; Samuels, JR., Jr.; Sutton, KL. An inverse problem statistical methodology summary, CRSC Technical Report, CRSC-TR08-01, NCSU, January, 2008; Chapter 11. In: Chowell, Gerardo; Hyman, Mac; Hengartner, Nick; Bettencourt, Luis M.A; Castillo-Chavez, Carlos, editors. Statistical Estimation Approaches in Epidemiology. Springer; Berlin Heidelberg New York: 2009. p. 249-302.

[4]. Banks HT, Dediu S, Ernstberger SE. Sensitivity functions and their uses in inverse problems. J. Inverse and Ill-posed Problems 2007;15:683–708.

[5]. Banks, HT.; Dediu, S.; Ernstberger, SL.; Kappel, F. CRSC Technical Report CRSC-TR08-12. NCSU; Sep. 2008 A new approach to optimal design problems.

[6]. Banks HT, Ernstberger SL, Grove SL. Standard errors and confidence intervals in inverse problems: sensitivity and associated pitfalls. J. Inverse and Ill-posed Problems 2007;15:1–18.

[7]. Banks, HT.; Kunsich, K. Estimation Techniques for Distributed Parameter Systems. Birkhaüser; Boston: 1989.

[8]. Banks, HT.; Tran, HT. Mathematical and Experimental Modeling of Physical and Biological Processes. CRC Press; Boca Raton: 2009.

[9]. Bates, DM. Nonlinear Regression Analysis and its Applications. J. Wiley & Sons; Somerset, NJ: 1988.

[10]. Baumeister, J. Stable Solutions of Inverse Problems. Vieweg; Braunschweig: 1987.

[11]. Bickel P, Freedman D. Some asymptotic theory for the bootstrap. The Annals of Statistics 1981;9 (6):1196–1217.

[12]. Carroll, RJ.; Ruppert, D. Transformation and Weighting in Regression. Chapman & Hall; London: 1988.

[13]. Carroll RJ, Wu CFJ, Ruppert D. The effect of estimating weights in Weighted Least Squares. J. Amer. Statistical Assoc 1988;83:1045–1054.

[14]. Casella, G.; Berger, RL. Statistical Inference. Duxbury; California: 2002.

[15]. Davidian, M. Chapters 9 and 11. Nonlinear Models for Univariate and Multivariate Response. 2007. ST 762 Lecture Noteshttp://www4.stat.ncsu.edu/ davidian/courses.html

[16]. Davidian, M.; Giltinan, DM. Nonlinear Models for Repeated Measurement Data. Chapman & Hall; London: 1995.

[17]. Efron, B.; Tibshirani, RJ. An Introduction to the Bootstrap. Chapman & Hall/CRC Press; Boca Raton: 1998.

[18]. Engl, HW.; Groetsch, CW., editors. Inverse and Ill-posed Problems. Academic Press; Orlando: 1987.

[19]. Groetsch, CW. Inverse Problems in the Mathematical Sciences. Vieweg; Braunschweig: 1993.

[20]. Groetsch, CW. The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind. Pitman; London: 1984.

[21]. Hoffman, B. Regularization for Applied Inverse and Ill-posed Problems. Teubner; Leipzig: 1986.

[22]. Kot, M. Elements of Mathematical Ecology. Cambridge University Press; Cambridge: 2001.

[23]. Matloff N, Rose R, Tai R. A comparison of two methods for estimating optimal weights in regression analysis. J. Statist. Comput. Simul 1984;19:265–274.

[24]. Rothenberg TJ. Approximate normality of generalized least squares estimates. Econometrica 1984;52(4):811–825.

[25]. Seber, GAF.; Wild, CJ. Nonlinear Regression. Wiley-Interscience; Hoboken, NJ: 2003.

[26]. Shao, J.; Tu, D. The Jackknife and Bootstrap. Springer-Verlag; New York: 1995.

[27]. Shi X, Chen J, Wu CFJ. Weak and strong representations for quantile processes from finite populations with applications to simulation size in resampling methods. Canadian J. Statist 1990;18:141–148.

[28]. Thomaseth K, Cobelli C. Generalized sensitivity functions in physiological system identification. Annals of Biomedical Engineering 1999;27:607–616. [PubMed: 10548330]

[29]. Tikhonov, AN.; Arsenin, VY. Solutions of Ill-posed Problems. Winston and Sons; Washington: 1977.

[30]. Vogel, CR. Computational Methods for Inverse Problems. SIAM FR23; Philadelphia: 2002.

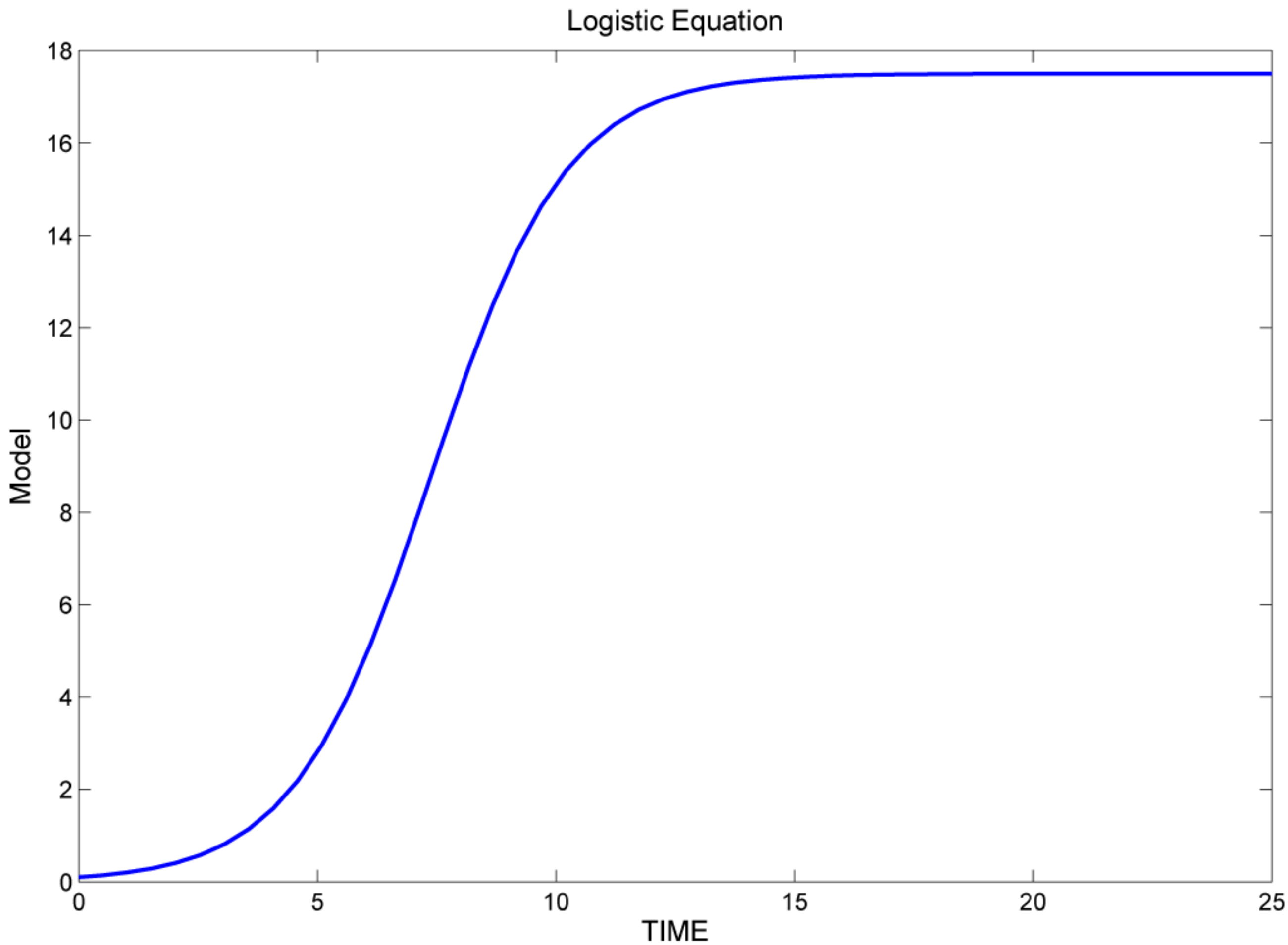

Figure 1.
Logistic curve with K = 17.5, r = .7 and x_0 = .1.

NIH-PA Author Manuscript
NIH-PA Author Manuscript

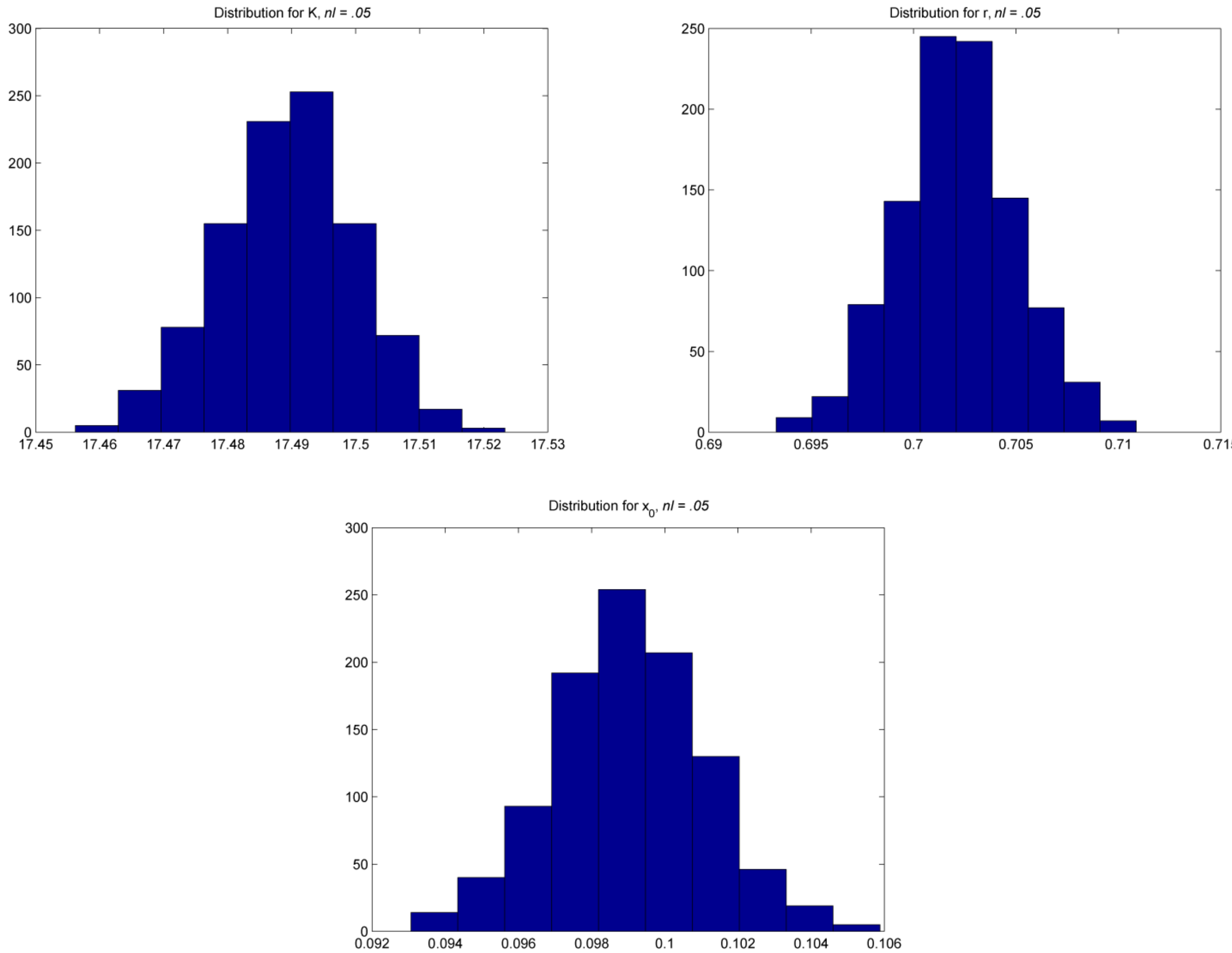

Figure 2.
Bootstrap Parameter Distributions Corresponding to 5% Noise with CV.

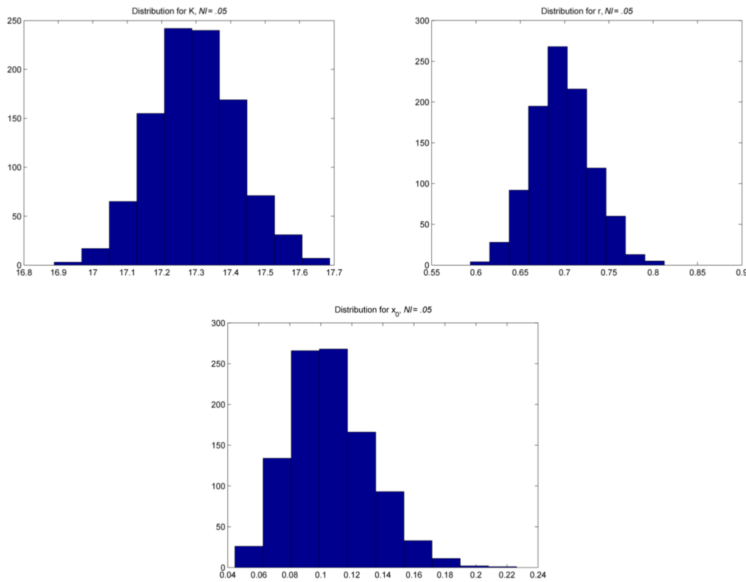

Figure 3.
Bootstrap Parameter Distributions for 5% Noise with NCV.

Table 1

Asymptotic and Bootstrap OLS Estimates for CV Data, *Nl* = 0.01

θ	$\hat{\theta}$	$SE(\hat{\theta})$	95% CI
$\hat{K}_{asy}$	17.498576	0.002021	(17.494615, 17.502537)
$\hat{r}_{asy}$	0.700186	0.000553	(0.699103, 0.701270)
$(\hat{K}_0)_{asy}$	0.100044	0.000407	(0.099247, 0.100841)
$\hat{K}_{boot}$	17.498464	0.001973	(17.494597, 17.502331)
$\hat{r}_{boot}$	0.700193	0.000548	(0.699118, 0.701268)
$(\hat{x}_0)_{boot}$	0.100034	0.000399	(0.099252, 0.100815)

Table 2

Asymptotic and Bootstrap OLS Estimates for CV Data, *Nl* = 0.05

θ	$\hat{\theta}$	$SE(\hat{\theta})$	95% CI
$\hat{K}_{asy}$	17.486571	0.010269	(17.466444, 17.506699)
$\hat{r}_{asy}$	0.702352	0.002825	(0.696815, 0.707889)
$(\hat{x}_0)_{asy}$	0.098757	0.002050	(0.0947386, 0.102775)
$\hat{K}_{boot}$	17.489658	0.010247	(17.469574, 17.509742)
$\hat{r}_{boot}$	0.702098	0.002938	(0.696339, 0.707857)
$(\hat{x}_0)_{boot}$	0.0990520	0.002152	(0.094834, 0.103270)

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 3

Asymptotic and Bootstrap OLS Estimates for CV Data, *Nl* = 0.1

θ	$\hat{\theta}$	$\mathrm{SE}(\hat{\theta})$	95% CI
$\hat{K}_{asy}$	17.528701	0.019091	(17.491283, 17.566120)
$\hat{r}_{asy}$	0.699335	0.005201	(0.689140, 0.709529)
$(\hat{x}_0)_{asy}$	0.100650	0.003851	(0.093103, 0.108198)
$\hat{K}_{boot}$	17.523374	0.019155	(17.485829, 17.560918)
$\hat{r}_{boot}$	0.699803	0.005092	(0.689824, 0.709783)
$(\hat{x}_0)_{boot}$	0.100317	0.003800	(0.092869 0.107764)

Table 4

Computation Times (sec) for Asymptotic Theory vs. Bootstrapping

Noise Level	Asymptotic Theory	Bootstrapping
1%	0.017320	4.285640
5%	0.009386	4.625428
10%	0.008806	4.914146

Table 5

Asymptotic and Bootstrap GLS Estimates for NCV Data, *Nl* = 0.01

θ	$\hat{\theta}$	$SE(\hat{\theta})$	95% CI
$\hat{K}_{asy}$	17.514706	0.028334	(17.459171, 17.570240)
$\hat{r}_{asy}$	0.70220	0.001156	(0.699934, 0.704465)
$(\hat{x}_0)_{asy}$	0.099145	0.000435	(0.098292,0.099999)
$\hat{K}_{boot}$	17.515773	0.027923	(17.461045, 17.570502)
$\hat{r}_{boot}$	0.702136	0.001110	(0.699960, 0.704311)
$(\hat{x}_0)_{boot}$	0.099160	0.000416	(0.098344, 0.099976)

Table 6

Asymptotic and Bootstrap GLS Estimates for NCV Data, *Nl* = 0.05

θ	$\hat{\theta}$	$SE(\hat{\theta})$	95% CI
$\hat{K}_{asy}$	17.322554	0.148891	(17.030728,17.614380)
$\hat{r}_{asy}$	0.699744	0.006126	(0.687736, 0.711752)
$(\hat{x}_0)_{asy}$	0.099256	0.002313	(0.094723, 0.103790)
$\hat{K}_{boot}$	17.329282	0.146030	(17.043064,17.615500)
$\hat{r}_{boot}$	0.700060	0.006003	(0.688294,0.711825)
$(\hat{x}_0)_{boot}$	0.099210	0.002329	(0.094645,0.103775)

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 7

Asymptotic and Bootstrap GLS Estimates for NCV Data, *Nl* = 0.1

θ	$\hat{\theta}$	$\mathrm{SE}(\hat{\theta})$	95% CI
$\hat{K}_{asy}$	17.233751	0.294422	(16.656683,17.810818)
$\hat{r}_{asy}$	0.676748	0.011875	(0.653473, 0.700024)
$\left(\hat{x}_0\right)_{asy}$	0.109710	0.005015	(0.099880, 0.119540)
$\hat{K}_{boot}$	17.241977	0.275328	(16.702335, 17.781619)
$\hat{r}_{boot}$	0.676694	0.011845	(0.653479, 0.699909)
$\left(\hat{x}_0\right)_{boot}$	0.109960	0.005031	(0.100098, 0.119821)

NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 8

Computation Times (sec) for Asymptotic Theory vs. Bootstrapping

Noise Level	Asymptotic Theory	Bootstrapping
1%	0.030065	16.869108
5%	0.032161	21.549255
10%	0.037183	24.530157

Table 9

Asymptotic and Bootstrap GLS Estimates for CV Data, *Nl* = 0.01

θ	$\hat{\theta}$	$SE(\hat{\theta})$	95% CI
$\hat{K}_{asy}$	17.473446	0.043866	(17.387468, 17.559424)
$\hat{r}_{asy}$	0.706720	0.001802	(0.703188, 0.710252)
$(\hat{x}_0)_{asy}$	0.096926	0.000662	(0.095628, 0.098224)
$\hat{K}_{boot}$	17.471946	0.041954	(17.389716, 17.554176)
$\hat{r}_{boot}$	0.706738	0.001766	(0.703277, 0.710199)
$(\hat{x}_0)_{boot}$	0.096932	0.000655	(0.095649, 0.098215)

Table 10

Asymptotic and Bootstrap GLS Estimates for CV Data, *Nl* = 0.05

θ	$\hat{\theta}$	$SE(\hat{\theta})$	95% CI
$\hat{K}_{asy}$	17.486405	0.169916	(17.153369, 17.819441)
$\hat{r}_{asy}$	0.696663	0.006939	(0.683063, 0.710263)
$(\hat{x}_0)_{asy}$	0.103291	0.002722	(0.097956, 0.108625)
$\hat{K}_{boot}$	17.477246	0.165693	(17.152487, 17.802004)
$\hat{r}_{boot}$	0.696828	0.006770	(0.683558, 0.710098)
$(\hat{x}_0)_{boot}$	0.103210	0.002548	(0.098215, 0.108205)

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 11

Asymptotic and Bootstrap GLS Estimates for CV Data, *Nl* = 0.1

θ	$\hat{\theta}$	$\mathrm{SE}(\hat{\theta})$	95% CI
$\hat{K}_{asy}$	17.648739	0.504870	(16.659193, 18.638285)
$\hat{r}_{asy}$	0.680706	0.019755	(0.641985, 0.719427)
$(\hat{x}_0)_{asy}$	0.106576	0.008143	(0.090616, 0.122537)
$\hat{K}_{boot}$	17.668510	0.490496	(16.707137, 18.629882)
$\hat{r}_{boot}$	0.681164	0.018922	(0.644077, 0.718251)
$(\hat{x}_0)_{boot}$	0.106784	0.007835	(0.091427, 0.122140)

Table 12

Computation Times (sec) for Asymptotic Theory vs. Bootstrapping

Noise Level	Asymptotic Theory	Bootstrapping
1%	0.021631	17.416247
5%	0.019438	21.073697
10%	0.044052	27.152972

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 13

Asymptotic and Bootstrap OLS Estimates for NCV Data, *Nl* = 0.01

θ	$\hat{\theta}$	$SE(\hat{\theta})$	95% CI
$\hat{K}_{asy}$	17.509101	0.024315	(17.461445, 17.556757)
$\hat{r}_{asy}$	0.706845	0.006750	(0.693615, 0.720076)
$(\hat{x}_0)_{asy}$	0.095928	0.004757	(0.086605, 0.105252)
$\hat{K}_{boot}$	17.511800	0.023465	(17.465808, 17.557791)
$\hat{r}_{boot}$	0.706690	0.006891	(0.693184, 0.720195)
$(\hat{x}_0)_{boot}$	0.096219	0.004871	(0.086671, 0.105767)

Table 14

Asymptotic and Bootstrap OLS Estimates for NCV Data, *Nl* = 0.05

θ	$\hat{\theta}$	$\mathrm{SE}(\hat{\theta})$	95% CI
$\hat{K}_{asy}$	17.301393	0.122316	(17.061653, 17.541133)
$\hat{r}_{asy}$	0.694052	0.033396	(0.628596, 0.759508)
$(\hat{x}_0)_{asy}$	0.105893	0.025879	(0.055171, 0.156615)
$\hat{K}_{boot}$	17.291459	0.118615	(17.058973, 17.523945)
$\hat{r}_{boot}$	0.697367	0.034670	(0.629413, 0.765320)
$(\hat{x}_0)_{boot}$	0.106239	0.026765	(0.053780, 0.158697)

Table 15

Asymptotic and Bootstrap OLS Estimates for NCV Data, *Nl* = 0.1

θ	$\hat{\theta}$	$SE(\hat{\theta})$	95% CI
$\hat{K}_{asy}$	17.081926	0.262907	(16.566629, 17.597223)
$\hat{r}_{asy}$	0.727602	0.078513	(0.573717, 0.881487)
$(\hat{x}_0)_{asy}$	0.082935	0.047591	(−0.010343, 0.176213)
$\hat{K}_{boot}$	17.095648	0.250940	(16.603807, 17.587490)
$\hat{r}_{boot}$	0.733657	0.081852	(0.573228, 0.894087)
$(\hat{x}_0)_{boot}$	0.094020	0.054849	(−0.013484, 0.201524)

Table 16

Computation Times (sec) for Asymptotic Theory vs. Bootstrapping

Noise Level	Asymptotic Theory	Bootstrapping
1%	0.008697	5.075417
5%	0.009440	6.002897
10%	0.013305	6.903042

Table 17

Average of 1000 Monte Carlo trials of Asymptotic Theory and Bootstrap (M = 250) Estimates, and Standard Errors, using GLS Estimates for NCV Data, Nl = 0.1 and n = 20.

θ	$\hat{\theta}_{avg}$	$\mathrm{SE}_{avg}(\hat{\theta})$
$\hat{K}_{asy}$	17.212625	0.533831
$\hat{r}_{asy}$	0.750631	0.021148
$(\hat{x}_0)_{asy}$	0.119780	0.008379
$\hat{K}_{boot}$	17.210893	0.556522
$\hat{r}_{boot}$	0.750690	0.022585
$(\hat{x}_0)_{boot}$	0.119753	0.008229

Table 18

Average of 1000 Monte Carlo trials of Asymptotic Theory and Bootstrap (M = 250) Estimates, and Standard Errors, using GLS Estimates for NCV Data, Nl = 0.1 and n = 25.

θ	$\hat{\theta}_{avg}$	$\mathrm{SE}_{avg}(\hat{\theta})$
$\hat{K}_{asy}$	17.233834	0.479995
$\hat{r}_{asy}$	0.750298	0.019315
$(\hat{x}_0)_{asy}$	0.120060	0.007727
$\hat{K}_{boot}$	17.234113	0.501584
$\hat{r}_{boot}$	0.750304	0.020572
$(\hat{x}_0)_{boot}$	0.120055	0.007571

Table 19

Average of 1000 Monte Carlo trials of Asymptotic Theory and Bootstrap (*M* = 250) Estimates, and Standard Errors, using GLS Estimates for NCV Data, *Nl* = 0.1 and *n* = 30.

θ	$\hat{\theta}_{avg}$	$\mathrm{SE}_{avg}(\hat{\theta})$
$\hat{K}_{asy}$	17.168470	0.436761
$\hat{r}_{asy}$	0.750720	0.017809
$(\hat{x}_0)_{asy}$	0.119995	0.007166
$\hat{K}_{boot}$	17.167208	0.456785
$\hat{r}_{boot}$	0.750792	0.018936
$(\hat{x}_0)_{boot}$	0.119964	0.007025

Table 20

Average of 1000 Monte Carlo trials of Asymptotic Theory and Bootstrap (M = 250) Estimates, and Standard Errors, using GLS Estimates for NCV Data, Nl = 0.1 and n = 35.

θ	$\hat{\theta}_{avg}$	$\mathrm{SE}_{avg}(\hat{\theta})$
$\hat{K}_{asy}$	17.205855	0.403379
$\hat{r}_{asy}$	0.750168	0.016522
$(\hat{x}_0)_{asy}$	0.119871	0.006678
$\hat{K}_{boot}$	17.206793	0.423147
$\hat{r}_{boot}$	0.750167	0.017589
$(\hat{x}_0)_{boot}$	0.119886	0.006548

Table 21

Average of 1000 Monte Carlo trials of Asymptotic Theory and Bootstrap (M = 250) Estimates, and Standard Errors, using GLS Estimates for NCV Data, Nl = 0.1 and n = 40.

θ	$\hat{\theta}_{avg}$	$\mathrm{SE}_{avg}(\hat{\theta})$
$\hat{K}_{asy}$	17.216539	0.379682
$\hat{r}_{asy}$	0.750290	0.015638
$(\hat{x}_0)_{asy}$	0.1198878	0.006340
$\hat{K}_{boot}$	17.215171	0.398576
$\hat{r}_{boot}$	0.750355	0.016641
$(\hat{x}_0)_{boot}$	0.119875	0.006212

Table 22

Average of 1000 Monte Carlo trials of Asymptotic Theory and Bootstrap (M = 250) Estimates, and Standard Errors, using GLS Estimates for NCV Data, Nl = 0.1 and n = 45.

θ	$\hat{\theta}_{avg}$	$\mathrm{SE}_{avg}(\hat{\theta})$
$\hat{K}_{asy}$	17.220679	0.359348
$\hat{r}_{asy}$	0.750266	0.014858
$(\hat{x}_0)_{asy}$	0.119829	0.006038
$\hat{K}_{boot}$	17.219544	0.377269
$\hat{r}_{boot}$	0.750351	0.015801
$(\hat{x}_0)_{boot}$	0.119809	0.005919

NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 23

Average of 1000 Monte Carlo trials of Asymptotic Theory and Bootstrap (M = 250) Estimates, and Standard Errors, using GLS Estimates for NCV Data, Nl = 0.1 and n = 50.

θ	$\hat{\theta}_{avg}$	$\mathrm{SE}_{avg}(\hat{\theta})$
$\hat{K}_{asy}$	17.187670	0.337956
$\hat{r}_{asy}$	0.749043	0.014056
$(\hat{x}_0)_{asy}$	0.120329	0.005734
$\hat{K}_{boot}$	17.187487	0.3539129
$\hat{r}_{boot}$	0.749089	0.014936
$(\hat{x}_0)_{boot}$	0.120308	0.005619

Table 24

Average Correction term: $\mathrm{SE}_{avg}\left(\widehat{\theta}_{\mathrm{boot}}\right)-\mathrm{SE}_{avg}\left(\widehat{\theta}_{asy}\right)$, from the Average Standard Errors of 1000 Monte Carlo trials of Asymptotic Theory and Bootstrap method (M = 250) using GLS Estimates for NCV Data, Nl = 0.1, for θ = {K, r}, and n time points.

n	$\mathrm{SE}_{avg}(\hat{K}_{boot})-\mathrm{SE}_{avg}(\hat{K}_{asy})$	$\mathrm{SE}_{avg}(\hat{r}_{boot})-\mathrm{SE}_{avg}(\hat{r}_{asy})$
20	0.022691	0.001437
25	0.021588	0.001257
30	0.020024	0.001127
35	0.019769	0.001067
40	0.018893	0.001003
45	0.017921	0.000943
50	0.015957	0.000880